



Article

Chloroplast Pan-Genomes and Comparative Transcriptomics Reveal Genetic Variation and Temperature Adaptation in the Cucumber

Lei Xia, Han Wang, Xiaokun Zhao, Hesbon Ochieng Obel , Xiaqing Yu , Qunfeng Lou , Jinfeng Chen * and Chunyan Cheng *

National Key Laboratory of Crop Genetics & Germplasm Enhancement and Utilization, College of Horticulture, Nanjing Agricultural University, Nanjing 210095, China; 2019204026@njau.edu.cn (L.X.)
* Correspondence: jfchen@njau.edu.cn (J.C.); chunyancheng@njau.edu.cn (C.C.);
Tel.: +86-025-8439-6370 (J.C. & C.C.)

Abstract: Although whole genome sequencing, genetic variation mapping, and pan-genome studies have been done on a large group of cucumber nuclear genomes, organelle genome information is largely unclear. As an important component of the organelle genome, the chloroplast genome is highly conserved, which makes it a useful tool for studying plant phylogeny, crop domestication, and species adaptation. Here, we have constructed the first cucumber chloroplast pan-genome based on 121 cucumber germplasms, and investigated the genetic variations of the cucumber chloroplast genome through comparative genomic, phylogenetic, haplotype, and population genetic structure analysis. Meanwhile, we explored the changes in expression of cucumber chloroplast genes under high- and low-temperature stimulation via transcriptome analysis. As a result, a total of 50 complete chloroplast genomes were successfully assembled from 121 cucumber resequencing data, ranging in size from 156,616–157,641 bp. The 50 cucumber chloroplast genomes have typical quadripartite structures, consisting of a large single copy (LSC, 86,339–86,883 bp), a small single copy (SSC, 18,069–18,363 bp), and two inverted repeats (IRs, 25,166–25,797 bp). Comparative genomic, haplotype, and population genetic structure results showed that there is more genetic variation in Indian ecotype cucumbers compared to other cucumber cultivars, which means that many genetic resources remain to be explored in Indian ecotype cucumbers. Phylogenetic analysis showed that the 50 cucumber germplasms could be classified into 3 types: East Asian, Eurasian + Indian, and Xishuangbanna + Indian. The transcriptomic analysis showed that *matK* were significantly up-regulated under high- and low-temperature stresses, further demonstrating that cucumber chloroplasts respond to temperature adversity by regulating lipid metabolism and ribosome metabolism. Further, *accD* has higher editing efficiency under high-temperature stress, which may contribute to the heat tolerance. These studies provide useful insight into genetic variation in the chloroplast genome, and established the foundation for exploring the mechanisms of temperature-stimulated chloroplast adaptation.

Keywords: chloroplast pan-genome; cucumber; genetic variation; RNA editing efficiency; temperature adversity



Citation: Xia, L.; Wang, H.; Zhao, X.; Obel, H.O.; Yu, X.; Lou, Q.; Chen, J.; Cheng, C. Chloroplast Pan-Genomes and Comparative Transcriptomics Reveal Genetic Variation and Temperature Adaptation in the Cucumber. *Int. J. Mol. Sci.* **2023**, *24*, 8943. <https://doi.org/10.3390/ijms24108943>

Academic Editor: Bartolome Sabater

Received: 4 April 2023

Revised: 1 May 2023

Accepted: 3 May 2023

Published: 18 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

The cucumber (*Cucumis sativus* L.) originated in India, and is one of the major vegetable crops [1]. As the first horticultural crop whose whole genome has been sequenced, cucumber has been studied extensively in several areas, including sexual differentiation and vascular bundle formation [2]. *Cucumis sativus* var. *hardwickii*, whose plant is diminutive and used as an herbal cure, has been identified as the wild ancestor of cucumber, according to earlier research [3]. Wild cucumber has been transformed via crop domestication into delectable vegetables that are widely cultivated worldwide [4]. Depending on geographical location, cucumber germplasm resources can be divided into four groups: the

Indian, East Asian, Eurasian, and Xishuangbanna groups [3,5,6]. To better understand the evolutionary mechanisms of phenotypic diversity in cucumber, a huge number of genome sequencing projects have been undertaken, such as the resequencing of 115 cucumber core germplasms [3] and the graph-based pan-genome [4] of cucumber. Although the cucumber genome has been extensively studied, very little research has focused on the organelle genome. Until now, only 13 cucumber organelle genomes have been reported, including 12 chloroplast genomes [7–13] and 1 mitochondrial genome [14]. Performing a chloroplast pan-genome study can deepen the understanding of cucumber genetics.

Chloroplasts are key organelles in plants that are involved in important biological processes such as photosynthesis and plant immunity [15]. There are three sets of genetic material in plant cells: the nuclear genome, the mitochondrial genome, and the chloroplast genome [16]. In contrast to the nuclear genome, the chloroplast and mitochondrial genomes are typically inherited maternally [17]. The chloroplast genome is a typical circular DNA module that ranges in size from 115 to 180 kb, multiple copies of which exist in the cell [18]. Chloroplast genomes have a typical quadripartite circular structure, including a large single copy (LSC) region, a small single copy (SSC) region and two inverted repeats (IRs) regions [19]. Based on the low mutation rate feature, chloroplast genomes provide valid information about plastid genome evolution as well as plant phylogeny [20,21]. In recent years, chloroplast genome sequencing has been completed in species such as *Ginger* [22], *Asteraceae* [23], and *Brassicaceae* [24], and different hotspots have been identified that can be used for phylogenetic analysis [25,26]. Furthermore, chloroplast genome sequences have been used for population genetics analyses. For example, 412 rice chloroplast and mitochondrial genomic analyses revealed that indica and japonica rice underwent different domestication processes [27]. With the continuous development of sequencing technology and improvement of chloroplast genome assembly technology, chloroplast pan-genome analysis has been carried out in several species such as oilseed rape [28], Japanese apricot [29], sweet potato [30], and pecan [31], which have explained their origin, evolutionary, phylogeography, and genetic diversity. Meanwhile, performing chloroplast pan-genome studies and analyzing the overall polymorphism of chloroplast genome sequences are important to explore candidate loci for genetically related agronomic traits in chloroplasts. For example, cold tolerance in cucumber is regulated by the F1FO-ATP synthase (CF1FO-ATPase) beta-subunit gene (*atpB*) [32].

In this study, the chloroplast pan-genome was constructed to explore the genetic diversity of 121 re-sequenced cucumber germplasm resources. Comparative transcriptomics clarifies how cucumber plastid genes respond to temperature stress. These results provide a basis for the study of chloroplast pan-genomes, genetic diversity, and adaptation to temperature stress in cucumbers.

2. Results

2.1. General Characteristics of 50 Cucumber Chloroplast Genome

A total of 50 complete chloroplast genomes were successfully assembled and annotated from 121 samples. These chloroplast genomes range in length from 156,616–157,641 bp, but have a common 37% GC content (Table S1, Figure 1). These chloroplast genomes have a typical quadripartite molecular structure, including LSC (86,339–86,883 bp), SSC (18,069–18,363 bp), and two IRs (25,166–25,797 bp). Comparative genome analysis with the reference cucumber variety “GY14” revealed good co-linearity of our assembled chloroplast genome (Figure S1), demonstrating that the assembly of all chloroplast genomes was accurate. The annotation information revealed that the chloroplast genome has 85 common protein-coding genes, 37 common tRNA genes, and 8 common rRNA genes (Table S2), where *ndhB*, *rps7*, *rps12*, *rpl2*, *rpl23*, and *ycf2* have two copies. Among the protein-coding genes, 12 genes contained introns, namely *rps16*, *atpF*, *rpoC1*, *ycf3*, *rps18*, *rps12*, *rpl2*, two copies of *ycf2*, two copies of *ndhB*, and *ndhA*.

