diversity

MDPI

Article

Characterization and Phylogenetic Analyses of the Complete Mitochondrial Genome of Sugarcane (*Saccharum* spp. Hybrids) Line A1

Dinggang Zhou [1,2,3,*], Ying Liu [1], Jingzuo Yao [1], Ze Yin [1], Xinwen Wang [1], Liping Xu [2], Youxiong Que [2], Ping Mo [4] and Xiaolan Liu [1,*]

1 Hunan Key Laboratory of Economic Crops Genetic Improvement and Integrated Utilization, School of Life Science, Hunan University of Science and Technology, Xiangtan 411201, China; ly1071457534@163.com (Y.L.); y412625112@163.com (J.Y.); yz18390246279@163.com (Z.Y.); wxw010521@163.com (X.W.)
2 Key Laboratory of Sugarcane Biology and Genetic Breeding, Ministry of Agriculture/National Engineering Research Center for Sugarcane, Ministry of Science and Technology, Fujian Agriculture and Forestry University, Fuzhou 350002, China; xlpmail@126.com (L.X.); queyouxiong@126.com (Y.Q.)
3 Key Laboratory of Ecological Remediation and Safe Utilization of Heavy Metal-polluted Soils, College of Hunan Province, Xiangtan 411201, China
4 College of Life and Environmental Sciences, Hunan University of Arts and Science, Changde 415000, China; moping2015@126.com
* Correspondence: dgzhoucn@hnust.edu.cn (D.Z.); liuxiaolan1018@126.com (X.L.)

Abstract: Modern sugarcane cultivars are highly polyploid with complex nuclear genomic genetic background, while their mitochondrion (mt) genomes are much simpler, smaller and more manageable and could provide useful phylogenetic information. In this study, the mt genome of a modern commercial cultivar A1 was sequenced via Illumina Hiseq XTen and PacBio Sequel platform. The assembled and annotated mitochondrial genomes of A1 were composed of two circular DNA molecules, one large and one small, which were named Chromosome 1 and Chromosome 2. The two distinct circular chromosomes of mitogenome construct is consisted with other sugarcane cultivars i.e., *Saccharum officinarum* Khon Kaen 3 and *Saccharum* spp. hybrids ROC22 and FN15. The Chromosome 1 of A1 mitogenome is 300,822 bp in length with the GC content of 43.94%, and 7.14% of Chromosome 1 sequences (21,468 nucleotides) are protein coding genes (PCGs) while 92.86% (279,354 nucleotides) are intergenic region. The length of Chromosome 2 is 144,744 bp with the GC content of 43.57%, and 8.20% of Chromosome 2 sequences (11,865 nucleotides) are PCGs while 91.80% (132,879 nucleotides) are intergenic region. A total of 43 genes are located on Chromosome 1, which contains 22 PCGs (six *nad* genes, four *rps* genes, four *atp* genes, three *ccm* genes, three *cox* genes, one *mat* gene and one *mtt* gene) and 21 non-coding genes including 15 tRNAs and 6 rRNAs. Chromosome 2 includes 18 genes in total, which contains 13 PCGs (four *nad* genes, three *rps* genes, two *atp* genes, one *ccm* gene, one *cob* gene, one *cox* gene and one *rpl* gene) and five non-coding genes (tRNA genes). Analysis of codon usage of 35 PCGs showed that codon ending in A/U was preferred. Investigation of gene composition indicated that the types and copy numbers of CDS genes, tRNAs and rRNAs of A1 and FN15 were identical. The *cox1* gene has two copies and the *trnP* gene has one copy in A1, FN15 and ROC22 three lines, while there is only one copy of *cox1* and two copies of *trnP* in *S. officinarum* Khon Kaen 3. In addition, *S. officinarum* Khon Kaen 3 have no *nad1* gene and *rps7* gene. 100 sequence repeats, 38 SSRs and 444 RNA editing sites in A1 mt genome were detected. Moreover, the maximum likelihood phylogenetic analysis found that A1 were more closely related to *S.* spp. hybrid (ROC22 and FN15) and *S. officinarum* (Khon Kaen 3). Herein, the complete mt genome of A1 will provide essential DNA molecular information for further phylogenetic and evolutionary analysis for *Saccharum* and Poaceae.

Keywords: sugarcane; *Saccharum* spp. hybrid; mitogenome; genomic structure; comparative analysis; phylogenetic tree

Citation: Zhou, D.; Liu, Y.; Yao, J.; Yin, Z.; Wang, X.; Xu, L.; Que, Y.; Mo, P.; Liu, X. Characterization and Phylogenetic Analyses of the Complete Mitochondrial Genome of Sugarcane (*Saccharum* spp. Hybrids) Line A1. *Diversity* **2022**, *14*, 333. https://doi.org/10.3390/d14050333

Academic Editor: Mario A. Pagnotta

Received: 22 March 2022
Accepted: 22 April 2022
Published: 25 April 2022

1. Introduction

Mitochondria (mt) are vital organelles in cell, which are involved in a large number of metabolic processes associated with energy production and the synthesis of some compounds [1,2]. Mitochondria are known as the energy factories which are major sites of ATP synthesis by oxidative phosphorylation [3]. Mitochondria are semiautonomous organelle due to they have their own genome and protein-synthesizing mechanism, and thus, mitochondria are involved in multitudinous other living processes, such as the biosynthesis of amino acids, vitamin cofactors and fatty acids [3].

Though plants can obtain energy from sunlight by the chloroplasts, plant mitochondria remain the energy centers of their metabolism through plant cell respiration [4]. Plant mitochondria are typically 1–3 μm in length and ~0.5 μm in diameter, meaning that they are the same size and shape as some common bacteria because they are developed from endosymbiotic bacteria [4]. Mitochondrial has long been thought to tend to integrate DNA from different sources, especially bacterial genomic DNAs, through intracellular and horizontal transfer [5]. It is difficult to observe by light microscopy due to tiny genome size and this postponed their recognition [4]. In addition, it is difficult to purify plant mitochondria and its purification is often interfered with by chloroplasts and other plastids, which also leads to a lag in the study of mitochondrial genome compared with that of animals.

Plant mitochondrial genomes range from 42 Kb (*Mesostigma viride*) to 11.3 Mb (*Silene conica*) in size [6,7]. The mt genomes of plants have distinct differences in length, GC content, sequence and gene content and could contain potential phylogenetic messages [8]. Moreover, the mt genome exhibits many features, such as a faster frequency of evolution than nuclear DNA, the characteristics of conservation gene functions and high AT content [9]. The mt genome is typically inherited maternally, and the maternal inheritance limits outcrossing recombination [10]. Thus, the mt genome sequences usually were employed to phylogenetic and evolutionary analysis and retention of maternal inheritance in crossbreeding [11,12].

