Plant Biotechnology Reports (2021) 15:471–482
https://doi.org/10.1007/s11816-021-00692-3

Online ISSN 1863-5474
Print ISSN 1863-5466

ORIGINAL ARTICLE

Tissue-specificity of RNA editing in plant: analysis of transcripts from three tobacco (*Nicotiana tabacum*) varieties

Jing Fang[1,3] · Xiao-Han Jiang[1,3] · Teng-Fei Wang[1,3] · Xiu-Jun Zhang[1,2,4] · Ai-Di Zhang[1,2,4]

Received: 28 April 2021 / Revised: 28 June 2021 / Accepted: 3 July 2021 / Published online: 8 July 2021

Abstract

RNA editing is a significant post-transcriptional molecular process of modifying the primary transcripts by editosome. In plants, it remains unknown whether and to what extent RNA editing contributes to tissue-specific regulation from a global perspective. To obtain an overview of RNA editing events in model plant tobacco (*Nicotiana tabacum*), we implemented a bioinformatics analysis of DNA-Seq and RNA-Seq data from roots and leaves of three tobacco varieties (TN90, Basma and K326). The results showed that hundreds of RNA editing sites were detected to be located in the protein-coding region of plastid/mitochondria for all three varieties. Among these sites, some of them were detected in leaves but not or reduced in roots. Interestingly, most of the disappeared editing sites in roots were located in plastid transcripts encoding subunits of NADH dehydrogenase. The average editing efficiencies in roots were reduced significantly compared with leaves across three varieties in both organelles. In addition, we found that the reduction of RNA editing efficiency in mitochondria is mild compared with plastid. Expression analysis further showed that an extraordinarily high percentage of RNA editing factors were down-regulated in roots, particularly for PPR, MORF proteins, indicating that the distinct editing patterns between roots and leaves might result from the differential regulation of RNA editing factors expression. This study provides references and insights into understanding the function of plant RNA editing in tissue-specific regulation mediated by editosomes.

Keywords RNA editing · Tissue specificity · Editosome · Plant organelles

Abbreviations

C-to-U	Cytosine-to-uracil
PPR	Pentatricopeptide repeat
NDH	NADH dehydrogenase
MORF	Multiple organelle RNA editing factors
ORRM	Organelle RNA recognition motif-containing
CET	Cyclic electron transport
ATP	Adenosine triphosphate synthase
cox	Cytochrome c oxidase
rps	Ribosomal protein

✉ Xiu-Jun Zhang
zhangxj@wbgcas.cn

✉ Ai-Di Zhang
zhangaidi@wbgcas.cn

[1] CAS Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430000, China

[2] Center of Economic Botany, Core Botanical Gardens, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430000, China

[3] University of Chinese Academy of Sciences, Beijing 100049, China

[4] Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China

Introduction

RNA editing event is a significant post-transcriptional modification process that results in the difference between RNA genetic information and genome template including nucleotide insertion, deletion, and replacement (Takenaka et al. 2013; Shikanai 2015; Yan et al. 2018). RNA editing has been found in primitive eukaryotes, vertebrates, plants, fungi and viruses (Zahn 2017). In plants, RNA editing only occurs in organelles, i.e. plastid and mitochondria (Takenaka et al. 2013; Edera et al. 2018; Yan et al. 2018). A typical vascular plant usually conducts about 200–500 editing events in mitochondria, while much less abundant in plastid with about 30–50 ones (Zahn 2017; Edera et al. 2018). RNA editing event commonly converts nucleotide base from cytidine (C) to uridine (U) in most plants, and also reverses nucleotide base from uridine (U) to cytidine (C) in some plant clades,

such as ferns, hornworts and isoetales (Grewe et al. 2011; Knie et al. 2016). The conversion of RNA editing, which is mainly located in the coding region, is thought to act as a corrective mechanism for DNA mutations by restoration of conserved amino acids to guarantee proper protein function (Takenaka et al. 2013; Shikanai 2015). Abnormal RNA editing can result in a series of plant developmental defects, such as reduced embryo and endosperm development, impaired plastid and mitochondria biogenesis, and retarded seedling growth (Sosso et al. 2012; Takenaka et al. 2013).