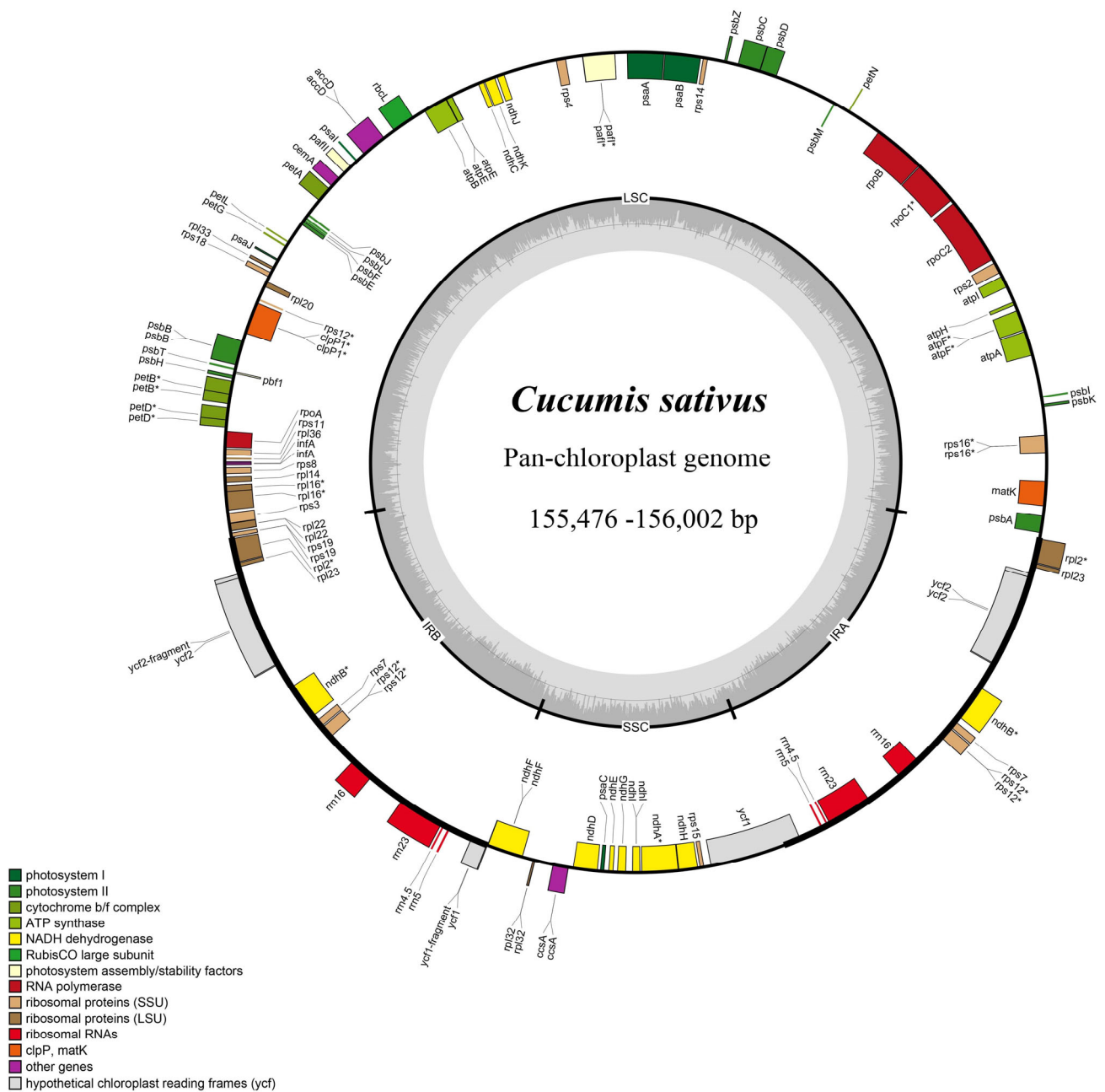


Figure 1. Pan-chloroplast genome map of *Cucumis sativus*. * Represents a gene containing an intron.

Based on chloroplast genome size differences, we classified these 50 cucumber chloroplast genomes into 23 categories (Table S3). In the East Asian-type cucumbers, 17 germplasm resources were classified into 4 categories: Huang Gua, Qiu Huang Gua, Liao Tong Mi Ci, and Jia Huang Gua. In the Eurasian-type cucumbers, the 11 germplasm resources were classified into 6 categories: Altaisky Ranny, SC 50, 8181, EC1, Ames 1208, and SC 53-B (6). In the Xishuangbanna-type cucumbers, 8 germplasm resources were classified into 4 categories: SWCC1, SWCC6, Man Wa Yuan Zong Huang Di Huang Gua, and SWCC8. In the Indian-type cucumbers, 14 germplasm resources were classified into 9 categories: CUS 155, CUS 233, USM 414, Hw 2, 8288, 11621, and Hw 3. Meanwhile, we found that the chloroplast genome of Indian-type cucumbers is rich in genetic variation compared to the other three types, which demonstrates the loss of genetic variation in cucumbers during domestication by artificial selection.

2.2. Comparative Analysis of Chloroplast Genome

Chloroplast DNA is usually used to explore genetic diversity within species, and expansion or narrowing of the IR region in the chloroplast genome can lead to size variation at the LSC/IRB/SSC/IRA boundary. In this study, we performed a comparative analysis of cucumber chloroplast genomes. When comparing the chloroplast genomes of different cucumber ecotypes, we found that *ycf1* was the only difference in the IRa-SSC (JSA) (Figure S2). When comparing the chloroplast genomes of the East Asian-type cucumbers (Figure 2a), we found that the four boundaries were relatively conserved, differing only in *ycf1* for the IRa-SSC (JSA) and IRb-SSC (JSB) boundaries. The results of Eurasian-type (Figure 2b) and Xishuangbanna-type (Figure 2c) cucumber chloroplast genome comparative analyses found that only the JSB boundary differs between the two. However, when performing chloroplast genome comparative analysis of Indian-type (Figure 2d) cucumbers, it is found that they differed in the four boundaries of JSA, JSB, IRa-LSC (JLA), and IRb-LSC (JLB). These results also indicate that there is expansion and contraction of chloroplast DNA boundary regions between different species of the same type studied, most notably in the Indian-type cucumbers.

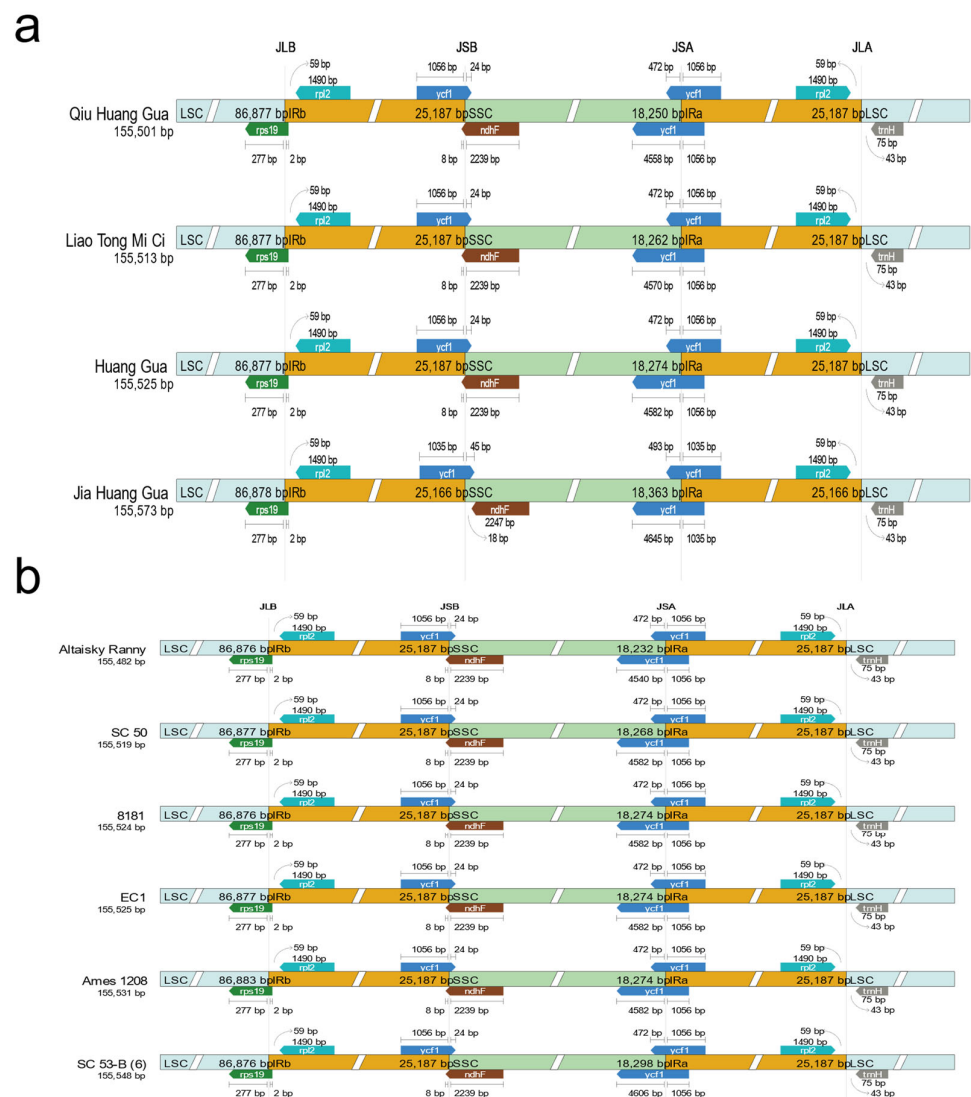


Figure 2. Cont.