Sugarcane, which ranks amongst the top ten crops worldwide, is the most important crop for sugar and biofuel [13,14]. It accounts for more than 80% of sugar in the world, approximately 90% of sugar in China, and about 60% of ethanol global production [15,16]. Modern sugarcane cultivars are complex hybrids at least from the three species: *Saccharum spontaneum*, *Saccharum robustum* and *Saccharum officinarum* [17]. *S. officinarum* (called "noble cane") has much sugar content and had the largest cultivated area in the world before the 1920s [18]. However, they were susceptible to disease and insects. To introduce high resistance characters and reserve high sugar content trait, noble canes were constantly crossbreeding with other Saccharum [19,20] such as *S. spontaneum*. Hence, *S. officinarum* contributes to high sugar content, and *S. spontaneum* provides disease resistance, stress resistance and ratooning ability [21]. In other words, modern sugarcane varieties have an extremely complex interspecific polyploid genome and the complete genome is still remaining to be explored, whether nuclear genome or organelle genome (chloroplast and mitochondrial genome) [22]. At present, the whole genome of AP85-441, which belongs to *S. spontaneum* (allele-defined genome of tetraploid), was assembled by Zhang et al. in 2018 [20]. Modern sugarcane cultivar R570 based on bacterial artificial chromosome (BAC) clones have been assembled [23]. The complete chloroplast genome of two sugarcane ancestors *S. officinarum* and *S. spontaneum* were assembled and analyzed [11]. With the development of next-generation sequencing technologies (NGS), a growing number of mt genomes have been assembled. At present, more than 300 complete mt genomes have been submitted to GenBank Organelle Genome Resources. The complete mitochondrial genome of modern commercial sugarcane cultivars such as *S.* spp. hybrid (ROC22 and FN15) and *S. officinarum* Khon Kaen 3 were obtained and analyzed [15,24,25].

In the present study, to better comprehend the evolutionary information of modern sugarcane, the complete mitochondrial genome of *S.* spp. hybrid line A1 were sequenced and assembled using a combination of Illumina Hiseq XTen and PacBio Sequel platform.

The mitogenome characteristics, codon usage bias, repetitive element during genome, the prediction of RNA editing and phylogenetic relationship with in Poaceae were investigated. The mitochondrial genome herein will facilitate the sugarcane breeding, molecular marker selection and further phylogenetic and evolutionary analysis.

2. Materials and Methods

2.1. Plant Sampling, Mitochondria DNA Sequencing and Genome Assembly

Sugarcane line A1 were provided by the Fujian Agriculture and Forestry University (geographic coordinates: 26°9′8″ N, 119°24′24″ E), Fujian, Fuzhou, China. The specimen of A1 was stored in the Key Laboratory of Sugarcane Biology and Genetic Breeding, Fujian Agriculture and Forestry University with store number A1-FJ2016004. A1 were nursed in the greenhouse at 32 °C. After one week of dark culture, fresh yellowing seedlings were harvested. The A1 fresh yellowing seedlings were frozen in liquid nitrogen immediately after sampling and stored at −80 °C until mtDNA extraction. MtDNA was extracted and purified by an improved protocol as described by Chen et al. [26].

After mtDNA isolation, one microgramme of purified DNA was fragmented to construct short-insert libraries (insert size 430 bp) according to the manufacturer's instructions (Illumina, San Diego, CA, USA), then sequenced on the Illumina Hiseq XTen and PacBio Sequel platform by Shanghai BIOZERON Co., Ltd. (Shanghai, China). The raw sequenced reads from Illumina Hiseq were filtered firstly [27]. Then the mitochondria genome was reconstructed and assembled by SPAdes v3.10.1 using a combination of both Illumina and Pacbio data [28].

2.2. Genome Annotation and Sequence Analyses

The mitochondria genome was annotated using GeSeq. Transfer RNA (tRNA) genes and Ribosome RNA (rRNA) genes were predicted by tRNAscan-SE and rRNAmmer 1.2 [29,30]. A whole mitochondria genome amino acid sequences BlastP search (E-value $\leq 1 \times 10^{-5}$, identity $\geq 40\%$) was performed against KEGG, COG, NR, Swiss-Prot, and GO databases [31–35]. The circular A1 mitochondria genome map was drawn using OrganellarGenomeDRAW v1.2 [36]. The codon usage frequency was calculated using the codonW 1.4.2.

2.3. Repetitive Elements Analysis

REPuter online software (https://bibiserv.cebitec.uni-bielefeld.de/reputer, accessed on 2 November 2021) were used to detect forward repetitive sequences (F), palindromic repetitive sequence (P), complement repeat sequences (C), and reverse repeat sequences (R) with the default settings (maximum computed repeats: 50; minimal repeat size 8; hamming distance: 3) [37]. Simple repeat sequences(SSR), including mononucleotide SSR, dinucleotide SSR, trinucleotide SSR, tetranucleotide SSR, pentanucleotide SSR and hexanucleotide SSR, were searched in sugarcane mitochondrial genomes by using MISA software (https://webblast.ipk-gatersleben.de/misa/, accessed on 2 November 2021), with the default settings (The repeats of 1, 2, 3, 4, 5 and 6 bases with 10, 6, 5, 5, 5 and 5 repeat numbers, respectively.) [38].

2.4. RNA Editing Analyses

Based on the proverbial principle of calculating predictive RNA editing sites, where editing in plant organelles increases protein preservation across species, PREP suite software (http://prep.unl.edu/, accessed on 4 November 2021) with a cut off value of 0.2 was performed [39].

2.5. Phylogenetic Analysis

To better understand the phylogenetic relationship within the Poaceae family, 11 species (two outgroups and nine Poaceae species) were selected to construct a phylogenetic tree. The whole mitchondrial genome DNA sequences of *S.* spp. line A1 (GenBank: MT921804 and MT921805), *S.* spp. hybrid ROC22 (GenBank: MT921808 and MT921809),

S. spp. hybrid FN15 (GenBank: MT411890 and MT411891), *Saccharum officinarum* Khon Kaen 3 (GenBank: LC107874.1 and LC107875.1), *Oryza sativa* (GenBank: NC_011033.1), *Zea mays* (GenBank: NC_007982), *Sorghum bicolor* (GenBank: NC_008360.1), *Triticum aestivum* (GenBank: NC_036024.1), *Arabidopsis thaliana* (GenBank: NC_037304.1), *Brassica napus* (GenBank: NC_008285.1) and *Hordeum vulgare* (GenBank: MN127982.1) were compared and a Maximum Likelihood phylogenetic tree was constructed using PhyML v3.0 (http://www.atgc-montpellier.fr/phyml/, accessed on 5 November 2022) with default parameters (1000 bootstrap replications; HKY85 model; optimize tree topology) [40].

