



Article

Assembly of the Complete Mitochondrial Genome of *Pereskia aculeata* Revealed That Two Pairs of Repetitive Elements Mediated the Recombination of the Genome

Xue Zhang ^{1,†}, Yuanyu Shan ^{1,†}, Jingling Li ¹ , Qiulin Qin ¹, Jie Yu ^{1,2,*} and Hongping Deng ^{3,*}

¹ College of Horticulture and Landscape Architecture, Southwest University, Chongqing 400716, China

² Key Laboratory of Agricultural Biosafety and Green Production of Upper Yangtze River, Ministry of Education, Southwest University, Chongqing 400715, China

³ Center for Biodiversity Conservation and Utilization, School of Life Sciences, Southwest University, Chongqing 400715, China

* Correspondence: yujie1982@swu.edu.cn (J.Y.); denghp@swu.edu.cn (H.D.)

† These authors contributed equally to this work.

Abstract: *Pereskia aculeata* is a potential new crop species that has both food and medicinal (antinociceptive activity) properties. However, comprehensive genomic research on *P. aculeata* is still lacking, particularly concerning its organelle genome. In this study, *P. aculeata* was studied to sequence the mitochondrial genome (mitogenome) and to ascertain the assembly, informational content, and developmental expression of the mitogenome. The findings revealed that the mitogenome of *P. aculeata* is circular and measures 515,187 bp in length with a GC content of 44.05%. It contains 52 unique genes, including 33 protein-coding genes, 19 tRNA genes, and three rRNA genes. Additionally, the mitogenome analysis identified 165 SSRs, primarily consisting of tetra-nucleotides, and 421 pairs of dispersed repeats with lengths greater than or equal to 30, which were mainly forward repeats. Based on long reads and PCR experiments, we confirmed that two pairs of long-fragment repetitive elements were highly involved with the mitogenome recombination process. Furthermore, there were 38 homologous fragments detected between the mitogenome and chloroplast genome, and the longest fragment was 3962 bp. This is the first report on the mitogenome in the family Cactaceae. The decoding of the mitogenome of *P. aculeata* will provide important genetic materials for phylogenetic studies of Cactaceae and promote the utilization of species germplasm resources.

Keywords: mitochondrial plastid DNAs; mitogenome; *Pereskia aculeata*; recombination; repetitive elements; RNA editing



Citation: Zhang, X.; Shan, Y.; Li, J.; Qin, Q.; Yu, J.; Deng, H. Assembly of the Complete Mitochondrial Genome of *Pereskia aculeata* Revealed That Two Pairs of Repetitive Elements Mediated the Recombination of the Genome. *Int. J. Mol. Sci.* **2023**, *24*, 8366. <https://doi.org/10.3390/ijms24098366>

Academic Editor: Frank M. You

Received: 13 April 2023

Revised: 1 May 2023

Accepted: 3 May 2023

Published: 6 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Pereskia aculeata Mill. is a climbing shrub plant [1] that belongs to the family Cactaceae. This species is native to South America and mainly in southern to northeastern Brazil [2]. Its leaves have a high nutritional value and are eaten by local people as a vegetable, known as “Barbados Gooseberry” [3]. For instance, the content of protein in the leaves of *P. aculeata* is much higher than that of other common vegetables [4]. In many economically underdeveloped areas, people often rely on it as the main source of protein in everyday life, so *P. aculeata* is also known as the “meat of the poor”. Moreover, in addition to minerals such as calcium, magnesium, manganese, and zinc, the leaves are also high in total dietary fiber, as well as vitamin A, vitamin C, and folate [5]. Apart from this, *P. aculeata* also has medical value. The leaves of *P. aculeata* are very high in mucus, which is often used in traditional Brazilian medicine to heal skin wounds and treat inflammation [6]. Furthermore, the functional mucilage extracted from *P. aculeata* has been shown to be rich in polysaccharides and proteins and can be used as an emulsifier and stabilizer in the food industry [7–9], also used as a corrosion inhibitor of steel [10].

The mitochondrion is a semi-autonomous eukaryotic organelle that has a small genome (mitogenome) that interacts with the nuclear and chloroplast genome to provide the biochemical machinery for energy conversion. Mitochondria synthesize Adenosine Triphosphate (ATP) through the tricarboxylic acid cycle and oxidative phosphorylation to provide energy for cells [11,12]. Mitochondria are also involved in cell differentiation, information transmission, and apoptosis [13]. Studies have shown that mitochondria are involved in maternally inherited cytoplasmic male sterility (CMS) [14]. Modern plant breeders focus on the process of rearranging the mitogenome to restore fertility in plants [15].

The plant mitogenome is more structurally complex and difficult to sequence compared to the chloroplast genome and is still in a booming stage of development [16]. By January 2023, The National Center for Biotechnology Information (NCBI) has released 9537 chloroplast genomes, 562 mitogenomes, and 1279 plastid genomes (<https://www.ncbi.nlm.nih.gov/genome/browse#!/organelles/>; accessed on 12 January 2023). Although the mitogenome of plants generally shows a circular genome structure [17], the physical organization of the mitogenomes includes a set of sub-genomic forms resulting from homologous recombination between repeats [18], with a mixture of linear, circular and branched structures [19]. For all organisms, even viruses, recombination is an important process in DNA replication [20–22] because recombination activities are involved in the rescuing of stalled or damaged replication forks. The recombination mechanism can be divided into two pathways: homologous recombination (HR) and non-homologous recombination [23]. HR is the main pathway of plant mitochondrial DNA repair [24], and it relies on the identification of sequences with high similarity to the region to be repaired. Non-homologous recombination pathways utilize very limited or non-sequence similarities for the repair of broken DNA. It mainly occurs in eukaryotes, which can lead to deletions or the duplications of sequences [25,26].

It is important to note that the regulation of HR in plant mitochondria is crucial to avoiding the occurrence of deleterious genomic rearrangements. Factors involved in HR or its regulation are sensitive and should be well understood to prevent instability and damage to the mitochondria. The endosymbiosis theory postulates that mitochondria arise from an endosymbiotic alpha-proteobacterium within an archaeal-derived host cell and finally evolved into organelles of eukaryotic cells [27]. Thus, the plant mitochondrial HR is similar to bacterial HR [28–30]. It has been suggested that mitochondrial HR is a frequent and reversible recombination involving large repeats. If these changes do not harm the function of the mitochondria, they will be retained, resulting in an overall increase in mitogenome size [31].

Although nearly 500 plant mitogenomes have been reported since 1990, there are no mitogenome studies of plants from the cactus family. Here, we completed the assembly and annotation of the mitogenome of *P. aculeata* and revealed its genomic characteristics and structural features. We analyzed relative synonymous codon usage (RSCU) and repeat sequences and discussed the possibility of chloroplast DNA transfer into the mitogenome. We also investigated RNA editing sites of protein-coding genes (PCGs) and explored their phylogenetic relationships. The results reported here provide a unique insight into the mitochondrial evolution of a cactus species, and it is also a basis for developing available genetic resources and the molecular marker-assisted breeding of *P. aculeata*.

2. Results

2.1. Mitogenome Structure of *P. aculeata*

When the mitogenome of *P. aculeata* was assembled, a unitig graph mediated by two repeats was generated (Figure 1A). Among them, the length of R1 is 9026 bp, and that of R2 is 7040 bp. The two repeat sequences each have four paths (p1–p4 and p5–p8, Figure 1B,E). We verified the existence of these paths in the mitogenome of *P. aculeata* using a polymerase chain reaction (PCR) experiment. In this experiment, we designed a total of six pairs of primers. The first three pairs of primers (F1 + R1, F2 + R1, F3 + R3, F3 + R2) are used to verify repeated sequence R1, and the rest (F4 + R4, F5 + R5, F6 + R4,

F5 + R6) are used to verify repeated sequence R2. The distribution diagram and sequence information of these primers are shown in Figure 1C,F and Table S2, respectively. The PCR products were consistent with the expected result (Figure 1D,G). The PCR-based experiments showed that the assembly here was correct and that there were two large repeating elements in this mitogenome.

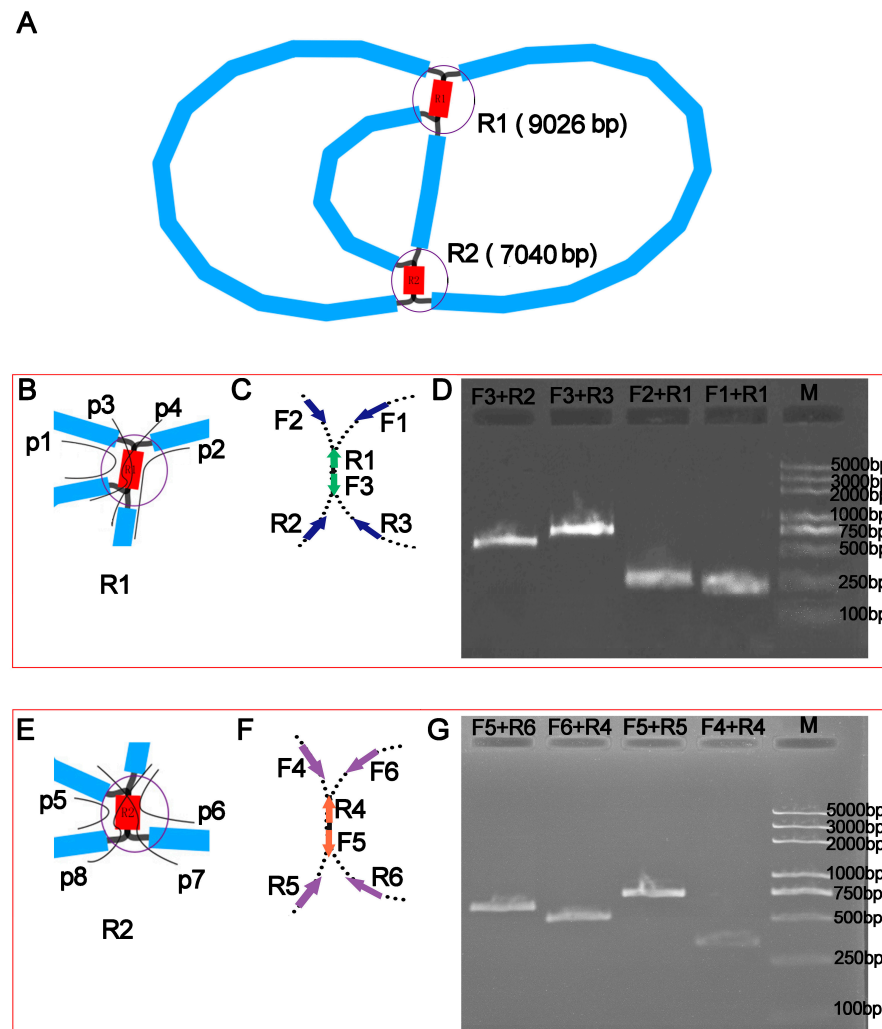


Figure 1. Assembly graph of *P. aculeata* mitogenome and schematic diagram of the designed specific primer pairs. (A) Assembly graph visualized using Bandage. The assembly contains six contigs that include two DRs (R1 and R2, highlighted in red). R1 has four paths (p1–p4), as shown in panel (B). Panel (C) shows the design of three pairs of primers for the PCR experiments. These three pairs are used to verify the repeat sequence R1 (9026 bp). Among them, F1, F2, R2 and R3 are represented by the blue arrow. On the contrary, R1 and F3 are represented by the green arrow. Panel (D) shows the results of agarose gel electrophoresis, as expected. R2 has four paths (p5–p8), as shown in panel (E). Panel (F) shows the design of three pairs of primers for the PCR experiments. These three pairs are used to verify the repeat sequence R2 (7040 bp). Among them, F4, F6, R5 and R6 are represented by the blue arrow. On the contrary, R4 and F5 are represented by the green arrow. Panel (G) are the results of agarose gel electrophoresis, also as expected.

