



Article

# Comparative Analysis of Chloroplast Pan-Genomes and Transcriptomics Reveals Cold Adaptation in *Medicago sativa*

Tianxiang Zhang <sup>1,†</sup>, Xiuhua Chen <sup>2,†</sup>, Wei Yan <sup>3</sup>, Manman Li <sup>1</sup> , Wangqi Huang <sup>4</sup>, Qian Liu <sup>3</sup>, Yanan Li <sup>3,\*</sup>, Changhong Guo <sup>1</sup> and Yongjun Shu <sup>1,\*</sup>

- <sup>1</sup> Key Laboratory of Molecular Cytogenetics and Genetic Breeding of Heilongjiang Province, College of Life Science and Technology, Harbin Normal University, Harbin 150025, China; hsdztx@stu.hrbnu.edu.cn (T.Z.); limanman1208@stu.hrbnu.edu.cn (M.L.); kaku3008@hrbnu.edu.cn (C.G.)
- <sup>2</sup> International Agriculture Research Institute, Yunnan Academy of Agricultural Sciences, Kunming 650200, China; cxh@yaas.org.cn
- <sup>3</sup> Institute of Tropical and Subtropical Cash Crops, Yunnan Academy of Agricultural Sciences, Baoshan 678000, China; rjsyw@yaas.org.cn (W.Y.); rjsslq@yaas.org.cn (Q.L.)
- <sup>4</sup> National Engineering Research Center for Ornamental Horticulture, Yunnan Flower Breeding Key Laboratory, Flower Research Institute, Yunnan Academy of Agricultural Sciences, Kunming 650200, China; hwq@yaas.org.cn
- \* Correspondence: rjslyn@yaas.org.cn (Y.L.); shuyji@hrbnu.edu.cn or syjun2003@126.com (Y.S.); Tel.: +86-451-88060576 (Y.S.)
- † These authors contributed equally to this work.

**Abstract:** Alfalfa (*Medicago sativa*) is a perennial forage legume that is widely distributed all over the world; therefore, it has an extremely complex genetic background. Though population structure and phylogenetic studies have been conducted on a large group of alfalfa nuclear genomes, information about the chloroplast genomes is still lacking. Chloroplast genomes are generally considered to be conservative and play an important role in population diversity analysis and species adaptation in plants. Here, 231 complete alfalfa chloroplast genomes were successfully assembled from 359 alfalfa resequencing data, on the basis of which the alfalfa chloroplast pan-genome was constructed. We investigated the genetic variations of the alfalfa chloroplast genome through comparative genomic, genetic diversity, phylogenetic, population genetic structure, and haplotype analysis. Meanwhile, the expression of alfalfa chloroplast genes under cold stress was explored through transcriptome analysis. As a result, chloroplast genomes of 231 alfalfa lack an IR region, and the size of the chloroplast genome ranges from 125,192 bp to 126,105 bp. Using population structure, haplotypes, and construction of a phylogenetic tree, it was found that alfalfa populations could be divided into four groups, and multiple highly variable regions were found in the alfalfa chloroplast genome. Transcriptome analysis showed that tRNA genes were significantly up-regulated in the cold-sensitive varieties, while *rps7*, *rpl32*, and *ndhB* were down-regulated, and the editing efficiency of *yef1*, *yef2*, and *ndhF* was decreased in the cold-tolerant varieties, which may be due to the fact that chloroplasts store nutrients through photosynthesis to resist cold. The huge number of genetic variants in this study provide powerful resources for molecular markers.

**Keywords:** alfalfa; chloroplast genome; population structure; genetic variation; cold stress



**Citation:** Zhang, T.; Chen, X.; Yan, W.; Li, M.; Huang, W.; Liu, Q.; Li, Y.; Guo, C.; Shu, Y. Comparative Analysis of Chloroplast Pan-Genomes and Transcriptomics Reveals Cold Adaptation in *Medicago sativa*. *Int. J. Mol. Sci.* **2024**, *25*, 1776. <https://doi.org/10.3390/ijms25031776>

Academic Editor: Manosh Biswas

Received: 20 December 2023

Revised: 23 January 2024

Accepted: 30 January 2024

Published: 1 February 2024



**Copyright:** © 2024 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

Chloroplasts are important organelles with autonomous genetic information within the cell [1]. Chloroplasts are involved in photosynthesis and energy conversion in plants and are commonly found in plants and algae [2]. The chloroplast genome is highly conserved in gene composition and structural arrangements [3,4]. Chloroplast genomes are structurally simple, have a relatively slow rate of evolution and low nucleotide substitution rates compared to nuclear genes, are mostly maternal in inheritance, and lack recombination [5,6]. Therefore, chloroplast genomes have been widely used for species identification and plant phylogenetic analysis [7,8]. Since the first tobacco chloroplast genome was determined

in 1986, there have been thousands of species with chloroplast genome sequence data identified in the NCBI database [9]. Chloroplasts of higher plants have a typical quadripartite circular structure, and the size of the chloroplast genomes ranges between 120 kb and 180 kb and usually include 100–130 genes [10]. They consist of a pair of inverted repeats (IRs) with the same sequence in opposite directions, a large single copy region (LSC), and a small single copy region (SSC) [11,12]. During the long-term evolution, there also exists a class of chloroplast genomes missing the IR region in nature [13,14], such as the chloroplast genomes of some legumes, including *Trifolium subterraneum*, *Medicago truncatula*, and *Cicer arietinum*. This type of chloroplast genome lacks an IR region, called the inverted repeat lacking clade (IRLC) [15,16]. However, these chloroplast genomes are only about 120 kb, which also proves that the IR region is not an essential region of the chloroplast genome [17]. Chloroplast genome structure variation can be used to explore evolutionary relationships between species [18]. The sequences of coding and non-coding regions of the chloroplast genome can not only be used for phylogenetic analysis but also are ideal material for species identification [19].

Alfalfa (*Medicago sativa*) is known as the “Queen of Forages”, with a cultivating area of approximately 32 million hectares globally [20,21]. As one of the most important forage crops in the world, alfalfa has excellent characteristics, such as high protein levels, high digestibility, rapid growth, and strong environmental adaptability [22]. However, alfalfa has the characteristics of wide distribution and variable life-types, which makes it difficult for the academic community to directly correlate the evolutionary history of alfalfa with its geographic and habitat characteristics [22,23]. Chloroplast genes play an important role in plant adaptation to environmental changes and can quickly respond to environmental changes. Chloroplast genes actively involved in the cold stress response of plants by encoding proteins participate in photosynthesis, transcriptional regulation, and signal transduction [24]. Research has shown that low temperatures can inhibit the biogenesis and development of chloroplasts during the early greening stage of rice (*Oryza sativa* L.) seedlings [25]. *Atpl* and *atpE* play an important role in the adaptation of species to environmental changes within the *Celastraceae* family, affecting electron transfer and energy synthesis processes [26]. Tobacco lacking the chloroplast gene *petL* displays impaired photosynthesis and bleached leaves in cold conditions [27]. These studies show that the chloroplast genome can provide a great deal of genetic information that can be used as a basis for studying plant cold tolerance.

In recent years, chloroplast genomes have made important contributions in exploring the evolution of species and the relationships between species [14]. The chloroplast genomes of different subspecies of some plants have been sequenced. For instance, research on three subspecies of rice, including *Oryza nivara* [28], *Oryza sativa indica* [29], and *Oryza sativa japonica* [30], revealed their phylogeny and different domestication processes based on chloroplast genome analyses [31]. In modern phylogenetic studies, more and more species have been analyzed for chloroplast genomes with the increasing improvement of sequencing technologies, and studies on *Cucumis sativus* L. [32], *Fragaria ananassa* [33], and *Carya cathayensis* [34] have identified their origin, evolution, and genetic diversity. The alfalfa genome has now been extensively sequenced, and studies on 220 alfalfa germplasms [35] and 137 alfalfa global core germplasms [36] have been based on the genome, but few data from the chloroplast genome are available to explore inter-species and intra-species relationships. However, legumes have experienced a large number of chloroplast genome rearrangements, and the study of alfalfa chloroplast genomes is necessary as well as conducive to the study of alfalfa varieties for further differentiation.

In this study, 231 alfalfa chloroplast genomes were successfully assembled. The genetic diversity of the chloroplast genome of alfalfa was explored by constructing the chloroplast pan-genome and comparative genome, haplotype analysis, phylogenetic analysis, and genetic structure analysis. Meanwhile, transcriptome analysis was used to reveal the response of alfalfa chloroplast genes under cold stress. The results provide the foundation

for the study of alfalfa chloroplast genomes, phylogenetic trees, genetic diversity, and adaptation of alfalfa to cold stress.