RNA editing in plant is mainly mediated by the editosome, which is an editing complex containing some editing factors, such as PLS-type pentatricopeptide repeat (PPR) proteins, organelle RNA recognition motif (ORRM) containing proteins, and multiple organellar RNA editing factors (MORF) (Yan et al. 2018). PPR protein family is exclusively expanded in plants, with over 450 members in *Arabidopsis* and rice (Yan et al. 2018). PPR editing factors have been reported to function as site-specific editing factors that specially bind to the *cis* element of target RNA, and played various roles in organelle gene expression, including RNA transcription, splicing, editing and cleavage, *PPR* mutants usually display various developmental defects (Sosso et al. 2012; Sun et al. 2015; Tang et al. 2017; Yan et al. 2018). MORF is a small protein family, with 10 members in *Arabidopsis* (Zehrmann et al. 2015), MORF mutants exhibited reduced efficiency at multiple sites (Takenaka et al. 2012). Disruption of *morf1*, *morf3* and *morf8* genes affects 19%, 26% and 72% of mitochondria editing events respectively, while mutants of either *morf2* or *morf9* exhibited reduced editing at nearly all sites in the plastid (Takenaka et al. 2012; Yan et al. 2018).

Recent studies have shown that RNA editing was modulated in a tissue- and stage-specific manner in certain genes. For instance, the editing of *photosystem II protein VI* (*psbF*) and *photosystem II protein L* (*psbL*) transcripts was modulated in a tissue- and stage-specific manner in spinach plastids (Bock et al. 1993). Similarly, in moss (*Physcomitrella patens*) plastids, the RNA efficiency of *ribosomal protein S14* (*rps14*) was 80% in the young protonemata, decreased to approximately 20% in old protonemata, and fully developed leafy shoots (Miyata and Sugita 2004). Editing of *ATP synthase F0 subunit 9* (*atp9*), *NADH dehydrogenase subunit 3* (*nad3*), and *cytochrome c oxidase subunit II* (*cox2*) transcripts in mitochondria was also investigated, the result showed RNA editing level varied under different developmental or growth conditions in maize seedlings (Grosskopf and Mulligan 1996). Developmental co-variation of RNA editing extent of 34 plastid editing sites was demonstrated in tobacco, a hypothesis that cluster-specific editing factors exist and their less abundance in roots might limit the editing extent of certain sites in roots plastids was proposed (Chateigner-Boutin and Hanson 2003). A recent study also demonstrated that plastid gene expression and RNA editing are specifically and differentially regulated in various types of *Arabidopsis* non-green tissues (Tseng et al. 2013). In addition to developmental and tissue-specific regulation, the extent of RNA editing is also affected by environmental stresses and inhibitor treatments (Nakajima and Mulligan 2001; Rodrigues et al. 2017; Xiong et al. 2017; Zhang et al. 2020). However, earlier researches only focused on selected genes of RNA editing in plastid based on experimental methods with the disadvantage of low throughput. With the rapid improvement of genomic and transcriptome sequencing technology, a great quantity of RNA-seq data combined with bioinformatics approaches offer an opportunity to examine the regulatory mechanism of RNA editing in both two organelles.

As an important worldwide economical plant, tobacco plays a key role in plant molecular research and provides a valuable model system for investigating the RNA editing in plants. Previous studies of RNA editing regulation in tobacco only focused on plastid transcripts, and little is known about the regulation of RNA editing in mitochondria transcripts. In this study, to obtain a global view of the regulation of RNA editing in tobacco, we used plenty of RNA-seq and genome re-sequencing data from three tobacco varieties (TN90, Basma, and K326) to examine the RNA editing profiles in roots and leaves in plastid and mitochondria. The results showed that the editing levels in leaves were higher than that in roots across three varieties. A number of editing sites were detected in leaves but not in roots, particularly for plastid genome. The down-regulated RNA editing factors detected in roots might limit its reduced editing efficiencies and extent. The results suggested that RNA editing in tobacco might function in the regulation of tissue development, the resultant organelle proteomes in turn are likely to be pertinent to the specific tissue functions.

Materials and methods

Data collection

The whole genome and transcriptome sequencing data of three main tobacco varieties (TN90, K326, Basma) were downloaded from NCBI (https://www.ncbi.nlm.nih.gov) with accession numbers PRJNA208209, PRJNA208210, and PRJNA208211 (Sierro et al. 2014). For each tobacco variety, the transcriptome of root and leaf were sequenced with three replicates at the same time. Detailed data information is listed in Table S1. The genome sequences of tobacco mitochondria and plastid, as well as their corresponding genome annotation files in "tbl" format, were downloaded from the NCBI data repository with accession numbers NC_006581.1 and NC_001879.2. All the

above-used reference genomes were included in the assembly of Ntab-TN90 deposited in NCBI.