3. Results

3.1. Characteristics of the Mitogenome of Sugarcane Line A1

Mitochondrial genomes of sugarcane line A1 were sequenced via Illumina sequencing (7099 Mb raw data, Q20 = 97.28%) and PacBio sequencing (188,008 subreads, N50 = 1553 bp). The assembled and annotated mitochondrial genomes of sugarcane line A1 were composed of two circular DNA molecules, one large and one small, which were named Chromosome 1 and Chromosome 2 (Chr1 and Chr2). The Chr1 is 300,822 bp in length with the GC content of 43.94%, and 7.14% of Chr1 genome sequences (21,468 nucleotides) are PCGs (protein-coding genes) while 92.86% (279,354 nucleotides) are intergenic region. The length of Chr2 is 144,744 bp with the GC content of 43.57%, and 8.20% of Chromosome 2 sequences (11,865 nucleotides) are PCGs while 91.80% (132,879 nucleotides) are intergenic region (Table S1).

A total of 43 genes are located on Chromosome 1, which contains 22 PCGs (six *nad* genes, four *rps* genes, four *atp* genes, three *ccm* genes, three *cox* genes, one *mat* gene and one *mtt* gene) and 21 non-coding genes including 15 tRNAs and 6 rRNAs. Chromosome 2 includes 18 genes in total, which contains 13 PCGs (four *nad* genes, three *rps* genes, two *atp* genes, one *ccm* gene, one *cob* gene, one *cox* gene and one *rpl* gene) and five non-coding genes (tRNA genes) (Figure 1).

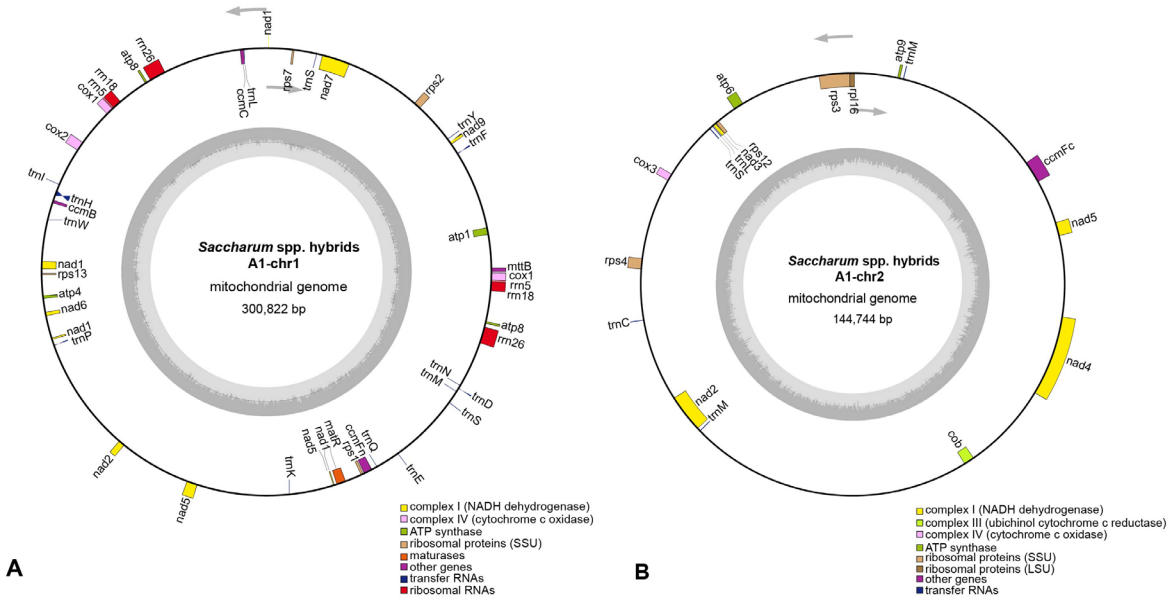

Figure 1. Circular maps of mitochondrial genome of *Saccharum* spp. hybrids line A1. (**A**) Chromosome 1 of *Saccharum* spp. hybrids line A1 mitochondrial genome; (**B**) Chromosome 2 of *Saccharum* spp. hybrids line A1 mitochondrial genome. Genes are represented by blocks of different colors. The genes outside the ring are located on the direct strand, while the genes inside the ring are located on the reverse strand. The gray color of inner circle represented the GC content of mt genome.

We further investigated the differences of gene composition between sugarcane line A1 and other *Saccharum* complexes (*S.* spp. hybrid ROC22, *S.* spp. hybrid FN15 and *S. officinarum* Khon Kaen 3) (Table S2). The types and copy numbers of CDS genes, tRNAs

and rRNAs of A1 and FN15 were identical. The *cox1* gene has two copies and the *trnP* gene has one copy in A1, FN15 and ROC22, while there is only one copy of *cox1* and two copies of *trnP* in *S. officinarum* Khon Kaen 3. Moreover, *S. officinarum* Khon Kaen 3 have no *nad1* gene and *rps7* gene.

In Chromosome 1, almost all the PCGs use ATG as the initiation codon, while gene *nad1* starts with ACG and *matR* begins with ATA. In Chromosome 2, however, all the PCGs use the initiation codon ATG except for *nad2* (TTG) and *nad5* (CCA). Regarding the stop codon, whether in Chromosome 1 or Chromosome 2, most of the PCGs stop with TAA, TAG and TGA, except for gene nad2 and nad5 in Chromosome 1 terminate with CGG and GTA, respectively (Table S3).

3.2. Codon Usage Bias

We analyzed the codon usage frequency of the A1 mitochondrial genome. A total of 11,320 codons were calculated, which included 20 amino acids and three stop codons, indicating a stronger coding capacity (Figure 2 and Table S4). The most abundant amino acid is leucine (Leu) and the next is isoleucine (Ile), with the number of 1246 and 1052 codons, respectively. The least amount is cysteine (Cys) with only 313 codons. Both methionine (AUG) and tryptophan (UGG) had only one codon type with number of 288 and 198, respectively, and showed no bias (RSCU = 1.00). The codon CAA for glutamine with the highest RSCU (relative synonymous codon usage) values (1.48). While the codon CGC for arginine with the minimal RSCU values (0.46). Spectacularly, all codons whose RSCU values were higher than one ended with U or A, except for AUC-isoleucine, UUG-leucine, UCC-serine, and UAG-termination codon.