## 2. Results

### 2.1. General Characteristics of 231 Alfalfa Chloroplast Genomes

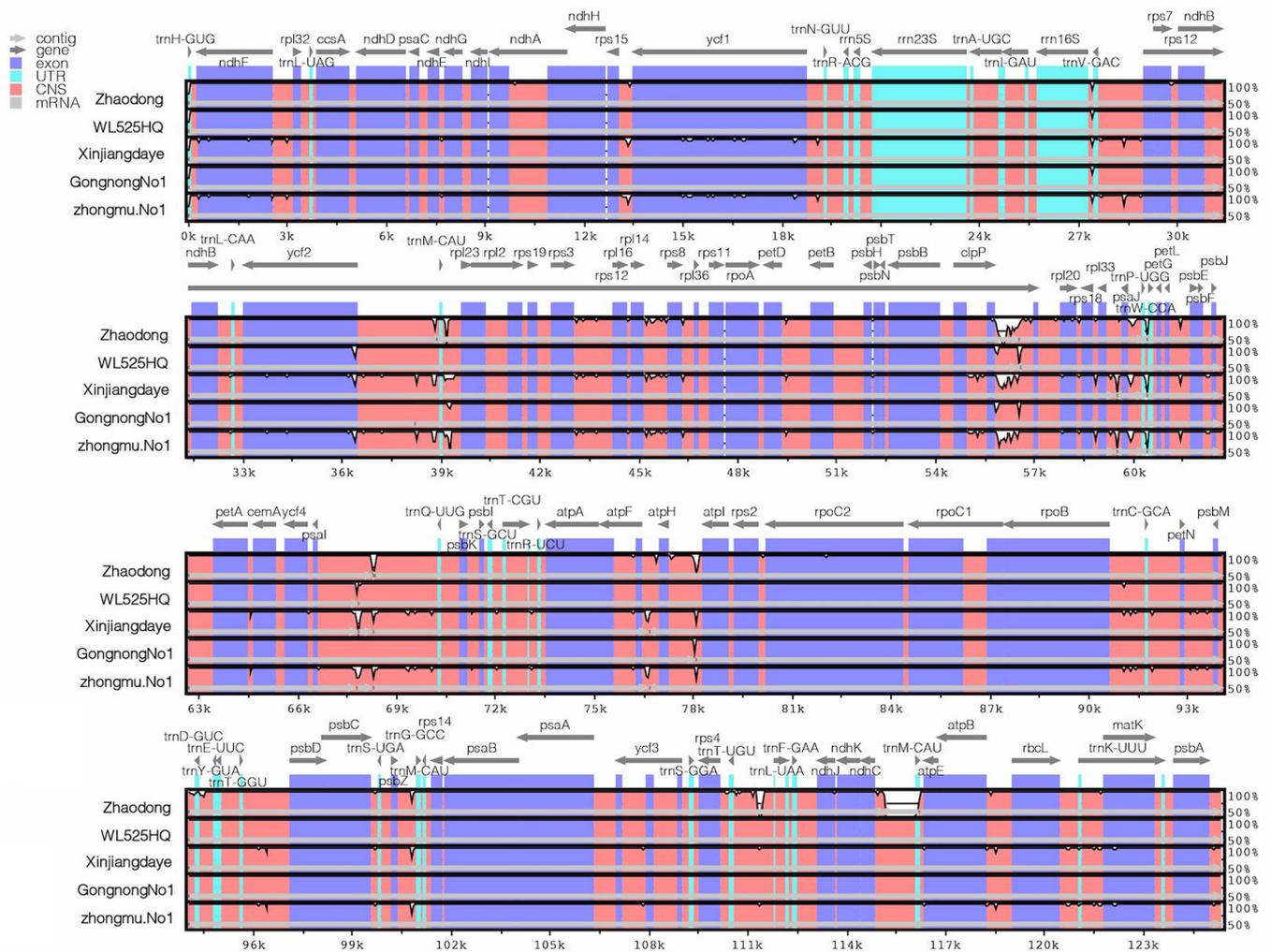
The chloroplast genome of *Medicago sativa* Deqin (MN218692) was selected as a reference genome, and the mapping reads were screened for assembly [37]. We successfully assembled and annotated a total of 231 chloroplast genomes by removing redundancy from 359 resequencing data (Figure S1). The sequence lengths of the 231 chloroplast genomes ranged from 125,192 bp to 126,105 bp. The average GC content of these chloroplast genomes was 33.78%, with a maximum of 33.9% and a minimum of 33.7% (Table S1). All of these chloroplast genomes are a typical circular structure and contain only one copy of the inverted repeat sequence (IR). Annotation results show that these chloroplast genomes have in common 109 genes, including 75 common protein-coding genes, 4 common ribosomal RNA (rRNA) genes, and 30 common transfer RNA (tRNA) genes (Figure 1). Consistent with the structure of angiosperms, these coding genes are divided into three categories. One is associated with photosynthesis, with 43 genes, of which *ndhA*, *ndhB*, *petB*, *petD*, and *atpF* contained an intron. Another group of 57 genes was related to self-replication, including 4 rRNA genes and 30 tRNA genes. The nine genes *rpl2*, *rps12*, *rpoC1*, *trnA-UGC*, *trnG-GCC*, *trnI-GAU*, *trnK-UUU*, *trnL-UAA*, and *trnV-UAC* all have an intron, where *rps12* is a trans-spliced gene. Finally, there four genes for conserved hypothetical chloroplast ORF of unknown function and five genes for other encoded proteins, and we found that *clpP* contained one intron, *ycf3* contained two introns, and *accD* had three introns. Among alfalfa chloroplast genomes, *rpl22*, *rps16*, and *infA* genes were absent (Table 1). The chloroplast genome contains a large amount of genetic information, and conservative gene composition is suitable for phylogenetic research in alfalfa. Photosynthesis-related genes and biosynthesis genes are of great significance for the plant resistance to environmental changes.

**Table 1.** Chloroplast gene annotation information of alfalfa.

Category	Gene Group	Gene Name
Photosynthesis	Subunits of photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>
	Subunits of photosystem II	<i>psbA</i> , <i>psbB</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbM</i> , <i>psbN</i> , <i>psbT</i> , <i>psbZ</i>
	Subunits of NADH dehydrogenase	<i>ndhA*</i> , <i>ndhB*</i> , <i>ndhC</i> , <i>ndhD</i> , <i>ndhE</i> , <i>ndhF</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i>
	Subunits of cytochrome b/f complex	<i>petA</i> , <i>petB</i> , <i>petD</i> , <i>petG</i> , <i>petL</i> , <i>petN</i>
	Subunits of ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF*</i> , <i>atpH</i> , <i>atpI</i>
	Large subunit of rubisco	<i>rbcL</i>
	Subunits of photochlorophyllide reductase	-
Self-replication	Proteins of large ribosomal subunit	<i>rpl14</i> , <i>rpl16</i> , <i>rpl2*</i> , <i>rpl20</i> , <i>rpl23</i> , <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>
	Proteins of small ribosomal subunit	<i>rps11</i> , <i>rps12*</i> , <i>rps14</i> , <i>rps15</i> , <i>rps18</i> , <i>rps19</i> , <i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> , <i>rps8</i>
	Subunits of RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1*</i> , <i>rpoC2</i>
	Ribosomal RNAs	<i>rri16S</i> , <i>rri23S</i> , <i>rri5S</i>
	Transfer RNAs	<i>trnA-UGC*</i> , <i>trnC-GCA</i> , <i>trnD-GUC</i> , <i>trnE-UUIC</i> , <i>trnF-GAA</i> , <i>trnG-GCC</i> , <i>trnH-GUG</i> , <i>trnI-GAU*</i> , <i>trnK-UUU*</i> , <i>trnL-CAA</i> , <i>trnL-UAA*</i> , <i>trnL-UAG</i> , <i>trnM-CAU(3)</i> , <i>trnN-GUU</i> , <i>trnP-UGG</i> , <i>trnQ-UUG</i> , <i>trnR-ACG</i> , <i>trnR-UCU</i> , <i>trnS-GCU</i> , <i>trnS-GGA</i> , <i>trnS-UGA</i> , <i>trnT-CGU*</i> , <i>trnT-GGU</i> , <i>trnT-UGU</i> , <i>trnV-GAC</i> , <i>trnW-CCA</i> , <i>trnY-GUA</i>



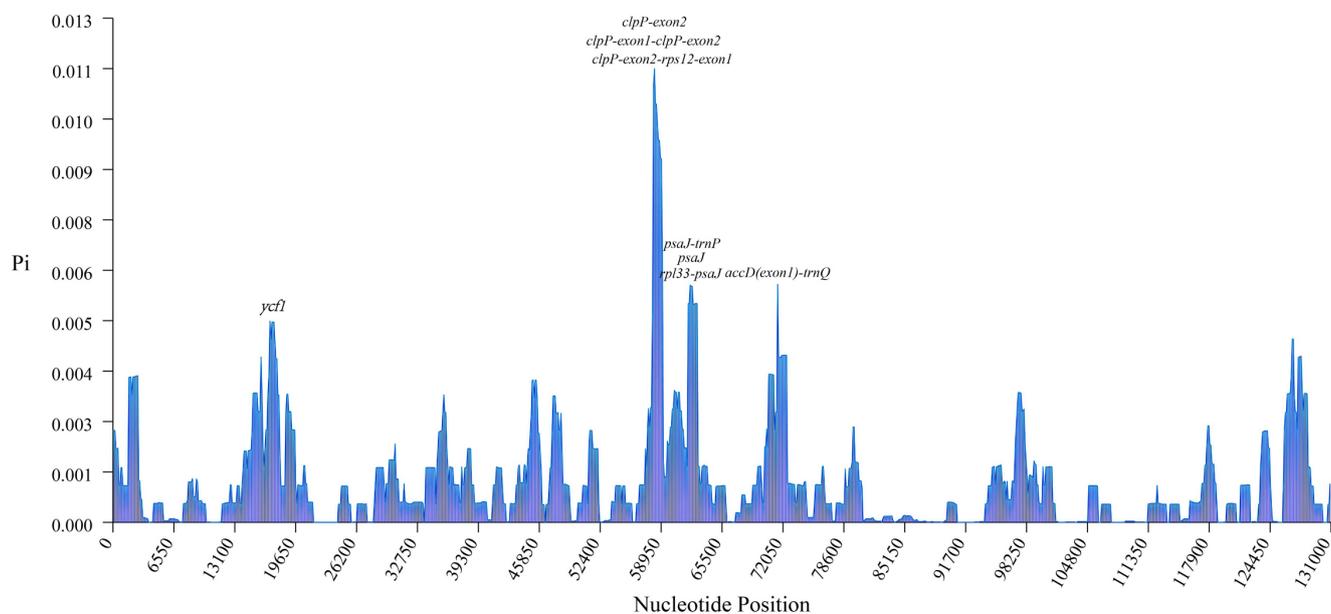
*sativa* Deqin was used as the reference sequence, and the whole sequence consisting of five genomes, including ZhaoDong, WL525HQ, ZhongmuNo1, GongnongNo1, and Xinjiangdaye, was compared (Figure 2). The analyses revealed small differences in these chloroplast genomes, which existed in both coding and non-coding regions, with significant sequence differences visible in non-coding regions in the 36–40 kb, 55–72 kb, and 76–78 kb ranges, and relatively high levels of variation in the *ndhF*, *clpP*, *ycf1*, and *ycf2* genes of the coding regions. The results showed that the level of variation in both coding and non-coding regions is higher in cold-tolerant varieties compared to the sensitive variety WL525HQ. The *ycf1*, *matK*, and *ndhF* regions in Xinjiangdaye and ZhongmuNo1 have particularly high levels of variation. Similarly, it was found that *rps12*, *trnW-CCA*, *clpP*, *rps8*, and *psaJ* regions show higher levels of variation in the cold-tolerant varieties. From the perspective of sequence differences, the sequence differences are lowest in the exonic regions, while the conserved non-coding sequences show the highest level of sequence differences. The highly similar sequences in the functional gene regions indicate a high conservation of chloroplast functional genes in alfalfa.



