Article

Complete Chloroplast Genome Sequence of Broomcorn Millet (*Panicum miliaceum* L.) and Comparative Analysis with Other Panicoideae Species

Xiaojun Nie ^{1,2,*} , Xian Zhao ¹, Sue Wang ¹, Ting Zhang ¹, Chong Li ¹, Hui Liu ¹, Wei Tong ¹ and Yuan Guo ^{1,*}

¹ College of Agronomy, Northwest A&F University, Yangling 712100, Shaanxi, China; jooxian568394@163.com (X.Z.); wangsue0518@163.com (S.W.); zhangting0825@outlook.com (T.Z.); lichong@nwsuaf.edu.cn (C.L.); ours2010@163.com (H.L.); tongw@nwsuaf.edu.cn (W.T.)

² State Key Laboratory of Crop Stress Biology in Arid Areas, Northwest A&F University, Yangling 712100, Shaanxi, China

* Correspondence: small@nwsuaf.edu.cn (X.N.); guoyuan210g@163.com (Y.G.); Tel.: +86-29-8708-2984 (X.N.); Fax: +86-29-8708-2203 (X.N.)

Received: 22 July 2018; Accepted: 13 August 2018; Published: 22 August 2018

Abstract: Broomcorn millet (*Panicum miliaceum* L.) is one of the earliest domesticated cereals worldwide, holding significant agricultural, historical, and evolutionary importance. However, our genomic knowledge of it is rather limited at present, hampering further genetic and evolutionary studies. Here, we sequenced and assembled the chloroplast genome (cp) of broomcorn millet and compared it with five other Panicoideae species. Results showed that the cp genome of broomcorn millet was 139,826 bp in size, with a typical quadripartite structure. In total, 108 genes were annotated and 18 genes were duplicated in the IR (inverted region) region, which was similar to other Panicoideae species. Comparative analysis showed a rather conserved genome structure between them, with three common regions. Furthermore, RNA editing, codon usage, and expansion of the IR, as well as simple sequence repeat (SSR) elements, were systematically investigated and 13 potential DNA markers were developed for Panicoideae species identification. Finally, phylogenetic analysis implied that broomcorn millet was a sister species to *Panicum virgatum* within the tribe Paniceae, and supported a monophyly of the Panicoideae. This study has reported for the first time the genome organization, gene content, and structural features of the chloroplast genome of broomcorn millet, which provides valuable information for genetic and evolutionary studies in the genus *Panicum* and beyond.

Keywords: broomcorn millet; chloroplasts; comparative genomics; Panicoideae; RNA editing

1. Introduction

Chloroplasts are organelles that are responsible for photosynthesis, which is vital for providing energy for plants and algae. It is believed that chloroplasts originated from cyanobacteria through endosymbiosis [1–3]. They have their own genetic replication mechanisms, have a mostly maternal inheritance, and transcribe their own genome relatively independently [4]. In higher plants, the chloroplast genome (cp) is generally a closed circular DNA in the form of multiple copies in a plant cell. The chloroplast genome of higher plants has a highly conserved quadripartite structure with approximately 110–130 genes, containing two inverted repeat sequences (IR), as well as one large single-copy region (LSC) and one small single-copy region (SSC) [5]. The size of the chloroplast genome

ranges from 115 to 165 kb, depending on species, and the genome size variation is mainly decided by the length variation of the IR region [6,7]. Due to its abundant genetic information, moderate nucleotide replacement rate, maternal inheritance, and highly conserved structures [5,8], the chloroplast genome is widely used for species identification, plant barcoding, phylogenetic classification, and molecular evolutionary studies [9–11]. Nowadays, rapid progress in next-generation sequencing technologies has resulted in a boom in completed chloroplast genomes [12,13], making phylogenomics analysis based on complete cp genome accessible. The phylogenomic analysis of cp genomes is of great value in providing resolution and support for evolutionary relationships of angiosperms. Furthermore, it will shed light on the cp genome evolution, such as sequence inversion [9,14], gene loss and gain [15,16], and structure variation [17,18].

Panicoideae is the second largest subfamily of the grass family (Poaceae), which is composed of approximately 3316 species in 12 tribes [19], such as lawn grasses (*Eremochloaophioides* and *Stenotaphrum secundatum*) [20], biofuel stocks (*Panicum virgatum*) [21], and some important crops like maize (*Zea mays* L.), sugarcane (*Saccharum officinarum* L.), sorghum (*Sorghum bicolor* L.), and broomcorn millet (*Panicum miliaceum* L.) as well as foxtail millet (*Setaria italica* L.) [22,23]. In light of its economic and evolutionary significance, extensive studies have been performed to investigate the phylogenetic relationships and genetic divergence in the Panicoideae subfamily [24,25]. Early studies showed that there were at least seven tribes in the Panicoideae subfamily [26]. Numerous studies were carried out to study the accurate phylogeny based on morphological characters and molecular markers, although the relationship could not be defined perfectly and there are low support values for phylogenetic nodes [27,28]. Recently, the phylogenomic analysis using 50 full plastomes provided better resolution to determine the phylogenetic relationships in the Panicoideae subfamily [29].

Broomcorn millet is one of the oldest domesticated cereals in China and its agricultural value in the northern region can be traced back to the Pleistocene and Holocene boundaries [30]. Before rice and wheat were domesticated, broomcorn millet was the main food staple of humans living in the semi-arid region of Asia. Broomcorn millet has many excellent agronomic traits, such as short growth cycle, high yield, lower nutrient and lower water requirements, and good resistance to diverse stresses [31]. In recent years, progress has been made in the study of the nuclear genome of broomcorn millet, such as the exploration and identification of stress-related WRKY gene family in the transcriptome of broomcorn millet and SSR analysis based on high-throughput sequencing [32,33]. However, the chloroplast genome of broomcorn millet has not been sequenced and the phylogenetic relationship between *Panicum miliaceum* L. with other Panicoideae subfamily species is still not well understood at present. In this study, we sequenced and assembled the complete chloroplast genome of *Panicum miliaceum* L. based on the Illumina sequencing technology. Then, a comparative analysis of broomcorn millet with other five closely related species, including sugarcane (KU214867.1), foxtail millet (KF646538.1), sorghum (EF115542.1), switchgrass (NC_015990.1), and maize (KF241981.1), was performed. This study not only provided useful information on genome organization, gene content, and structure variation in the broomcorn millet cp genome, but also provided important clues to its phylogenetic relationships, which will contribute to genetic and evolutionary studies in broomcorn millet and beyond.