Long reads were subsequently used to further explore the authenticity of these paths (p1–p8) and their specific proportions. As shown in Table S3, the eight paths supported by long reads are as follows: p1 (27), p2 (23), p3 (23), p4 (14), p5 (34), p6 (35), p7 (32) and p8 (18). The number in parentheses indicates the number of long reads that support this path. This further confirmed the existence of multiple mitogenome conformations. Based on this,

we drew the possible isomers, as shown in Figure 2. In this study, we use the structure shown in Figure 2A for the downstream analysis, which is supported by most long reads.

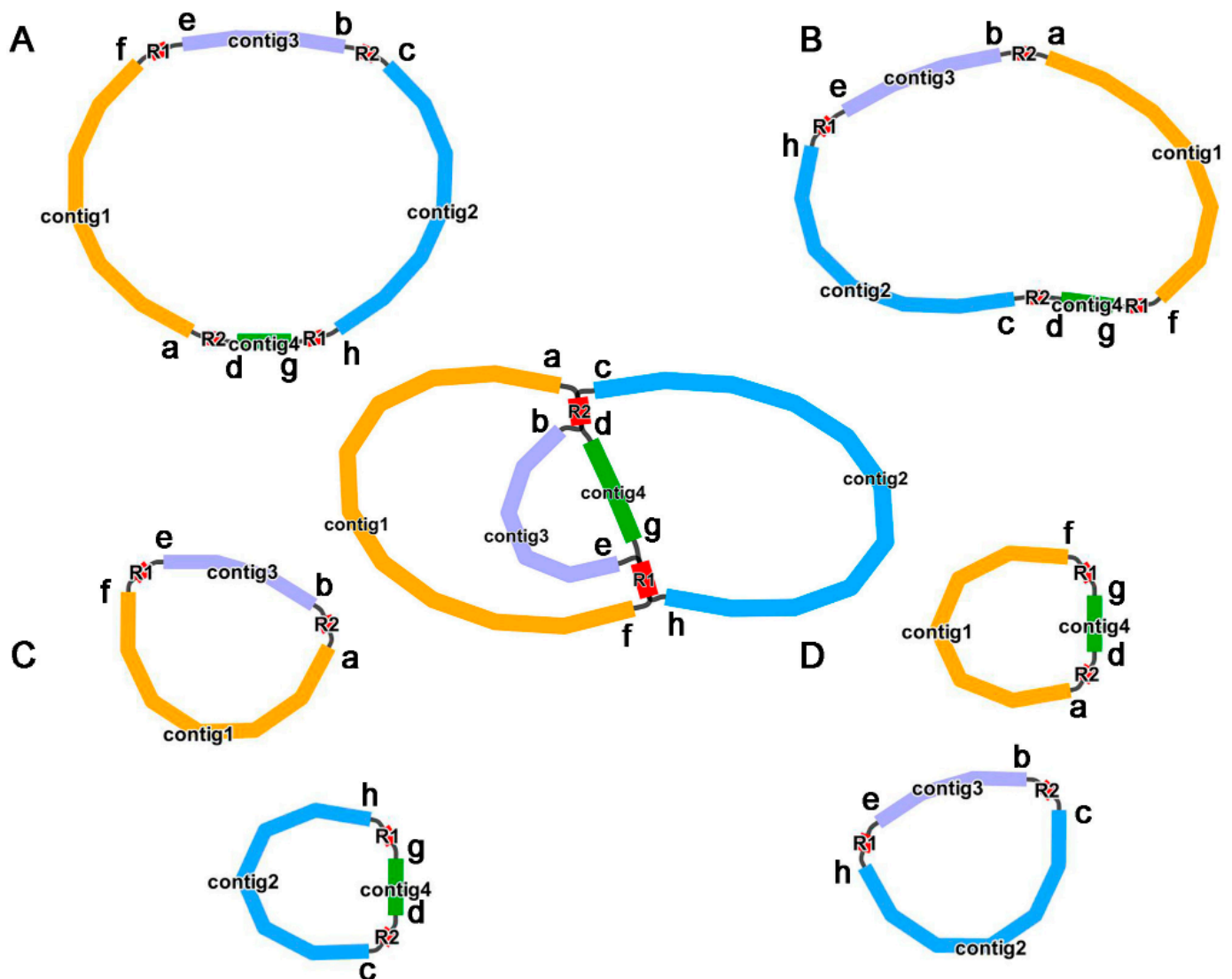


Figure 2. Potential isomers of *P. aculeata* mitogenome inferred from PCR experiments and long reads. The initial assembly is shown in the middle of the figure, with (A–D) representing the four possible isomers formed after solving the paths of the two pairs of repeating regions (R1 and R2). We use lowercase letters (a–h) to mark breakpoints at several nodes to describe several paths, where (e,f) stands for p1, (g,h) stands for p2, (e–h) stands for p3, (f,g) stands for p4, (a,b) stands for p5, (c,d) stands for p6, (a–d) stands for p7 and (b,c) stands for p8. These isomers contain two ‘master circle’ structures (A,B), which differ in sequence order, and two structures of two separate small circles (C,D). In this study, we use the structure shown in panel (A) for the downstream analysis, which is supported by most long reads.

Overall, *P. aculeata* mitogenome is a putative circular DNA with a length of 515,187 bp, and the GC content is 44.05% (Figure 3). Two direct repeats (DRs) with lengths of 9026 bp and 7040 bp. The accuracy of the mitogenome assembly was confirmed by mapping the ONT long reads (average 75-fold depth) onto the assembly (Figure S1).

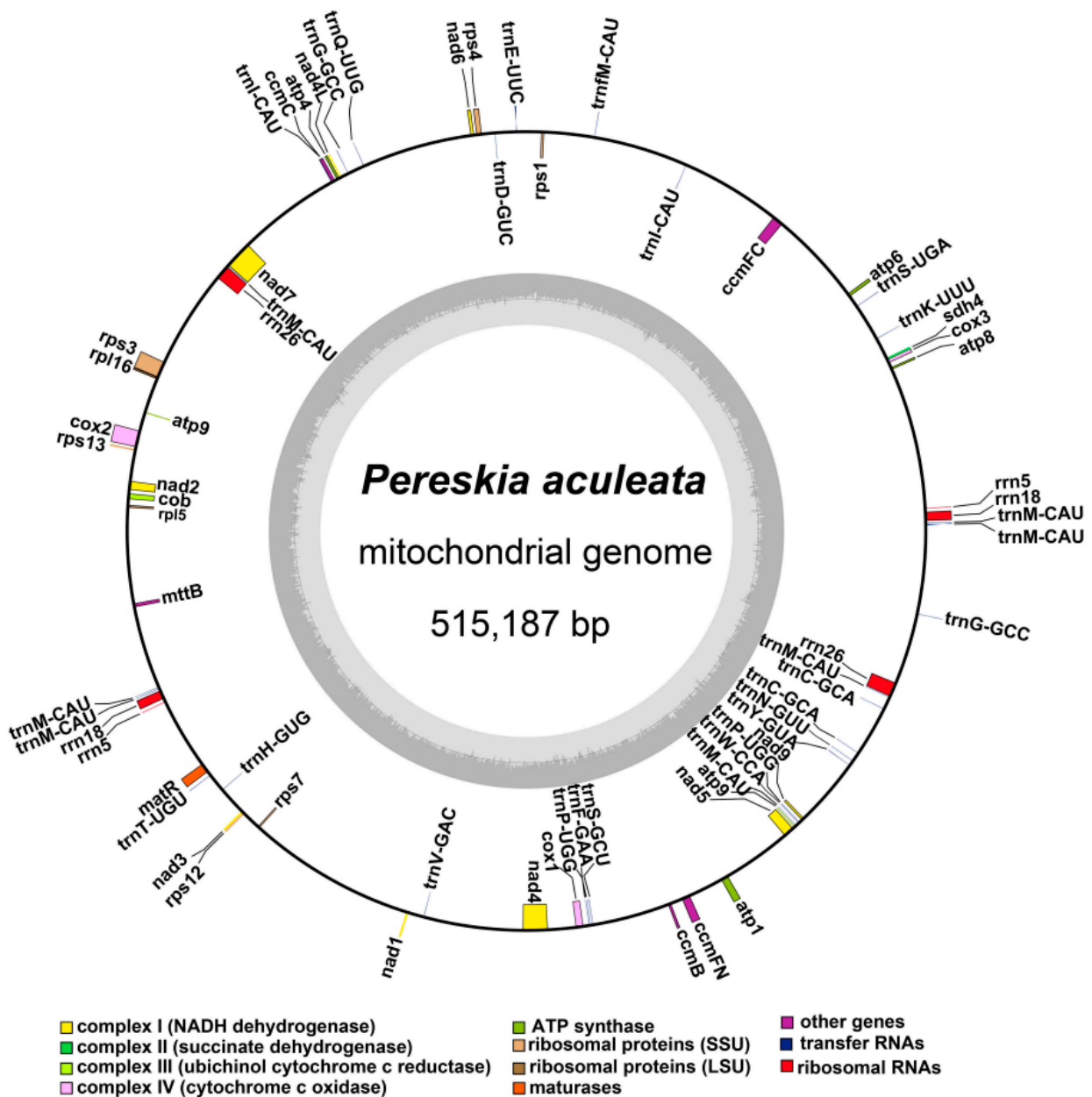


Figure 3. The putative circular mitogenome maps of *P. aculeata*. The genomic features inside and outside the circle represent the clockwise and counterclockwise chains on the transcription, respectively. Different color blocks represent different functional gene groups.

2.2. Mitochondrial Gene Content

A total of 33 unique protein-coding genes (PCGs) were annotated (Table 1), including 24 core genes and nine variable genes. The core genes included five ATP synthase genes (*atp1*, *atp4*, *atp6*, *atp8*, and *atp9*), nine NADH dehydrogenase genes (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, and *nad9*), four cytochrome c biogenesis genes (*ccmB*, *ccmC*, *ccmFC*, and *ccmFN*), three cytochrome c oxidase genes (*cox1*, *cox2*, and *cox3*), one membrane transport protein gene (*mttB*), one mature enzyme gene (*matR*), and one ubiquinol cytochrome c reductase gene (*cob*). The variable genes included two large subunits of ribosomal protein (*rpl5* and *rpl16*), six small subunits of ribosomal protein (*rps1*, *rps3*, *rps4*, *rps7*, *rps12*, and *rps13*), and one succinate dehydrogenase (*sdh4*). Among these, the genes *atp9* and *nad4* had two copies.

Table 1. Gene composition in the mitogenome of *P. aculeata*.