**Figure 2.** Comparison of chloroplast genome sequences of six varieties of Alfalfa. The ordinate indicates the proportion of agreement (50–100%), the horizontal indicates the chloroplast genome position, and annotated genes are listed at the top. Genomic regions in pink indicate Conservative Non-Coding Sequences (CNS); exons are in dark blue; rRNAs and tRNAs are in light blue; grey arrows indicate gene orientation.

We also calculated the nucleotide diversity (Pi) values for the chloroplast genomes of the alfalfa germplasm, and the Pi values for the nucleotide diversity of these 231 cpDNAs

ranged from 0 to 0.01126 (Table S2, Figure 3). There is high nucleotide diversity in the *clpP*, *psaJ*, and *ycf1* genes, and the maximum Pi value is for an exon of the *clpP* gene. We found higher nucleotide diversity in the intergenic region compared to the coding region. We found that the highly variable regions were clustered between 57 kb and 72 kb, which is consistent with the range of regions in our comparative genome analysis, and these spacers were *clpP* (*exon1*)-*clpP* (*exon2*), *psaJ*-*trnP*, *clpP* (*exon2*)-*rps12* (*exon1*), *rpl33*-*psaJ*, and *accD* (*exon1*)-*trnQ*, with *clpP* (*exon1*)-*clpP* (*exon2*) having the highest nucleotide diversity.

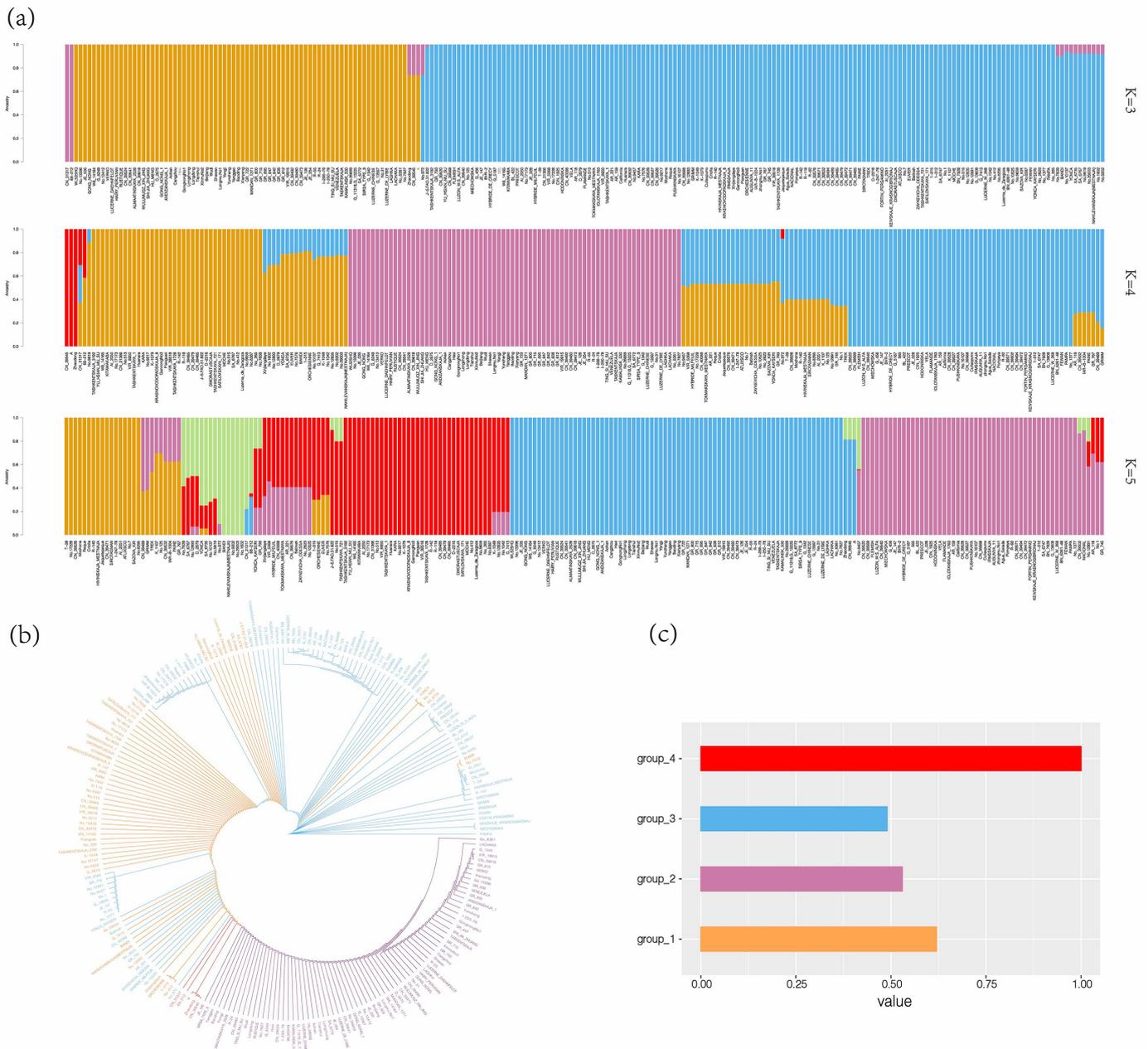


**Figure 3.** Statistics of nucleotide diversity (Pi) in 231 alfalfa chloroplast genomes. The X-axis represents the position of nucleotides; the Y-axis represents the nucleotide diversity in each window. Parameters are set to a window length of 1000 bp and a step size of 100 bp.

### 2.3. Phylogenetic Tree and Population Genetic Structure Analysis Based on SNPs

To investigate the genetic structure among alfalfa varieties, we extracted SNPs from the alfalfa chloroplast genome. SNP loci were used to construct the ML phylogenetic tree and conduct population structure analyses. Population structure is the existence of varying degrees of genetic relatedness between certain subgroups within species. The cross-validation (CV) error values were calculated for different K values (K = 2 to 9) when alfalfa population structure analysis was performed using admixture. When K = 4, group\_1 and group\_3 formed group\_2 together, and the varieties of group\_4 originated from the divergence of group\_1. The chloroplast genomes of 231 alfalfa can be divided into four groups: most of the samples are divided into the three largest groups, while the smallest group consisted of just five varieties. We divided the alfalfa population into four major branches by choosing the CV with the lowest value of K = 4 (Figure 4a, Table S4). However, there were differences in the proportion of cold-tolerant varieties among the four groups, and all varieties in group\_4 were cold-tolerant. (Figure 4c). Over an extended period of evolution and natural selection, some varieties may have developed stronger cold tolerance genes. The transition from K = 3 to K = 4 suggests that group\_2 may have evolved from a common ancestor of group\_1 and group\_3. The results of constructing a phylogenetic tree based on SNPs indicate that all varieties cluster on one branch (Figure 4b). Group\_1 and group\_3 have a close genetic relationship. Group\_1 has incorporated varieties from group\_3, and this result may be attributed to the widespread mutual introduction and cross-fertilization of diverse varieties of alfalfa across different regions. The clustering of group\_2 and group\_4 is consistent with the results of the population structure analysis.

Taken together, the findings suggested that it was better to divide the alfalfa population into four groups.