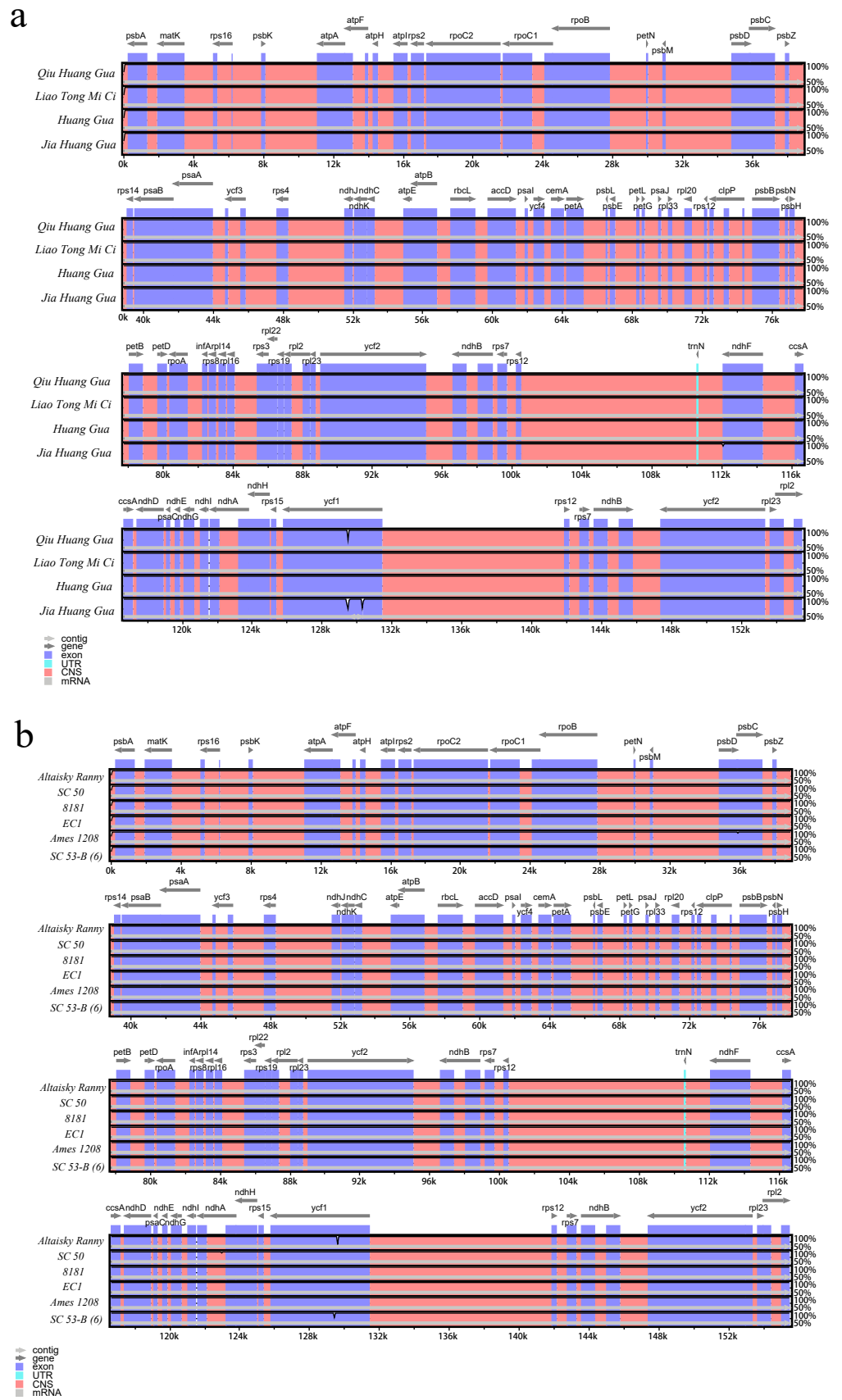


Figure 3. Cont.

2.3. Phylogenetic Analysis and Species Delimitation

An ML tree was constructed based on the whole chloroplast genome to study the phylogenetic relationships of cucumbers. The tree shows that the 50 cucumber germplasm resources were divided into 3 main branches (Figure 4a): East Asian, Eurasian + Indian, and Xishuangbanna + Indian. The phylogenetic analysis found that Indian-type cucumbers are more closely related to Xishuangbanna-type cucumbers. Interestingly, we found three species excluded from their original ecotype classification: EC1, SWCC6, and 8288. This result might be caused by the extensive mutual introduction and cross-fertilization of different ecotypes of cucumber. Meanwhile, haplotype analysis was performed using 50 cucumber chloroplast genomes. The results classified 50 cucumbers into 13 categories (Figure 4b). Among them, East Asian and Xishuangbanna were divided into 2 groups, and Eurasian and Indian-type were divided into 3 and 7 groups, respectively. The results of EC1, SWCC6, and 8288 divisions were consistent with the results of evolutionary tree analysis (Figure 4c), proving the accuracy and robustness of the results obtained in this study. On the other hand, it turned out that the Indian type has more haplotypic material compared to other ecotypes, demonstrating its rich genetic diversity.

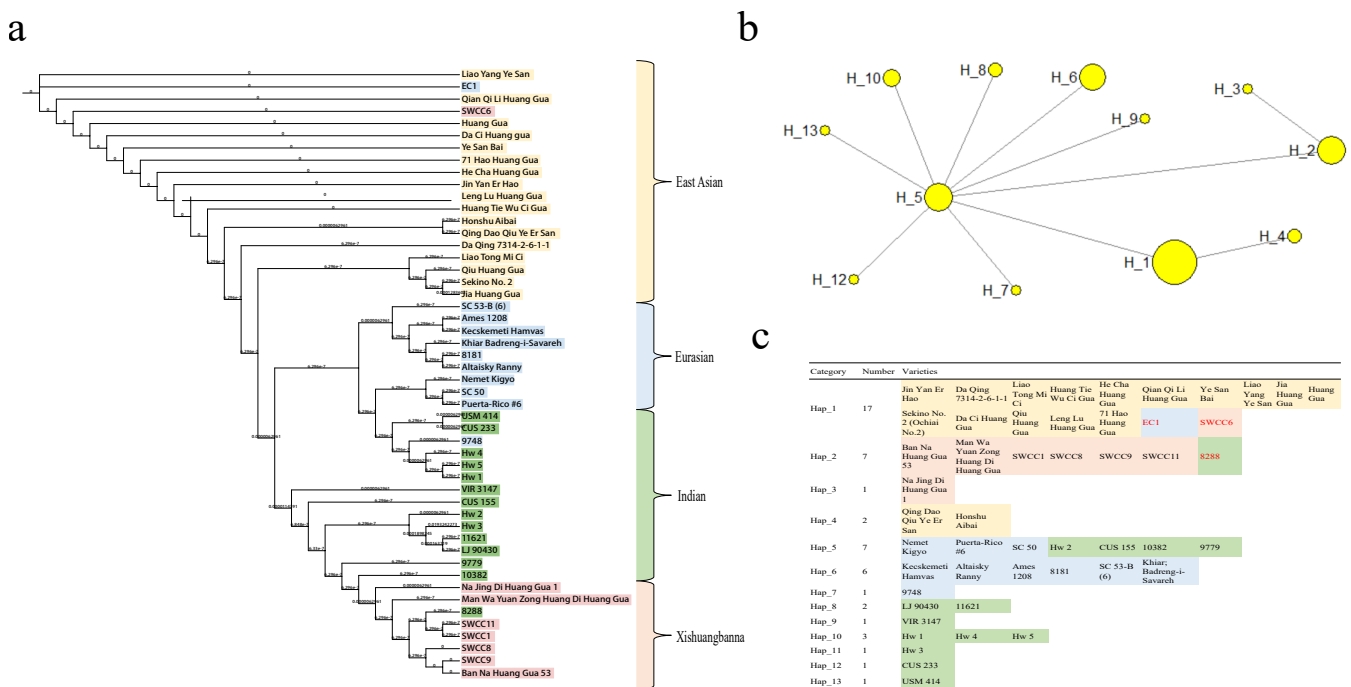


Figure 4. Phylogenetic analysis and species delimitation in cucumbers. (a) The ML phylogenetic tree based on complete chloroplast genomes. Haplotype analysis (b) and classification (c) of 50 cucumber chloroplast genomes.

2.4. Genetic Variation and Structure Analysis

To investigate the genetic diversity of cucumber chloroplast genomes, we analyzed the nucleotide diversity (Pi) values and variants of 50 cucumber chloroplast genomes. Highly variable and significantly higher Pi values > 0.002 were found in *ycf1*, *accD*, *clpP*. Among them, the *ycf1* region had the highest divergence value, which was 0.01. The gene *ycf1* harbored the most variants, which corresponded to the results of the comparative genome analysis. Variants in cucumber chloroplast DNA were detected in each ecotype (Figure 5b). Among the East Asian, Eurasian, and Xishuangbanna types, 4 (4 Indels), 14 (8 SNPs + 6 Indels), and 6 (4 SNPs + 2 Indels) variants were found, respectively. Unlike the aforementioned types, a large number of genetic variants were found in the Indian type, including 2695 SNPs and 376 Indels. These results demonstrate that Indian cucumbers have more genetic variation compared to other cucumbers.