Group of Genes	Name of Genes
ATP synthase	<i>atp1, atp4, atp6, atp8, atp9</i> (2)
NADH dehydrogenase	<i>nad1, nad2, nad3, nad4</i> (2), <i>nad4L, nad5, nad6, nad7, nad9</i>
Ubichinol cytochrome c reductase	<i>cob</i>
Cytochrome c biogenesis	<i>ccmB, ccmC, ccmFC, ccmFN</i>
Cytochrome c oxidase	<i>cox1, cox2, cox3</i>
Maturases	<i>matR</i>
Transport membrane protein	<i>mttB</i>
Large subunit of ribosome	<i>rpl5, rpl16</i>
Small subunit of ribosome	<i>rps1, rps3, rps4, rps12, rps13, rps7</i>
Succinate dehydrogenase	<i>sdh4</i>
Ribosome RNAs	<i>rrn18</i> (2), <i>rrn26</i> (2), <i>rrn5</i> (2)
Transfer RNAs	<i>trnC-GCA</i> (2), <i>trnD-GUC, trnE-UUC, trnF-GAA, trnM-CAU</i> (7), <i>trnH-GUG, trnI-CAU</i> (2), <i>trnK-UUU, trnN-GUU, trnP-UGG</i> (2), <i>trnQ-UUG, trnS-GCU, trnS-UGA, trnV-GAC, trnW-CCA, trnY-GUA,</i> <i>trnG-GCC</i> (2), <i>trnT-UGU, trnfM-CAU</i>

Note. The numbers in parentheses represent copy numbers of genes.

A total of 29 tRNA genes were annotated based on tRNAscan-SE. Based on sequence similarity and a previous report [32], we distinguished these tRNA genes transferred from plastid to mitochondria in the *P. aculeata* mitogenome. Firstly, a total of 13 tRNA genes are native to mitochondria, including *trnC-GCA* (mt), *trnE-UUC* (mt), *trnF-GAA* (mt), two copies of *trnG-GCC* (mt), *trnK-UUU* (mt), *trnI-CAU* (mt), *trnfM-CAU* (mt), *trnP-UGG* (mt), *trnQ-UUG* (mt), *trnS-GCU* (mt), *trnS-UGA* (mt), and *trnY-GUA* (mt). Secondly, nine tRNA genes were highly homologous to the chloroplast's tRNA genes in sequences, including *trnD-GUC* (cp), *trnH-GUG* (cp), *trnN-GUU* (cp), *trnV-GAC* (cp), *trnW-CCA* (cp), two copies of *trnM-CAU* (cp), *trnP-UGG* (cp), and *trnI-CAU* (cp). This suggests that these sequences were transferred between organelles during evolution. Notably, we also found a bacterial-origin tRNA gene (*trnC-GCA*) with a high degree of sequence homology with previously reported genes [33,34]. Lastly, gene *trnT-UGU* and the remaining five copies of *trnM-CAU* are from unknown sources.

Based on the BLASTn program, we annotated three unique rRNA genes, *rrn5*, *rrn18*, and *rrn26*, all have two copies. The detailed location of each gene is given in Table S4. Figure 3 shows the genome map of *P. aculeata*.

Previous studies have shown that the number of intron-contained genes varies among the mitogenomes of land plants [35]. The introns of mitochondrial genes were analyzed, and 10 genes were found to have introns (Table S4). Among them, gene *ccmFC*, *rps3*, *trnI-CAU*, and *trnT-UGU* had one *cis*-intron; the genes *cox2* and *nad4* had two *cis*-introns; the gene *nad7* had four *cis*-introns; The gene *nad1* had one *cis*-intron and three *trans*-introns; the gene *nad5* had two *cis*-introns and two *trans*-introns; the gene *nad2* had three *cis*-introns and one *trans*-intron.

2.3. Codon Usage Analysis of PCGs

The eukaryotic genome has 64 codons that encode 20 different amino acids. All amino acids except Met and Try are encoded by multiple codons. Because of the degeneracy of codons, the usage of codons in different species varies greatly. Codon usage analysis was performed on 35 PCGs of *P. aculeata* mitogenome (Figure 4 and Table S5). Codons with an RSCU greater than 1 are used preferentially by amino acids [36,37], and it can be found that mitochondrial PCGs have a universal codon usage preference. For example, leucine (Leu) prefers UUA codons, and its RSCU value is the highest among mitochondrial PCGs, reaching 1.64. The second was alanine (Ala), with a preference for GCU (RSCU = 1.61). In addition, UUU (Phe), AUU (Ile), and UUA (Leu) are the three most common codons in *P. aculeata*. This is possibly a kind of preference produced by *P. aculeata* after a long period of evolutionary selection.

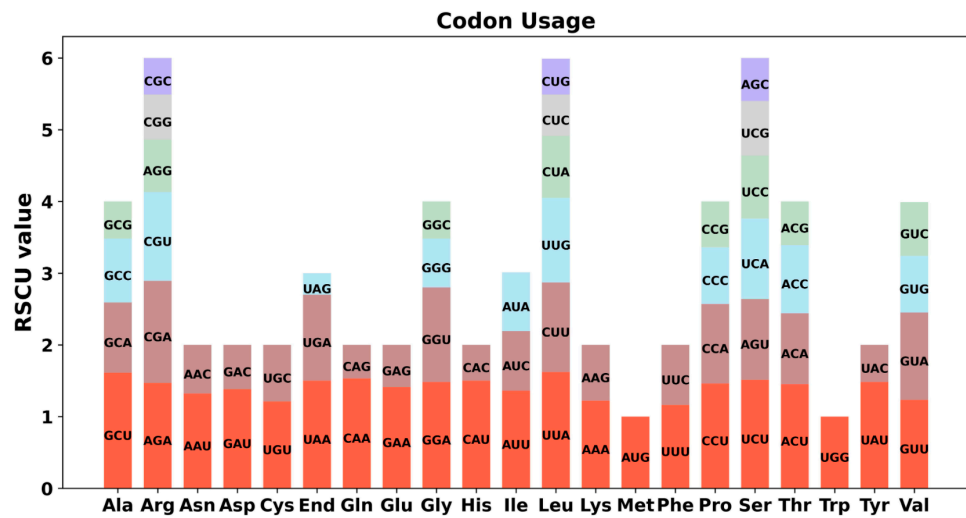


Figure 4. Codon usage bias of mitochondrial PCGs of *P. aculeata*. The RSCU refers to relative synonymous codon usage.

2.4. Repeat Elements

The SSRs are special tandem repeats, usually no more than 6 bp [38]. The SSRs are often used for molecular marker development [39,40]. A total of 165 SSRs were identified in *P. aculeata* mitogenome (Table S6), and 60 of them were mainly repeat units consisting of tetranucleotides, accounting for 36.36% of the total SSRs. The second was mononucleotide consisting of 44 repeat units, accounting for 26.67% of the total SSRs. In addition, 28 dinucleotide repeat units, 23 trinucleotide repeat units, eight pentanucleotide repeat units, and two hexanucleotide repeat units were detected. The number of SSRs is shown in Figure 5A.

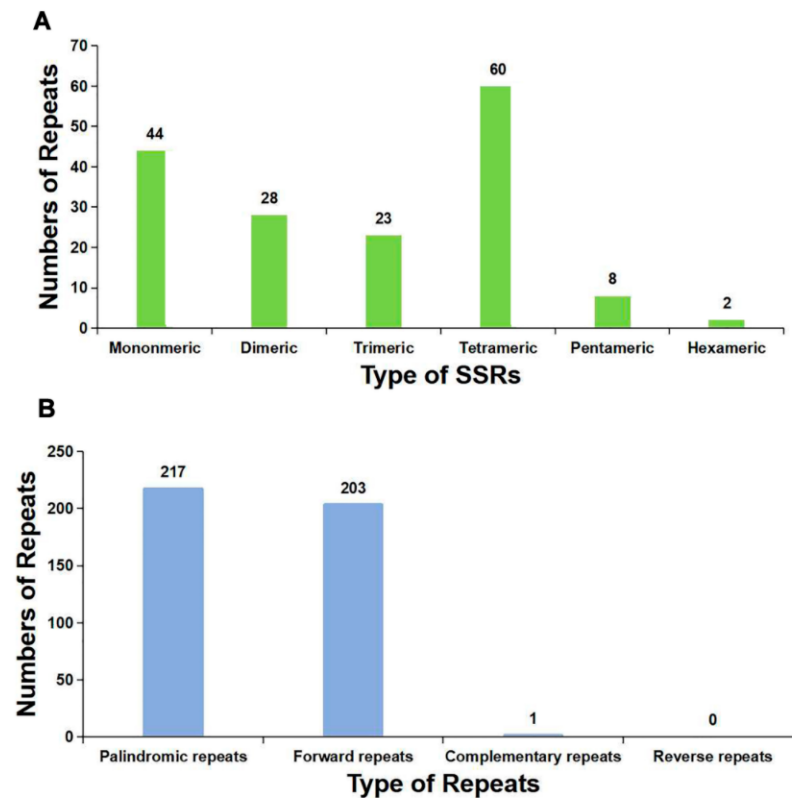


Figure 5. The SSRs and dispersed repeats identified in the mitogenomes of *P. aculeata*. (A) The SSRs identified in the *P. aculeata* mitogenomes. Each column represents different nucleotide repeated units displayed in different colors. (B) Dispersed repeats (≥ 30 bp) identified in the *P. aculeata* mitogenomes.

We identified the dispersed repeats in the mitogenome of *P. aculeata*. The forward and palindromic repeats account for the vast majority. A total of 421 pairs of repetitive sequences with lengths greater than or equal to 30 were observed, including 203 pairs of forward repeat, 217 pairs of palindromic repeat, and one pair of the complementary repeat (Table S7 and Figure 5B). No reverse repeats were detected in the mitogenome of *P. aculeata*. It has been reported that these repeats may promote genome rearrangements and affect genome size [41]. In the present study, most of these repeats are less than 100 bp in length, and only 23 repeats are longer than 100 bp. Among them, the longest forward repeat was 9026 bp, followed by 7040 bp, and they were both confirmed to be associated with the alternative conformations.

The total length of these dispersed repeats is 35,740 bp, accounting for 6.94% of the whole mitogenome of *P. aculeata*, which showed an abundance of repeats. These dispersed repeats can not only cause genome recombination but also induce mutation [42].

2.5. Plastid DNA Insertion in Mitogenome

We observed a lot of sequences being transferred from the chloroplast genome to its mitogenome in *P. aculeata*. According to sequence similarity analysis, there were 38 homologous fragments between the mitogenome and chloroplast genome, with a total length of 26,721 bp, accounting for 5.19% of the total mitogenome (Table S8). Among them, 10 fragments exceeded 1000 bp, we numbered them according to their length, and each started with 'MTPT'. MTPT1 is the longest at 3962 bp. These results indicate that the chloroplast genome provides abundant non-native sequences for its mitogenome. By annotating these homologous sequences, 15 complete genes were also found on 38 homologous fragments, including six plastidial PCGs (*rpoC1*, two *rpl2*, *petN*, *psbD*, and *ndhB*) and nine tRNA genes (*trnD-GUC*, two *trnM-CAU*, *trnI-CAU*, *trnW-CCA*, *trnP-UGG*, *trnH-GUG*, *trnV-GAC*, *trnN-GUU*). The schematic representation of the MTPTs is shown in Figure 6. In addition, there are also some plastid gene fragments found in the mitogenome, such as *rpl23*, *rbcL*, *rpoC2*, *rpoB*, *psbC*, *ycf3*, *atpA*, and *atpF*. However, no gene has been annotated on homologous fragments such as MTPT4, MTPT8, MTPT13, and so on. Detailed information about transferred genes can be found in Table S8.