2. Results and Discussion

2.1. Chloroplast Genome Assembly and Sequence Analysis of Broomcorn Millet

Based on the assembly method described by Nie et al. [9], we obtained the complete chloroplast genome sequence of broomcorn millet. After annotation and modification, the genome sequence was submitted to the GenBank database with accession no. KU343177.1. The size of this genome was 139,826 bp, with the typical quadripartite structure containing a pair of inverted repeats (IRA and IRB, 22,785 bp in size), one SSC region (12,574 bp in size), and one LSC region (81,682 bp in size), which is similar to other Panicoideae cp genomes [22,29] (Table 1). The GC content of the IR

region (43.95%) was much higher than that of the LSC (36.47%) and SSC regions (33.09%) in the broomcorn millet cp genome. The higher GC content in the IR region was probably the result of the presence of four ribosomal RNA (rRNA) genes in this region, which is consistent with previous analyses in other plants [18,34] (Table 1 and Figure 1). Additionally, the GC content in the overall cp genome and IR region was nearly the same as in other Panicoideae cp genomes, suggesting that the cp genome in this subfamily had rather conserved genome organization [15,29]. A total of 108 genes were annotated in broomcorn millet cp genome, including 76 protein-coding genes, 28 transfer RNAs (tRNA) genes, and four rRNA genes. Among them, six protein-coding genes (namely *rps19*, *rps7*, *rpl23*, *rpl2*, *rps15*, and *ycf15*), four rRNA genes (*rRNA23*, *rRNA16*, *rRNA5*, and *rRNA4.5*) and eight tRNA genes (*trnM-CAU*, *trnH-GUG*, *trnI-CAU*, *trnI-GAU*, *trnL-CAA*, *trnM-CAU*, *trnN-GUU*, *trnR-ACG* and *trnV-GAC*) were duplicated as they located in the IR region. At the same time, the SSC region consisted of 10 protein-coding genes and one tRNA, while the LSC region consisted of 61 protein-coding genes and 22 tRNA (Table 2). Additionally, 28 unique tRNA genes, which encoded 20 amino acids, were distributed unevenly within the genome (one in the SSC region, 19 in the LSC region, and eight in the IR region). Furthermore, out of 108 genes, 13 genes each have one intron (namely *rps16*, *atpF*, *petB*, *petD*, *rpl2*, *ndhB*, *ndhA*, *trnK-UUUU*, *trnA-UGC*, *trnL-UAA*, *trnV-UAC*, *rpl16*, and *trnI-GAU*) and one gene (*ycf3*) has two introns. Out of the 14 genes with introns, nine genes were located in the LSC (five protein-coding genes, three tRNA with one intron, and one *ycf* gene with two introns), one in the SSC (a protein-coding gene with one intron), and four in the IR region (two protein-coding genes and two tRNAs with one intron). Like most other land plants, *rps12* is a trans-spliced gene with the 5' end exon located in the LSC region and the duplicated 3' end exon located in the IR regions. The *trnK-UUUU* has the largest intron (2472 bp), where another important gene, *matK*, was located (Table 3).

Table 1. Comparisons of CpDNA features among the six Panicoideae species.

Genome Characteristics	<i>P. miliaceum</i>	<i>P. virgatum</i> [15]	<i>Z. mays</i> [34]	<i>S. bicolor</i> [22]	<i>S. officinarum</i> [18]	<i>S. italic</i> [35]
Genome Size (bp)	139,826	139,619	140,447	140,754	141,182	135,516
LSC length (bp)	81,682	81,659	82,375	82,688	83,048	79,896
SSC length (bp)	12,574	12,564	12,540	12,502	12,544	12,012
IR length (bp)	22,785	22,698	22,766	22,782	22,795	21,804
Genome length with one copy of IR (bp)	117,068	116,921	117,681	117,972	118,387	113,712
No. of different genes	108	111	110	110	110	110
No. of different protein-coding genes	76	77	76	78	76	76
No. of different tRNA genes	28	30	30	30	30	30
No. of rRNA genes	4	4	4	4	4	4
No. of genes duplicated by IR	15	16	18	18	20	18
No. of different genes with introns	15	16	15	16	16	15
coding sequence length (bp)	60,903	60,315	60,348	60,966	61,272	60,168
No. of codons	20,301	20,105	20,116	20,322	20,424	20,056
% GC content in LSC	36.47	33.1	36.26	36.38	36.28	36.82
% GC content in SSC	33.09	36.43	32.83	32.63	32.86	33.27
% GC content in IR	43.95	44.01	43.93	43.99	43.91	43.15
% GC content of genome	38.61	38.59	38.44	38.26	38.44	38.97
Coding % GC content	39.07	39.00	39.19	39.10	39.15	39.14

Figure 1. Chloroplast genome map of broomcorn millet. Genes located in the outside of the outer circle are transcribed clockwise, and the genes located inside are transcribed counterclockwise. Genes with different functions were divided into different groups with various colors.

Table 2. Genes found in the assembled broomcorn millet chloroplast genome.

Category	Group of Genes	Name of Genes
Transcription and translation related genes (27)	Large subunit ribosomal proteins	<i>rpl2b</i> , <i>c</i> , 14, 16 <i>b</i> , 20, 22, 23 <i>c</i> , 32, 33, 36
	Small subunit ribosomal proteins	<i>rps2</i> , 3, 4, 7 <i>c</i> , 8, 11, 12 <i>b</i> , <i>c</i> , <i>d</i> , 14, 15 <i>c</i> , 16 <i>b</i> , 18, 19 <i>c</i>
	translation initiation factor	<i>infA</i>
	DNA dependent RNA polymerase	<i>rpoA</i> , <i>B</i> , <i>C1</i> , <i>C2</i>
RNA genes (32)	Ribosomal RNAs	<i>rrn23c</i> , 16 <i>c</i> , 5 <i>c</i> , 4.5 <i>c</i>
	Transfer RNAs	<i>trnC</i> -GCA, <i>trnD</i> -GUC, <i>trnE</i> -UUC, <i>trnF</i> -GAA, <i>trnG</i> -CAU, <i>trnH</i> -UCC, <i>trnI</i> -CAU, <i>trnL</i> -GAU, <i>trnK</i> -UUU, <i>trnL</i> -CAA, <i>trnL</i> -UAA, <i>trnL</i> -UAG, <i>trnM</i> -CAU, <i>trnN</i> -GUU, <i>trnP</i> -UGG, <i>trnQ</i> -UUG, <i>trnR</i> -ACG, <i>trnR</i> -UCU, <i>trnS</i> -GCU, <i>trnS</i> -GGA, <i>trnS</i> -UGA, <i>trnT</i> -GGU, <i>trnT</i> -UGU, <i>trnV</i> -GAC, <i>trnV</i> -UAC, <i>trnW</i> -CCA, <i>trnY</i> -GUA
photosynthesis related genes (46)	Photosystem I	<i>psaA</i> , <i>B</i> , <i>C</i> , <i>I</i> , <i>J</i> , <i>ycf3a</i> , <i>ycf4</i>
	Photosystem II	<i>psbA</i> , <i>B</i> , <i>C</i> , <i>D</i> , <i>E</i> , <i>F</i> , <i>H</i> , <i>I</i> , <i>J</i> , <i>K</i> , <i>L</i> , <i>M</i> , <i>N</i> , <i>T</i> , <i>Z</i>
	NADH oxidoreductase	<i>ndhA</i> , <i>B</i> , <i>c</i> , <i>C</i> , <i>D</i> , <i>E</i> , <i>F</i> , <i>G</i> , <i>H</i> , <i>I</i> , <i>J</i> , <i>K</i>
	Cytochrome <i>b6/f</i>	<i>petA</i> , <i>B</i> , <i>Db</i> , <i>G</i> , <i>L</i> , <i>N</i>
	ATP synthase	<i>atpA</i> , <i>B</i> , <i>E</i> , <i>Fb</i> , <i>H</i> , <i>I</i>
Other genes (4)	Rubisco	<i>rbcL</i>
	fatty acid synthesis	<i>ccsA</i>
	carbon metabolism	<i>cemA</i>
	Proteolysis	<i>clpP</i>
Proteins of unknown function (3)	RNA processing	<i>matK</i>
	conserved open reading frames	<i>ycf15c</i> , <i>ycf68c</i> , <i>orf42</i>

a Gene containing two intron. b Gene containing a single intron. c Two gene copies in the IRs. d Gene divided into two independent transcription units.