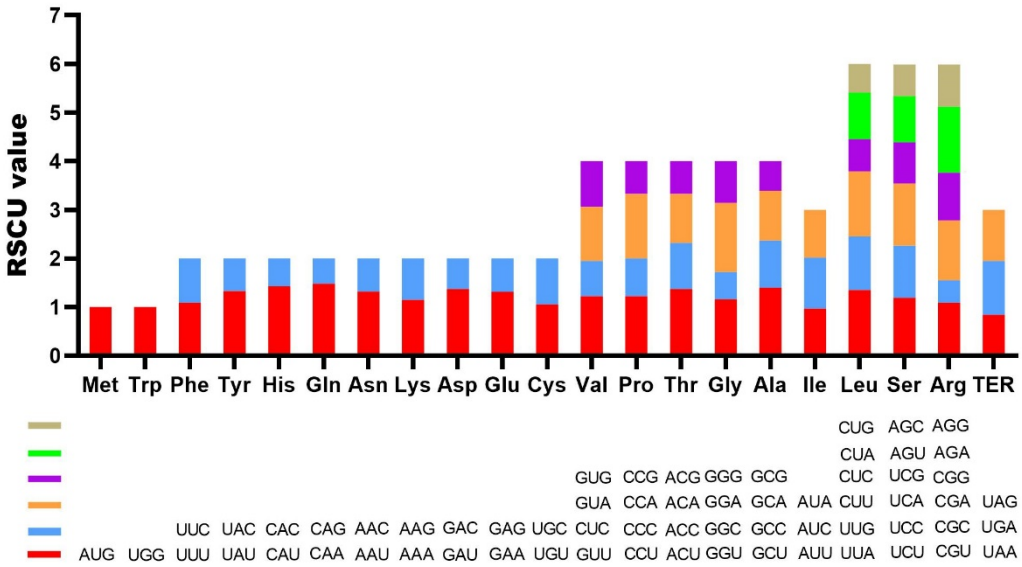

Figure 2. Codon usage analysis of the mitochondrial genome from sugarcane line A1. Bar diagram in different colors and sizes reflects codon usage bias and RSCU value, respectively.

3.3. Repetitive Element and Simple Sequence Repeat (SSR) Analysis

A total of 100 repetitive sequences of three types were detected in the A1 mitochondrial genome. F (forward repeat), P (palindromic repeat) and R (reverse repeat) accounted for 64%, 19% and 17%, respectively (Table S5). There were no complement repeat sequences, while reverse repeat sequences only existed on chromosome 2. In addition, the longest forward repeat was 12,241 bp and the longest palindromic repeat was 4058 bp. It also showed that the 30–90 bp repeats are most common in the two chromosomes.

A total of 38 SSRs were found in the mt genome of A1, including five types of SSR, and most were located in intergenic regions (Table 1). The mono-nucleotide repeats were the most common (27, 71.05%), followed by penta-nucleotide (four, 10.54%) and compound-

nucleotide (three, 7.89%). Two additional types of SSRs were less abundant: di-nucleotide (two, 5.26%) and tri-nucleotide (two, 5.26%). In addition, A/T-containing motifs are most frequent among the SSRs, accounting for 76.3%. This phenomenon is consistent with the previous study in plant mitochondrial genomes [39].

Table 1. Distribution of SSRs in mitochondrial genome from sugarcane line A1.

ID	SSR Type	SSR	Start	End	Location
A1-chr1	p1	(T)10	21,434	21,443	IGS
A1-chr1	p5	(ATAGA)12	28,425	28,484	IGS
A1-chr1	p1	(T)10	42,775	42,784	IGS
A1-chr1	p3	(CTA)5	65,115	65,129	IGS
A1-chr1	p1	(T)12	65,595	65,606	IGS
A1-chr1	p1	(T)10	74,767	74,776	*nad1*
A1-chr1	p3	(ATA)5	76,342	76,356	IGS
A1-chr1	p1	(T)14	138,736	138,749	IGS
A1-chr1	c	(T)11acttattaaattctctgtcttgctaaa-cacaaatccttctttttcttgtatagacg(A)11	145,252	145,328	IGS
A1-chr1	p1	(A)11	153,201	153,211	IGS
A1-chr1	p1	(T)10	166,287	166,296	IGS
A1-chr1	p1	(A)13	178,747	178,759	IGS
A1-chr1	p1	(T)10	184,327	184,336	IGS
A1-chr1	p1	(T)10	208,192	208,201	IGS
A1-chr1	p2	(GA)6	220,084	220,095	IGS
A1-chr1	p1	(A)10	221,396	221,405	IGS
A1-chr1	p1	(A)11	245,842	245,852	IGS
A1-chr1	p1	(T)13	246,511	246,523	*rps1*
A1-chr1	p1	(A)10	256,149	256,158	IGS
A1-chr1	p1	(A)12	262,049	262,060	IGS
A1-chr1	p1	(T)10	280,624	280,633	IGS
A1-chr2	p1	(G)10	12,943	12,952	IGS
A1-chr2	p1	(T)10	26,230	26,239	IGS
A1-chr2	p1	(T)10	30,036	30,045	IGS
A1-chr2	p1	(T)10	34,409	34,418	IGS
A1-chr2	p1	(A)10	37,040	37,049	*rps3*
A1-chr2	p1	(T)11	61,305	61,315	IGS
A1-chr2	p1	(A)10	61,669	61,678	IGS
A1-chr2	p5	(TAATA)14	67,753	67,822	IGS
A1-chr2	p1	(T)12	80,494	80,505	IGS
A1-chr2	c	(T)10ctctccta(T)10	81,051	81,078	IGS
A1-chr2	c	(T)10agttcgcactgctctttctctct-aaattgcatcaaagaaaat(AG)6	88,702	88,765	IGS
A1-chr2	p5	(ATAGA)20	90,935	91,034	IGS
A1-chr2	p2	(AT)7	105,454	105,467	IGS
A1-chr2	p5	(GTATA)9	108,891	108,935	IGS
A1-chr2	p1	(T)10	109,400	109,409	IGS
A1-chr2	p1	(A)10	122,745	122,754	IGS
A1-chr2	p1	(T)10	140,466	140,475	IGS

Note: "p" represents the SSR type of pure. "c" represents the SSR type of compound. "IGS" means intergenic regions.

3.4. The Prediction of RNA Editing

RNA editing widely exists in eukaryotes, including higher plants. In mitochondrion, the conversion of specific cytosine into uridine changes the genomic information. In this analysis, the software PREP was employed to predict the RNA edit site, 444 RNA editing sites within 35 PCGs were predicted in the two chromosomes of A1, using the PREP-MT program (Table 2 and Figure 3). Among these PCGs, *nad5* have no possible editing site, while *ccmC* (36) has the maximum editing sites on Chromosome 1. In addition, 35.36% (157) were located at the first position of the triplet codes, 61.49% (273) occurred with the second base of the triplet codes. There were two particular editing cases in which the first and

second positions of the triplet codes were all edited, resulting in two amino acids changing from the proline (CCT, CCC) to phenylalanine (TTT, TTC).