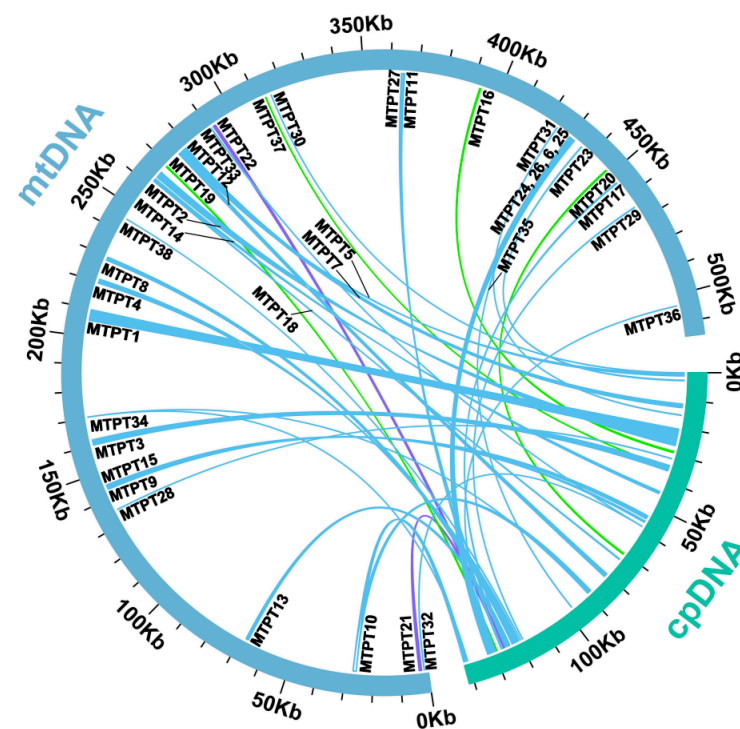


Figure 6. Schematic representation of homologous sequences between chloroplast genome and mitogenomes in *P. aculeata*. The blue arcs represent mitogenomes, the green arcs represent chloroplast

genomes, and the lines between arcs correspond to homologous genome segments. A total of 38 MTPTs were identified, and their numbers are given in the figure. Among them, lines with a percentage of sequence similarity above 90 are blue, those between 80 and 90 are green, and those less than 80 are purple. A total of 17 complete genes migrated to the mitogenome, including 8 PCGs and 9 tRNA genes. These complete genes were labeled next to the corresponding MTPT. Some homologous sequences are in the IR region of the plastome. We counted them according to unique locations on the mitogenome.

2.6. RNA Editing

Based on Deepred-mt, the RNA editing events of 33 PCGs from mitogenome were identified. The 25 species on the page were selected as the database. Using this criterion, 362 potential RNA editing sites were identified (Figure 7 and Table S9). There are great differences in the number of sites edited by different genes. Among mitochondrial genes, the *ccmB* gene identified 30 RNA editing sites, the most among all genes. This was followed by *ccmC*, with 28. However, the *rps13* and *rps7* genes only identified one RNA editing site.

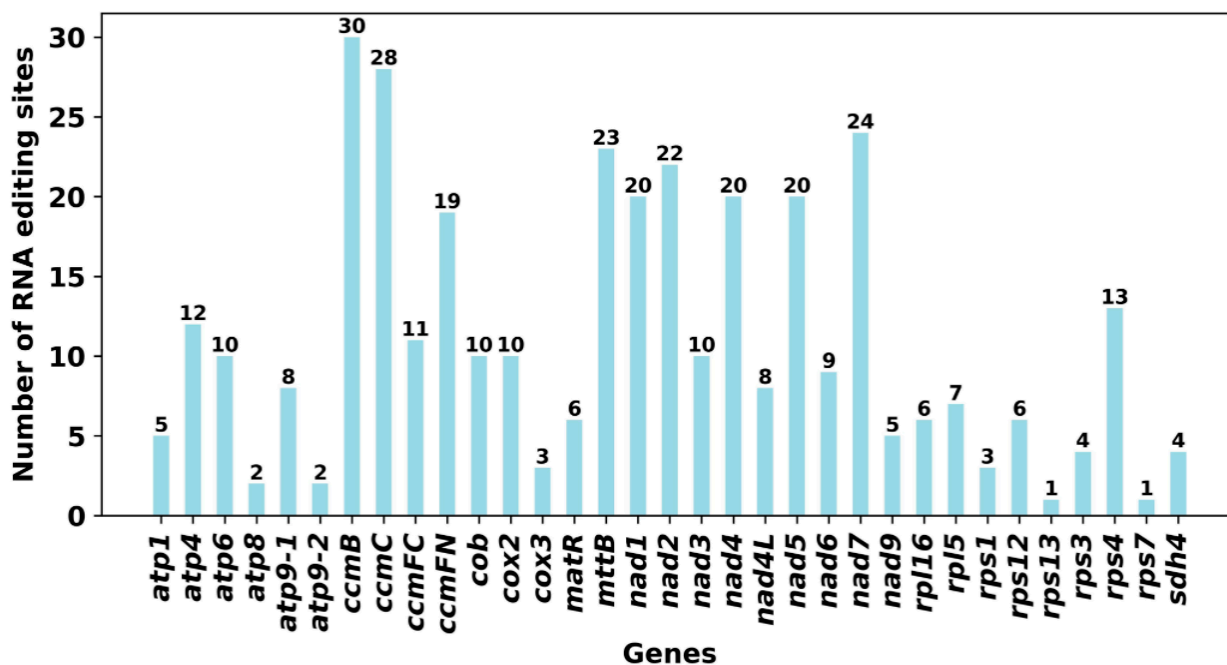


Figure 7. Characteristics of the RNA editing sites identified in PCGs of *P. aculeata* mitogenome. Number of RNA editing sites predicted by individual PCGs using Deepred-mt. The abscissa shows the name of the gene, and the ordinate shows the number of edited sites.

Studies have shown that a large number of C to U substitutions occur during mitochondrial DNA transcription in higher plants [43]. By examining the types of RNA editing events, it was found that all the editing sites in the mitogenome of *P. aculeata* were C-U editing. This is consistent with previous statements. Notably, we found that the stop codons of two genes (*atp6* and *atp9*) and the start codons of four genes (*cox2*, *nad1*, *nad4L*, and *nad7*) were created through RNA editing events and were supported by the high confidence of Deepred-mt. The new start and stop codons are usually generated to encode proteins that are more conserved and homologous to the corresponding proteins of other species, thus allowing better expression of the genes in the mitochondria [44].

According to the result of Deepred-mt, the C to U RNA editing events generated three stop codons in two genes, namely *atp6* and *atp9*. The *atp6* is CAA to UAA, and the *atp9* is CGA to UGA. The start codons of *cox2*, *nad1*, *nad7*, and *nad4L* are created by RNA editing, which was achieved by editing ACG to AUG. RNA editing in plants may lead to deleterious phenotypes and even death [45]. Therefore, it is very important to study RNA editing in plants.

2.7. Phylogenetic Inference

To investigate the phylogenetic position of *P. aculeata*, we downloaded mitogenome sequences of 24 species from GenBank, including 21 closely related species and two outgroups. The phylogenetic tree (Figure 8) showed that Caryophyllaceae, Amaranthaceae, and Chenopodiaceae were clustered into one clade, and Aizoaceae, Nyctaginaceae, and Cactaceae were clustered into another clade. Polygonaceae and Nepenthaceae each form a separate clade. The phylogenetic tree is consistent with the Angiosperm Phylogeny Group IV (APG IV) classification system; that is, *P. aculeata* belongs to the Cactaceae family of Caryophyllales, which is closely related to Aizoaceae and Nyctaginaceae. Most of the nodes of the phylogenetic tree in this study had relatively high bootstrap values, indicating that the phylogenetic analysis had high reliability.

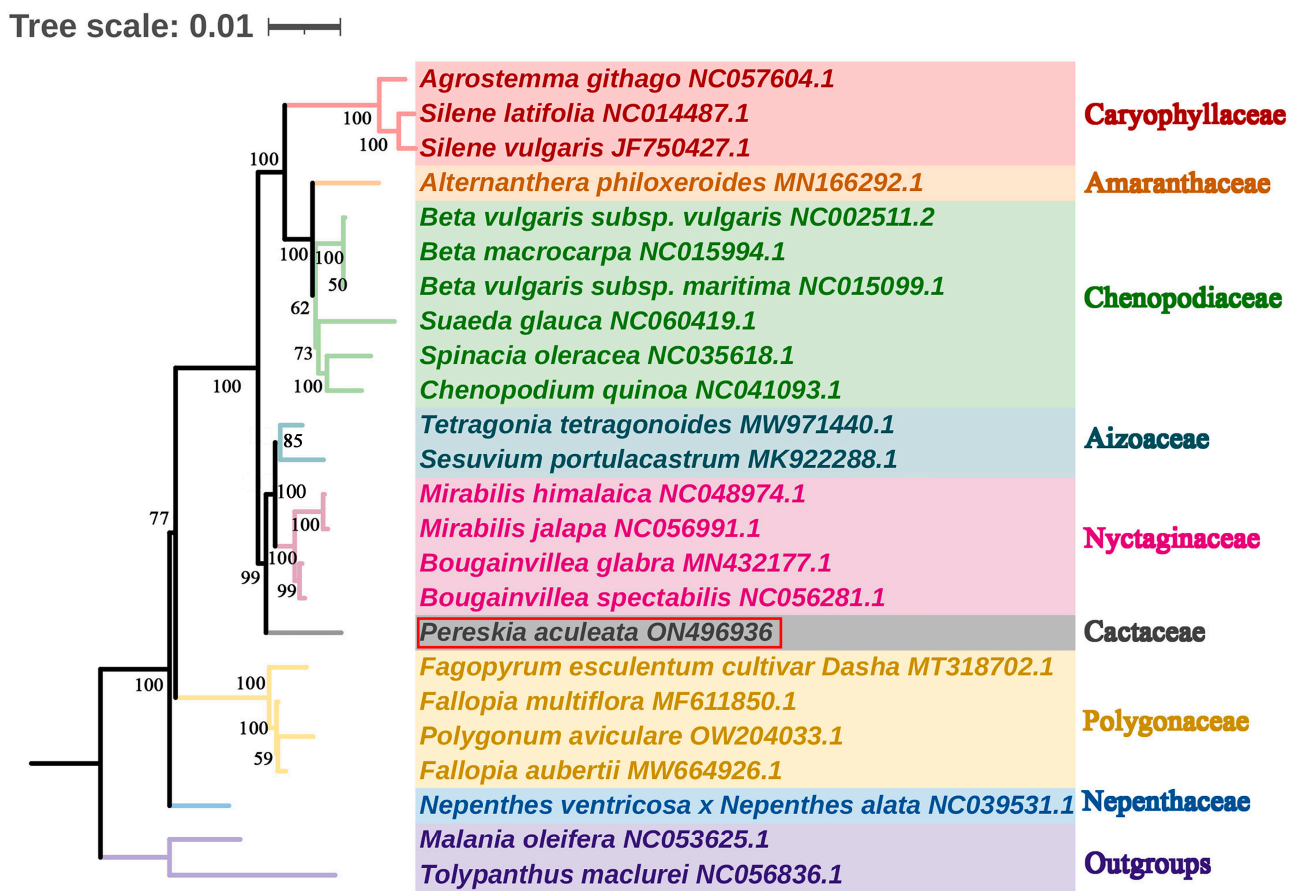


Figure 8. Phylogenetic tree constructed from the shared mitochondrial PCGs of 24 species. The *Tolypanthus maclurei*, and *Malania oleifera* were used as the outgroups. The number at each node showing the bootstrap support value.