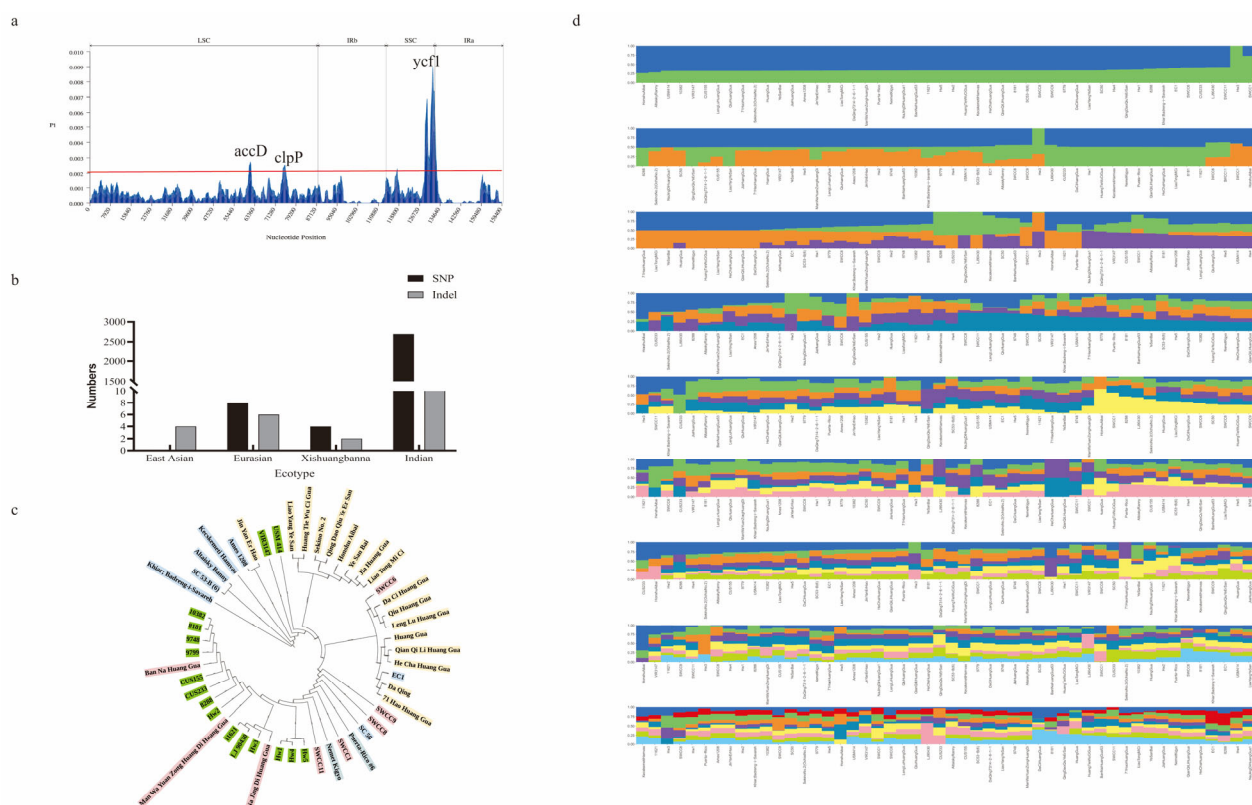


Figure 5. Genetic variation and structure analysis of 50 cucumber chloroplast genomes. (a) Statistics of nucleotide diversity (P_i) in 50 cucumber chloroplast genomes with parameters of a window length of 1000 bp and a step size of 100 bp. (b) The variations in 50 cucumber chloroplast genomes. The phylogenetic tree (c) and population structure analyses (d) are based on SNP.

Phylogenetic tree and population structure analyses were performed by extracting SNPs. The phylogenetic tree analysis showed that four major groups were clustered (Figure 5c), which is generally in agreement with the chloroplast genome phylogenetic tree results. However, Jin Yan Er Hao did not cluster with the East Asian cucumber, which may be caused by differences in mutation rates between coding and non-coding regions of the chloroplast genome. Population structures were analyzed, with the K value ranging from 2 to 10 (Figure 5d). The cross-validation (CV) error was the lowest with $K = 9$ (Figure S3). Taken together, the findings suggested that the cucumber population could be divided into nine groups.

2.5. Response of Plastid Genes to Temperature Stress in Cucumber

Chloroplasts are thought to be involved in temperature sensing and adaptive regulation, which are key components of the response to temperature stimuli. We use transcriptome data to mine how temperature stresses affect the development of plant plastids. After the high-temperature treatment, we found that a large number of chloroplast genes were down-regulated in expression (Figure 6a, Table S4). Among them, genes related to ribosome metabolism in photosynthesis were significantly inhibited, such as *rps3*, *rps8*, *rps9*, and *rpl14*, among others. This indicates that ribosomal metabolism is hindered under high-temperature conditions and is an important factor affecting chloroplast development and photosynthesis in plants. Interestingly, we found that the expression of nine genes were up-regulated, including *accD* and *matK*, both of which are closely related to lipid synthesis. It is hypothesized that plants under high-temperature stress resist adversity stress by regulating lipid metabolism. After low-temperature treatment, we found significant inhibition of genes related to ribosome metabolism (Figure 6b, Table S5), such as *rps18* and *rpl33*. These results suggest that ribosome metabolism is the single most important factor

affecting plant chloroplast development and photosynthesis under temperature adversity stress. Similarly, we found that *matK* was significantly up-regulated at low temperatures, demonstrating that chloroplast resistance to temperature stress through the regulation of lipid metabolism is an important pathway for plant stress tolerance.

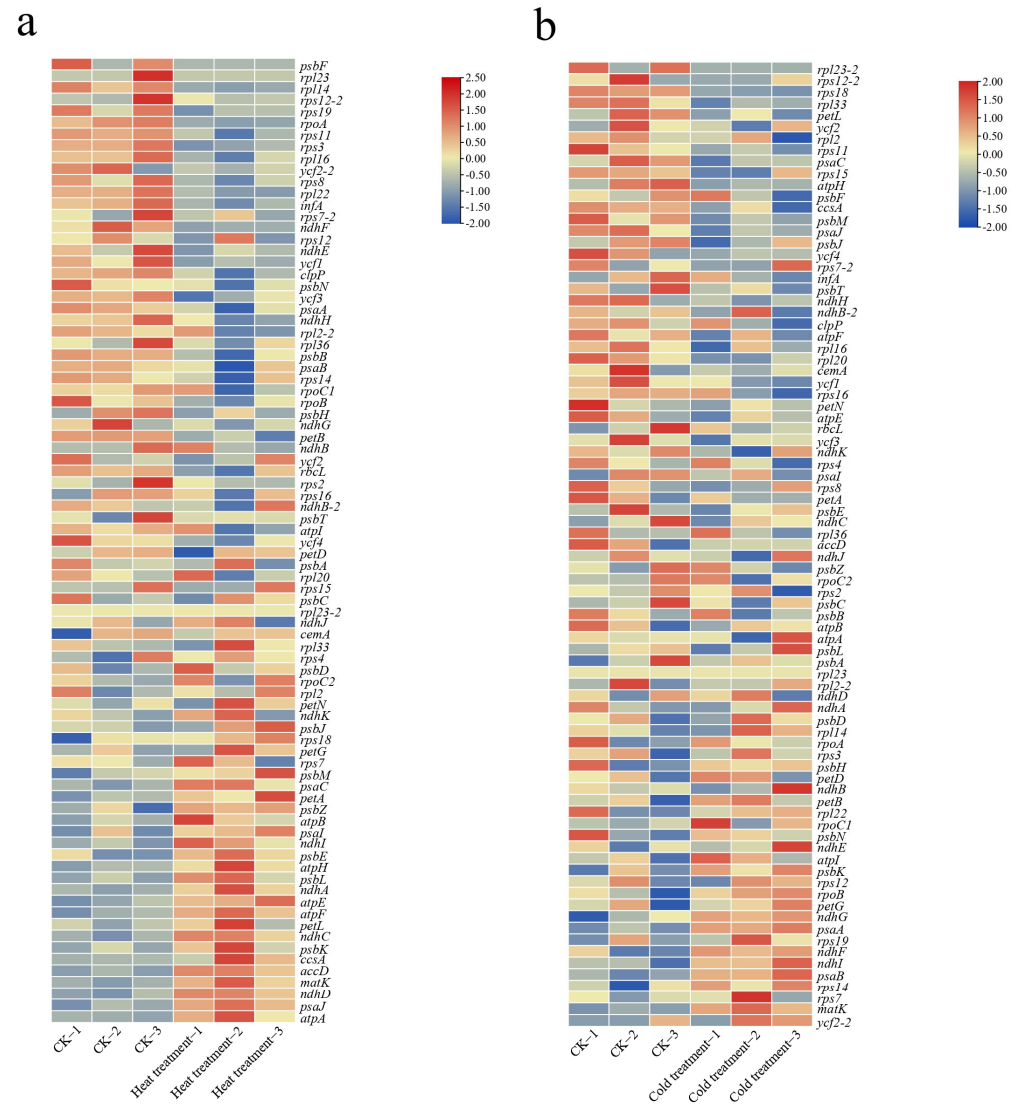


Figure 6. Heatmap of plastid gene expression under temperature stress in the cucumber. (a) Heatmap of plastid gene expression under heat stress and at normal temperature (CK). (b) Heatmap of plastid gene expression under cold stress and at normal temperature (CK).