Table 3. The genes having introns in the broomcorn cp genome, and the length of the exons and introns. The length units for the exons and introns are base pair (bp).

Gene	Location	ExonI	IntronI	ExonII	IntronII	ExonIII
<i>rps16</i>	LSC	218	840	40		
<i>atpF</i>	LSC	159	816	408		
<i>ycf3</i>	LSC	159	732	228	739	132
<i>petB</i>	LSC	6	764	645		
<i>rps12</i> *	LSC	114	–	29	541	232
<i>petD</i>	LSC	8	740	475		
<i>rpl2</i>	IR	390	661	429		
<i>ndhB</i>	IR	753	711	777		
<i>ndhA</i>	SSC	534	1022	537		
<i>trnK</i> -UUU	LSC	33	2472	38		
<i>trnA</i> -UGC	IR	38	811	35		
<i>trnL</i> -UAA	LSC	35	541	50		
<i>trnV</i> -UAC	LSC	37	601	39		
<i>rpl16</i>	LSC	402	1001	9		
<i>trnI</i> -GAU	IR	42	947	35		

* *rps12* is a trans-spliced gene with the 5' end exon located in the LSC region and the duplicated 3' end exon located in IR regions.

Sequence analysis revealed that the 43.5%, 2.9%, and 6.6% of broomcorn millet cp genome encoded protein, tRNAs, and rRNAs respectively, and the remaining 47.0% belonged to non-coding regions. Among the coding regions, there were 76 protein-coding genes categorized into different functional groups, including nine genes for large ribosomal proteins, 12 genes for small subunit ribosomal proteins, seven genes for photosystem I, 15 genes for photosystem II, and six genes for ATP synthase. A similar gene composition was also found in other monocots [34] (Table 2). Furthermore, the 82 protein-coding genes comprise of 60,903 bp in length coding for 20,301 codons. Then, the codon usage frequency of broomcorn millet cp genome was further calculated. Results showed that there

were 2199 codons (10.8%) encoding for leucine, representing the most used amino acid, while there were 228 codons (1.12%) for cysteine, which was least used (Table 4). Among them, ATT (823) for isoleucine was the most used codon. The AT content was 52.17%, 60.43%, and 69.33% at the 1st, 2nd, and 3rd codon positions of these codons, respectively. The preference for A or T at the 3rd position suggested codon bias towards A or T at the third codon position, which was consistent with that of switchgrass, barley, sorghum, and foxtail millet [35,36].

Table 4. The codon–anticodon recognition pattern and codon usage for broomcorn cp genome.

Amino Acid	Codon	No. *	tRNA	Amino Acid	Codon	No.	tRNA	Amino Acid	Codon	No.	tRNA
Phe	UUU	719	trnF-GAA	Ser	UCU	383	trnS-GGA	Tyr	UAU	578	trnY-GUA
Phe	UUC	405		Ser	UCC	291		Tyr	UAC	149	
Leu	UUA	734	trnL-UAA	Ser	UCA	235		stop	UAA	43	
Leu	UUG	396	trnL-CAA	Ser	UCG	129		stop	UAG	23	
Leu	CUU	463	trnL-UAG	Pro	CCU	344	trnP-UGG	His	CAU	341	trnH-GUG
Leu	CUC	159		Pro	CCC	208		His	CAC	120	
Leu	CUA	324		Pro	CCA	239		Gln	CAA	525	trnQ-UUG
Leu	CUG	123		Pro	CCG	94		Gln	CAG	151	
Ile	AUU	823	trnI-GAU	Thr	ACU	444	trnT-GGU	Asn	AAU	598	trnN-GUU
Ile	AUC	324		Thr	ACC	209		Asn	AAC	206	
Ile	AUA	524	trnI-CAU	Thr	ACA	302	trnT-UGU	Lys	AAA	754	trnK-UUU
Met	AUG	464	trnM-CAU	Thr	ACG	132		Lys	AAG	291	
			trnM-CAU								
Val	GUU	444	trnV-GAC	Ala	GCU	555	trnA-UGC	Asp	GAU	570	trnD-GUC
Val	GUC	145		Ala	GCC	191		Asp	GAC	161	
Val	GUA	447	trnV-UAC	Ala	GCA	376		Glu	GAA	801	trnE-UUC
Val	GUG	159		Ala	GCG	144		Glu	GAG	264	
Cys	UGU	170	trnC-GCA	Arg	CGU	294	trnR-ACG	Ser	AGU	293	trnS-GCU
Cys	UGC	58		Arg	CGC	111		Ser	AGC	110	
stop	UGA	21		Arg	CGA	275		Arg	AGA	370	trnR-UCU
Trp	UGG	350	trnW-CCA	Arg	CGG	94		Arg	AGG	128	
Gly	GGU	487	trnG-UCC	Gly	GGC	160					
Gly	GGA	602		Gly	GGG	274					

* Numerals indicate the frequency of usage of each codon in 20,301 codons in 78 potential protein-coding genes.

2.2. Comparison of cp Genomes between Broomcorn Millet and Other Five Panicoideae Species

Five complete cp genomes within the Panicoideae subfamily were chosen to compare with broomcorn millet. The results showed that some small variation was present in the size of these cp genomes. The genome size of broomcorn millet, sugarcane, foxtail millet, sorghum, switchgrass and maize was 139,826 bp, 141,182 bp, 135,516 bp, 141,279 bp, 139,619 bp, and 140,447 bp, respectively. The size variation among them was mainly due to the difference in the length of the LSC and IR regions (Table 1). Furthermore, the general gene structure, gene organization, and gene content were highly conserved among these six cp genomes, and a total of 75 protein-coding genes were shared by them. However, we found that three genes, namely *accD*, *ycf1*, and *ycf2*, were all lost in these six cp genomes, which is similar to previous results regarding other Panicoid grasses [37,38]. Similarly, two genes, including *ycf15* and *orf42*, were only found in broomcorn millet, and one gene, *ycf68*, was observed in broomcorn millet, switchgrass, and maize. In addition, the loss of an intron in *clpP* and *rpoC1* was also identified in the broomcorn millet cp genome, which was also found in switchgrass, barley, and sorghum cp genomes [22]. Further study is needed to reveal the molecular mechanism and impact of intron loss events in these genes.

Furthermore, the detailed IR-SSC and IR-LSC borders, together with the adjacent genes, were systematically compared across the six cp genomes. Results showed that the border of IRb/SSC was located in the coding region of *ndhF* gene, resulting in a pseudogene of *ndhF* in the IRa region with the same size as IRb region expanded into *ndhF* gene in all six Panicoideae cp genomes. However, different lengths of IRb expanding into *ndhF* were found; the size in broomcorn millet was 112 bp, while it was 29 bp for the other five Panicoideae species (Figure 2). The expansion of *ndhF* was found to be unique to Panicoideae, which could be considered a unique feature to barcode the members of this subfamily [29]. At the same time, we found that there was no expansion of the IRa/SSC boundary

into *ndhH* genes in all six Panicoideae species, an difference from Ehrhartoideae and Pooideae species, where the partial duplication of *ndhH* genes were found. Additionally, the size of the intergenic spacer regions between *rps19* and the end of the IRa region was consistently 35 bp in these six cp genomes, while the sizes of the start site of LSC to *psbA*, and *rpl22* to the end of LSC, as well as the start site of IRb to *rps19*, showed some minor variations. Overall, the contraction or expansion of IR regions only partially contributed to the variation of the total size of the Panicoideae plastid genome, although they showed diverse contraction or expansion of IR. The variations in intergenic spacer regions might also exert an impact on the Panicoideae plastid genome size [29].