2.8. Synteny Analysis

Synteny analysis deals with the arrangement of homologous genes or sequences. Based on Mauve (version 20150226) software [46], we performed synteny analysis on the mitogenomes of *P. aculeata* and four related species (*Tetragonia tetragonoides*, *Mirabilis jalapa*, *Sesuvium portulacastrum*, and *Bougainvillea spectabilis*) (Table S10). These species exhibit many common colinear mitogenomic sequences. However, these collinear sequences have an irregular arrangement, showing a rearrangement of the mitogenome. These results also indicate that the conformation of mitogenome is extremely non-conserved. As shown in Figure 9, there are only 86 shared collinear blocks among these five mitogenomes, with lengths ranging from 115 bp to 3780 bp. In contrast, a large number of regions (mainly non-coding regions) show high species-specific levels.

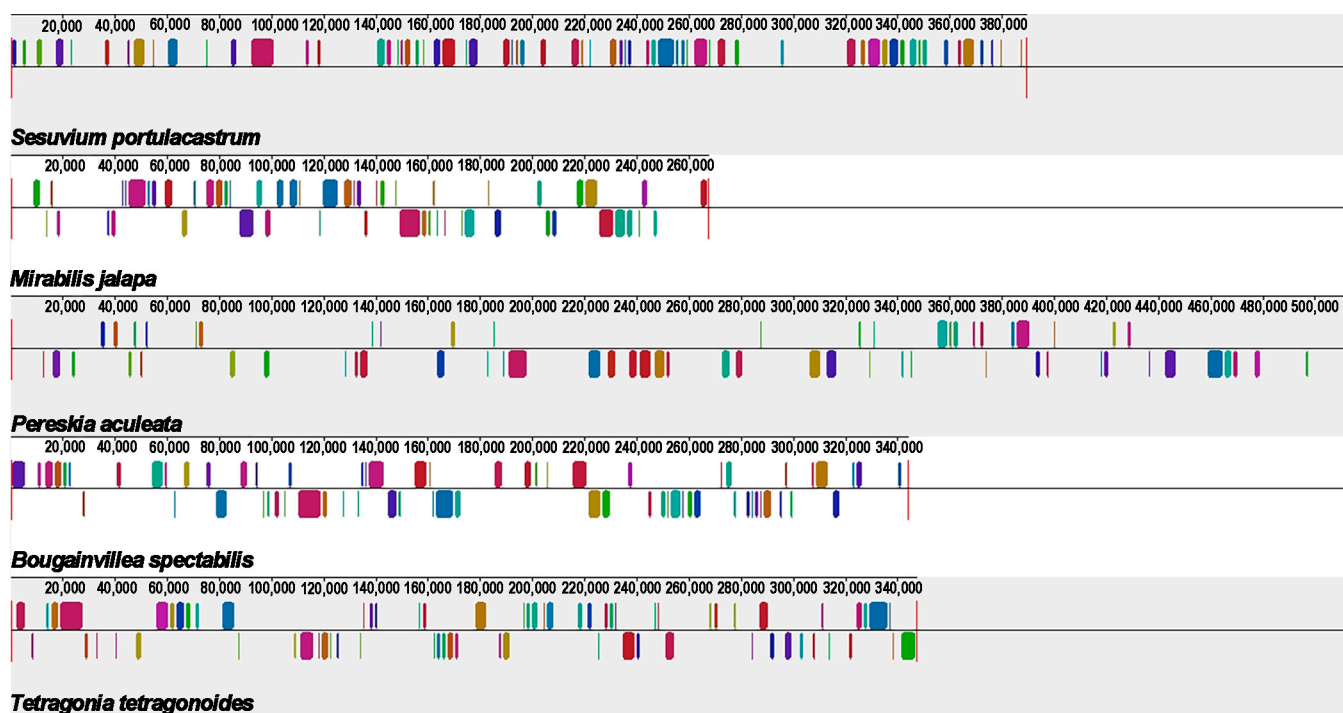


Figure 9. Synteny analysis of five mitogenomes. Only collinear blocks over 0.1 kb in length are retained.

3. Discussion

Our study involves the analysis of the *P. aculeata* mitogenome, which is 515,187 bp long (represented in Figure 3). It is important to note that there is a possibility of alternative configurations and a multi-chromosomal architecture. Although Unicycler only obtained one independent and circular contig by resolving repeats with long reads, we showed that there are potential multiple conformations based on PCR and long reads. The 9026 bp and 7040 bp repetitive elements could possibly be the key to recombination. Furthermore, we found that the similarities between the two repetitive units of the above dispersed repeats are 100% (Table S7). A high similarity could be indicative of ongoing gene conversion and concerted evolution [47]. The plant mitogenome undergoes recombination with the aid of large repetitive segments present in multiple locations within mitochondrial DNA. These large repetitive elements were found in many plant mitogenomes. For example, the mitogenome of *Thuja sutchuenensis* consists of one circular contig and three linear contigs (1,390,975 bp, 519,836 bp, and 293,570 bp) with overlapping regions, which was confirmed through PCR [48]. A 21,809 bp repetitive element was also found in the *Taraxacum mongolicum* mitogenome, which was also confirmed through PCR [49]. A total of five pairs of repetitive elements have been shown to mediate recombination in dandelion mitogenomes [49]. Short repetitive elements also confirmed such genome recombination. According to research, *Ipomoea batatas* [50], *Nymphaea colorata* [51], *Salvia miltiorrhiza* [52], and *Ginkgo biloba* [53] can all produce different alternative configurations through short repeats. It was shown that repetitive sequences mediate the reorganization of the mitochondrial genome with different frequencies [51]. Similar phenomena were observed in *Silene latifolia* [54] and *Mimulus guttatus* [55]. Here, we confirmed that two pairs of long repeats were involved in genome recombination based on long reads and PCR experiments in the *P. aculeata* mitogenome. This recombination involving long repeats may not have caused functional damage to the mitogenome of *P. aculeata*, and this recombination is likely to be reversible. However, it is not known whether some short-repeats-mediated recombination exists, and in the future, high sequencing depth data will be required to achieve low-frequency recombination detection.

A total of 165 SSRs were identified in the mitochondrial genome of *P. aculeata*. The number of SSRs in the form of a single-nucleotide (A/T) polymer was the largest, accounting for 26.06% of all SSRs. The same was true in the mitogenome of *Phaseolus vulgaris* [16]. This may be related to the low GC content in mitochondria. Secondly, we found that the mitogenome contained a large number of dispersed repeats. They are mainly forward repeats and palindromic repeats. The ratio between them is close to 1:1. A similar pattern was found in *Gleditsia sinensi*, which had 26 forward repeats and 24 palindromic repeats [36].

The plant mitogenome is more prone to accept and integrate foreign DNA because of the special properties of its genome structure and evolutionary process [31]. Foreign DNA is widely found in the plant mitogenome [56]. Some studies have shown that the division and partial gene transfer of the ribosomal protein gene *rpl2* has been found in the plant mitogenome, which provides direct evidence for the easy acceptance and integration of foreign DNA in the plant mitotic genome [57]. For angiosperms, there is extensive communication of genetic material between the two organelles in the cell, mainly the migration of DNA from the chloroplast to the mitogenome [17]. Mitochondrial DNA transfer into the chloroplast genome has been found only in *Anacardium occidentale* [58]. In *P. aculeata*, we found 38 MTPTs. Among them, the MTPT1 is the longest at 3962 bp. Research on MTPTs has been widely reported. In *Suaeda glauca*, 26.87 kb of MTPTs accounted for 5.18% of its mitogenome [17]. These large MTPTs were thought to influence the evolution of eukaryotes extensively and promote genetic diversity. Additionally, one study showed that the origin of tRNAs in plant mitochondria is divided into two parts: one is inherited from the mitochondrial ancestors, and the other is acquired from chloroplasts by HGT [59].

Based on sequence similarity and a previous report [32], we can trace the tRNA genes transferred from plastid to mitochondria in the *P. aculeata* mitogenome. In *P. aculeata*, *trnD-GUC*, *trnH-GUG*, *trnN-GUU*, *trnV-GAC*, *trnW-CCA*, *trnM-CAU*, *trnP-UGG*, and *trnI-CAU* were transferred from chloroplast genes. It makes up 42.11 percent of all tRNAs. Of these tRNA genes, *trnW-CCA* (cp) is broadly present in the mitogenomes of other angiosperms, and they appear to be homologous to chloroplasts. However, *Amborella trichopoda* has mitochondrial native *trnW-CCA*, and this is not seen in other angiosperms [34]. Additionally, *trnP-UGG* (cp) may also be functional, as reported by Richardson et al. [32]. Furthermore, we also found a bacterial-origin tRNA gene (*trnC-GCA*) in the mitogenome of *P. aculeata*. The bacteria-derived *trnC-GCA* has been reported previously [33,34], but it is not very commonly found in plant mitogenomes.

In *P. aculeata*, we found that MTPTs contained 15 complete genes, including six PCGs and nine tRNA genes. In addition, some chloroplast PCGs lost their integrity during evolution and migrated from the chloroplast genome to the mitogenome in the form of fragments. There is a similar situation in the mitogenome of *Suaeda glauca* [17]: chloroplast genome PCGs may migrate among organelles in the form of partial sequences. Therefore, it was found that tRNA genes are more conserved than PCGs and play an important role in the evolution of plants. In general, the chloroplast genome of *P. aculeata* provides abundant foreign sequences for its mitogenome, which increases the diversity of mitochondrial DNA sequence sources.

In eukaryotes, the coding regions for transcribed RNA undergo a process called RNA editing by adding, missing, or replacing bases [60]. Studies have found that RNA editing is widespread in higher plant mitochondria, and it is one of the essential steps for gene expression in the plant mitogenome [61,62]. For most angiosperms, the chemistry of RNA editing is a deamination reaction in which a site-specific cytosine (C) becomes uracil (U) [63]. Studies have shown that RNA editing of mitochondrial genes may regulate plant cytoplasmic inheritance-related traits [64]. In *P. aculeata*, we recorded a total of 362 RNA editing events, and most of the RNA editing sites occurred at the first position or second position, which is similar to the situation in most plants [16,36,65,66]. By identifying the RNA editing sites, the gene function of the new code can be effectively predicted. Notably, we found that the stop codons of two genes (*atp6* and *atp9*) and the start codons of four genes (*cox2*, *nad1*, *nad7*, and *nad4L*) were created by RNA editing events and were supported by the

high confidence of Deepred-mt. These new start and stop codons encode more conserved and homologous proteins, promoting better gene expression in mitochondria [44].