RNA editing plays an important role in plant growth, development, and evolutionary adaptation, and environmental stress can affect the efficiency of RNA editing. We used transcriptome data to analyze the RNA editing efficiency of plant plastid genes under temperature stress. After high-temperature treatment, we found RNA editing events in 21 genes and identified 36 editing sites (Figure 7a, Table S6), which far exceeds the 24 editing sites detected at normal temperature (Table S7). Of these, a total of 25 were C-to-U, accounting for 69.44% of RNA editing events, which is lower than the C-to-U editing efficiency of 83.33% at normal temperature. Compared with normal temperature, the RNA editing efficiency of *matK*, *accD*, *atpB*, *rpoC2*, and *petA* was significantly improved under high-temperature conditions. Among them, we found that the increased efficiency of *accD* and *matK* RNA editing may promote the expression of *accD* and *matK* in response to high-temperature stress. After low-temperature treatment, RNA editing events occurred

in 21 genes, and 34 editing sites were identified (Figure 7b, Table S8). Of those 34 editing events, 29 were C-to-U, accounting for 85.29% of RNA editing events, which were not significantly different from the 27 C-to-U editing events detected at normal temperature (Table S9). Compared with normal temperature, we found that the RNA editing efficiency of *atpA*, *rps2* and *rps4* was significantly improved, while the RNA editing efficiency of *rpoB*, *psaA*, *rbcL* and *accD* was significantly reduced in low-temperature conditions. These results suggested that low temperature leads to a decrease in editing efficiency of photosynthetic genes (*rpoB*, *psaA*, and *rbcL*), which may further affect plant photosynthesis.

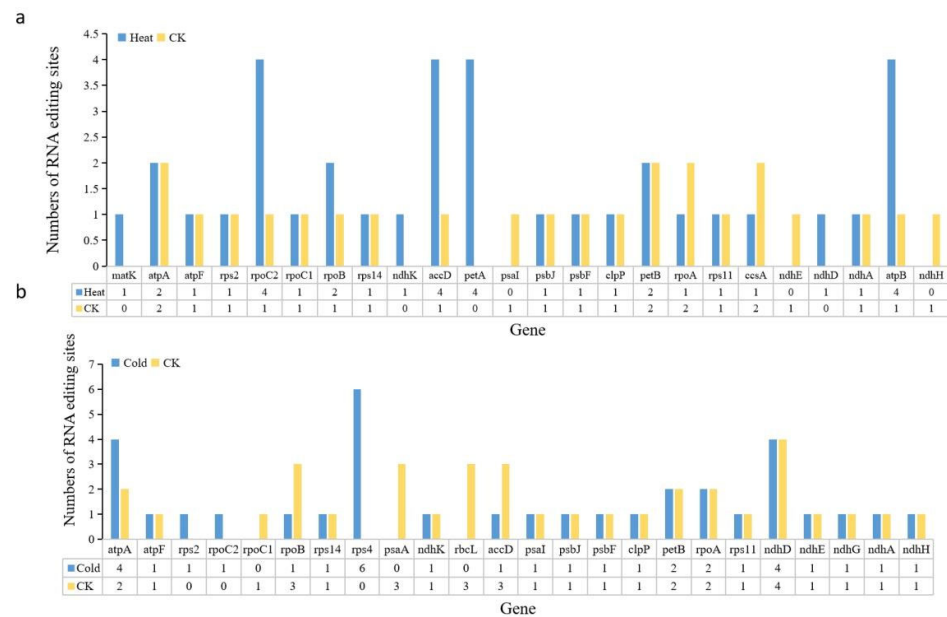


Figure 7. Analysis of RNA editing efficiency in plastid genes under temperature stress in cucumber. (a) Analysis of RNA editing efficiency of heat stress and normal temperature (CK). (b) Analysis of RNA editing efficiency of cold stress and normal temperature (CK).

3. Discussion

In this study, 50 cucumber chloroplast genomes were assembled and annotated. This is the first report on the cucumber chloroplast pan-genome, which is important for exploring genetic diversity and chloroplast genome evolution in cucumbers. The 50 cucumber chloroplast genomes showed high similarity in genome structure, GC content, and gene composition with previously reported cucumber chloroplast genomes [9,11]. However, compared to other cucumbers, the chloroplast genomes of Indian cucumbers have more variations, such as genome size. Comparative analysis of chloroplast genome size in Indian cucumbers revealed that the differences were mainly in the LSC and SSC regions, which is consistent with the results of *Eriocaulon* [33]. Expansion and contraction of the chloroplast genome is a common evolutionary phenomenon in plants, which can be altered by IR contraction and expansion [34]. Analysis of 50 cucumber germplasm IRB/SSC/IRA/LSC boundaries revealed that the IR region is conserved and most of the substitutions occur in the SSC and LSC regions. This is similar to findings in the plastid genomes of species such as *Yam* [35], *Nicotiana* [36], and *Mukdenia* [37].

The chloroplast genome regions (*atpF-H*, *matK*, and *rbcL*) have been used as candidate markers for DNA barcoding in plants to facilitate rapid identification of species [38,39]. Based on the results of mVISTA, nucleotide diversity, and comparative chloroplast genome analyses, 8 highly variable regions were identified in 50 cucumbers, including 3 intergenic regions (*rbcL-accD*, *atpB-ycf3*, and *ndhF-trnN*) and 5 gene regions (*rbcL*, *psaB*, *clpP*, *accD*, and *ycf1*). Previous studies have shown that *ycf1* shows large variations in different species and is used as a chloroplast DNA barcode to identify different species [40]. In this study,

ycf1 was highly variable between different cucumbers (Figure S4), whose coding sequences can be used to distinguish different cucumber ecotype germplasm resources.

The chloroplast genome has been used as an effective marker for studying species diversity in many plants [21]. Genetic diversity of cucumber germplasm was detected based on 37 SSR markers in non-coding regions of the chloroplast genome [11]. The conduct of chloroplast pangenome studies could provide a new tool to explore the genetic diversity of cucumber germplasms. Based on phylogenetic tree results, the 50 cucumber germplasms were divided into 4 major branches: East Asian, Eurasian, Xishuangbanna, and Indian, which are the same as the previous phylogenetic relationship of the nuclear genome [3]. However, some varieties were not well-differentiated at the chloroplast genome level, indicating that the chloroplast genome of cucumber is conserved within the species. Among these, Indian wild cucumbers were more distantly related to the East Asian cultivated cucumbers, which could be the result of artificial domestication selection [4]. In terms of phylogenetic relationships, the Indian cucumbers and the Xishuangbanna cucumbers are closely related, likely due to their close geographical location and similar ecological environment [41]. Interestingly, we found that the three varieties—EC1, SWCC6, and 8288—did not cluster together with their original ecotype classification. The conflict of phylogenetic trees between chloroplast genomes and nuclear genes is also common in angiosperms [42,43]. This result may be caused by extensive mutual introgression and cross-fertilization of different cucumber ecotypes, which is consistent with citrus results [43].

Genetic diversity and population structure of cucumber varieties have been studied using resequencing data from 115 cucumbers. The results of phylogenetic evolution in the nuclear genome classified 115 cucumbers into 4 groups (East Asian, Eurasian, Xishuangbanna, and Indian), consistent with the chloroplast genome phylogenetic findings [3]. Molecular markers such as SSR and SNP can be used for variety identification and genetic evolution studies [44–48]. In this study, we extracted SNPs and performed a genetic structure analysis on resequencing data from 50 cucumbers. The haplotype and population genetic structure results found that the Indian cucumber is rich in genetic diversity and has more haplotypes, which is the same as the previous resequencing results [6]. However, the chloroplast genomes of different germplasms of the same ecotype are indistinguishable, such as “Huang Gua and Da Ci Huang Gua” in the East Asian type, likely due to characteristics such as a highly conserved chloroplast genome and little intraspecific genetic variation within species.

Chloroplasts play a central role in the perception and integration of temperature stresses in plants [49]. Significant down-regulation of ribosomal protein genes in the chloroplast genome of cucumber was also identified in this study through the analysis of transcriptome data from high- and low-temperature stresses. Plastid ribosomal proteins are an important component of the protein synthesis machinery and have multiple roles in plant growth and development [50]. Previous studies have found that overexpression of the *rps5* gene improves plant cold tolerance in *Arabidopsis* [51], while reduced expression of *rps1* causes a significant reduction in plant heat tolerance [52]. RNA editing regulates the expression of organelle genes in plants in response to different environmental stress [53]. Our study found that *accD* exhibited significantly higher expression in high- and low-temperature treatments, which may be a defense measure of plant chloroplasts in response to temperature stimuli. Meanwhile, *accD* showed a remarkable increase in RNA editing efficiency under high-temperature conditions. Fatty acid biosynthesis is initiated in the chloroplast under the control of the *accD* gene [54]. Improving the editing efficiency of *accD* in plants significantly enhanced heat tolerance in *Arabidopsis* [55], implying that tolerance in plants such as cucumber could be improved through plastid transgenesis. C-to-U RNA editing events occur frequently in organelle genes of plants. Previous studies have found that the RNA editing efficiency of C-to-U in the *Arabidopsis* organelle genomes is reduced in response to heat stress, which is associated with stalling of ribosomal genes [56]. Similarly, it was found that the editing efficiency of C-to-U in the chloroplast genome of cucumber

was reduced after heat treatment, which implied that the reduced rate of C-to-U RNA editing might be involved in regulating the stress response of plants.