Figure 2. Comparison of the border positions of SSC, LSC, and IR regions among six Panicoideae chloroplast genomes.

Pairwise alignment was also conducted between broomcorn millet and the other five Panicoideae plastids, and the mVISTA tool was used to plot the comprehensive sequence identity of these six Panicoideae cp genomes, using broomcorn millet as a reference (Figure 3). Results demonstrated that the SSC and LSC regions showed more divergence than IRa and IRb regions. Also, non-coding regions were more divergent than coding regions. Among these six cp genomes, there were some highly divergent regions found, including *rpoC2*, *ndhB*, *trnL*, *rpl22*, *rpl23*, *psbK* and *matK*, of which some have been commonly used as makers for plant identification and phylogenetic analysis [25,39]. In addition, these newly identified regions could be considered as novel candidates for speciation identification and phylogenetic studies in Panicoideae and beyond.

Figure 3. Percent identity analysis for the comparison of six Panicoideae chloroplast genomes using mVISTA program. The top line displays genes in order with arrows. Sequence similarity of aligned regions between broomcorn millet and the other five species is shown in the direction of horizontal level implying average percent identity between 50–100% (shown on y-axis of graph). The x-axis is considered the coordinate in the cp genome.

Gene order between the broomcorn millet and other five Panicoideae species showed similar patterns, while it was different from that of *Nicotiana tabacum* as a result of three inversions of 28 kb, 6 kb, and less than 1 kb that were found in these Panicoideae species and have been identified by previous studies [40,41] (Figure 4). The endpoints of the first inversion occurred between *trnG-UCC* and *trnR-UCU* at one end and *rps14* and *trnfM-CAU* at the other end when compared to *N. tabacum*, and the second inversion (about 6 kb) had endpoints between *trnS* and *psbD* on one end and *trnG-UCC* and *trnT-GGU* on the other end [42]. The third inversion has endpoints between *trnG-UCC* and *trnT-GGU* to *trnT-GGU* and *trnE-UUC*. This inversion is very small and explained the inverted orientation of *trnT-GGU*.

Figure 4. Comparison of the genome structure of six Panicoideae chloroplast genomes, with *N. tabacum* as the reference. The boxes above the line represent DNA sequences in a clockwise direction, and those below the line represent DNA sequences in the counterclockwise direction.

2.3. SSR Loci Identified in Panicoideae cp Genomes

Simple sequence repeats (SSRs), also known as microsatellites, are 1–6 bp nucleotide and widely distributed in eukaryotes genome. Due to its abundant polymorphism, high stability, codominant inheritance and ease of use, SSRs have gradually been a popular and robust tool for population genetic analysis in non-model organisms [43]. It is believed that the threshold for SSRs screening should be over 8 bp, as SSRs of 8 bp or longer tend towards slip-strand mispairing (SSM), which is considered the main mutational mechanism that has some influence on SSRs [7,16]. In the present study, we set parameters over 8 bp to identify the potential SSRs in the chloroplast genome of broomcorn millet, together with other five related species. In total, 97 perfect SSRs were detected in the broomcorn millet cp genome, unevenly located at both genic and intergenic regions (Figure 5).

Figure 5. Analysis of simple sequence repeat (SSR) distribution in broomcorn millet and five other Panicoideae species.

Similarly, 94, 78, 87, 89, and 98 SSRs were identified in the maize, switchgrass, sorghum, foxtail millet, and sugarcane cp genome, respectively. The majority of the SSRs in these six cp genome were trinucleotides, ranging from 42 in foxtail millet to 50 in maize, and mononucleotides were the second most abundant, varying from 21 in foxtail millet to 36 in the broomcorn millet. Additionally, SSRs with pentanucleotide and hexanucleotide motifs in these genomes were scarce: only one hexanucleotide was found in the sorghum and one pentanucleotide was found in foxtail millet, sugarcane, and switchgrass (Figure 5 and Table S1). Among all the SSR motifs identified in the broomcorn millet cp genome, 85.9% of mononucleotides were comprised of A or T, and a similar proportion of trinucleotides (73.8%) was composed of A or T. Similarly, the motif tended to have a higher percentage of A or T in the other five related species. Our results were consistent with previous studies that reported that potential SSR markers identified in the cp genome commonly consisted of polyA or polyT repeats, and seldom included C and G repeats, which partially explained why broomcorn millet had a higher frequency of AT content [34,36]. Based on the SSR dynamics in these six cp genomes, a total of 13 potential SSR markers were developed, which could be used for molecular identification of Panicoideae and other species (Table S2).

2.4. Prediction of RNA Editing Sites in Panicoideae cp Genomes

RNA editing is an important way to regulate gene expression at a post-transcriptional level, which generally changes original genetic information by insertion, deletion, and base substitution of nucleotides within mRNA molecules [44]. Various studies in the chloroplasts of higher plants have suggested that identification of RNA editing sites could be very useful for enriching the genetic information, which makes it easier for organisms to adapt to the environment and also uncover the evolutionary patterns of RNA editing [18,45]. Here, a total of 175 editing sites were predicted in these six Panicoideae cp genomes, and further analysis found that all of these sites were C to U conversions, which was consistent with a previous study of RNA editing in cp genomes in *N. tabacum* [46]. Among them, 31 editing sites in 14 genes were identified for broomcorn millet, 30 editing sites in 15 genes for maize, 29 editing sites in 14 genes for switchgrass, 28 editing sites in 14 genes for sorghum, 28 editing sites in 14 genes for foxtail millet, and 29 editing sites in 15 genes for sugarcane. Furthermore, 11 sites in switchgrass, two sites in sorghum, one site in foxtail millet, and nine sites in maize were validated by EST alignment analysis. In addition, we compared the RNA editing sites patterns in these plastids, and 15 sites in nine genes were found to be shared by these six Panicoideae cp genomes (Table 5). It has been proven that the RNA editing was evolutionarily conserved and that species having a closer relationship generally shared more editing sites [45]. In this study, broomcorn

millet was found to share 29 editing sites with switchgrass, which was more overlap than for the other four species. Although they seemed to have rather similar features in RNA editing, there were still some specific editing sites, such as *rpl2*, that were only identified in maize and sugarcane. Thus, the conserved and species-specific RNA editing sites might provide some valuable information to trace the origin and evolution of Panicoideae species.

Table 5. RNA editing sites predicted in the chloroplast genome of broomcorn millet and five other Panicoideae species.