4. Materials and Methods

4.1. Plant Sampling, DNA Extraction, and Sequencing

We collected the fresh leaves of *P. aculeata* from Nanshan Botanical Garden, Chongqing (Geospatial coordinates: N29.561209, E106.635764). The samples were identified by Professor Jie Yu based on morphological characteristics, and they have been deposited in the Herbarium of Southwest University in Chongqing, with the accession number SWU-YJ-009. The total genomic DNA was extracted using the CTAB method [67], and the total DNA was fragmented using ultrasound. The DNA library with an insertion size of 350 bp was constructed with the NEBNext[®] library construction kit (NEB, Ipswich, MA, USA) [68] and was sequenced by using the HiSeq Xten PE150 sequencing platform (Novogene, Beijing, China). Sequencing produced a total of 5.35 Gb of raw data, including 17,835,750 raw reads. Clean data were obtained using Trimmomatic (v0.36) [69]: low-quality sequences with more than 5% base N were eliminated, and more than 5% base quality values $Q < 19$. Finally, we obtained a total of 17,631,944 clean reads. The same DNA sample was also used for Oxford Nanopore sequencing based on the promethION platform. A total of 13.4 Gb of sequence reads were obtained. The mean read length of sequencing was 17.52 kb, and the N50 was 26.50 kb.

4.2. Genome Assembly

To begin, we utilized Flye software (v2.9.1-b1780) [70] to conduct de novo assembly of the Oxford Nanopore long reads derived from *P. aculeata*. The parameters used for this assembly were ‘-min-overlap 8000’. We then employed the BLASTn (v2.2.30+) [71] program to identify the draft mitogenome of the *P. aculeata* based on the assembled contigs. To achieve this, we constructed a database for the assembled sequences using makeblastdb and selected the conserved mitochondrial genes from *Liriodendron tulipifera* (NC_021152.1) as a query sequence to identify contigs that contain conserved mitochondrial genes. A total of six unitigs were identified. These six unitigs are interconnected to form a branching genome structure. Finally, we mapped both the short reads and long reads onto these contigs and retained all the mapped reads using BWA (v0.7.12-r1039) [72] and SAMTools software (v0.1.19) [73].

Due to the possibility of chloroplast-homologous regions of the mitogenome being replaced by chloroplast counterparts during polishing, we used a combination of Illumina short-reads and Nanopore long reads to perform hybrid assembly using Unicycler (v0.4.8) [74], with the parameters of ‘-kmers 27, 53, 67, 71, 89, 99, 105’. Initially, the mapped Illumina short-reads were assembled using spades [75]. Then, repetitive regions of the assembly were resolved using minimap2 [76] with the Nanopore long reads. The GFA format files produced by Unicycler were visualized using Bandage (v0.8.1) [77]. Finally, Unicycler generated one circular contig, and as this contig was assembled using short reads, no additional polish was necessary.

4.3. Validating of Genome Structure

We use two strategies to determine the potential structure of the mitogenome. First, we use long reads to detect whether the two long repetitive elements mediate the genome’s recombination. Specifically, we extracted the sequences of these two pairs of repetitive elements and their flanking regions of 500 bp from the genome assembled by Unicycler, which represents the reference sequence supporting this master circle structure. Subsequently, we simulate the sequence of possible recombination by swapping the flanking regions. Next, we used makeblastdb to construct a database of long reads, and the major and recombinant sequences were together used as query sequences using the BLASTn program with the parameter ‘-evalue 1e-10 -outfmt 6 -num_threads 10 -max_hsps 1’. Finally, we counted

those long reads that completely spanned these query sequences and used them as reads that support this structure.

The second strategy is based on PCR experiments. We designed specific primers at several joints to verify the authenticity of this bifurcated structure. The primers were designed using the Primer designing tool on NCBI (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>; accessed on 12 January 2023), and the parameters are default. We extracted the total DNA, and the amplifications were carried out in a Pro-Flex PCR system (Applied Biosystems, Waltham, MA, USA). The final volume of PCR amplification was 25 μ L, including 2 μ L template DNA, 1 μ L forward primer, 12.5 μ L 2 \times Taq PCR Master Mix, and 9.5 μ L ddH₂O. The following amplification conditions: denaturation at 94 °C for 5 min, 30 cycles for 30 s at 94 °C, 30 s at 58 °C, 60 s at 72 °C and 72 °C for 5 min as the final extension. The PCR amplicons were visualized using 1% agarose gel electrophoresis, and the single bright bands were cut and sent to the Sangon Biotech (Shanghai, China) Co., Ltd. for Sanger sequencing.

4.4. Genome Annotation

The chloroplast genome of *P. aculeata* [78] was published before with the accession number MW553064 (<https://www.ncbi.nlm.nih.gov/nuccore/MW553064>; accessed on 12 January 2023). We used CPGAVAS2 [79] with the customized option; the *P. aculeata* chloroplast genome (NC_062888.1) was used as the reference genome. The mitogenomes of *Liriodendron tulipifera* (NC_021152.1) and *Beta macrocarpa* (NC_015994.1) were used as the reference genome, and Geseq (<https://chlorobox.mpimp-golm.mpg.de/geseq.html>; accessed on 12 January 2023) [80] was used to annotate the mitogenome of *P. aculeata*. Apollo software (v1.11.8) [81] was used to manually edit and correct problematic annotations. In addition, all tRNA genes in the mitogenome were annotated using tRNAscan-SE software (v2.0) [82], and all rRNA were annotated using BLASTn software [71].

4.5. Codon Usage of Mitochondrial Genes

PhyloSuite software (v1.2.2) [83] was used to parse the GenBank format file of the *P. aculeata* mitogenome, and we extracted the protein-coding gene (PCGs). Mega 7.0 software [84] was used to analyze the codon usage of mitochondrial PCGs by calculating RSCU values.

4.6. Analysis of Repeat Elements

The simple sequence repeats (SSRs) were obtained via the online website MISA (<https://webblast.ipk-gatersleben.de/misa/>; accessed on 12 January 2023) [85]; the parameters of the minimum numbers of mono-, di-, tri-, tetra-, penta-, and hexanucleotides were set as 10, 5, 4, 3, 3, and 3, respectively. In addition, we used REPuter (<https://bibiserv.cebitec.uni-bielefeld.de/reputer/>; accessed on 12 January 2023) [86] to calculate forward, reverse, palindromic, and complementary repeat sequences, with the following settings: hamming distance of three and minimal repeat size of 30 bp, and e-value was limited to less than 1×10^{-5} . The visualization of the repeat elements was accomplished by using the Circos [87] package.

4.7. Identification of Homologous Sequences among Organelle Genomes

The BLASTn program was performed on the two organelle genomes of *P. aculeata*, and the mitogenome was used for building databased with makeblastdb, and the chloroplast genome was used as query sequence with the following parameters: ‘-evalue 1e-10-word_size 7-outfmt 6’. Finally, the BLASTn results were visualized using the ‘Advanced Circos’ module in TBtools (v1.18) [88].

4.8. The Prediction of RNA Editing Sites

We used Deepred-mt [89], a tool for predicting C to U RNA editing sites based on the convolutional neural network (CNN) model, to predict C to U RNA editing sites. All

mitochondrial protein-coding genes were extracted from the mitogenome and used for prediction, and we selected results with probability values greater than 0.9.

4.9. Phylogenetic Inference

According to the genetic relationship, the related species of *P. aculeata* were selected, and their complete mitogenome sequences were downloaded. The species information is listed in Table S1. We downloaded the mitogenomes of 24 closely related species and two outgroups (*Tolypanthus maclurei* and *Malaria oleifera*) on NCBI (<https://www.ncbi.nlm.nih.gov/>; accessed on 12 January 2023) nucleotide database. A total of 31 orthologous PCGs among the analyzed species were identified and extracted by using PhyloSuite (v1.2.2) [83]. The corresponding nucleotide sequences were aligned using MAFFT (v7.471) [90]. Next, these aligned sequences were concatenated and used to construct the phylogenetic trees. The maximum likelihood (ML) method was implemented in RAxML (v8.2.4). The parameters were “raxmlHPC-PTHREADS-SSE3 -f a -N 1000 m GTRGAMMA -x 551314260 -p 551314260”. The bootstrap analysis was performed with 1000 replicates. Finally, the result was uploaded to the ITOL software (v4) [91] for beautification.

4.10. Synteny Analysis

We chose four closed species: *Bougainvillea spectabilis* (NC_056281.1), *Tetragonia tetragonoides* (MW971440.1), *Sesuvium portulacastrum* (MK922288.1), and *Mirabilis jalapa* (NC_056991.1) for the colinear analysis with *P. aculeata*. Based on the progressive Mauve module embedded in Mauve (v.20150226) software [46], the above five mitogenomes were analyzed with default parameters. For the calculated conserved collinear blocks, we only reserved the blocks with lengths greater than 100 bp.

5. Conclusions

In this study, we report the complete mitogenome of *P. aculeata*. The genome length is 515,187 bp and its GC content is 44.05%. Long reads and PCR confirmed that there are two repetitive sequences in the *P. aculeata* mitogenome that promote the rearrangement of the mitogenomes and generate multiple isomers. The mitogenome of *P. aculeata* was studied in detail from many aspects, which provided important genetic resources for the phylogenetic study of *P. aculeata* and laid a foundation for the mitogenome study of the Cactaceae family.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms24098366/s1>.

Author Contributions: Conceptualization, J.Y.; Data curation, J.L.; Formal analysis, Q.Q.; Funding acquisition, J.Y.; Investigation, X.Z.; Methodology, X.Z.; Project administration, J.Y. and H.D.; Resources, Y.S.; Software, X.Z. and Y.S.; Supervision, X.Z. and Y.S.; Validation, J.L.; Visualization, Q.Q.; Writing—original draft, X.Z. and Y.S.; Writing—review and editing, J.L., J.Y. and H.D. All authors have read and agreed to the published version of the manuscript.

Funding: Natural Science Foundation of Chongqing, China (CSTB2022NSCQ-MSX1190). The funders were not involved in the study design, data collection, analysis, publication decision, or manuscript preparation.

Institutional Review Board Statement: We collected fresh leaf materials of *P. aculeata* for this study. The study, including plant samples, complies with relevant institutional, national, and international guidelines and legislation. No specific permits were required for plant collection.

Informed Consent Statement: Not applicable.

Data Availability Statement: The mitogenome sequence is available in the nucleotide database of NCBI (<https://www.ncbi.nlm.nih.gov/>; accessed on 12 January 2023) with accession numbers: NC_067638.1 or ON496936.1. The sequencing reads used for mitogenome assembly in this study have been released on the NCBI with those accession numbers: PRJNA715621 (BioProject), SAMN18357766 (BioSample), and SRR21847255 (SRA). The sample has been deposited in the Herbarium of Southwest University in Chongqing, with the accession number SWU-YJ-009. The mitochondrial long reads can be found in Supplementary File S1.

Acknowledgments: We would sincerely thank Wuhan Benagen Tech Solutions Company Limited (www.benagen.com; accessed on 10 January 2023) and the MitoRun research group for their assistance in genomic sequencing and bioinformatics analysis.