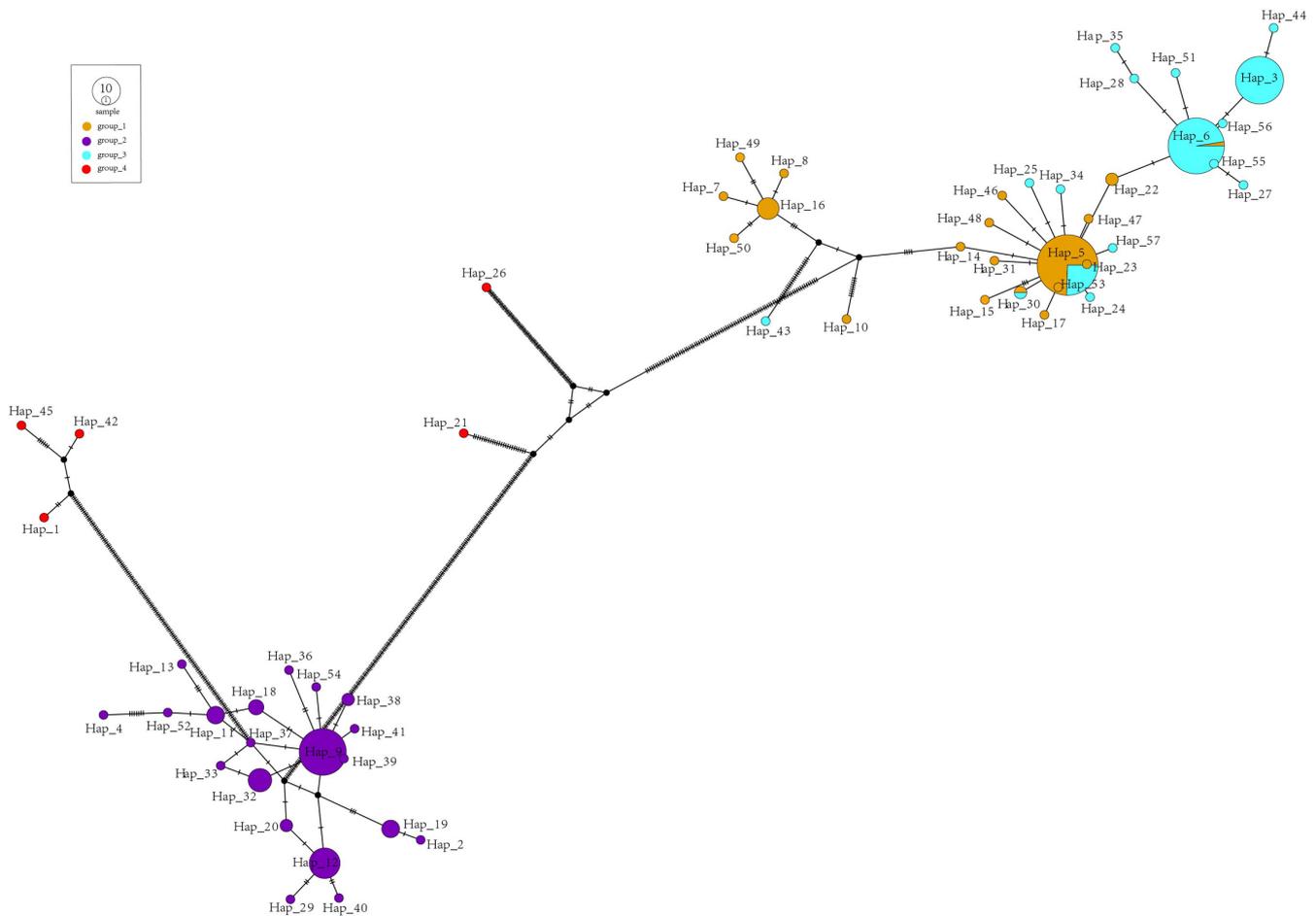


**Figure 4.** Phylogenetic tree and population genetic structure analysis based on SNP: (a) population structure analysis of chloroplast genomes of 231 alfalfa is based on SNP; (b) phylogenetic tree of chloroplast genomes of 231 alfalfa are constructed based on SNP; (c) the proportion of cold-tolerant varieties in different groups. Each color represents a different group.

#### 2.4. Haplotype and Phylogenetic Analysis

We used 231 alfalfa chloroplast genomes for haplotype analysis, which classified 57 haplotypes (Hap\_1-Hap\_57) (Table S3). Based on the population structure, we divided the alfalfa population into four groups (Figure 5). Group\_1 contained 19 haplotypes, and group\_2 and group\_3 contained 20 and 16 haplotypes. Group\_4 had five haplotypes. Haplotype network diagram results show that both group\_2 and group\_4 form a monophyletic group, while group\_1 and group\_3 cluster together. Hap\_9 is the main haplotype distributed in the group\_2 cluster, and it is only distributed in group\_2. Most of the

haplotypes appear singly. Hap\_5, Hap\_6, and Hap\_30 are shared haplotypes in group\_1 and group\_3, and the main distributed haplotypes in group\_1 and group\_3 are Hap\_5 and Hap\_6, respectively. This may be caused by the abundant mutual introduction and cross-fertilization of alfalfa between different varieties. Hap\_5 has 35 varieties in group\_1 and another 12 varieties in group\_3. In Hap\_6, we find that only one variety is assigned to group\_1, while group\_3 contains the remaining 39 varieties. Hap\_30 contains only one variety that is assigned to both group\_1 and group\_3. Based on the chloroplast whole genome, we constructed a phylogenetic tree, divided into four groups according to the distribution of haplotypes, and found that the distribution of varieties corresponded to the results of the population structure analysis.

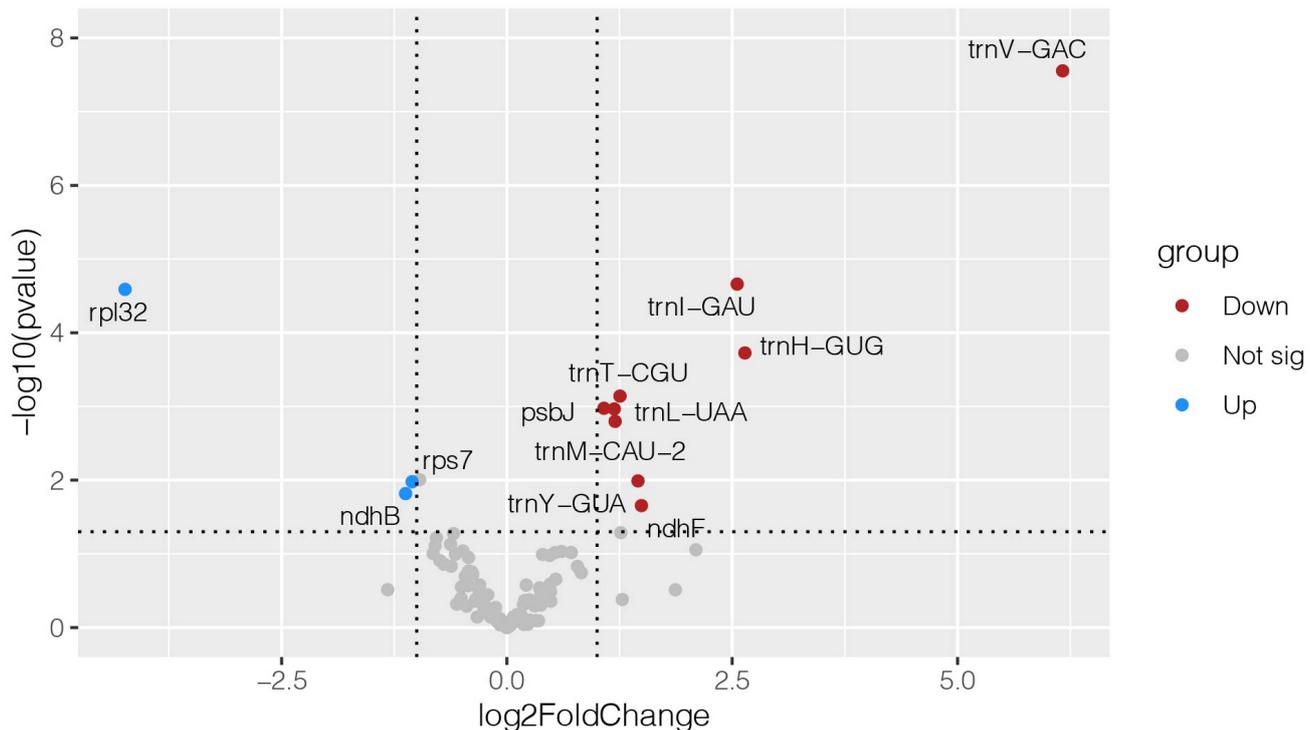


**Figure 5.** Haplotype network of alfalfa chloroplasts. The size of the circles represents the number of haplotypes. Dots represent putative haplotypes. Group\_1 is shown in orange, group\_2 in purple, group\_3 in blue, and group\_4 in red.

To explore the phylogenetic relationships among different alfalfa germplasms, we employed the chloroplast genomes of 231 alfalfa to construct an ML phylogenetic tree (Figure 6). The phylogenetic results based on the chloroplast genome sequences differed from the clustering based on the SNP construction mentioned earlier, forming distinct branches for alfalfa. In the comprehensive chloroplast genome phylogenetic tree, one branch included all varieties from group\_1 and group\_3, while group\_2 clustered on another branch. Both the SNP-constructed tree and the chloroplast genome-constructed tree indicated a close affinity between group\_1 and group\_3, making it challenging to clearly separate them. This could be attributed to genetic infiltration between the varieties of the two groups, suggesting potential genetic admixture or hybridization, leading to individuals with mixed genetic traits. Varieties from group\_4 did not cluster together,