Gene	Site	Nucleotide Position	Amino Acid Position	Edited Codon	Amino Acid Change	<i>Panicum miliaceum</i> L.	<i>Panicum virgatum</i>	Sorghum	Seteria	Saccharum	Zea
<i>matK</i>	1 ^a	331	111	Ctt-Ttt	L-F	+	+	+	+	+	+
	2	1015	339	Cca-Tca	P-S	+	+	(-)	(-)	(-)	(-)
	3	1258	420	Cat-Tat	H-Y	+	+	+	+	+	(-)
	4 ^b	424	142	Ctc-Ttc	L-F	(-)	(-)	(-)	(-)	(-)	+
	5 ^b	1351	451	Cat-Tat	H-Y	(-)	(-)	(-)	(-)	(-)	+
<i>ropB</i>	1 ^a	467	156	tCg-tTg	S-L	+	+ ^c	+	+	+	+ ^c
	2 ^a	545	182	tCa-tTa	S-L	+	+ ^c	+	+	+	+ ^c
	3	560	187	tCg-tTg	S-L	(-)	(-)	+	+	+	+ ^c
	4 ^a	617	206	cCg-cTg	P-L	+	+	+	+	+	+ ^c
<i>rpoC2</i>	1	2209	737	Cat-Tat	H-Y	+	(-)	(-)	(-)	(-)	(-)
	2	2312	771	tCa-tTa	S-L	+	+	(-)	(-)	(-)	(-)
	3	2804	935	tCg-tTg	S-L	+	+	(-)	(-)	(-)	(-)
	4	2753	918	tCg-tTg	S-L	(-)	(-)	+	+	(-)	(-)
	5	2795	932	tCg-tTg	S-L	(-)	(-)	(-)	(-)	+	(-)
	6	1732	578	Ctc-Ttc	L-F	(-)	(-)	(-)	(-)	(-)	+
	7	2774	925	tCg-tTg	S-L	(-)	(-)	(-)	(-)	(-)	+ ^c
<i>atpA</i>	1 ^a	1148	383	tCa-tTa	S-L	+	+ ^c	+ ^c	+	+	+
<i>rps14</i>	1 ^a	80	27	tCa-tTa	S-L	+	+ ^c	+	+	+	+
<i>ycf3</i>	1 ^a	44	15	tCc-tTc	S-F	+	+ ^c	+	+	+	+ ^c
	2	191	64	aCg-aTg	S-F	+	+ ^c	+	+	(-)	(-)
	3	185	62	aCg-aTg	T-M	(-)	(-)	(-)	(-)	+	+ ^c
<i>rpl20</i>	1 ^a	308	103	tCa-tTa	S-L	+	+	+	+	+	+ ^c
<i>petB</i>	1	617	206	cCa-cTa	P-L	+	+ ^c	(-)	(-)	(-)	(-)
	2	668	223	cCa-cTa	P-L	(-)	(-)	+	+	(-)	+
	3	611	204	cCa-cTa	P-L	(-)	(-)	(-)	(-)	+	(-)
<i>rps8</i>	1 ^a	182	61	tCa-tTa	S-L	+	+ ^c	+	+	+	+
<i>rpl2</i>	1	2	1	aCg-aTg	T-M	(-)	(-)	(-)	(-)	+	+
<i>ndhB^a</i>	1	467	156	cCa-cTa	P-L	+	+	+	+	+	+
	2	586	196	Cat-Tat	H-Y	+	+	+	+	+	+
	3	611	204	tCa-tTa	S-L	+	+	+	+	+	+
	4	737	246	cCa-cTa	P-L	+	+	+	+	+	+
	5	830	277	tCa-tTa	S-L	+	+	+	+	+	+ ^c
	6	1481	494	cCa-cTa	P-L	+	+	+	+	+	+
<i>ndhF</i>	1 ^a	62	21	tCa-tTa	S-L	+	+	+	+	+	+
	2	1834	612	Ctt-Ttt	L-F	(-)	(-)	(-)	+	(-)	(-)
	3	1840	614	Ctt-Ttt	L-F	+	+	(-)	(-)	(-)	(-)
	4 ^b	1915	639	Ctt-Ttt	L-F	+	+	+	(-)	+	+
	5	1909	637	Ctt-Ttt	L-F	(-)	(-)	(-)	+	(-)	(-)
<i>ccsA</i>	1	647	216	aCt-aTt	T-I	+	(-)	(-)	(-)	(-)	(-)
	2	644	215	aCc-aTc	T-I	(-)	+	+	(-)	+	+
	3	641	214	aCt-aTt	T-I	(-)	(-)	(-)	+	(-)	(-)
<i>ndhD</i>	1 ^a	878	293	tCa-tTa	S-L	+	+ ^c	+	+ ^c	+	+
<i>ndhA</i>	1	35	12	tCg-tTg	S-L	+	+	(-)	(-)	(-)	(-)
	2	458	153	tCa-tTa	S-L	+	+ ^c	(-)	(-)	(-)	(-)
	3	548	183	tCa-tTa	S-L	+	+ ^c	(-)	(-)	(-)	(-)
	4	1055	352	tCc-tTc	S-F	+	(-)	(-)	(-)	(-)	(-)
	5	50	17	tCg-tTg	S-L	(-)	(-)	+	(-)	+	+
	6	473	158	tCa-tTa	S-L	(-)	(-)	+	(-)	+	+
	7	563	188	tCa-tTa	S-L	(-)	(-)	+	(-)	+	+
	8	1070	357	tCc-tTc	S-F	(-)	(-)	+ ^c	(-)	+	+
	9 ^b	47	16	tCg-tTg	S-L	(-)	(-)	(-)	+	(-)	(-)
	10 ^b	470	157	tCa-tTa	S-L	(-)	(-)	(-)	+	(-)	(-)
	11 ^b	560	187	tCa-tTa	S-L	(-)	(-)	(-)	+	(-)	(-)
Number of genes						14	14	14	14	15	15
Number of edited sites						31	29	28	28	29	30

Capital letters in codon triplets indicate target nucleotides; (-), no editing. a conserved genes/sites; b unique edited site in only one species; c validated by EST.

2.5. Phylogenetic Analysis

Poaceae is one of the most species-rich families of the angiosperms, with approximately 12,000 species and 771 genera [28]. Among them, there are 3560 species in the subfamily Panicoideae, which are further divided into 12 tribes. The chloroplast genome is a vital resource for addressing diverse phylogenetic questions across the Panicoideae family and analyzing the evolutionary relationships within the family. Various studies have been carried out to study the taxonomic status in order to have a better understanding of the phylogenetic relationship in the Panicoideae family; the single gene *matK* was chosen to conduct a molecular phylogenetic study [39]. At the same time, *trnL-trnF*, *psbJ-petA*, and *trnQ-rps16* gene sites were combined to analyze the evolutionary relationships among highly variable and controversial taxonomic groups of Panicoideae [42]. In order to further study the phylogenetic relationship of broomcorn millet within the Panicoideae family, 81 genes shared by the 66 cp genomes including five Panicoideae members, were used for multiple alignment based on the MSWAT database. After gap removal, there were 55,628 characters remaining in the final database. MP analysis based on MEGA software 7.0 [47] with *Nuphar advena* and *Nymphaea alba* as outgroup generated a single tree with a length of 146,741, a consistency index of 0.3838, and a retention index of 0.6016. Bootstrap analysis implied that 51 of 63 nodes have bootstrap values >95% and 42 of these with bootstrap values of 100%. Maximum likelihood (ML) analysis generated the same topology tree with the MP method. ML Bootstrap values were high and all 63 nodes had 100% bootstrap values (Figure 6). Overall, the phylogenetic tree was divided into two major clades: monocots and eudicots. All five Panicoideae species were clustered into the Panicoideae group and placed within the monocots. It was observed that *P. virgatum* was sister to *P. miliaceum* L. In a previous study, *P. virgatum* was clustered into subtribe Panicinae, which was the subfamily of tribe Paniceae, and *S. bicolor* was sister to *S. officinarum*, though *S. bicolor* belonged to subtribe Sorghinase and *S. officinarum* belonged to subtribe Saccharinae. In addition, *Z. mays* was clustered into subtribe Tripsacinae. All three subtribes belonged to the tribe Andropogoneae, which was similar to previous reports [44]. All five species are clustered into the PACMAD clade, and *P. virgatum* shows a closer relationship with *P. miliaceum* than the other three species.