4. Materials and Methods

4.1. Plant Material and Sequencing

The 121 cucumber germplasm resources used in this study were derived from 115-core cucumber germplasm collections and 6-core germplasms from our lab (Table S10). Among the germplasm collections, the resequenced data of 115-core germplasm resources were downloaded from the SRA database (SRA056480) on the NCBI website. The 6-core germplasms were grown in a greenhouse at Baima Teaching and Research Base of Nanjing Agricultural University. Total genomic DNA was extracted from fresh leaves using a modified CTAB method [57]. DNA concentration and quality were measured using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Carlsbad, CA, USA). The qualified DNA samples were sequenced on the Illumina NovaSeq 6000 platform (Illumina Inc., San Diego, CA, USA).

4.2. Assembly and Annotation of Chloroplast Genomes

The fastq-dump in the SRA toolkit (<https://ftp-trace.ncbi.nlm.nih.gov/sra/sdk/2.8.0/>, accessed on 3 April 2023) was used to convert the SRA data from NCBI to the fastq file. The raw reads were trimmed and quality-controlled by Trimmomatic [58] to harvest the clean reads, which were used for the chloroplast genome assembly. The chloroplast genome was assembled by GetOrganelle software to obtain chloroplast circular DNA molecules [59]. The complete chloroplast genome sequences were compared with the reference genome GY14 (DQ865975.1) using TBtools software [60] to verify the accuracy of assembly. Chloroplast genes were annotated by module GeSeq [61] and manually corrected. Organellar Genome DRAW was used to draw the chloroplast map.

4.3. Comparative Analysis of Chloroplast Genomes

The IRscope [62] was used to detect LSC/IRB/SSC/IRA boundaries between the chloroplast genome sequences of different cucumbers for comparative analysis. The chloroplast genomes of different cucumbers were compared using the mVISTA program [63]. The annotation of GY14 (DQ865975.1) and *Cucumis sativus* var. *hardwickii* (KT852702.1) was used as a reference in the Shuffle-LAGAN model, where *Cucumis sativus* var. *hardwickii* (KT852702.1) was used for comparison with Indian-type cucumbers.

4.4. Phylogenetic Tree and Haplotype Analysis

This phylogenetic tree analysis was constructed based on the complete chloroplast genome sequences of 50 cucumber germplasms. All sequences were aligned using MAFFT software with default settings [64]. The phylogenetic tree was created using IQ-TREE 2 [65] by selecting the maximum likelihood (ML) technique and 1000 bootstrap replications through Ultrafast bootstrap parameters. It was visualized and modified by iTOL (interactive tree of life). Nucleic acid diversity and haplotype analyses of the chloroplast genome were performed with DnaSP 6 software [66]. Haplotype network diagrams were drawn using network software (<https://www.fluxus-engineering.com/sharenet.htm>, accessed on 3 April 2023).

4.5. Variants Calling

Clean reads were aligned to the reference chloroplast genome by BWA to obtain comparison files [67]. Used SAMtools and BCFtools to call the variants from the comparison files [68,69]. SNPs and InDels were filtered using VCFtools [70], setting parameters for deletion rates below 50% and minor allele frequencies above 0.05.

4.6. Phylogenetic Tree and Population Genetic Structure Based on Variants

Plink [71] was used to convert the vcf format into the phylip format, which was used for the phylogenetic tree analysis. FastTree [72] was selected to construct the phylogenetic tree by the ML method. CV errors were assessed using ADMIXTURE [73], setting parameters from $K = 2$ to $K = 10$. The visualization was performed by the R package (bar graph).

4.7. RNA-Seq Analysis

RNA-Seq data were downloaded from the SRA database (SRP305598 and SRP262962) on the NCBI website. The fastq-dump in the SRA toolkit (<https://ftp-trace.ncbi.nlm.nih.gov/sra/sdk/2.8.0/>, accessed on 3 April 2023) was used to convert the data to the fastq file. The raw reads' quality was trimmed by Trimmomatic software [58]. The quality-controlled reads were aligned to the chloroplast genome using Hisat2 [74]. Mapped reads were counted using featureCounts [75]. The comparative analysis of gene expression levels between samples was performed using DESeq2 [76]. Fragments per kilobase million (FPKM) were calculated, and differentially expressed genes (DEGs) were defined as those with a fold change (FC) > 2 between samples and a p -value < 0.05 .

4.8. Analysis of RNA Editing Efficiency

RNA-Seq data were downloaded from the SRA database (SRP305598 and SRP262962) on the NCBI website. The fastq-dump in the SRA toolkit (<https://ftp-trace.ncbi.nlm.nih.gov/sra/sdk/2.8.0/>, accessed on 3 April 2023) was used to convert the data to the fastq file. The raw reads' quality was trimmed by Trimmomatic software [58]. The quality-controlled reads were aligned to the chloroplast genome using BWA [67]. Use GATK [77] to mine variation information such as SNP and Indel. REDO [78] was used to detect RNA editing sites in the chloroplast genome in the variant calling files. If there were 2 or more RNA editing site events in 3 replicate samples, RNA editing was considered to have occurred.

5. Conclusions

In this study, we assembled and compared 50 cucumber chloroplast genomes, and found 8 highly variable regions that can be used as potential sources of molecular markers for species identification. The haplotype and population genetic structure results revealed that the Indian-type cucumbers have more genetic variation compared to other cucumbers, which means that many genetic resources are remaining to be explored in the Indian-type cucumbers. Transcriptome results showed that *accD*, *matK*, and ribosomal protein genes in the chloroplast genome expression were disrupted under temperature stimulation. Taken together, our results provide useful information on genetic variation and adaptation to temperature stress in the chloroplast genome.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms24108943/s1>.

Author Contributions: Conceptualization, J.C. and C.C.; methodology, L.X. and C.C.; software, L.X. and X.Z.; formal analysis, L.X. and H.W.; investigation, Q.L. and C.C.; data curation, C.C., H.O.O. and X.Y.; writing—original draft preparation, L.X.; writing—review and editing, J.C., L.X. and C.C.; project administration, X.Y.; funding acquisition, J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key R&D Program of China (2021YFD1200200), Jiangsu Agricultural Innovation of New Cultivars (PZCZ201719), Asia Cooperation Fund project-2021, Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institution.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All relevant data analyzed during this study are included in this article and additional files. The SRR numbers of all resequencing and transcriptome data analyzed during

this study are in Supplementary File S1. The 50 chloroplast genome sequences were released to GenBank and their accession numbers are available in Supplementary File S1.