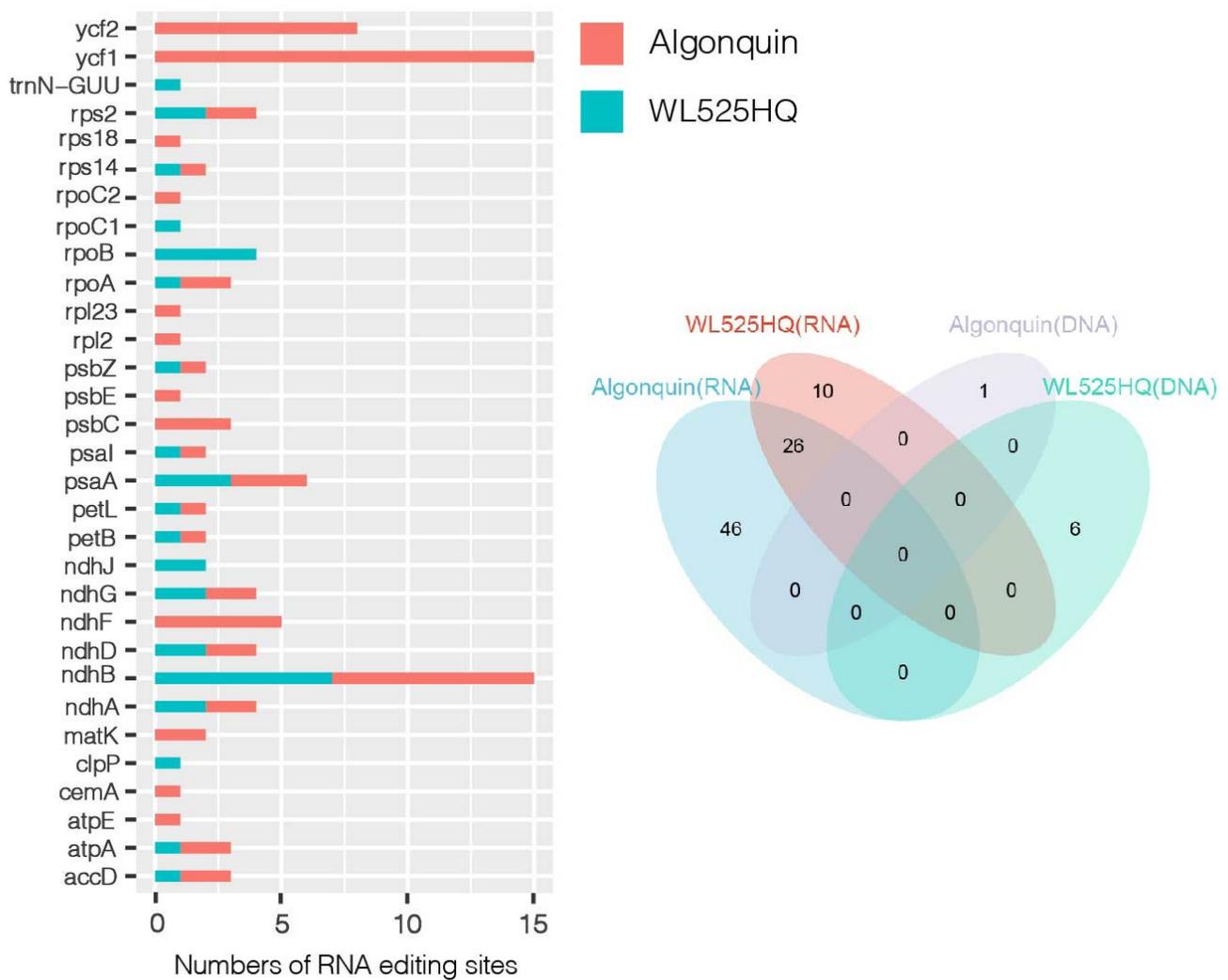


that low temperatures affected the normal splicing of the chloroplast gene *ndhB*, which results in abnormal chloroplast development. The expression of genes related to ribosomal subunit proteins was down-regulated in cold-sensitive alfalfa *rpl32* and *rps7*; conversely, these genes were up-regulated in cold-tolerant varieties. This may be a protective mechanism of plant chloroplasts in response to cold stress through ribosome metabolism. We also found that *trnV-GAC*, *trnI-GAU*, and *trnH-GUG* showed mostly significant down-regulation in cold-tolerant varieties. This may be attributed to the adaptive response of the cold-tolerant varieties to cold environments, where the reduction in the expression of *trn* genes slows down biochemical reactions and protein synthesis, thereby lowering intracellular energy consumption.



**Figure 7.** Volcano plot of chloroplast gene expression under cold stress in alfalfa. Two alfalfa varieties, Algonquin and WL525HQ, were studied under cold stress, with three biological replications for each variety. Plot the volcano for the chloroplast genome of alfalfa under cold stress using the average transcriptome data from three biological replicates. The two vertical lines represent genes with a fold change in expression of two or more on the outer side, and the line above indicates a  $p$ -value less than 0.05.

We used data from these six transcriptomes to analyze RNA editing events in chloroplasts (Figure 8, Table S5). We identified a total of 31 genes that underwent editing events. We found C to T editing was most common in chloroplast genomes. In the chloroplast genome, we identified a total of eight genes with more than two editing sites. *Ycf1* had the most RNA editing sites, with 15, followed by *ndhB*, *ycf2*, and *ndhF*. Interestingly, we found that the reduced editing efficiency of *ndhF* may promote *ndhF* expression under cold stress. In cold-sensitive varieties, reduced editing sites were observed in *ycf1*, *ycf2*, and *ndhF*. Algonquin and WL525HQ shared 17 editing sites, which were mainly concentrated in *ycf2* and *ndhF* genes. The low temperature led to a decrease in the editing efficiency of these sites, which may slow down the efficiency of electron transfer and thus affect the photosynthesis of chloroplasts. In cold-tolerant varieties, the increased editing efficiency at sites may be a key adaptation to cold environments.



**Figure 8.** RNA editing events under low-temperature stress in alfalfa varieties: (a) analysis of RNA editing efficiency in plastid genes under low-temperature stress in alfalfa; (b) number of RNA edits in different cultivars.

### 3. Discussion

In this study, 231 complete alfalfa chloroplast genomes were assembled and annotated, with an average length of 125,622 bp (Figure 3). These 231 chloroplast genomes have a highly conserved structure, and the GC content of the chloroplast genomes was also high, similar to that previously reported [37,38]. The size of chloroplast genome in higher plants is usually 120 kb to 180 kb. The expansion and contraction of the IR region is one of the important factors affecting the size of the chloroplast genome. The chloroplast genome of alfalfa and most of the legume structures are highly similar to the IR region, and this group is called IRLC [16,39,40]. One example of IR contraction is the IRLC taxon, and it is now believed that large IR contractions may be associated with abnormal genomic rearrangements [41]. It was found that the new IR region is adjacent to *ycf1* and *clpP* in *Medicago minima*, and these two genes may be major regions of genetic recombination due to their high variation [16]. Based on the comparative analysis of chloroplasts and nucleotide diversity, we identified eight highly variable regions, among which *clpP* (exon2) has the highest nucleotide diversity. A large number of studies have used some highly variable regions, such as *clpP* (exon1)-*psbB*, *rpl33*-*rps18*, *psaJ*-*rpl33*, and *ndhF*-*rpl32*, as DNA barcodes for variety identification [42,43]. In the present study, the intergenic region of *clpP* (exon1)-*clpP* (exon2), *psaJ*-*trnP*, and *clpP* (exon2)-*rps12* (exon1) genes is a highly variable

fragment, which is expected to be a specific fragment to distinguish alfalfa germplasm resources. In addition to the deletion of the IR region, chloroplast gene structure variation is also affected by the transfer and loss of genes, and it has been found in the study of chloroplast genomes that genes such as *accD*, *rpl23*, *rpl22*, *ycf1*, *ycf2*, *rps16*, *ycf4*, and *infA* have been partially or completely deleted from the chloroplast genomes of leguminous plants, and some genes (e.g., *infA*) have even been lost repeatedly in a variety of plants [44], which was well verified in the chloroplast genome of alfalfa, where deletions of the *rpl22*, *rps16*, and *infA* genes were found and could potentially be characteristic of alfalfa varieties.

In higher plants, chloroplasts are organelles with autonomous inheritance and a large amount of genetic information, and chloroplast genome sequences have made significant contributions in exploring the genetic and evolutionary relationships [45,46]. The germplasm resources of alfalfa have similar characteristics, and the evolutionary classification between them is not obvious. Alfalfa is vigorous and can adapt to most geographical environments, which also results in the plasticity of the alfalfa phenotype and genetic diversity, resulting in germplasm diversity. Although a large number of studies have been carried out, it is still difficult to determine the transmission and evolution mechanism of alfalfa [22,23]. We constructed haplotype analyses of 231 alfalfa germplasm resources, and the results were divided into four groups. A total of 88 varieties were closely clustered at the chloroplast genome level. Long et al. divided alfalfa into three clusters based on SNPs in the genome, but these three clusters exhibited that were weakly grouped together [35]. SNPs can be used as molecular markers to study evolutionary relationships; they have high genetic diversity and can be used for variety identification and genetic evolution studies [47,48]. We analyzed the population structure using SNPs from 231 germplasm resources and constructed an exclusive phylogenetic tree of SNPs. The results showed that 231 alfalfa germplasm resources were classified into four groups, and there were still varieties that were closely related. It is presumed that group\_1 and group\_3 formed group\_2 together and continued its performance. Because of the self-incompatibility of alfalfa, genes in different regions are exchanged, forming a highly heterozygous alfalfa. However, the chloroplast genomes of different germplasms are indistinguishable. We found that only the alfalfa varieties in group\_4 were cold-tolerant, and other groups were mixed with different phenotypes (Table S4). The population structure clustering information clarified the genetic data divided between different varieties, which was clearer than the results of studies using genomes of alfalfa. Alfalfa involves a large number of genes flowing among species in terms of domestication and adaptation to environmental conditions, resulting in many varieties not able to be classified according to ecotype. It is difficult for the present study to completely solve the problem, but the chloroplast-based classification results provide some useful information for the evolution of alfalfa.