Figure 6. The MP phylogenetic tree is based on 81 genes from 66 plant taxa. The MP tree has the length of 146741 and consistency index of 0.3838, and retention index of 0.6016. The numbers above the node are the bootstrap support values. The ML tree, which has the same topology, is not shown.

3. Materials and Methods

3.1. Genome Assembly and Genome Annotation

The raw Illumina sequence data of broomcorn millet downloaded from the NCBI database (accession no. SRR2163427) were used in this study. The chloroplast genome of broomcorn millet was assembled using the method reported by Nie et al. with some modifications [9]. The lower-quality reads were filtered out with the trimmomatic-0.36 tool (<http://www.usadellab.org/cms/?page=trimmomatic>). The obtained clean data were used to capture the chloroplast reads by mapping with

switchgrass and corn cp genomes as a reference based on BWA software [48]. Then, the captured cp reads were directly assembled into contigs with a minimum length of 100 bp using the SOAPdenovo2 software [49]. Furthermore, these contigs were aligned to the switchgrass cp genome (considered the reference genome) by the BLAST program (<http://blast.ncbi.nlm.nih.gov/>), and these aligned contigs were ordered according to the reference genome. Finally, the clean data were first mapped against the obtained sequence to close gaps and then the consensus sequence was used to close the gaps again to obtain the candidate genome sequence.

The online program DOGMA was used to predict and annotate the protein-coding genes, tRNA genes, and rRNA genes in the Broomcorn millet cp genome with default parameters [50], coupled with manual corrections. The position of codons, or intron/exon junctions of the protein-coding gene, were determined following Sugita and Sugiura [51], with those of the switchgrass cp genome as a reference. A circular map of the broomcorn millet cp genome was drawn using OGDRAW v1.2 software [52]. The GC content and codon usage were calculated manually.

3.2. Sequence and Structural Comparison in Six Panicoideae cp Genomes

The Shuffle-LAGAN mode of the mVISTA online program (<http://genome.lbl.gov/vista/mvista/>) [53] was used to compare the sequence similarity of the full chloroplast genome of broomcorn millet with other five Panicoideae chloroplast genomes, including sugarcane (KU214867.1), foxtail millet (KF646538.1), sorghum (EF115542.1), switchgrass (NC_015990.1), and maize (KF241981.1) using the annotation information of broomcorn millet. A structural comparison of these genomes was performed using the MAUVE tool [54]. Finally, the IR expansion or contraction, together with the adjacent genes, was compared across the six Panicoideae cp genomes manually based on the annotation files.

3.3. SSRs Identification and RNA Editing

MISA software (<http://pgrc.ipk-gatersleben.de/misa>) was used to detect simple sequence repeats (SSRs) within the six cp genome, with the parameters setting as ≥ 10 repeats of mononucleotides, ≥ 5 repeats of dinucleotides, trinucleotides and tetranucleotides, as well as ≥ 3 repeats of pentanucleotide and hexanucleotide SSRs, respectively. All the protein-coding genes of the six Panicoideae cp genomes were obtained from the NCBI database (<http://www.ncbi.nlm.nih.gov>) based on their assembled and annotated genome information. The RNA editing sites in these genes were predicted by Predictive RNA editor for plants (PREP)-Cp tool (<http://prep.unl.edu/cgi-bin/cp-input.pl>) with default parameters. Significant hits were examined manually, and only the base-pair transversion C to T was viewed as the RNA-edited sites. In order to validate the prediction, the expressed sequence tags (EST) sequences of six species obtained from the Compositae Genome Project Database (<http://compgenomics.ucdavis.edu/>) were searched against the predicted genes using BLAST tools.

3.4. Phylogenetic Analysis

A total of 81 cp genes from 66 taxa comprising five Panicoideae species, 58 other angiosperm lineages, and three gymnosperms were used to analyze the phylogenetic relationships among these species. All sequences were aligned using the database MSWAT (<http://mswat.cccb.utexas.edu/>) and edited manually. *Nuphar advena* and *Nymphaea alba* were designated as out-groups, and MEGA version 7 software [47] was used to perform maximum parsimony (MP) analysis with 1000 replicates and tree bisection–reconnection (TBR) branch swapping with the MulTrees option, and the GTR + I + G nucleotide substitution model was selected for the maximum likelihood (ML) analysis using PhyMLv3.0 [55].

4. Conclusions

This study reported the complete cp genome of broomcorn millet and a comparative analysis of six Panicoideae cp genomes was performed to reveal the genome features and evolutionary significance. Furthermore, the SSR composition, RNA editing sites, and phylogenetic relationships of these cp genomes were systematically investigated and compared. This not only provided important information on the genome organization, gene content, and structural variation of broomcorn millet and other Panicoideae cp genomes, but also provided important clues and tools for phylogenetic studies in the Panicoideae subfamily, which could be useful for molecular phylogeny and population studies within this subfamily and beyond.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4395/8/9/159/s1>. Table S1: The SSR abundance and composition in broomcorn millet and other five Panicoideae species; Table S2: The potential SSR marker developed for Panicoideae species identification.

Author Contributions: X.N. performed the entire analysis and drafted the paper. X.Z. and S.W. contributed to genome assembly and comparative analysis. C.L. and H.L. performed the RNA editing analysis. W.T. contributed reagents/materials/analysis tools and revised the manuscript. Y.G. conceived the entire study. All authors read and approved the manuscript.

Acknowledgments: We are grateful to Mengxing Wang and Hong Yue for helping to draw the gene map of broomcorn millet chloroplast genome, and also to Zhonghe Du for his help with the phylogenetic analysis. This work was mainly supported by the Fundamental Research Funds for the Central Universities of China (Grant No. 2452018101).