Table 2. Prediction of RNA editing sites.

Type	RNA-Editing	Number	Percentage
hydrophobic	CCA (P) = > CTA (L)	38	30.85%
	CCG (P) = > CTG (L)	19	
	CCT (P) = > CTT (L)	24	
	CCT (P) = > TTT (F)	6	
	CCC (P) = > TTC (F)	8	
	GCC (A) = > GTC (V)	3	
	GCG (A) = > GTG (V)	2	
	GCT (A) = > GTT (V)	3	
	GCA (A) = > GTA (V)	2	
	CTT (L) = > TTT (F)	22	
	CTC (L) = > TTC (F)	10	
hydrophilic	CAT (H) = > TAT (Y)	16	11.49%
	CAC (H) = > TAC (Y)	5	
	CGT (R) = > TGT (C)	25	
	CGC (R) = > TGC (C)	5	
hydrophobic-hydrophilic	CCG (P) = > TCG (S)	9	12.16%
	CCT (P) = > TCT (S)	12	
	CCA (P) = > TCA (S)	8	
	CCC (P) = > CTC (L)	10	
	CCC (P) = > TCC (S)	15	
hydrophilic-hydrophobic	CGG (R) = > TGG (W)	28	45.05%
	TCC (S) = > TTC (F)	31	
	TCT (S) = > TTT (F)	32	
	TCA (S) = > TTA (L)	51	
	TCG (S) = > TTG (L)	38	
	ACC (T) = > ATC (I)	2	
	ACT (T) = > ATT (I)	7	
	ACA (T) = > ATA (I)	7	
	ACG (T) = > ATG (M)	4	
hydrophilic-stop	CGA (R) = > TGA (X)	1	0.45%
	CAG (Q) = > TAG (X)	1	

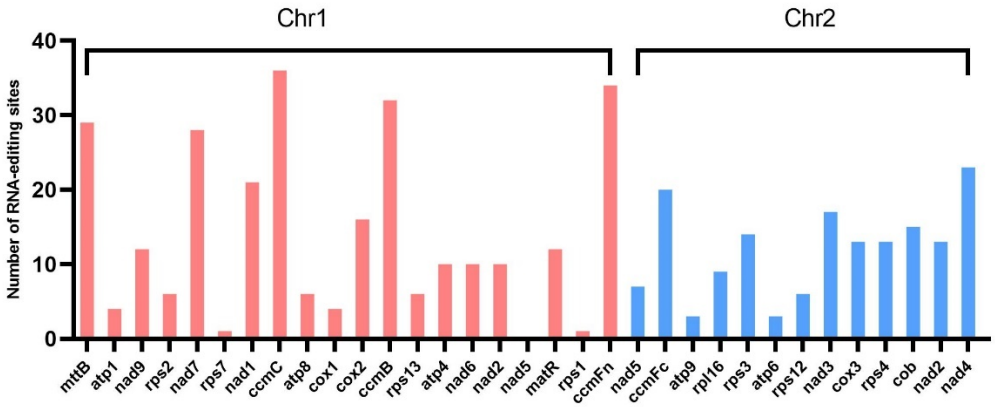

Figure 3. Number of RNA-editing sites. *atp8* and *cox1* have two copies in Chromosome 1.

After the RNA editing, 42.34% of amino acids did not change hydrophobicity or hydrophilicity. The proportion of amino acids that changed from hydrophilic to hydrophobic was 45.05%, while the proportion of amino acids that changed from hydrophobic to hydrophilic was 2.16%. In addition, the amino acid of predicted editing codons showed a leucine (170 sites) bias after RNA editing.

3.5. Phylogenetic Analysis

To understand the genetic status of A1, the phylogenetic tree was performed based on the whole mtDNAs of 11 species (nine Poaceae species and two outgroups) using PhyML v3.0 (Figure 4). The phylogenetic tree showed that *S.* spp. hybrid A1 is very close to *S.* spp. hybrid (FN15 and ROC22) and *S. officinarum* (Khon Kaen 3). The complete mitochondrial genome herein will provide important and fundamental DNA molecular information of evolutionary analysis for *Saccharum* and Poaceae.

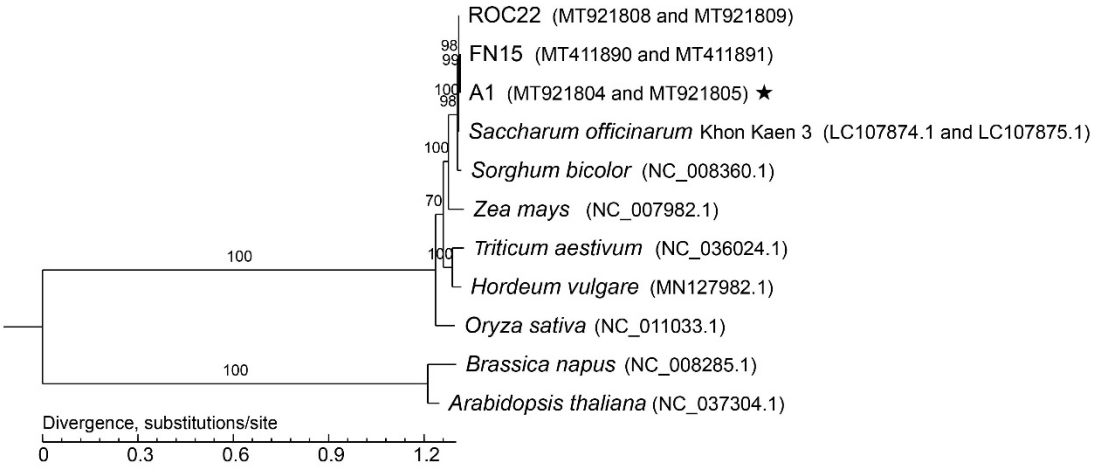

Figure 4. Maximum likelihood tree based on the complete mt genome sequences of 11 species. The bootstrap values are located on the branches. The number is the GenBank accession number after the species name. Pentagram, the mt genome sequences of A1 in this study.