The chloroplast participates in plant metabolism and some important physiological processes. Cold stress has great effects on chloroplast structure and physiological metabolism [49]. With further research, it was found that cold stress also seriously affects chloroplast protein synthesis and gene expression [50]. The transcriptome data of cold stress can effectively determine how cold temperature affects the expression of plant chloroplast genes, and we found that the expression of *rpl32* and *rps7* is up-regulated in cold-tolerant varieties. This suggests that alfalfa may need increased activity of ribosomes to promote protein synthesis and enhance adaptation to cold conditions. Overexpression of ribosomal subunit protein genes in *Arabidopsis thaliana* directly improved its cold tolerance [51]. Photosynthesis is a crucial biological process in plants, providing energy for all life processes. In cold treatment, the expression of *psbJ* is upregulated in sensitive varieties, participating in the capture and conversion of light energy in photosynthesis to promote growth. *PsbS* and *PsbH* have been shown to be upregulated in *Brassica campestris* L. under cold stress, enhancing the role of light-harvesting chlorophyll a/b-binding (*Lhc*) proteins in plant growth and development [52]. Meanwhile, most of the transfer RNA *trnV-GAC*, *trnL-GAU*, and *trnI-GUC* genes were significantly downregulated in cold-tolerant varieties and significantly up-regulated in cold-sensitive varieties. The expression of tRNA

genes in cold-tolerant alfalfa was suppressed, which slowed the growth of alfalfa and stored nutrition in winter by reducing the consumption of nutrients in autumn [53,54]. The upregulation of *psbJ* and tRNA genes in sensitive varieties indicates that cold-sensitive alfalfa continues to grow and develop in autumn, and the available sugars for maintaining alfalfa survival decrease in winter, thus reducing the winter survival rate. Compared with cold-sensitive varieties, the *ndhF*, *ycf1*, and *ycf2* genes of cold-tolerant varieties also showed higher RNA editing efficiency under cold stress. There are considerable variations in *ndhF*, *ycf1*, and *ycf2*, and the gene expression of *ndhF* in two varieties is also different. The protein encoded by the unknown function gene *ycf1* may involve protein transport as a component of the complex, and the transported protein may be directly related to cold and directly affect the differentiation of varieties. The change in the *ndhF* gene means that it directly affects the adaptability of varieties to cold conditions, which can be used as molecular markers to identify alfalfa varieties. Noncoding regions are also highly variable regions and have higher nucleotide diversity compared to coding regions [55]. Screening of highly variable sequences for the identification of varieties has been successfully used in *Dendrobium* and *Bupleurum* [55,56]. Highly variable regions are suitable for the construction of DNA barcodes *clpP(exon1-clpP(exon2))*, *psaJ-trnP*, *clpP-rps12*, *rpl33-psaJ*, and *ndhB-ndhF* and have great potential in the identification of alfalfa varieties and winter tolerance.

#### 4. Materials and Methods

##### 4.1. Alfalfa Materials and Plastid Isolation

The 359 alfalfa germplasm resources used in this study were derived from the National Genome Sciences Data Centre (<https://ngdc.cncb.ac.cn>) (accessed on 15 June 2023). The resequenced data of 137 and 220 core germplasm resources were downloaded from the GSA databases (CRA002917) and (CRA003659) of the NGDC website, respectively. Two alfalfa germplasms, Algonquin and WL525HQ, were grown at Harbin Normal University experiment station. One large seeding was selected from each germplasm, leaves were collected from the selected seedings, and they were rapidly frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ . The samples were sent to BGI Genomics (Shenzhen, China) (<http://www.genomics.cn>) (accessed on 14 November 2023), where the construction of genomic libraries and sequencing were completed in accordance with the described protocols. The genomic library was constructed with about 420 bp insertion size, and sequencing was performed using the Illumina HiSeq platform with a paired-end sequencing model. All the adapter sequences and low qualitative reads were removed from the raw sequence data; then, the cleaned reads of two alfalfa germplasms were submitted to the GSA database (CRA014561, URL: <https://ngdc.cncb.ac.cn/gsa/s/5c2aM2hX>) (accessed on 21 January 2024).

##### 4.2. Assembly and Annotation of Chloroplast Genome

The chloroplast genome was assembled using GetOrganelle software (version 1.7.5) using *Medicago sativa* Deqin (MN218692) as a reference, and a total of 231 complete chloroplast circular DNAs were obtained after removing redundancy [37]. Collinearity analysis was performed using mummer to verify the accuracy of the assembly [57]. Based on the GenBank format, the file of *Medicago sativa* Deqin was annotated using CPGAVAS2 software (version 2.0) [58], and CPGview-RSG software (<http://47.96.249.172:16100/cpgviewer/home>) (accessed on 16 November 2023) was used to detect the annotation errors. The sequence length and GC content of the samples were counted by seqkit [58,59]. Finally, the physical mapping of the alfalfa chloroplast genome was performed by the online software CHLOROPLLOT (<https://irscope.shinyapps.io/Chloroplplot/>) (accessed on 16 November 2023) [60].

##### 4.3. Comparative Analysis of Chloroplast Genome

Sequence differences between the reference chloroplast genome *Medicago sativa* Deqin (MN218692) and the chloroplast genomes of five closely related varieties were compared

using the Shuffle-LAGAN model in the online mVISTA tool (<https://genome.lbl.gov/vista/index.shtml>) (accessed on 18 November 2023) [61]. Nucleic acid diversity was analyzed using DnaSP 6 (DNA Sequence Polymorphism) software (version 6.12.03) with a sliding window length of 1000 bp, a step size of 100 bp, and other parameters set to defaults [62].

#### 4.4. Phylogenetic Tree and Haplotype Analysis

The chloroplast genome of 231 alfalfa sequences was aligned using MAFFT software (version 7.487) with default settings [63]. The phylogenetic tree was constructed using IQ-TREE (version 2.0) by selecting the maximum likelihood (ML) technique [64]; the bootstrap value was set as 1000, and the phylogenetic tree was visualized using figtree software (version 1.4.4) [65]. The results of alignment were imported into DnaSP 6 software (version 6.12.03) for the counting of the number of haplotypes (Hap), and POPART software (version 1.7) was used for the construction of the haplotype network diagram [62,66].

#### 4.5. Variant Calling

The same reference genome *Medicago sativa* Deqin (MN218692) GenBank format file was selected; the variants of the chloroplast genomes were called using Snippy (<https://github.com/tseemann/snippy>) (accessed on 21 November 2023), and VCF formats were generated. The SNPs were extracted by VCFtools (version 0.1.17), and Plink (version 1.9) filtered the variant information of SNPs, with the minor allele frequency set higher than 0.05 to output the VCF format file [67,68].

#### 4.6. Comparative Analysis Based on SNP

The VCF format file was converted to the phylip format file using Plink [68] and then compared using MAFFT software (version 7.487) [63]. IQ-TREE 2 set the bootstrap value to 1000 to construct the ML phylogenetic tree [64]. The treefile file was imported into figtree (version 1.4.4) to visualize the phylogenetic tree [65]. We then evaluated 231 genetic lineages of alfalfa using ADMIXTURE (version 1.3.0), setting parameters from  $K = 2$  to  $K = 10$ , where  $K = 3$ ,  $K = 4$ , and  $K = 5$  used R visualization (version 4.3.0) [69].

#### 4.7. Analysis of RNA-Seq and RNA Editing Efficiency

The RNA-seq experiments were performed as described previously [70]. In briefly, seeds of two alfalfa germplasms were selected and germinated on filter paper and then transferred to pots with a mixture of perlite and sand (3:1,  $v/v$ ). The seedlings were grown in a chamber and were irrigated with half-strength Hoagland solution every two days. The conditions of the chamber were set as 14 h light conditions at 24 °C (as day) and 10 h dark conditions at 18 °C (as night). Eight weeks later, all alfalfa seedlings were transferred to a new chamber at 4 °C. Three hours later, three samples were taken from the cold stress groups for each germplasm; each sample was leaf tissue collected from five individual seedlings. The collected samples were immediately frozen in liquid nitrogen and stored in a −80 °C freezer, and total RNA was extracted from the plants of the two alfalfa germplasms using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions. The extracted total RNA from six samples was then quantified using the NanoDrop 2000 spectrophotometer (Thermo Fisher, Waltham, MA, USA) and the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). As previously described, the RNA samples were sent to BGI Genomics (Shenzhen, China) (<http://www.genomics.cn>) (accessed on 18 June 2023) for RNA-seq, utilizing paired-end sequencing with a read length of 150 bp. All the adapter sequences and low qualitative reads were removed from the raw sequence data; then, the cleaned reads of these samples were submitted to the GSA database (CRA014537, URL: <https://ngdc.cnbc.ac.cn/gsa/s/G8X9s3Q1>) (accessed on 21 January 2024). Clean RNA-seq sequences were mapped to *Medicago sativa* Deqin using Samlon software (version 0.14.2). TPM values were calculated and then visualized using R heatmaps [71]. Quality-controlled reads were aligned to the *Medicago sativa* Deqin chloroplast genome using Burrows-Wheeler Aligner (BWA) software (version 0.7.17) [72]

and then spliced with GATK software (version 4.4.0.0) to remove duplicates to mine SNPs for variant information [73]. REDO (version 1.0) detects RNA editing sites in the chloroplast genome based on VCF files from alfalfa [74].

## 5. Conclusions

In this study, we assembled and compared 231 alfalfa chloroplast genomes and identified seven highly variable regions that can be used as molecular markers for alfalfa variety identification. Haplotype, population genetic structure, and phylogenetic tree results revealed that alfalfa can be divided into four groups, but there will still be 88 alfalfa varieties clustered. The results of transcriptome analysis showed that 11 genes were differently expressed in Algonquin and WL525HQ under cold stress, and the results of RNA editing showed that *ndhF*, *ycf1*, and *ycf2* would be edited significantly less efficiently in WL525HQ, which was cold-sensitive; thus, these genes are beneficial for the identification of cold-tolerant varieties. In conclusion, our results provide useful information on the genetic diversity of alfalfa and adaptation to cold stress in the chloroplast genome.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms25031776/s1>.

**Author Contributions:** Y.S. and Y.L. conceived the study; T.Z., X.C. and W.Y. conducted the experiments; Y.S., T.Z., M.L., Y.L., W.H., Q.L. and C.G. analyzed the data; Y.S., T.Z. and Y.L. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Natural Science Foundation of Heilongjiang Province (grant number LH2022C050), the Innovative Project for Postgraduate Students of Harbin Normal University (grant number HSDSSCX2023-42), the Open Fund of Yunnan Province Flower Breeding Key Laboratory (FKL-202203), the Agriculture Joint Special Project of Science and Technology Plan Project of Yunnan Science and Technology Department (202301BD070001-208), the Green Food Brand Build a Special Project (Floriculture) supported by Yunnan Provincial Finance Department (53000021000000013742), the Natural and Science Foundation of China (grant number U21A20182), and the Construction of Tengchong Rural Revitalization Technological Innovation County.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** None of the plant materials used in this study included any wild species at risk of extinction. No specific permits were required for sample collection in this study. We complied with relevant institutional, national, and international guidelines and legislation on plant study.