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- Gray, M.W. The evolutionary origin of organelles. *Trends Genet.* **1989**, *5*, 294–299. [[CrossRef](#)]
- Neuhaus, H.E.; Emes, M.J. Nonphotosynthetic metabolism in plastids. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2000**, *51*, 111–140. [[CrossRef](#)] [[PubMed](#)]
- Howe, C.J.; Barbrook, A.C.; Koumandou, V.L.; Nisbet, R.E.R.; Symington, H.A.; Wightman, T.F. Evolution of the chloroplast genome. *Phil. Trans. R. Soc. Lond. B* **2003**, *358*, 99–107. [[CrossRef](#)] [[PubMed](#)]
- Fu, P.C.; Zhang, Y.Z.; Geng, H.M.; Chen, S.L. The complete chloroplast genome sequence of *Gentiana lawrencei* var. *farreri* (Gentianaceae) and comparative analysis with its congeneric species. *PeerJ* **2016**, *4*, e2540. [[CrossRef](#)] [[PubMed](#)]
- Jansen, R.K.; Raubeson, L.A.; Boore, J.L.; Chumley, T.W.; Haberle, R.C.; Wyman, S.K.; Alverson, A.J.; Peery, R.; Herman, S.J.; Fourcade, H.M.; et al. Methods for obtaining and analyzing whole chloroplast genome sequences. In *Methods in Enzymology*; Elsevier Academic Press: Amsterdam, The Netherlands, 2005; Volume 395, pp. 348–384.
- 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 × 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](#)]
- Cai, Z.; Guisinger, M.; Kim, H.G.; Ruck, E.; Blazier, J.C.; McMurtry, V.; Kuehl, J.V.; Boore, J.; Jansen, R.K. Extensive reorganization of the plastid genome of *Trifolium subterraneum* (Fabaceae) is associated with numerous repeated sequences and novel DNA insertions. *J. Mol. Evol.* **2008**, *67*, 696–704. [[CrossRef](#)] [[PubMed](#)]
- Provan, J.; Powell, W.; Hollingsworth, P.M. Chloroplast microsatellites: New tools for studies in plant ecology and evolution. *Trends Ecol. Evol.* **2001**, *16*, 142–147. [[CrossRef](#)]
- Nie, X.; Lv, S.; Zhang, Y.; Du, X.; Wang, L.; Biradar, S.S.; Tan, X.; Wan, F.; Weining, S. Complete chloroplast genome sequence of a major invasive species, crofton weed (*Ageratina adenophora*). *PLoS ONE* **2012**, *7*, e36869. [[CrossRef](#)] [[PubMed](#)]
- Ruhfel, B.R.; Gitzendanner, M.A.; Soltis, P.S.; Soltis, D.E.; Burleigh, J.G. From algae to angiosperms—inferring the phylogeny of green plants (*Viridiplantae*) from 360 plastid genomes. *BMC Evol. Biol.* **2014**, *14*, 23. [[CrossRef](#)] [[PubMed](#)]

11. Williams, A.V.; Miller, J.T.; Small, I.; Nevill, P.G.; Boykin, L.M. Integration of complete chloroplast genome sequences with small amplicon datasets improves phylogenetic resolution in *Acacia*. *Mol. Phylogenet. Evol.* **2016**, *96*, 1–8. [[CrossRef](#)] [[PubMed](#)]
12. Moore, M.J.; Dhingra, A.; Soltis, P.S.; Shaw, R.; Farmerie, W.G.; Folta, K.M.; Soltis, D.E. Rapid and accurate pyrosequencing of angiosperm plastid genomes. *BMC Plant Biol.* **2016**, *6*, 1–13.
13. Saarela, J.M.; Wysocki, W.P.; Barrett, C.F.; Soreng, R.J.; Davis, J.I.; Clark, L.G.; Kelchner, S.A.; Pires, J.C.; Edger, P.P.; Mayfield, D.R.; et al. Plastid phylogenomics of the cool-season grass subfamily: Clarification of relationships among early-diverging tribes. *Aob Plants* **2015**, *7*, plv046. [[CrossRef](#)] [[PubMed](#)]
14. Hiratsuka, J.; Shimada, H.; Whittier, R.; Ishibashi, T.; Sakamoto, M.; Mori, M.; Kondo, C.; Honji, Y.; Sun, C.R.; Meng, B.Y.; et al. The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **1989**, *217*, 185–194. [[CrossRef](#)] [[PubMed](#)]
15. Young, H.A.; Lanzatella, C.L.; Sarath, G.; Tobias, C.M. Chloroplast genome variation in upland and lowland switchgrass. *PLoS ONE* **2011**, *6*, e23980. [[CrossRef](#)] [[PubMed](#)]
16. Huang, J.; Chen, R.; Li, X. Comparative Analysis of the Complete Chloroplast Genome of Four Known *Ziziphus* Species. *Genes* **2017**, *8*, 340. [[CrossRef](#)] [[PubMed](#)]
17. Asaf, S.; Khan, A.L.; Khan, A.R.; Waqas, M.; Kang, S.M.; Khan, M.A.; Lee, S.M.; Lee, I.J. Complete chloroplast genome of *Nicotiana glauca* and its comparison with related species. *Front. Plant Sci.* **2016**, *7*, 843. [[CrossRef](#)] [[PubMed](#)]
18. Asano, T.; Tsudzuki, T.; Takahashi, S.; Shimada, H.; Kadowaki, K.I. Complete nucleotide sequence of the sugarcane (*Saccharum officinarum*) chloroplast genome: a comparative analysis of four monocot chloroplast genomes. *DNA Res.* **2004**, *11*, 93. [[CrossRef](#)] [[PubMed](#)]
19. Kellogg, E.A. Subfamily Panicoideae Link. In *Flowering Plants. Monocots: Poaceae*; Springer international publishing: Zurich, Switzerland, 2015; Volume 13, pp. 271–345.
20. Hitchcock, A.S.; Chase, A. *Manual of the grasses of the United States*; Dover Publications: New York, NY, USA, 1971; Volume 2.
21. Palmer, N.A.; Saathoff, A.J.; Waters, B.M.; Donze, T.; Heng-Moss, T.M.; Twigg, P.; Tobias, C.M.; Sarath, G. Global changes in mineral transporters in tetraploid switchgrasses (*Panicum virgatum* L.). *Front. Plant Sci.* **2014**, *4*, 549. [[CrossRef](#)] [[PubMed](#)]
22. Saski, C.; Lee, S.B.; Fjellheim, S.; Guda, C.; Jansen, R.K.; Luo, H.; Tomkins, J.; Rognli, O.A.; Daniell, H.; Clarke, J.L. Complete chloroplast genome sequences of *Hordeum vulgare*, *Sorghum bicolor* and *Agrostis stolonifera*, and comparative analyses with other grass genomes. *Theor. Appl. Genet.* **2007**, *115*, 571–590. [[CrossRef](#)] [[PubMed](#)]
23. Tsang, C.H.; Li, K.T.; Hsu, T.F.; Tsai, Y.C.; Fang, P.H.; Hsing, Y.I.C. Broomcorn and foxtail millet were cultivated in Taiwan about 5000 years ago. *Bot. Stud.* **2017**, *58*, 3. [[CrossRef](#)] [[PubMed](#)]
24. Sánchez-Ken, J.G.; Clark, L.G.; Kellogg, E.A.; Kay, E.E. Reinstatement and emendation of subfamily Micrairoideae (Poaceae). *Syst. Bot.* **2007**, *32*, 71–80. [[CrossRef](#)]
25. Washburn, J.D.; Schnable, J.C.; Davidse, G.; Pires, J.C. Phylogeny and photosynthesis of the grass tribe Paniceae. *Am. J. Bot.* **2015**, *102*, 1493–1505. [[CrossRef](#)] [[PubMed](#)]
26. Giussani, L.M.; Cota-Sánchez, J.H.; Zuloaga, F.O.; Kellogg, E.A. A molecular phylogeny of the grass subfamily Panicoideae (Poaceae) shows multiple origins of C4 photosynthesis. *Am. J. Bot.* **2001**, *88*, 1993–2012. [[CrossRef](#)] [[PubMed](#)]
27. Sánchez-Ken, J.G.; Clark, L.G. Phylogeny and a new tribal classification of the Panicoideae s.l. (Poaceae) based on plastid and nuclear sequence data and structural data. *Am. J. Bot.* **2010**, *97*, 1732–1748. [[CrossRef](#)] [[PubMed](#)]
28. Soreng, R.J.; Peterson, P.M.; Romaschenko, K.; Davidse, G.; Zuloaga, F.O.; Judziewicz, E.J.; Filgueiras, T.S.; Davis, J.I.; Morrone, O. A worldwide phylogenetic classification of the Poaceae (Gramineae). *J. Syst. Evol.* **2015**, *53*, 117–137. [[CrossRef](#)]
29. Burke, S.V.; Wysocki, W.P.; Zuloaga, F.O.; Craine, J.M.; Pires, J.C.; Edger, P.P.; Mayfield-Jones, D.; Clark, L.G.; Kelchner, S.A.; Duvall, M.R. Evolutionary relationships in Panicoid grasses based on plastome phylogenomics (Panicoideae; Poaceae). *BMC Plant Biol.* **2016**, *16*, 1–11. [[CrossRef](#)] [[PubMed](#)]