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

Abbreviations

PCR	Polymerase chain reaction
DR	Direct repeat
SSR	Simple sequence repeat
ML	Maximum likelihood
NCBI	National Center for Biotechnology Information
BLAST	Basic Local Alignment Search Tool
PCCGs	Protein-coding gene sequences
MTPT	Mitochondrial plastid DNA

References

- Duarte, M.R.; Hayashi, S.S. Estudo anatômico de folha e caule de *Pereskia aculeata* Mill. (Cactaceae). *Rev. Bras. Farmacogn.* **2005**, *15*, 103–109. [[CrossRef](#)]
- Paterson, I.D.; Downie, D.A.; Hill, M.P. Using molecular methods to determine the origin of weed populations of *Pereskia aculeata* in South Africa and its relevance to biological control. *Biol. Control.* **2009**, *48*, 84–91. [[CrossRef](#)]
- Pinto, N.D.C.; Duque, A.P.D.; Pacheco, N.R.; Mendes, R.D.; Motta, E.V.D.; Bellozi, P.M.Q.; Ribeiro, A.; Salvador, M.J.; Scio, E. *Pereskia aculeata*: A plant food with antinociceptive activity. *Pharm. Biol.* **2015**, *53*, 1780–1785. [[CrossRef](#)]
- Merce, A.L.R.; Landaluze, J.S.; Mangrich, A.S.; Szpoganicz, B.; Sierakowski, M.R. Complexes of arabinogalactan of *Pereskia aculeata* and Co^{2+} , Cu^{2+} , Mn^{2+} , and Ni^{2+} . *Bioresour. Technol.* **2001**, *76*, 29–37. [[CrossRef](#)] [[PubMed](#)]
- Takeiti, C.Y.; Antonio, G.C.; Motta, E.M.P.; Collares-Queiroz, F.P.; Park, K.J. Nutritive evaluation of a non-conventional leafy vegetable (*Pereskia aculeata* Miller). *Int. J. Food Sci. Nutr.* **2009**, *60*, 148–160. [[CrossRef](#)] [[PubMed](#)]
- Pinto, N.C.C.S.R.; Machado, D.C. Cytotoxic and antioxidant activity of *Pereskia aculeata* Miller. *Pharmacologyonline* **2012**, *3*, 63–69.
- Cruz-Rubio, J.M.; Mueller, M.; Viernstein, H.; Loeppert, R.; Praznik, W. Prebiotic potential and chemical characterization of the poly and oligosaccharides present in the mucilage of *Opuntia ficus-indica* and *Opuntia joconostle*. *Food Chem.* **2021**, *362*, 130167. [[CrossRef](#)]
- de Andrade Vieira, E.; Alves Alcantara, M.; Albuquerque Dos Santos, N.; Duarte Gondim, A.; Iacomini, M.; Mellinger, C.; Tribuzy de Magalhaes Cordeiro, A.M. Mucilages of cacti from Brazilian biodiversity: Extraction, physicochemical and technological properties. *Food Chem.* **2021**, *346*, 128892. [[CrossRef](#)]
- Silva, S.H.; Neves, I.C.O.; Oliveira, N.L.; de Oliveira, A.C.F.; Lago, A.M.T.; Giarola, T.M.D.; de Resende, J.V. Extraction processes and characterization of the mucilage obtained from green fruits of *Pereskia aculeata* Miller. *Ind. Crops Prod.* **2019**, *140*, 111716. [[CrossRef](#)]
- Torres-Acosta, A.A.; González-Calderón, P.Y. *Opuntia ficus-Indica* (OFI) Mucilage as Corrosion Inhibitor of Steel in CO_2 -Contaminated Mortar. *Materials* **2021**, *14*, 1316. [[CrossRef](#)]
- Bonner, J.; Millerd, A. Oxidative phosphorylation by plant mitochondria. *Arch. Biochem. Biophys.* **1953**, *42*, 135–148. [[CrossRef](#)] [[PubMed](#)]
- Daloso, D.M.; Müller, K.; Obata, T.; Florian, A.; Tohge, T.; Bottcher, A.; Riondet, C.; Bariat, L.; Carrari, F.; Nunes-Nesi, A.; et al. Thioredoxin, a master regulator of the tricarboxylic acid cycle in plant mitochondria. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E1392–E1400. [[CrossRef](#)] [[PubMed](#)]
- Valero, T. Mitochondrial Biogenesis: Pharmacological Approaches. *Curr. Pharm. Des.* **2014**, *20*, 5507–5509. [[CrossRef](#)] [[PubMed](#)]
- Levings, C.S.; Pring, D.R. Restriction Endonuclease Analysis of Mitochondrial-DNA from Normal and Texas Cytoplasmic Male-Sterile Maize. *Science* **1976**, *193*, 158–160. [[CrossRef](#)]
- Guo, J.; Wang, P.; Cheng, Q.; Sun, L.; Wang, H.; Wang, Y.; Kao, L.; Li, Y.; Qiu, T.; Yang, W.; et al. Proteomic analysis reveals strong mitochondrial involvement in cytoplasmic male sterility of pepper (*Capsicum annuum* L.). *J. Proteom.* **2017**, *168*, 15–27. [[CrossRef](#)]
- Bi, C.W.; Lu, N.; Xu, Y.Q.; He, C.P.; Lu, Z.H. Characterization and Analysis of the Mitochondrial Genome of Common Bean (*Phaseolus vulgaris*) by Comparative Genomic Approaches. *Int. J. Mol. Sci.* **2020**, *21*, 3778. [[CrossRef](#)]

17. Cheng, Y.; He, X.; Priyadarshani, S.V.G.N.; 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](#)]
18. Liu, D.; Guo, H.; Zhu, J.; Qu, K.; Chen, Y.; Guo, Y.; Ding, P.; Yang, H.; Xu, T.; Jing, Q.; et al. Complex Physical Structure of Complete Mitochondrial Genome of *Quercus acutissima* (Fagaceae): A Significant Energy Plant. *Genes* **2022**, *13*, 1321. [[CrossRef](#)]
19. 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](#)]
20. Heyer, W.D.; Ehmsen, K.T.; Liu, J. Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.* **2010**, *44*, 113–139. [[CrossRef](#)]
21. Rocha, E.P.; Cornet, E.; Michel, B. Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genet.* **2005**, *1*, e15. [[CrossRef](#)] [[PubMed](#)]
22. Kreuzer, K.N. Recombination-dependent DNA replication in phage T4. *Trends Biochem. Sci.* **2000**, *25*, 165–173. [[CrossRef](#)] [[PubMed](#)]
23. Yeeles, J.T.; Poli, J.; Mariani, K.J.; Pasero, P. Rescuing stalled or damaged replication forks. *Cold Spring Harb. Perspect. Biol.* **2013**, *5*, a012815. [[CrossRef](#)] [[PubMed](#)]
24. Cappadocia, L.; Marechal, A.; Parent, J.S.; Lepage, E.; Sygusch, J.; Brisson, N. Crystal structures of DNA-Whirly complexes and their role in *Arabidopsis organelle* genome repair. *Plant Cell* **2010**, *22*, 1849–1867. [[CrossRef](#)]
25. Puchta, H. Double-strand break-induced recombination between ectopic homologous sequences in somatic plant cells. *Genetics* **1999**, *152*, 1173–1181. [[CrossRef](#)] [[PubMed](#)]
26. Chang, H.H.Y.; Pannunzio, N.R.; Adachi, N.; Lieber, M.R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell. Biol.* **2017**, *18*, 495–506. [[CrossRef](#)] [[PubMed](#)]
27. Roger, A.J.; Munoz-Gomez, S.A.; Kamikawa, R. The Origin and Diversification of Mitochondria. *Curr. Biol.* **2017**, *27*, R1177–R1192. [[CrossRef](#)]
28. Morimatsu, K.; Kowalczykowski, S.C. RecQ helicase and RecJ nuclease provide complementary functions to resect DNA for homologous recombination. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E5133–E5142. [[CrossRef](#)]
29. Edmondson, A.C.; Song, D.; Alvarez, L.A.; Wall, M.K.; Almond, D.; McClellan, D.A.; Maxwell, A.; Nielsen, B.L. Characterization of a mitochondrially targeted single-stranded DNA-binding protein in *Arabidopsis thaliana*. *Mol. Genet. Genom.* **2005**, *273*, 115–122. [[CrossRef](#)]
30. Garcia-Medel, P.L.; Baruch-Torres, N.; Peralta-Castro, A.; Trasvina-Arenas, C.H.; Torres-Larios, A.; Brieba, L.G. Plant organellar DNA polymerases repair double-stranded breaks by microhomology-mediated end-joining. *Nucleic Acids Res.* **2019**, *47*, 3028–3044. [[CrossRef](#)]
31. Christensen, A.C. Plant Mitochondrial Genome Evolution Can Be Explained by DNA Repair Mechanisms. *Genome Biol. Evol.* **2013**, *5*, 1079–1086. [[CrossRef](#)] [[PubMed](#)]
32. 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](#)]
33. Kitazaki, K.; Kubo, T.; Kagami, H.; Matsumoto, T.; Fujita, A.; Matsuhira, H.; Matsunaga, M.; Mikami, T. A horizontally transferred tRNA(Cys) gene in the sugar beet mitochondrial genome: Evidence that the gene is present in diverse angiosperms and its transcript is aminoacylated. *Plant J.* **2011**, *68*, 262–272. [[CrossRef](#)] [[PubMed](#)]
34. Rice, D.W.; Alverson, A.J.; Richardson, A.O.; Young, G.J.; Sanchez-Puerta, M.V.; Munzinger, J.; Barry, K.; Boore, J.L.; Zhang, Y.; DePamphilis, C.W.; et al. Horizontal Transfer of Entire Genomes via Mitochondrial Fusion in the *Angiosperm amborella*. *Science* **2013**, *342*, 1468–1473. [[CrossRef](#)]
35. Mower, J.P. Variation in protein gene and intron content among land plant mitogenomes. *Mitochondrion* **2020**, *53*, 203–213. [[CrossRef](#)]
36. Yang, H.X.; Li, W.H.; Yu, X.L.; Zhang, X.Y.; Zhang, Z.Y.; Liu, Y.X.; Wang, W.X.; Tian, X.X. Insights into molecular structure, genome evolution and phylogenetic implication through mitochondrial genome sequence of *Gleditsia sinensis*. *Sci. Rep.* **2021**, *11*, 14850. [[CrossRef](#)]
37. Sau, K.; Gupta, S.K.; Sau, S.; Mandal, S.C.; Ghosh, T.C. Factors influencing synonymous codon and amino acid usage biases in Mimivirus. *Biosystems* **2006**, *85*, 107–113. [[CrossRef](#)]
38. Lovin, D.D.; Washington, K.O.; deBruyn, B.; Hemme, R.R.; Mori, A.; Epstein, S.R.; Harker, B.W.; Streit, T.G.; Severson, D.W. Genome-based polymorphic microsatellite development and validation in the mosquito *Aedes aegypti* and application to population genetics in Haiti. *BMC Genom.* **2009**, *10*, 590. [[CrossRef](#)]
39. Morgante, M.; Hanafey, M.; Powell, W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat. Genet.* **2002**, *30*, 194–200. [[CrossRef](#)]
40. Naranpanawa, D.N.U.; Chandrasekara, C.H.W.M.R.B.; Bandaranayake, P.C.G.; Bandaranayake, A.U. Raw transcriptomics data to gene specific SSRs: A validated free bioinformatics workflow for biologists. *Sci. Rep.* **2020**, *10*, 18236. [[CrossRef](#)]
41. Kawata, M.; Harada, T.; Shimamoto, Y.; Oono, K.; Takaiwa, F. Short inverted repeats function as hotspots of intermolecular recombination giving rise to oligomers of deleted plastid DNAs (ptDNAs). *Curr. Genet.* **1997**, *31*, 179–184. [[CrossRef](#)] [[PubMed](#)]
42. Lopez, L.; Barreiro, R.; Fischer, M.; Koch, M.A. Mining microsatellite markers from public expressed sequence tags databases for the study of threatened plants. *BMC Genom.* **2015**, *16*, 781. [[CrossRef](#)] [[PubMed](#)]