**Data Availability Statement:** The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

**Acknowledgments:** We are grateful to Jinyue Zhao for alfalfa DNA sequencing and also are grateful to the high-performance computing center of Harbin Normal University for supporting our analysis.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Neuhaus, H.E.; Emes, M.J. Nonphotosynthetic metabolism in plastids. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2000**, *51*, 111–140. [[CrossRef](#)] [[PubMed](#)]
2. Odintsova, M.S.; Iurina, N.P. Plastidic genome of higher plants and algae: Structure and function. *Mol. Biol.* **2003**, *37*, 768–783. [[CrossRef](#)]
3. Dong, W.; Xu, C.; Cheng, T.; Lin, K.; Zhou, S. Sequencing angiosperm plastid genomes made easy: A complete set of universal primers and a case study on the phylogeny of saxifragales. *Genome Biol. Evol.* **2013**, *5*, 989–997. [[CrossRef](#)] [[PubMed](#)]
4. Liang, H.; Zhang, Y.; Deng, J.; Gao, G.; Ding, C.; Zhang, L.; Yang, R. The complete chloroplast genome sequences of 14 curcuma species: Insights into genome evolution and phylogenetic relationships within Zingiberales. *Front. Genet.* **2020**, *11*, 802. [[CrossRef](#)] [[PubMed](#)]
5. Clegg, M.T.; Gaut, B.S.; Learn, G.H., Jr.; Morton, B.R. Rates and patterns of chloroplast DNA evolution. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6795–6801. [[CrossRef](#)] [[PubMed](#)]

6. Irisarri, I.; Strasser, J.F.H.; Burki, F. Phylogenomic insights into the origin of primary plastids. *Syst. Biol.* **2021**, *71*, 105–120. [[CrossRef](#)] [[PubMed](#)]
7. Ahmed, I.; Matthews, P.J.; Biggs, P.J.; Naeem, M.; McLenachan, P.A.; Lockhart, P.J. Identification of chloroplast genome loci suitable for high-resolution phylogeographic studies of *Colocasia esculenta* (L.) Schott (Araceae) and closely related taxa. *Mol. Ecol. Resour.* **2013**, *13*, 929–937. [[CrossRef](#)] [[PubMed](#)]
8. Fu, X.; Xie, D.F.; Zhou, Y.Y.; Cheng, R.Y.; Zhang, X.Y.; Zhou, S.D.; He, X.J. Phylogeny and adaptive evolution of subgenus *Rhizirideum* (Amaryllidaceae, Allium) based on plastid genomes. *BMC Plant Biol.* **2023**, *23*, 70. [[CrossRef](#)] [[PubMed](#)]
9. Shinozaki, K.; Ohme, M.; Tanaka, M.; Wakasugi, T.; Hayashida, N.; Matsubayashi, T.; Zaita, N.; Chunwongse, J.; Obokata, J.; Yamaguchi-Shinohara, K.; et al. The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. *EMBO J.* **1986**, *5*, 2043–2049. [[CrossRef](#)]
10. 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](#)]
11. Jansen, R.K.; Raubeson, L.A.; Boore, J.L.; de Pamphilis, C.W.; Chumley, T.W.; Haberle, R.C.; Wyman, S.K.; Alverson, A.J.; Peery, R.; Herman, S.J.; et al. Methods for obtaining and analyzing whole chloroplast genome sequences. *Methods Enzymol.* **2005**, *395*, 348–384
12. Zhao, F.; Chen, Y.P.; Salmaki, Y.; Drew, B.T.; Wilson, T.C.; Scheen, A.C.; Celep, F.; Bräuchler, C.; Bendiksby, M.; Wang, Q.; et al. An updated tribal classification of Lamiaceae based on plastome phylogenomics. *BMC Biol.* **2021**, *19*, 2. [[CrossRef](#)]
13. Wicke, S.; Müller, K.F.; de Pamphilis, C.W.; Quandt, D.; Wickert, N.J.; Zhang, Y.; Renner, S.S.; Schneeweiss, G.M. Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. *Plant Cell* **2013**, *25*, 3711–3725. [[CrossRef](#)] [[PubMed](#)]
14. Song, Y.; Yu, W.B.; Tan, Y.; Liu, B.; Yao, X.; Jin, J.; Padmanaba, M.; Yang, J.B.; Corlett, R.T. Evolutionary comparisons of the chloroplast genome in Lauraceae and insights into loss events in the Magnoliids. *Genome Biol. Evol.* **2017**, *9*, 2354–2364. [[CrossRef](#)]
15. Cai, Z.; Guisinger, M.; Kim, H.G.; Ruck, E.; Blazier, J.C.; McMurtry, V.; Kuehl, J.V.; Boore, J.; Jansen, R.K. Extensive reorganization of the plastid genome of *Trifolium subterraneum* (Fabaceae) is associated with numerous repeated sequences and novel DNA insertions. *J. Mol. Evol.* **2008**, *67*, 696–704. [[PubMed](#)]
16. Choi, I.S.; Jansen, R.; Ruhlman, T. Lost and found: Return of the inverted repeat in the legume clade defined by its absence. *Genome Biol. Evol.* **2019**, *11*, 1321–1333. [[CrossRef](#)] [[PubMed](#)]
17. Ruhlman, T.A.; Zhang, J.; Blazier, J.C.; Sabir, J.S.M.; Jansen, R.K. Recombination-dependent replication and gene conversion homogenize repeat sequences and diversify plastid genome structure. *Am. J. Bot.* **2017**, *104*, 559–572. [[CrossRef](#)] [[PubMed](#)]
18. Niu, Z.; Xue, Q.; Zhu, S.; Sun, J.; Liu, W.; Ding, X. The complete plastome sequences of four orchid species: Insights into the evolution of the Orchidaceae and the utility of plastomic mutational hotspots. *Front. Plant Sci.* **2017**, *8*, 715. [[CrossRef](#)] [[PubMed](#)]
19. Shaw, J.; Shafer, H.L.; Leonard, O.R.; Kovach, M.J.; Schorr, M.; Morris, A.B. Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: The tortoise and the hare IV. *Am. J. Bot.* **2014**, *101*, 1987–2004. [[CrossRef](#)] [[PubMed](#)]
20. Small, E.; Jomphe, M. A synopsis of the genus *Medicago* (Leguminosae). *Can. J. Bot.* **1989**, *67*, 3260–3294. [[CrossRef](#)]
21. Small, E. *Alfalfa and Relatives: Evolution and Classification of Medicago*; NRC Research Press: Ottawa, ON, Canada, 2011.
22. Li, X.; Wang, X.; Wei, Y.; Brummer, E.C. Prevalence of segregation distortion in diploid alfalfa and its implications for genetics and breeding applications. *Theor. Appl. Genet.* **2011**, *123*, 667–679. [[CrossRef](#)]
23. Steele, K.P.; Ickert-Bond, S.M.; Zarre, S.; Wojciechowski, M.F. Phylogeny and character evolution in *Medicago* (Leguminosae): Evidence from analyses of plastid trnK/matK and nuclear GA3ox1 sequences. *Am. J. Bot.* **2010**, *97*, 1142–1155. [[CrossRef](#)]
24. Wei, X.; Liu, S.; Sun, C.; Xie, G.; Wang, L. Convergence and divergence: Signal perception and transduction mechanisms of cold stress in *Arabidopsis* and rice. *Plants* **2021**, *10*, 1864. [[CrossRef](#)]
25. Zhao, Y.; Han, Q.; Ding, C.; Huang, Y.; Liao, J.; Chen, T.; Feng, S.; Zhou, L.; Zhang, Z.; Chen, Y.; et al. Effect of low temperature on chlorophyll biosynthesis and chloroplast biogenesis of rice seedlings during greening. *Int. J. Mol. Sci.* **2020**, *21*, 1390. [[CrossRef](#)] [[PubMed](#)]
26. Cai, H.; Gu, X.; Li, Y.; Ren, Y.; Yan, S.; Yang, M. Cold resistance of *Euonymus japonicus* beihaidao leaves and its chloroplast genome structure and comparison with Celastraceae species. *Plants* **2022**, *11*, 2449. [[CrossRef](#)]
27. Gao, Y.; Thiele, W.; Saleh, O.; Scossa, F.; Arabi, F.; Zhang, H.; Sampathkumar, A.; Kühn, K.; Fernie, A.; Bock, R.; et al. Chloroplast translational regulation uncovers nonessential photosynthesis genes as key players in plant cold acclimation. *Plant Cell* **2022**, *34*, 2056–2079. [[CrossRef](#)] [[PubMed](#)]
28. Shahid Masood, M.; Nishikawa, T.; Fukuoka, S.; Njenga, P.K.; Tsudzuki, T.; Kadowaki, K. The complete nucleotide sequence of wild rice (*Oryza nivara*) chloroplast genome: First genome wide comparative sequence analysis of wild and cultivated rice. *Gene* **2004**, *340*, 133–139. [[CrossRef](#)] [[PubMed](#)]
29. Tang, J.; Xia, H.; Cao, M.; Zhang, X.; Zeng, W.; Hu, S.; Tong, W.; Wang, J.; Wang, J.; Yu, J.; et al. A comparison of rice chloroplast genomes. *Plant Physiol.* **2004**, *135*, 412–420. [[CrossRef](#)]
30. Hiratsuka, J.; Shimada, H.; Whittier, R.; Ishibashi, T.; Sakamoto, M.; Mori, M.; Kondo, C.; Honji, Y.; Sun, C.R.; Meng, B.Y.; et al. The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **1989**, *217*, 185–194. [[CrossRef](#)]