4. Discussion

Mitochondria produce the energy for life processes and are named as the powerhouses. Plant mitochondria have more complicated genomes than animals with a wide range of size variations, sequence arrangement, and a dynamic structure with various conformations [41,42]. In this study, after the algorithm predicts, the mt genome of a modern cultivated sugarcane line A1 was assembled into two distinct circular chromosomes. According to the published data, most mt genomes are single circular, while the mt genome of sugarcane has two circular chromosomes. For example, *S.* spp. hybrid ROC22, *S.* spp. hybrid FN15 and *Saccharum officinarum* Khon Kaen 3 all have two distinct circular chromosomes [15,24,25]. In addition, in some species like *Cucumis sativus* [43], *Glycine max* [44] and *Allium cepa* [45], mitochondria also consist of two or more circular chromosomes. However, there are no detailed studies to elucidate why the mitochondrial genome of sugarcane has two chromosomes.

Codons play a vital role in the course of transformation of genetic information. There are 11,320 codons in A1 mt genome. The most and least abundant amino acids are leucine and cysteine. This phenomenon is also found in *Mangifera* mitochondrial genomes [46]. In addition, codon usage bias is universal in plant species. In this study, codon usage bias was calculated by the RSCU value. The results suggested an intense A or U bias in the third position of the codon in the PCGs and were different with *Mangifera* mitochondrial genomes and *Solanales* (A or T bias in the third position), indicating that different plants have formed different codon predilection during the long periods of evolution [46,47].

The repeat sequences diffusely exist in the mt genome. The majority of differences in the mt genome size can be elucidated by distinction in the size of the repeat sequences in plants [48]. Previous studies also have shown that repeats in mitochondria are connected with rearrangement and recombination of the mt genome [49]. In this study, there are numerous repeats in sugarcane mitochondrial genome, which may imply that intermolecular recombination occurs hourly in sugarcane mitochondrial genome. Moreover, a total of 38 SSRs were found in the mt genome of A1, and most were located in IGS, which is

similar to *Suaeda glauca* mt genome [50]. Single base repeats were the most, accounting for 76.3%, which was different with the sunflower. While the result of A/T-containing motifs are most frequent among the SSRs was consistent with sunflower [51]. RNA-editing is a post-transcriptional step that exists in the chloroplast and mitochondrial genome of plants, conducing to the proteins fold better [52]. The investigation of RNA editing sites contributes to understanding the expression of CP and MT genes in plants. In rice mitochondria, 491 editing sites have been identified [53]. In this study, 444 RNA-editing sites within 35 genes were identified, which can provide available information for forecasting gene functions with novel codons.

We also made a series of comparisons with related cultivars (FN15, ROC22, and Khon Kaen 3), including chromosomes structure, gene composition, codon usage and repetitive elements. The results showed that the differences between them were extremely small. Firstly, A1, FN15, ROC22, and Khon Kaen 3 four lines all have two distinct circular chromosomes (a larger one and a smaller one). Secondly, the gene composition differences of them were tiny. The CDS genes compositions of A1, FN15 and ROC22 were coincident, while Khon Kaen 3 has no *nad1* gene and *rps7* gene and only one copy of *cox1* gene. Four lines all have two *5S rRNA*, two *18S rRNA* and two *26S rRNA*. In addition, the tRNA compositions of them are identical except the *trnP*, A1 and FN15 only have one copy of *trmP*, while ROC22 and Khon Kaen 3 have two copies (Table S2). Thirdly, the CDS genes of A1, FN15 and ROC22 were identical, so the codon usage analysis of them have no difference, while Khon Kaen 3 have differences caused by the three genes lost (Table S4). Fourthly, repetitive element analysis showed little difference between the A1, FN15, and ROC22. Forward repeat, Palindromic repeat and Reverse repeat account for 64, 19 and 17, respectively in line A1 and FN15, while ROC22 have 63 Forward repeats, 19 Palindromic repeats and 16 Reverse repeats, Khon Kaen 3 have 66 Forward repeats, 17 Palindromic repeats and 17 Reverse repeats (Table S5).

Plant mitochondria genomes are usually dynamic, resulting in heterogeneity, large-scale genomic reorganization, and gene mosaicism in the mitochondrial genomes of various species [54,55]. Size and structural changes of plant mt genomes are individual. Here, the whole genome sequences are used to construct phylogenetic tree to explore the lineage between Saccharum and Gramineae species. The result showed the phylogenetic relationship of *S.* spp. hybrid A1 is very close to FN15, ROC22 and Khon Kaen 3.

5. Conclusions

Here, the mt genome of a modern commercial cultivar A1 was sequenced, assembled, and annotated. Interestingly, A1 mitogenome contains two distinct circular chromosomes, one large and one small, which were named Chromosome 1 and Chromosome 2. The two distinct circular chromosomes of mitogenome construct is consisted with other sugarcane cultivars i.e., *Saccharum officinarum* Khon Kaen 3 and *Saccharum* spp. hybrids ROC22 and FN15. Investigation of gene composition indicated that the CDS gene contents of A1 mt genome were identical with ROC22 and FN15, while Khon Kaen 3 lacks three CDS genes (*cox1*, *nad1* and *rps7*). Analysis of codon usage of 35 PCGs showed that codon ending in A/U was preferred. RNA editing sites and SSRs analysis showed that 100 sequence repeats, 38 SSRs and 444 RNA editing sites in A1 mt genome were detected. Maximum likelihood phylogenetic analysis found that A1 were more closely related to *S.* spp. hybrid ROC22, *S.* spp. hybrid FN15 and *S. officinarum* Khon Kaen 3. The complete mitochondrial genome herein will provide essential genetic resources for further phylogenetic and evolutionary analysis for Saccharum and Poaceae.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/d14050333/s1; Table S1: Characteristics of mitochondrial genome of sugarcane line A1; Table S2: The genes composition of A1 and related species. Table S3: Initiation codon and termination codon in mitochondrial genome of sugarcane line A1; Table S4: Codon usage analysis of the A1 mitochondrial genome and related species; Table S5: The distribution of repeats in the A1 mt genome and related species.

Author Contributions: Conceptualization, Y.L., L.X. and D.Z.; methodology, Y.Q.; software, Y.L., J.Y., X.W., P.M. and Z.Y.; writing—original draft preparation, Y.L., X.L. and D.Z.; writing—review and editing, D.Z., L.X., Y.Q. and X.L.; project administration, L.X., X.L., Y.Q. and D.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by "National Natural Science Foundation of China, grant number U19A2029 and 31701491", "the open funding of National Engineering Research Center for Sugarcane, Fujian Agriculture and Forestry University, grant number 2017.6.2", "the Open Funding of Hunan Key Laboratory of Economic Crops Genetic Improvement and Integrated Utilization, grant number E22009 and E22011" and "National College Student Innovation and Entrepreneurship Training Project, grant number 202110534015)".