43. Simpson, L.; Emerson, R.B. RNA editing. *Annu. Rev. Neurosci.* **1996**, *19*, 27–52. [[CrossRef](#)] [[PubMed](#)]
44. Edera, A.A.; Gandini, C.L.; Sanchez-Puerta, M.V. Towards a comprehensive picture of C-to-U RNA editing sites in angiosperm mitochondria. *Plant Mol. Biol.* **2018**, *97*, 215–231. [[CrossRef](#)]
45. Chateigner-Boutin, A.L.; Small, I. Plant RNA editing. *Rna Biol.* **2010**, *7*, 213–219. [[CrossRef](#)]
46. 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](#)]
47. Carson, A.R.; Scherer, S.W. Identifying concerted evolution and gene conversion in mammalian gene pairs lasting over 100 million years. *BMC Evol. Biol.* **2009**, *9*, 156. [[CrossRef](#)]
48. Xia, C.; Li, J.; Zuo, Y.; He, P.; Zhang, H.; Zhang, X.; Wang, B.; Zhang, J.; Yu, J.; Deng, H. Complete mitochondrial genome of Thuja sutchuenensis and its implications on evolutionary analysis of complex mitogenome architecture in Cupressaceae. *BMC Plant Biol.* **2023**, *23*, 84. [[CrossRef](#)]
49. Jiang, M.; Ni, Y.; Li, J.; Liu, C. Characterisation of the complete mitochondrial genome of *Taraxacum mongolicum* revealed five repeat-mediated recombinations. *Plant Cell Rep.* **2023**, *42*, 775–789. [[CrossRef](#)]
50. Yang, Z.; Ni, Y.; Lin, Z.; Yang, L.; Chen, G.; Nijjati, N.; Hu, Y.; Chen, X. De novo assembly of the complete mitochondrial genome of sweet potato (*Ipomoea batatas* [L.] Lam) revealed the existence of homologous conformations generated by the repeat-mediated recombination. *BMC Plant Biol.* **2022**, *22*, 285. [[CrossRef](#)]
51. Dong, S.; Zhao, C.; Chen, F.; Liu, Y.; Zhang, S.; Wu, H.; Zhang, L.; Liu, Y. The complete mitochondrial genome of the early flowering plant *Nymphaea colorata* is highly repetitive with low recombination. *BMC Genom.* **2018**, *19*, 614. [[CrossRef](#)] [[PubMed](#)]
52. Yang, H.; Chen, H.; Ni, Y.; Li, J.; Cai, Y.; Ma, B.; Yu, J.; Wang, J.; Liu, C. De Novo Hybrid Assembly of the *Salvia miltiorrhiza* Mitochondrial Genome Provides the First Evidence of the Multi-Chromosomal Mitochondrial DNA Structure of Salvia Species. *Int. J. Mol. Sci.* **2022**, *23*, 14267. [[CrossRef](#)] [[PubMed](#)]
53. Guo, W.; Grewe, F.; Fan, W.; Young, G.J.; Knoop, V.; Palmer, J.D.; Mower, J.P. Ginkgo and Welwitschia Mitogenomes Reveal Extreme Contrasts in Gymnosperm Mitochondrial Evolution. *Mol. Biol. Evol.* **2016**, *33*, 1448–1460. [[CrossRef](#)] [[PubMed](#)]
54. Sloan, D.B.; Alverson, A.J.; Chackalovcak, 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](#)]
55. Mower, J.P.; Case, A.L.; Floro, E.R.; Willis, J.H. Evidence against equimolarity of large repeat arrangements and a predominant master circle structure of the mitochondrial genome from a monkeyflower (*Mimulus guttatus*) lineage with cryptic CMS. *Genome Biol. Evol.* **2012**, *4*, 670–686. [[CrossRef](#)]
56. Alverson, A.J.; Wei, X.X.; Rice, D.W.; Stern, D.B.; Barry, K.; Palmer, J.D. Insights into the Evolution of Mitochondrial Genome Size from Complete Sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Mol. Biol. Evol.* **2010**, *27*, 1436–1448. [[CrossRef](#)] [[PubMed](#)]
57. Adams, K.L.; Ong, H.C.; Palmer, J.D. Mitochondrial gene transfer in pieces: Fission of the ribosomal protein gene rpl2 and partial or complete gene transfer to the nucleus. *Mol. Biol. Evol.* **2001**, *18*, 2289–2297. [[CrossRef](#)] [[PubMed](#)]
58. Rabah, S.O.; Lee, C.; Hajrah, N.H.; Makki, R.M.; Alharby, H.F.; Alhebshi, A.M.; Sabir, J.S.M.; Jansen, R.K.; Ruhlman, T.A. Plastome Sequencing of Ten Nonmodel Crop Species Uncovers a Large Insertion of Mitochondrial DNA in Cashew. *Plant Genome* **2017**, *10*, 1–14. [[CrossRef](#)]
59. Sprinzl, M.; Vassilenko, K.S. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **2005**, *33*, D139–D140. [[CrossRef](#)]
60. Lukes, J.; Kaur, B.; Speijer, D. RNA Editing in Mitochondria and Plastids: Weird and Widespread. *Trends Genet.* **2021**, *37*, 99–102. [[CrossRef](#)]
61. Brennicke, A.; Marchfelder, A.; Binder, S. RNA editing. *FEMS Microbiol. Rev.* **1999**, *23*, 297–316. [[CrossRef](#)] [[PubMed](#)]
62. Ma, Q.; Wang, Y.; Li, S.; Wen, J.; Zhu, L.; Yan, K.; Du, Y.; Ren, J.; Li, S.; Chen, Z.; et al. Assembly and comparative analysis of the first complete mitochondrial genome of *Acer truncatum* Bunge: A woody oil-tree species producing nervonic acid. *BMC Plant Biol.* **2022**, *22*, 29. [[CrossRef](#)] [[PubMed](#)]
63. Gerke, P.; Szovenyi, P.; Neubauer, A.; Lenz, H.; Gutmann, B.; McDowell, R.; Small, I.; Schallenberg-Rudinger, M.; Knoop, V. Towards a plant model for enigmatic U-to-C RNA editing: The organelle genomes, transcriptomes, editomes and candidate RNA editing factors in the hornwort *Anthoceros agrestis*. *New Phytol.* **2020**, *225*, 1974–1992. [[CrossRef](#)] [[PubMed](#)]
64. Liu, Y.J.; Xiu, Z.H.; Meeley, R.; Tan, B.C. Empty Pericarp5 Encodes a Pentatricopeptide Repeat Protein That Is Required for Mitochondrial RNA Editing and Seed Development in Maize. *Plant Cell* **2013**, *25*, 868–883. [[CrossRef](#)]
65. Grewe, F.; Edger, P.P.; Keren, I.; Sultan, L.; Pires, J.C.; Ostersetzer-Biran, O.; Mower, J.P. Comparative analysis of 11 Brassicales mitochondrial genomes and the mitochondrial transcriptome of *Brassica oleracea*. *Mitochondrion* **2014**, *19*, 135–143. [[CrossRef](#)]
66. Kovar, L.; Nageswara-Rao, M.; Ortega-Rodriguez, S.; Dugas, D.V.; Straub, S.; Cronn, R.; Strickler, S.R.; Hughes, C.E.; Hanley, K.A.; Rodriguez, D.N.; et al. PacBio-Based Mitochondrial Genome Assembly of *Leucaena trichandra* (Leguminosae) and an Intrageneric Assessment of Mitochondrial RNA Editing. *Genome Biol. Evol.* **2018**, *10*, 2501–2517. [[CrossRef](#)]
67. Arseneau, J.R.; Steeves, R.; Laflamme, M. Modified low-salt CTAB extraction of high-quality DNA from contaminant-rich tissues. *Mol. Ecol. Resour.* **2017**, *17*, 686–693. [[CrossRef](#)]

68. Emerman, A.B.; Bowman, S.K.; Barry, A.; Henig, N.; Patel, K.M.; Gardner, A.F.; Hendrickson, C.L. NEBNext Direct: A Novel, Rapid, Hybridization-Based Approach for the Capture and Library Conversion of Genomic Regions of Interest. *Curr. Protoc. Mol. Biol.* **2017**, *119*, 7–30. [[CrossRef](#)]
69. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
70. Kolmogorov, M.; Yuan, J.; Lin, Y.; Pevzner, P.A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **2019**, *37*, 540–546. [[CrossRef](#)]
71. Chen, Y.; Ye, W.C.; Zhang, Y.D.; Xu, Y.S. High speed BLASTN: An accelerated MegaBLAST search tool. *Nucleic Acids Res.* **2015**, *43*, 7762–7768. [[CrossRef](#)] [[PubMed](#)]
72. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
73. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)]
74. Wick, R.R.; Judd, L.M.; Gorrie, C.L.; Holt, K.E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **2017**, *13*, e1005595. [[CrossRef](#)]
75. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Pribelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [[CrossRef](#)] [[PubMed](#)]
76. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **2018**, *34*, 3094–3100. [[CrossRef](#)] [[PubMed](#)]
77. Wick, R.R.; Schultz, M.B.; Zobel, J.; Holt, K.E. Bandage: Interactive visualization of de novo genome assemblies. *Bioinformatics* **2015**, *31*, 3350–3352. [[CrossRef](#)]
78. Yu, J.; Li, J.; Zuo, Y.; Qin, Q.; Zeng, S.; Rennenberg, H.; Deng, H. Plastome variations reveal the distinct evolutionary scenarios of plastomes in the subfamily Cereoideae (Cactaceae). *BMC Plant Biol.* **2023**, *23*, 132. [[CrossRef](#)]
79. Shi, L.C.; Chen, H.M.; Jiang, M.; Wang, L.Q.; Wu, X.; Huang, L.F.; Liu, C. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* **2019**, *47*, W65–W73. [[CrossRef](#)]
80. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq—Versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [[CrossRef](#)]
81. Lewis, S.E.; Searle, S.M.J.; Harris, N.; Gibson, M.; Iyer, V.; Richter, J.; Wiel, C.; Bayraktaroglu, L.; Birney, E.; Crosby, M.A.; et al. Apollo: A sequence annotation editor. *Genome Biol.* **2002**, *3*, 1–14. [[CrossRef](#)] [[PubMed](#)]
82. Chan, P.P.; Lowe, T.M. tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. *Methods Mol. Biol.* **2019**, *1962*, 1–14. [[PubMed](#)]
83. Zhang, D.; Gao, F.L.; Jakovlic, I.; Zou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2020**, *20*, 348–355. [[CrossRef](#)] [[PubMed](#)]
84. 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](#)]
85. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [[CrossRef](#)]
86. Kurtz, S.; Choudhuri, J.V.; Ohlebusch, E.; Schleiermacher, C.; Stoye, J.; Giegerich, R. REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **2001**, *29*, 4633–4642. [[CrossRef](#)]
87. Zhang, H.E.; Meltzer, P.; Davis, S. RCircos: An R package for Circos 2D track plots. *BMC Bioinform.* **2013**, *14*, 244. [[CrossRef](#)]
88. Chen, C.J.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.H.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202. [[CrossRef](#)]
89. Edera, A.A.; Small, I.; Milone, D.H.; Sanchez-Puerta, M.V. Deepred-Mt: Deep representation learning for predicting C-to-U RNA editing in plant mitochondria. *Comput. Biol. Med.* **2021**, *136*, 104682. [[CrossRef](#)]
90. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
91. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.