31. Cheng, L.; Nam, J.; Chu, S.H.; Rungnana, P.; Min, M.H.; Cao, Y.; Yoo, J.M.; 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](#)]
32. 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. [[CrossRef](#)]
33. Song, Y.; Li, C.; Liu, L.; Hu, P.; Li, G.; Zhao, X.; Zhou, H. The population genomic analyses of chloroplast genomes shed new insights on the complicated ploidy and evolutionary history in *Fragaria*. *Front. Plant Sci.* **2022**, *13*, 1065218. [[CrossRef](#)] [[PubMed](#)]
34. 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](#)] [[PubMed](#)]
35. Long, R.; Zhang, F.; Zhang, Z.; Li, M.; Chen, L.; Wang, X.; Liu, W.; Zhang, T.; Yu, L.X.; He, F.; et al. Genome assembly of alfalfa cultivar Zhongmu-4 and identification of SNPs associated with agronomic traits. *Genom. Proteom. Bioinform.* **2022**, *20*, 14–28. [[CrossRef](#)]
36. Shen, C.; Du, H.; Chen, Z.; Lu, H.; Zhu, F.; Chen, H.; Meng, X.; Liu, Q.; Liu, P.; Zheng, L.; et al. The chromosome-level genome sequence of the autotetraploid alfalfa and resequencing of core germplasms provide genomic resources for alfalfa research. *Mol. Plant* **2020**, *13*, 1250–1261. [[CrossRef](#)] [[PubMed](#)]
37. Zhao, Y.; Zeng, W.; Li, W.; Bi, Y. Complete chloroplast genome sequence of the drought and heat-resistant Chinese alfalfa landrace, *Medicago sativa* ‘Deqin’. *Mitochondrial DNA Part B Resour.* **2021**, *6*, 1488–1489. [[CrossRef](#)] [[PubMed](#)]
38. Tao, X.; Ma, L.; Zhang, Z.; Liu, W.; Liu, Z. Characterization of the complete chloroplast genome of alfalfa (*Medicago sativa*) (Leguminosae). *Gene Rep.* **2016**, *6*, 67–73. [[CrossRef](#)]
39. Kolodner, R.; Tewari, K.K. Inverted repeats in chloroplast DNA from higher plants. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 41–45. [[CrossRef](#)] [[PubMed](#)]
40. Palmer, J.D.; Thompson, W.F. Rearrangements in the chloroplast genomes of mung bean and pea. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 5533–5537. [[CrossRef](#)]
41. Blazier, J.C.; Jansen, R.K.; Mower, J.P.; Govindu, M.; Zhang, J.; Weng, M.L.; Ruhlman, T.A. Variable presence of the inverted repeat and plastome stability in *Erodium*. *Ann. Bot.* **2016**, *117*, 1209–1220. [[CrossRef](#)]
42. Li, J.; Xie, D.F.; Guo, X.L.; Zheng, Z.Y.; He, X.J.; Zhou, S.D. Comparative analysis of the complete plastid genome of five *Bupleurum* species and new insights into dna barcoding and phylogenetic relationship. *Plants* **2020**, *9*, 543. [[CrossRef](#)]
43. Yang, J.-P.; Zhu, Z.-L.; Fan, Y.-J.; Zhu, F.; Chen, Y.-J.; Niu, Z.-T.; Ding, X.-Y. Comparative plastomic analysis of three *Bulbophyllum* medicinal plants and its significance in species identification. *Acta Pharm. Sin.* **2020**, 2736–2745. [[CrossRef](#)]
44. Millen, R.S.; Olmstead, R.G.; Adams, K.L.; Palmer, J.D.; Lao, N.T.; Heggie, L.; Kavanagh, T.A.; Hibberd, J.M.; Gray, J.C.; Morden, C.W.; et al. Many parallel losses of *infA* from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. *Plant Cell* **2001**, *13*, 645–658. [[CrossRef](#)] [[PubMed](#)]
45. Zhou, J.; Chen, X.; Cui, Y.; Sun, W.; Li, Y.; Wang, Y.; Song, J.; Yao, H. Molecular structure and phylogenetic analyses of complete chloroplast genomes of two *Aristolochia* medicinal species. *Int. J. Mol. Sci.* **2017**, *18*, 1839. [[CrossRef](#)] [[PubMed](#)]
46. D’Andrea, W.J.; Theroux, S.; Bradley, R.S.; Huang, X. Does phylogeny control U37K-temperature sensitivity? Implications for lacustrine alkenone paleothermometry. *Geochim. Cosmochim. Acta* **2016**, *175*, 168–180. [[CrossRef](#)]
47. Huang, J.; Liu, Y.; Han, F.; Fang, Z.; Yang, L.; Zhuang, M.; Zhang, Y.; Lv, H.; Wang, Y.; Ji, J. Genetic diversity and population structure analysis of 161 broccoli cultivars based on SNP markers. *Hortic. Plant J.* **2021**, *7*, 423–433. [[CrossRef](#)]
48. 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](#)]
49. Sun, F.; Yang, L.; Xie, X.; Liu, G.; Li, Y. Effect of chilling stress on physiological metabolism in chloroplasts of seedlings of sugarcane varieties with different chilling resistance. *Acta Agron. Sin.* **2012**, *38*, 732–739.
50. Goulas, E.; Schubert, M.; Kieselbach, T.; Kleczkowski, L.A.; Gardeström, P.; Schröder, W.; Hurry, V. The chloroplast lumen and stromal proteomes of *Arabidopsis thaliana* show differential sensitivity to short- and long-term exposure to low temperature. *Plant J.* **2006**, *47*, 720–734. [[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. Yuan, L.; Zheng, Y.; Nie, L.; Zhang, L.; Wu, Y.; Zhu, S.; Hou, J.; Shan, G.L.; Liu, T.K.; Chen, G.; et al. Transcriptional profiling reveals changes in gene regulation and signaling transduction pathways during temperature stress in wucai (*Brassica campestris* L.). *BMC Genom.* **2021**, *22*, 687. [[CrossRef](#)]
53. Xu, H.; Li, Z.; Tong, Z.; He, F.; Li, X. Metabolomic analyses reveal substances that contribute to the increased freezing tolerance of alfalfa (*Medicago sativa* L.) after continuous water deficit. *BMC Plant Biol.* **2020**, *20*, 15. [[CrossRef](#)]
54. Erice, G.; Irigoyen, J.J.; Sánchez-Díaz, M.; Avicé, J.C.; Ourry, A. Effect of drought, elevated CO<sub>2</sub> and temperature on accumulation of N and vegetative storage proteins (VSP) in taproot of nodulated alfalfa before and after cutting—ScienceDirect. *Plant Sci.* **2007**, *172*, 903–912. [[CrossRef](#)]
55. Niu, Z.; Pan, J.; Xue, Q.; Zhu, S.; Liu, W.; Ding, X. Plastome-wide comparison reveals new SNV resources for the authentication of *Dendrobium huoshanense* and its corresponding medicinal slice (Huoshan Fengdou). *Acta Pharm. Sin. B* **2018**, *8*, 466–477. [[CrossRef](#)] [[PubMed](#)]
56. Han, W.-J.; Liu, X.; Wan, H.-F.; Liao, B.-S.; Han, J.-P. Screening on DNA barcodes for discriminating *Bupleurum* species. *J. Hortic. Sci. For.* **2021**, *3*, 102.

57. Marçais, G.; Delcher, A.L.; Phillippy, A.M.; Coston, R.; Salzberg, S.L.; Zimin, A. MUMmer4: A fast and versatile genome alignment system. *PLoS Comput Biol.* **2018**, *14*, e1005944. [[CrossRef](#)] [[PubMed](#)]
58. Shi, L.; Chen, H.; Jiang, M.; Wang, L.; Wu, X.; Huang, L.; Liu, C. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* **2019**, *47*, W65–W73. [[CrossRef](#)]
59. Shen, W.; Le, S.; Li, Y.; Hu, F. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* **2016**, *11*, e0163962.
60. Zheng, S.; Poczai, P.; Hyvönen, J.; Tang, J.; Amiryousefi, A. Chloroplot: An online program for the versatile plotting of organelle genomes. *Front. Genet.* **2020**, *11*, 576124. [[CrossRef](#)]
61. 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](#)]
62. 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](#)]
63. 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](#)]
64. 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](#)]
65. Rambaut, A. FigTree, a Graphical Viewer of Phylogenetic Trees. Institute of Evolutionary Biology University of Edinburgh. 2009. Available online: <http://tree.bio.ed.ac.uk/software/figtree> (accessed on 25 November 2023).
66. Bandelt, H.J.; Forster, P.; Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **1999**, *16*, 37–48. [[CrossRef](#)]
67. 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](#)] [[PubMed](#)]
68. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.; 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](#)] [[PubMed](#)]
69. Alexander, D.H.; Novembre, J.; Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **2009**, *19*, 1655–1664. [[CrossRef](#)] [[PubMed](#)]
70. Zhang, X.; Yang, H.; Li, M.; Bai, Y.; Chen, C.; Guo, D.; Guo, C.; Shu, Y. A Pan-Transcriptome Analysis Indicates Efficient Downregulation of the FIB Genes Plays a Critical Role in the Response of Alfalfa to Cold Stress. *Plants* **2022**, *11*, 3148. [[CrossRef](#)] [[PubMed](#)]
71. Patro, R.; Duggal, G.; Love, M.I.; Irizarry, R.A.; Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **2017**, *14*, 417–419. [[CrossRef](#)] [[PubMed](#)]
72. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
73. 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](#)]
74. 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](#)] [[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.