Acknowledgments: This project is supported by the Bioinformatics Center of Nanjing Agricultural University.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Sebastian, P.; Schaefer, H.; Telford, I.R.H.; Renner, S.S. Cucumber (*Cucumis sativus*) and melon (*C. melo*) have numerous wild relatives in Asia and Australia, and the sister species of melon is from Australia. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14269–14273. [[CrossRef](#)] [[PubMed](#)]
- Huang, S.; Li, R.; Zhang, Z.; Li, L.; Gu, X.; Fan, W.; Lucas, W.J.; Wang, X.; Xie, B.; Ni, P.; et al. The genome of the cucumber, *Cucumis Sativus* L. *Nat. Genet.* **2009**, *41*, 1275–1281. [[CrossRef](#)] [[PubMed](#)]
- Qi, J.; Liu, X.; Shen, D.; Miao, H.; Xie, B.; Li, X.; Zeng, P.; Wang, S.; Shang, Y.; Gu, X.; et al. A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. *Nat. Genet.* **2013**, *45*, 1510–1515. [[CrossRef](#)] [[PubMed](#)]
- Li, H.; Wang, S.; Chai, S.; Yang, Z.; Zhang, Q.; Xin, H.; Xu, Y.; Lin, S.; Chen, X.; Yao, Z.; et al. Graph-based pan-genome reveals structural and sequence variations related to agronomic traits and domestication in cucumber. *Nat. Commun.* **2022**, *13*, 682. [[CrossRef](#)]
- Lv, J.; Qi, J.; Shi, Q.; Shen, D.; Zhang, S.; Shao, G.; Li, H.; Sun, Z.; Weng, Y.; Shang, Y.; et al. Genetic diversity and population structure of cucumber (*Cucumis sativus* L.). *PLoS ONE* **2012**, *7*, e46919. [[CrossRef](#)]
- Wang, X.; Bao, K.; Reddy, U.K.; Bai, Y.; Hammar, S.A.; Jiao, C.; Wehner, T.C.; Ramírez-Madera, A.O.; Weng, Y.; Grumet, R.; et al. The USDA cucumber (*Cucumis sativus* L.) collection: Genetic diversity, population structure, genome-wide association studies, and core collection development. *Hortic. Res.* **2018**, *5*, 64. [[CrossRef](#)]
- Lee, S.C.; Lee, H.O.; Joh, H.J.; Kim, I.; Lee, W.K.; Yang, T.J.; Song, K. The complete chloroplast genome sequence with a novel 24-bp deletion of a Korean solid green-type cucumber variety (*Cucumis sativus* var. *sativus*). *Mitochondrial DNA Part B Resour.* **2017**, *2*, 755–756. [[CrossRef](#)]
- Park, H.S.; Lee, W.K.; Lee, S.C.; Lee, H.O.; Joh, H.J.; Park, J.Y.; Kim, S.; Song, K.; Yang, T.J. Inheritance of chloroplast and mitochondrial genomes in cucumber revealed by four reciprocal F1 hybrid combinations. *Sci. Rep.* **2021**, *11*, 2506. [[CrossRef](#)]
- Liu, B.; Zhang, D.; Gao, L.Z. The complete chloroplast genome sequence of *Cucumis Sativus* var. *Hardwickii*, the wild progenitor of cultivated cucumber. *Mitochondrial DNA Part A* **2016**, *27*, 4627–4628. [[CrossRef](#)]
- Gao, J.; Liu, B.; Jiang, H.H. The complete chloroplast genome sequence of wild cucumber (*Cucumis sativus* var. *Hardwickii*). *Mitochondrial DNA Part A* **2016**, *27*, 4524–4525. [[CrossRef](#)]
- Plader, W.; Yukawa, Y.; Sugiura, M.; Malepszy, S. The complete structure of the cucumber (*Cucumis sativus* L.) chloroplast genome: Its composition and comparative analysis. *Cell. Mol. Biol. Lett.* **2007**, *12*, 584–594. [[CrossRef](#)] [[PubMed](#)]
- Kim, J.S.; Jung, J.D.; Lee, J.A.; Park, H.W.; Oh, K.H.; Jeong, W.J.; Choi, D.W.; Liu, J.R.; Cho, K.Y. Complete sequence and organization of the cucumber (*Cucumis sativus* L. cv. *Baekmibaekdadagi*) chloroplast genome. *Plant Cell Rep.* **2006**, *25*, 334–340. [[CrossRef](#)] [[PubMed](#)]
- Chung, S.M.; Gordon, V.S.; Staub, J.E. Sequencing cucumber (*Cucumis sativus* L.) chloroplast genomes identifies differences between chilling-tolerant and -susceptible cucumber lines. *Genome* **2007**, *50*, 215–225. [[CrossRef](#)] [[PubMed](#)]
- Alverson, A.J.; Rice, D.W.; Dickinson, S.; Barry, K.; Palmer, J.D. Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. *Plant Cell* **2011**, *23*, 2499–2513. [[CrossRef](#)]
- Odintsova, M.S.; Yurina, N.P. Plastid genomes of higher plants and algae: Structure and functions. *Mol. Biol.* **2003**, *37*, 649–662. [[CrossRef](#)]
- Heldt, H.-W.; Piechulla, B. *20—A Plant Cell Has Three Different Genomes*, 4th ed.; Heldt, H.-W., Piechulla, B.B.T.-P.B., Eds.; Academic Press: San Diego, CA, USA, 2011; pp. 487–526. ISBN 978-0-12-384986-1.
- Zhang, Q.; Liu, Y. Sodmergen examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species. *Plant Cell Physiol.* **2003**, *44*, 941–951. [[CrossRef](#)]
- Wicke, S.; Schneeweiss, G.M.; Depamphilis, C.W.; Müller, K.F.; Quandt, D. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol. Biol.* **2011**, *76*, 273–297. [[CrossRef](#)]
- Dobrogojski, J.; Adamiec, M.; Luciński, R. The chloroplast genome: A review. *Acta Physiol. Plant.* **2020**, *42*, 98. [[CrossRef](#)]
- Lu, M.; Zhang, H.; An, H. Chloroplast DNA-based genetic variation of *Rosa Roxburghii* in southwest China: Phylogeography and conservation implications. *Hortic. Plant J.* **2021**, *7*, 286–294. [[CrossRef](#)]
- Fu, X.; Xie, D.F.; Zhou, Y.Y.; Cheng, R.Y.; Zhang, X.Y.; Zhou, S.; He, X.J. Phylogeny and adaptive evolution of subgenus rhizirideum (*Amaryllidaceae*, *Allium*) based on plastid genomes. *BMC Plant Biol.* **2023**, *23*, 70. [[CrossRef](#)]
- Jiang, D.; Cai, X.; Gong, M.; Xia, M.; Xing, H.; Dong, S.; Tian, S.; Li, J.; Lin, J.; Liu, Y.; et al. Complete chloroplast genomes provide insights into evolution and phylogeny of *Zingiber* (*Zingiberaceae*). *BMC Genomics* **2023**, *24*, 30. [[CrossRef](#)] [[PubMed](#)]

23. Kim, G.B.; Lim, C.E.; Kim, J.S.; Kim, K.; Lee, J.H.; Yu, H.J.; Mun, J.H. Comparative chloroplast genome analysis of *Artemisia* (Asteraceae) in East Asia: Insights into evolutionary divergence and phylogenomic implications. *BMC Genomics* **2020**, *21*, 415. [[CrossRef](#)] [[PubMed](#)]
24. Javaid, N.; Ramzan, M.; Khan, I.A.; Alahmadi, T.A.; Datta, R.; Fahad, S.; Danish, S. The chloroplast genome of *Farsetia Hamiltonii* Royle, phylogenetic analysis, and comparative study with other members of clade C of Brassicaceae. *BMC Plant Biol.* **2022**, *22*, 384. [[CrossRef](#)] [[PubMed](#)]
25. Fu, Y.; Liu, F.; Li, S.; Tian, D.; Dong, L.; Chen, Y.; Su, Y. Genetic diversity of the wild Asian lotus (*Nelumbo Nucifera*) from northern China. *Hortic. Plant J.* **2021**, *7*, 488–500. [[CrossRef](#)]
26. Perumal, S.; Waminal, N.E.; Lee, J.; Koo, H.-J.; Choi, B.; Park, J.Y.; Ahn, K.; Yang, T.-J. Nuclear and chloroplast genome diversity revealed by low-coverage whole-genome shotgun sequence in 44 *Brassica Oleracea* breeding lines. *Hortic. Plant J.* **2021**, *7*, 539–551. [[CrossRef](#)]
27. Cheng, L.; Nam, J.; Chu, S.H.; Rungnapa, P.; Min, M.; Cao, Y.; Yoo, J.; Kang, J.S.; Kim, K.W.; Park, Y.J. Signatures of differential selection in chloroplast genome between *Japonica* and *Indica*. *Rice* **2019**, *12*, 65. [[CrossRef](#)]
28. Liu, H.; Zhao, W.; Hua, W.; Liu, J. A large-scale population based organelle pan-genomes construction and phylogeny analysis reveal the genetic diversity and the evolutionary origins of chloroplast and mitochondrion in *Brassica napus* L. *BMC Genom.* **2022**, *23*, 339. [[CrossRef](#)]
29. Huang, X.; Coulibaly, D.; Tan, W.; Ni, Z.; Shi, T.; Li, H.; Hayat, F.; Gao, Z. The analysis of genetic structure and characteristics of the chloroplast genome in different Japanese apricot germplasm populations. *BMC Plant Biol.* **2022**, *22*, 354. [[CrossRef](#)]
30. Xiao, S.; Xu, P.; Deng, Y.; Dai, X.; Zhao, L.; Heider, B.; Zhang, A.; Zhou, Z.; Cao, Q. Comparative analysis of chloroplast genomes of cultivars and wild species of sweetpotato (*Ipomoea Batatas* [L.] Lam). *BMC Genom.* **2021**, *22*, 262. [[CrossRef](#)]
31. Shen, J.; Li, X.; Chen, X.; Huang, X.; Jin, S. The complete chloroplast genome of *Carya Cathayensis* and phylogenetic analysis. *Genes* **2022**, *13*, 369. [[CrossRef](#)]
32. Oravec, M.W.; Havey, M.J. Polymorphism in the chloroplast atp synthase beta-subunit is associated with a maternally inherited enhanced cold recovery in cucumber. *Plants* **2021**, *10*, 1092. [[CrossRef](#)] [[PubMed](#)]
33. Li, E.; Liu, K.; Deng, R.; Gao, Y.; Liu, X.; Dong, W.; Zhang, Z. Insights into the phylogeny and chloroplast genome evolution of *Eriocaulon* (Eriocaulaceae). *BMC Plant Biol.* **2023**, *23*, 32. [[CrossRef](#)] [[PubMed](#)]
34. Guo, Y.Y.; Yang, J.X.; Bai, M.Z.; Zhang, G.Q.; Liu, Z.J. The chloroplast genome evolution of Venus slipper (*Paphiopedilum*): IR expansion, SSC contraction, and highly rearranged SSC regions. *BMC Plant Biol.* **2021**, *21*, 248. [[CrossRef](#)] [[PubMed](#)]
35. Cao, J.; Jiang, D.; Zhao, Z.; Yuan, S.; Zhang, Y.; Zhang, T.; Zhong, W.; Yuan, Q.; Huang, L. Development of chloroplast genomic resources in Chinese Yam (*Dioscorea Polystachya*). *Biomed Res. Int.* **2018**, *2018*, 6293847. [[CrossRef](#)]
36. Wang, S.; Gao, J.; Chao, H.; Li, Z.; Pu, W.; Wang, Y.; Chen, M. Comparative chloroplast genomes of *Nicotiana* Species (*Solanaceae*): Insights into the genetic variation, phylogenetic relationship, and polyploid speciation. *Front. Plant Sci.* **2022**, *13*, 1–15. [[CrossRef](#)] [[PubMed](#)]
37. Liu, L.; Wang, Y.; He, P.; Li, P.; Lee, J.; Soltis, D.E.; Fu, C. Chloroplast genome analyses and genomic resource development for epilithic sister genera *Oresitrophe* and *Mukdenia* (*Saxifragaceae*), using genome skimming data. *BMC Genomics* **2018**, *19*, 235. [[CrossRef](#)]
38. Ramesh, G.A.; Mathew, D.; John, K.J.; Ravisanakar, V. Chloroplast gene *matK* holds the barcodes for identification of *Momordica* (*Cucurbitaceae*) species from Indian subcontinent. *Hortic. Plant J.* **2022**, *8*, 89–98. [[CrossRef](#)]
39. Dong, W.; Liu, J.; Yu, J.; Wang, L.; Zhou, S. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS ONE* **2012**, *7*, e35071. [[CrossRef](#)]
40. Dong, W.; Xu, C.; Li, C.; Sun, J.; Zuo, Y.; Shi, S.; Cheng, T.; Guo, J.; Zhou, S. Ycf1, the most promising plastid DNA barcode of land plants. *Sci. Rep.* **2015**, *5*, 8348. [[CrossRef](#)]
41. Bo, K.; Ma, Z.; Chen, J.; Weng, Y. Molecular mapping reveals structural rearrangements and quantitative trait loci underlying traits with local adaptation in semi-wild Xishuangbanna cucumber (*Cucumis sativus* L. Var. *Xishuangbannanensis* Qi et Yuan). *Theor. Appl. Genet.* **2015**, *128*, 25–39. [[CrossRef](#)]
42. Tsutsui, K.; Suwa, A.; Sawada, K.; Kato, T.; Ohsawa, T.A.; Watano, Y. Incongruence among mitochondrial, chloroplast and nuclear gene trees in *Pinus* Subgenus *Strobus* (Pinaceae). *J. Plant Res.* **2009**, *122*, 509–521. [[CrossRef](#)]
43. Wang, N.; Li, C.; Kuang, L.; Wu, X.; Xie, K.; Zhu, A.; Xu, Q.; Larkin, R.M.; Zhou, Y.; Deng, X.; et al. Pan-mitogenomics reveals the genetic basis of cytonuclear conflicts in citrus hybridization, domestication, and diversification. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2206076119. [[CrossRef](#)] [[PubMed](#)]
44. Huang, J.; Liu, Y.; Han, F.; Fang, Z.; Yang, L.; Zhuang, M.; Zhang, Y.; Lv, H.; Wang, Y.; Ji, J.; et al. Genetic diversity and population structure analysis of 161 broccoli cultivars based on SNP markers. *Hortic. Plant J.* **2021**, *7*, 423–433. [[CrossRef](#)]
45. Xue, Y.; Liu, R.; Xue, J.; Wang, S.; Zhang, X. Genetic diversity and relatedness analysis of nine wild species of tree peony based on simple sequence repeats markers. *Hortic. Plant J.* **2021**, *7*, 579–588. [[CrossRef](#)]
46. Laosatit, K.; Amkul, K.; Chankaew, S.; Somta, P. Molecular genetic diversity of winged bean gene pool in Thailand assessed by SSR markers. *Hortic. Plant J.* **2022**, *8*, 81–88. [[CrossRef](#)]
47. Yang, J.; Zhang, J.; Du, H.; Zhao, H.; Mao, A.; Zhang, X.; Jiang, L.; Zhang, H.; Wen, C.; Xu, Y. Genetic relationship and pedigree of Chinese watermelon varieties based on diversity of perfect snps. *Hortic. Plant J.* **2022**, *8*, 489–498. [[CrossRef](#)]