30. Hunt, H.V.; Denyer, K.; Packman, L.C.; Jones, M.K.; Howe, C.J. Molecular basis of the waxy endosperm starch phenotype in broomcorn millet (*Panicum miliaceum* L.). *Mol. Biol. Evol.* **2010**, *27*, 1478–1494. [[CrossRef](#)] [[PubMed](#)]
31. Liu, M.X.; Xu, Y.; He, J.H.; Zhang, S.; Wang, Y.Y.; Lu, P. Genetic diversity and population structure of broomcorn millet (*Panicum miliaceum* L.) cultivars and landraces in china based on microsatellite markers. *Int. J. Mol. Sci.* **2016**, *17*, 370. [[CrossRef](#)] [[PubMed](#)]
32. Yue, H.; Wang, L.; Liu, H.; Yue, W.J.; Du, X.H.; Song, W.N.; Nie, X.J. De novo assembly and characterization of the transcriptome of broomcorn millet (*Panicum miliaceum* L.) for gene discovery and marker development. *Front. Plant Sci.* **2016**, *7*, 1083. [[CrossRef](#)] [[PubMed](#)]
33. Yue, H.; Wang, M.; Liu, S.Y.; Du, X.H.; Song, W.N.; Nie, X.J. Transcriptome-wide identification and expression profiles of the WRKY transcription factor family in Broomcorn millet (*Panicum miliaceum* L.). *BMC Genom.* **2016**, *17*, 343. [[CrossRef](#)] [[PubMed](#)]
34. Bosacchi, M.; Gurdon, C.; Maliga, P. Plastid Genotyping Reveals Uniformity of cms-T Maize Cytoplasms. *Plant Physiol.* **2015**, *169*, 2129–2137. [[CrossRef](#)] [[PubMed](#)]
35. Wang, S.; Gao, L.Z. Complete chloroplast genome sequence of green foxtail (*Setaria viridis*), a promising model system for C₄ photosynthesis. *Mitochondrial DNA Part A* **2016**, *27*, 3707–3708. [[CrossRef](#)] [[PubMed](#)]
36. Zhang, Y.R.; Nie, X.J.; Jia, X.O.; Zhao, C.Z.; Biradar, S.S.; Wang, L.; Du, X.H.; Weining, S. Analysis of codon usage patterns of the chloroplast genomes in the Poaceae family. *Aust. J. Bot.* **2012**, *60*, 461–470. [[CrossRef](#)]
37. Maier, R.M.; Neckermann, K.; Igloi, G.L.; Kössel, H. Complete sequence of the maize chloroplast genome: Gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *J. Mol. Biol.* **1995**, *251*, 614–628. [[CrossRef](#)] [[PubMed](#)]
38. Guisinger, M.M.; Chumley, T.W.; Kuehl, J.V.; Boore, J.L.; Jansen, R.K. Implications of the Plastid Genome Sequence of Typha (Typhaceae, Poales) for Understanding Genome Evolution in Poaceae. *J. Mol. Evol.* **2010**, *70*, 149–166. [[CrossRef](#)] [[PubMed](#)]
39. Blaner, A.; Schneider, J. Phylogenetic relationships in the grass family (Poaceae) based on the nuclear single copy locus topoisomerase 6 compared with chloroplast DNA. *Syst. Biodivers.* **2014**, *12*, 111–124. [[CrossRef](#)]
40. Howe, C.J.; Barker, R.F.; Bowman, C.M.; Dyer, T.A. Common features of three inversions in wheat chloroplast DNA. *Curr. Genet.* **1988**, *13*, 343–349. [[CrossRef](#)] [[PubMed](#)]
41. Doyle, J.J.; Davis, J.I.; Soreng, R.J.; Garvin, D.; Anderson, M.J. Chloroplast DNA inversions and the origin of the grass family (Poaceae). *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 7722–7726. [[CrossRef](#)] [[PubMed](#)]
42. Kellogg, E.A.; Aliscioni, S.; Morrone, O.; Pensiero, J.; Zuloaga, F. A Phylogeny of Setaria (Poaceae, Panicoideae, Paniceae) and Related Genera Based on the Chloroplast Gene ndhF. *Int. J. Plant Sci.* **2009**, *170*, 117–131. [[CrossRef](#)]
43. Nybom, H.; Weising, K.; Rotter, B. DNA fingerprinting in botany: Past, present, future. *Investig. Genet.* **2014**, *5*, 1–35. [[CrossRef](#)] [[PubMed](#)]
44. Hein, A.; Polsakiewicz, M.; Knoop, V. Frequent chloroplast RNA editing in early-branching flowering plants: Pilot studies on angiosperm-wide coexistence of editing sites and their nuclear specificity factors. *BMC Evol. Biol.* **2016**, *16*, 23. [[CrossRef](#)] [[PubMed](#)]
45. Wang, M.X.; Liu, H.; Ge, L.Q.; Xing, G.W.; Wang, M.; Song, W.N.; Nie, X.J. Identification and analysis of RNA editing sites in the chloroplast transcripts of *Aegilops tauschii* L. *Genes* **2017**, *8*, 13. [[CrossRef](#)] [[PubMed](#)]
46. Hirose, T.; Kusumegi, T.; Tsudzuki, T.; Sugiura, M. RNA editing sites in tobacco chloroplast transcripts: Editing as a possible regulator of chloroplast RNA polymerase activity. *Mol. Gen. Genet.* **1999**, *262*, 462–467. [[CrossRef](#)] [[PubMed](#)]
47. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)] [[PubMed](#)]
48. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)] [[PubMed](#)]
49. Luo, R.; Liu, B.; Xie, Y.; Li, Z.; Huang, W.; Yuan, J.; He, G.; Chen, Y.; Pan, Q.; Liu, Y.; et al. SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *GigaScience* **2012**, *1*, 18. [[CrossRef](#)] [[PubMed](#)]
50. Wyman, S.K.; Jansen, R.K.; Boore, J.L. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* **2014**, *20*, 3252–3255. [[CrossRef](#)] [[PubMed](#)]

51. Sugita, M.; Sugiura, M. Regulation of gene expression in chloroplasts of higher plants. *Plant Mol. Biol.* **1996**, *32*, 315–326. [[CrossRef](#)] [[PubMed](#)]
52. 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](#)] [[PubMed](#)]
53. 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](#)] [[PubMed](#)]
54. 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](#)] [[PubMed](#)]
55. 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](#)] [[PubMed](#)]

© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).