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The original contributions presented in the present study are publicly available. The datasets generated for this study can be found in GenBank, the accession number: A1-chr1 is MT921804 and A1-chr2 is MT921805. The associated BioProject ID, Bio-Sample accession and SRA are PRJNA613602, SAMN14410478 and SRS6346145 for raw sequencing data, respectively.

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

References

1. Shtolz, N.; Mishmar, D. The mitochondrial genome–on selective constraints and signatures at the organism, cell, and single mitochondrion levels. *Front. Ecol.* **2019**, *7*, 342. [CrossRef]
2. 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]
3. Logan, D.C. The mitochondrial compartment. *J. Exp. Bot.* **2007**, *57*, 1225–1243. [CrossRef] [PubMed]
4. Møller, I.M.; Rasmusson, A.G.; Van Aken, O. Plant mitochondria—Past, present and future. *Plant J.* **2021**, *108*, 912–959. [CrossRef] [PubMed]
5. Bergthorsson, U.; Richardson, A.O.; Young, G.J.; Goertzen, L.R.; Palmer, J.D. Massive horizontal transfer of mitochondrial genes from diverse land plant donors to the basal angiosperm Amborella. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17747–17752. [CrossRef]
6. Oda, K.; Katsuyuki, Y.; Ohta, E.; Nakamura, Y.; Takemura, M.; Nozato, N.; Akashi, K.; Kanegae, T.; Ogura, Y.; Kohchi, T.; et al. Gene organization deduced from the complete sequence of liverwort Marchantia polymorpha mitochondrial DNA. A primitive form of plant mitochondrial genome. *J. Mol. Biol.* **1992**, *223*, 1–7. [CrossRef]
7. Sloan, D.B.; Alverson, A.J.; Chuckalovcak, J.P.; Wu, M.; McCauley, D.E.; Palmer, J.D.; Taylor, D.R. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol.* **2012**, *10*, e1001241. [CrossRef] [PubMed]
8. Richardson, A.O.; Rice, D.W.; Young, G.J.; Alverson, A.J.; Palmer, J.D. The "fossilized" mitochondrial genome of Liriodendron tulipifera: Ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. *BMC Biol.* **2013**, *11*, 29. [CrossRef] [PubMed]
9. Williams, S.T.; Foster, P.G.; Littlewood, D.T. The complete mitochondrial genome of a turbinid vetigastropod from MiSeq Illumina sequencing of genomic DNA and steps towards a resolved gastropod phylogeny. *Gene* **2014**, *533*, 38–47. [CrossRef] [PubMed]
10. Kivisild, T. Maternal ancestry and population history from whole mitochondrial genomes. *Investig. Genet.* **2015**, *6*, 3. [CrossRef] [PubMed]
11. Xu, F.; He, L.; Gao, S.; Su, Y.; Xu, L. Comparative analysis of two sugarcane ancestors *Saccharum officinarum* and *S. spontaneum* based on complete chloroplast genomesequences and photosynthetic ability in cold stress. *Int. J. Mol. Sci.* **2019**, *20*, 3828. [CrossRef] [PubMed]
12. Hisano, H.; Tsujimura, M.; Yoshida, H.; Terachi, T.; Sato, K. Mitochondrial genome sequences from wild and cultivated barley (*Hordeum vulgare*). *BMC Genom.* **2016**, *17*, 824. [CrossRef]
13. Zhou, D.; Liu, X.; Gao, S.; Guo, J.; Su, Y.; Ling, H.; Wang, C.; Li, Z.; Xu, L.; Que, Y. Foreign *cry1Ac* gene integration and endogenous borer stress-related genes synergistically improve insect resistance in sugarcane. *BMC Plant Biol.* **2018**, *18*, 342. [CrossRef] [PubMed]
14. Li, Z.; Hua, X.; Zhong, W.; Yuan, Y.; Wang, Y.; Wang, Z.; Ming, R.; Zhang, J. Genome-wide identification and expression profile analysis of WRKY family genes in the autopolyploid *Saccharum spontaneum*. *Plant Cell Physiol.* **2019**, *3*, 616–630. [CrossRef] [PubMed]
15. Liu, X.; Yin, Z.; Liu, Y.; Li, Z.; Zhou, D. The complete mitochondrial genome of sugarcane (*Saccharum* spp.) variety FN15. *Mitochondrial DNA B Resour.* **2020**, *5*, 2163–2165. [CrossRef]
16. Xu, F.; Wang, Z.; Lu, G.; Zeng, R.; Que, Y. Sugarcane ratooning ability:research satus, shortcomings, and prospects. *Biology* **2021**, *10*, 1052. [CrossRef]
17. Irvine, J. *Saccharum* species as horticultural classes. *Theor. Appl. Genet.* **1999**, *98*, 186–194. [CrossRef]
18. Bremer, G. Problems in breeding and cytology of sugar cane. *Euphytica* **1961**, *10*, 59–78. [CrossRef]