48. Zhang, J.; Yang, J.; Fu, S.; Ren, J.; Zhang, X.; Xia, C.; Zhao, H.; Yang, K.; Wen, C. Comparison of DUS testing and SNP fingerprinting for variety identification in cucumber. *Hortic. Plant J.* **2022**, *8*, 575–582. [[CrossRef](#)]
49. Hu, S.; Ding, Y.; Zhu, C. Sensitivity and responses of chloroplasts to heat stress in plants. *Front. Plant Sci.* **2020**, *11*, 375. [[CrossRef](#)]
50. Robles, P.; Quesada, V. Unveiling the functions of plastid ribosomal proteins in plant development and abiotic stress tolerance. *Plant Physiol. Biochem.* **2022**, *189*, 35–45. [[CrossRef](#)]
51. Zhang, J.; Yuan, H.; Yang, Y.; Fish, T.; Lyi, S.M.; Thannhauser, T.W.; Zhang, L.; Li, L. Plastid ribosomal protein S5 is involved in photosynthesis, plant development, and cold stress tolerance in *Arabidopsis*. *J. Exp. Bot.* **2016**, *67*, 2731–2744. [[CrossRef](#)]
52. Yu, H.D.; Yang, X.F.; Chen, S.T.; Wang, Y.T.; Li, J.K.; Shen, Q.; Liu, X.L.; Guo, F.Q. Downregulation of chloroplast *RPS1* negatively modulates nuclear heat-responsive expression of *HsfA2* and its target genes in *Arabidopsis*. *PLoS Genet.* **2012**, *8*, e1002669. [[CrossRef](#)] [[PubMed](#)]
53. Zhang, A.; Jiang, X.; Zhang, F.; Wang, T.; Zhang, X. Dynamic Response of RNA editing to temperature in Grape by RNA deep sequencing. *Funct. Integr. Genom.* **2020**, *20*, 421–432. [[CrossRef](#)] [[PubMed](#)]
54. Hölzl, G.; Dörmann, P. Chloroplast lipids and their biosynthesis. *Annu. Rev. Plant Biol.* **2019**, *70*, 51–81. [[CrossRef](#)] [[PubMed](#)]
55. Huang, C.; Liu, D.; Li, Z.A.; Molloy, D.P.; Luo, Z.F.; Su, Y.; Li, H.O.; Liu, Q.; Wang, R.Z.; Xiao, L.T. The PPR protein RARE1-mediated editing of chloroplast *accD* transcripts is required for fatty acid biosynthesis and heat tolerance in *Arabidopsis*. *Plant Commun.* **2023**, *4*, 100461. [[CrossRef](#)] [[PubMed](#)]
56. Chu, D.; Wei, L. Reduced C-to-U RNA editing rates might play a regulatory role in stress response of *Arabidopsis*. *J. Plant Physiol.* **2020**, *244*, 153081. [[CrossRef](#)]
57. Wright, R.M. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **1981**, *8*, 4321–4325. [[CrossRef](#)]
58. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
59. Jin, J.J.; Yu, W.B.; Yang, J.B.; Song, Y.; DePamphilis, C.W.; Yi, T.S.; Li, D.Z. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* **2020**, *21*, 241. [[CrossRef](#)]
60. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [[CrossRef](#)]
61. Tillich, M.; Lehwick, 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](#)]
62. Amiryousefi, A.; Hyvönen, J.; Poczai, P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018**, *34*, 3030–3031. [[CrossRef](#)] [[PubMed](#)]
63. 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](#)]
64. Katoh, K.; Misawa, K.; Kuma, K.; Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [[CrossRef](#)] [[PubMed](#)]
65. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; Von Haeseler, A.; Lanfear, R. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **2020**, *37*, 1530–1534. [[CrossRef](#)] [[PubMed](#)]
66. Rozas, J.; Ferrer Mata, A.; Sánchez DelBarrio, J.C.; Guirao Rico, S.; Librado, P.; Ramos Onsins, S.E.; Sánchez Gracia, A. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302. [[CrossRef](#)]
67. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997. [[CrossRef](#)]
68. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. Subgroup, =1000 Genome Project Data Processing The sequence alignment/map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
69. Danecek, P.; Bonfield, J.K.; Liddle, J.; Marshall, J.; Ohan, V.; Pollard, M.O.; Whitwham, A.; Keane, T.; McCarthy, S.A.; Davies, R.M.; et al. Twelve years of SAMtools and BCFtools. *Gigascience* **2021**, *10*, giab008. [[CrossRef](#)]
70. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [[CrossRef](#)]
71. Purcell, S.; Neale, B.; Todd Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)]
72. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **2009**, *26*, 1641–1650. [[CrossRef](#)]
73. Alexander, D.H.; Novembre, J.; Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **2009**, *19*, 1655–1664. [[CrossRef](#)] [[PubMed](#)]
74. Kim, D.; Paggi, J.M.; Park, C.; Bennett, C.; Salzberg, S.L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-Genotype. *Nat. Biotechnol.* **2019**, *37*, 907–915. [[CrossRef](#)] [[PubMed](#)]
75. Liao, Y.; Smyth, G.K.; Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **2014**, *30*, 923–930. [[CrossRef](#)] [[PubMed](#)]
76. Varet, H.; Brillat Guéguen, L.; Coppée, J.Y.; Dillies, M.A. SARTools: A DESeq2- and EdgeR-Based R Pipeline for comprehensive differential analysis of RNA-Seq data. *PLoS ONE* **2016**, *11*, e0157022. [[CrossRef](#)] [[PubMed](#)]

77. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The genome analysis toolkit: A mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [[CrossRef](#)] [[PubMed](#)]
78. Wu, S.; Liu, W.; Aljohi, H.A.; Alromaih, S.A.; Alanazi, I.O.; Lin, Q.; Yu, J.; Hu, S. REDO: RNA editing detection in plant organelles based on variant calling results. *J. Comput. Biol.* **2018**, *25*, 509–516. [[CrossRef](#)]

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.