19. Sobhakumari, V.P. New determinations of somatic chromosome number in cultivated and wild species of *Saccharum*. *Caryologia* **2013**, *66*, 268–274. [CrossRef]
20. Zhang, J.; Zhang, X.; Tang, H.; Zhang, Q.; Hua, X.; Ma, X.; Zhu, F.; Jones, T.; Zhu, X.; Bowers, J.; et al. Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. *Nat. Genet.* **2018**, *50*, 1565–1573. [CrossRef]
21. Zhang, J.; Zhang, Q.; Li, L.; Tang, H.; Zhang, Q.; Chen, Y.; Arrow, J.; Zhang, X.; Wang, A.; Miao, C.; et al. Recent polyploidization events in *three saccharum* founding species. *Plant Biotechnol. J.* **2018**, *17*, 847–852. [CrossRef] [PubMed]
22. D'Hont, A. Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. *Cytogenet Genome Res.* **2005**, *109*, 27–33. [CrossRef] [PubMed]
23. Garsmeur, O.; Droc, G.; Antonise, R.; Grimwood, J.; Potier, B.; Aitken, K.; Jenkins, J.; Martin, G.; Charron, C.; Hervouet, C.; et al. A mosaic monoploid reference sequence for the highly complex genome of sugarcane. *Nat. Commun.* **2018**, *9*, 2638. [CrossRef] [PubMed]
24. Zhou, D.; Yin, Z.; Liu, X.; Li, Z.; Yan, M.; Que, Y.; Xu, L. The complete mitochondrial genome sequence and phylogenetic analysis of sugarcane (*Saccharum* spp.) cultivar ROC22. *Mitochondrial DNA Part B* **2020**, *5*, 1915–1916. [CrossRef]
25. Shearman, J.R.; Sonthirod, C.; Naktang, C.; Pootakham, W.; Yoocha, T.; Sangsrakru, D.; Jomchai, N.; Tragoonrung, S.; Tangphatsornruang, S. The two chromosomes of the mitochondrial genome of a sugarcane cultivar: Assembly and recombination analysis using long PacBio reads. *Sci. Rep.* **2016**, *6*, 31533. [CrossRef]
26. Chen, J.; Guan, R.; Chang, S.; Du, T.; Zhang, H.; Xing, H. Substoichiometrically different mitotypes coexist in mitochondrial genomes of *Brassica napus* L. *PLoS ONE* **2011**, *6*, e17662. [CrossRef]
27. Borgström, E.; Lundin, S.; Lundeberg, J. Large scale library generation for high throughput sequencing. *PLoS ONE* **2011**, *6*, e19119. [CrossRef]
28. Antipov, D.; Korobeynikov, A.; McLean, J.S.; Pevzner, P.A. hybridSPAdes: An algorithm for hybrid assembly of short and long reads. *Bioinformatics* **2016**, *32*, 1009–1015. [CrossRef]
29. Lagesen, K.; Hallin, P.; Rødland, E.A.; Stærfeldt, H.-H.; Rognes, T.; Ussery, D.W. RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **2007**, *35*, 3100–3108. [CrossRef]
30. 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]
31. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]
32. Tatusov, R.L.; Fedorova, N.D.; Jackson, J.D.; Jacobs, A.R.; Kiryutin, B.; Koonin, E.V.; Krylov, D.M.; Mazumder, R.; Mekhedov, S.L.; Nikolskaya, A.N.; et al. The COG database: An updated version includes eukaryotes. *BMC Bioinform.* **2003**, *4*, 41. [CrossRef] [PubMed]
33. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **2000**, *25*, 25–29. [CrossRef] [PubMed]
34. Kanehisa, M.; Goto, S.; Kawashima, S.; Okuno, Y.; Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **2004**, *32*, D277–D280. [CrossRef] [PubMed]
35. Magrane, M.; UniProt Consortium. UniProt Knowledgebase: A hub of integrated protein data. *Database* **2011**, bar009. [CrossRef] [PubMed]
36. Lohse, M.; Drechsel, O.; Bock, R. OrganellarGenomeDRAW (OGDRAW): A tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr. Genet.* **2007**, *52*, 267–274. [CrossRef]
37. Benson, G. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580. [CrossRef]
38. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [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. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol.* **2010**, *59*, 307–321. [CrossRef]
41. Kozik, A.; Rowan, B.A.; Lavelle, D.; Berke, L.; Schranz, M.E.; Michelmore, R.W.; Christensen, A.C. The alternative reality of plant mitochondrial DNA: One ring does not rule them all. *PLoS Genet.* **2019**, *15*, e1008373. [CrossRef]
42. Morley, S.A.; Nielsen, B.L. Plant mitochondrial DNA. *Front. Biosci.* **2017**, *22*, 1023–1032.
43. Alverson, A.J.; Rice, D.W.; Dickinson, S.; Barry, K.; Palmer, J.D. Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. *Plant Cell* **2011**, *23*, 2499–2513. [CrossRef] [PubMed]
44. Liu, H.; Yu, J.; Yu, X.; Zhang, D.; Chang, H.; Li, W.; Song, H.; Cui, Z.; Wang, P.; Luo, Y.; et al. Structural variation of mitochondrial genomes sheds light on evolutionary history of soybeans. *Plant J.* **2021**, *108*, 1456–1472. [CrossRef] [PubMed]
45. Tsujimura, M.; Kaneko, T.; Sakamoto, T.; Kimura, S.; Shigyo, M.; Yamagishi, H.; Terachi, T. Multichromosomal structure of the onion mitochondrial genome and a transcript analysis. *Mitochondrion* **2019**, *46*, 179–186. [CrossRef] [PubMed]
46. Niu, Y.; Gao, C.; Liu, J. Complete mitochondrial genomes of three Mangifera species, their genomic structure and gene transfer from chloroplast genomes. *BMC Genom.* **2022**, *23*, 147. [CrossRef]

47. Shidhi, P.R.; Biju, V.C.; Anu, S.; Vipin, C.L.; Deelip, K.R.; Achuthsankar, S.N. genome characterization, comparison and phylogenetic analysis of complete mitochondrial genome of evolvulus alsinoides reveals highly rearranged gene order in *Solanales*. *Life* **2021**, *11*, 769. [CrossRef] [PubMed]
48. Guo, W.; Zhu, A.; Fan, W.; Mower, J.P. Complete mitochondrial genomes from the ferns Ophioglossum californicum and Psilotum nudum are highly repetitive with the largest organellar introns. *New Phytol.* **2016**, *213*, 391–403. [CrossRef]
49. Mileshina, D.; Koulintchenko, M.; Konstantinov, Y.; Dietrich, A. Transfection of plant mitochondria and in organello gene integration. *Nucleic Acids Res.* **2011**, *39*, e115. [CrossRef] [PubMed]
50. Cheng, Y.; He, X.; Priyadarshani, S.; Wang, Y.; Ye, L.; Shi, C.; Ye, K.; Zhou, Q.; Luo, Z.; Deng, F.; et al. Assembly and comparative analysis of the complete mitochondrial genome of *Suaeda glauca*. *BMC Genom.* **2021**, *22*, 167. [CrossRef] [PubMed]
51. Makarenko, M.S.; Omelchenko, D.O.; Usatov, A.V.; Gavrilova, V.A. The insights into mitochondrial genomes of sunflowers. *Plants* **2021**, *10*, 1774. [CrossRef] [PubMed]
52. Bi, C.; Paterson, A.H.; Wang, X.; Xu, Y.; Wu, D.; Qu, Y.; Jiang, A.; Ye, Q.; Ye, N. Analysis of the complete mitochondrial genome sequence of the diploid cotton *Gossypium raimondii* by comparative genomics approaches. *BioMed Res. Int.* **2019**, *2019*, 5040598. [CrossRef] [PubMed]
53. Zheng, P.; Wang, D.X.; Huang, Y.Q.; Chen, H.; Du, H.; Tu, J.M. Detection and analysis of C-to-U RNA Editing in rice mitochondria-encoded ORFs. *Plants* **2020**, *9*, 1277. [CrossRef] [PubMed]
54. Wu, Z.; Cuthbert, J.M.; Taylor, D.R.; Sloan, D.B. The massive mitochondrial genome of the angiosperm *Silene noctiflora* is evolving by gain or loss of entire chromosomes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10185–10191. [CrossRef]
55. Gualberto, J.M.; Newton, K.J. Plant Mitochondrial Genomes: Dynamics and Mechanisms of Mutation. *Annu. Rev. Plant Biol.* **2017**, *68*, 225–252. [CrossRef] [PubMed]