Alignment of sequencing data and variants calling

The transcriptome data were aligned against the tobacco mitochondria/plastid genome, and variants were called for RNA editing sites detection. To eliminate false-positive sites that derived from genomic variation, the genome re-sequencing data were also aligned against the tobacco mitochondria/plastid genome, SNPs were called for RNA editing sites filtration. To ensure the sequencing data with a higher depth, the data with three replicates from the same tissue were merged. In variants calling procedure, the sequencing data were first evaluated using 'fastaqc' tool to ensure sequence quality (Brown et al. 2017). The transcriptome data of each tissue was aligned against reference by 'HISAT2' software (Kim et al. 2015). The genome re-sequencing data was aligned against reference using 'bwa' tool (Jo and Koh 2015). The aligned reads were sorted, repeat-removed and indexed using 'samtool' tool (Li et al. 2009). The 'bcftools' tool was used to identify variants/SNPs and generate VCF files (Danecek and McCarthy 2017). Finally, RNA editing sites were filtered based on the results variants from genome sequencing data and transcriptome sequencing data.

Identification of RNA editing sites

Based on the SNP-calling results (in "VCF" format) and genome annotation files (in "tbl" format), RNA editing sites were identified by using REDO tool (Wu et al. 2018). REDO is a comprehensive application tool for identifying RNA editing events in plant organelles based on variant calling format files from RNA-seq data. To reduce the false positives, REDO tool implemented a series of comprehensive rule dependent and statistical filters including (1) quality control filter, (2) depth filter (DP > 4), (3) alt proportion filter (alt proportion < 0.1), (4) multiple alt filter, (5) distance filter, (6) spliced junction filter, (7) indel filter, likelihood ratio (LLR) test filter (LLR < 10), (8) Fisher's exact test filter (p value < 0.01). As a result, REDO used a complicated filtering model based on the prior information of experiment-validated RNA editing sites and the attributes of codon table to divide RNA editing sites into five tiers, such as codon position and amino acid changes (Wu et al. 2018). Finally, all RNA editing sites were identified with annotation information files. To further minimize false-positive sites, we manually examined all mismatches and excluded SNPs derived from genomic variation.

Comparison of RNA editing efficiency

For each sample, the filtered editing sites were used for further characteristic statistics including editing numbers, editing types, editing efficiencies, codon positions, amino acid changes, and involved genes. For each site, RNA editing efficiency was quantified by the proportion of edited transcripts in total covered transcripts. For the three tobacco varieties, we compared the distribution of RNA editing efficiency of roots with that of leaves. Hence, RNA editing sites with statistical significance (p value < 0.05) were identified. Aiming to decipher the tendency of RNA editing efficiency for different tissues, cluster analysis and heatmap plotting were also performed based on matrix of RNA editing efficiency, which was normalized by subtracting the row-wise mean from the values in each row of data and divided by standard deviation value of each row. 'Pheatmap' function in R was used to plot the heatmap, 'dist' function was used to calculate the distance matrix of different samples with the default Euclidean method, 'hclust' function was used to compute the hierarchical clustering.

Expression analyses of RNA editing factors

To explore the expression of RNA editing factors, three classes of RNA editing factors including PPR proteins, MORF proteins and ORRM proteins were chosen according to previous studies (Yan et al. 2018), their expression levels in roots versus leaves were compared. The representative protein sequences with UniProtKB IDs Q9SAD9, O49429, Q9FNR1 from three RNA editing factors classes of *Arabidopsis* were used as queries to search against the tobacco protein database using 'BLASTP' with e value 1e − 05. Transcriptome analysis was implemented by the protocol in a previous study (Pertea et al. 2016). Gene expression levels were measured by FPKM (fragments per kilobase of transcript per million mapped reads), and were also normalized by the method mentioned above. 'EdgeR' was used to determine the differentially expressed genes between roots and leaves (Dai et al. 2014). A heatmap with all samples was plotted using 'pheatmap' function in R.

Statistical analysis

The two-tailed Wilcoxon rank-sum test was used to perform the pairwise comparison of RNA editing efficiency between roots and leaves. The two-sample t test tool 'ttest2' in R was used for the pairwise comparison of each editing site.

Results

Alignment of transcriptome data

In this study, for the three tobacco varieties (TN90, K326, Basma), there were a total of 38 samples of RNA-seq data and 10 samples of DNA-seq data were used for the analysis, each tissue has no less than three replicates. The detailed data information is listed in Table S1. The tobacco mitochondria genome that encodes 183 genes is 430,597 bp in size with accession number NC_006581.1, and plastid genome that encodes 122 genes is 155,943 bp in size with accession number NC_001879.2. Both of them were derived from TN90 variety. For each tobacco variety, the data in all the replicates of each tissue were merged together to make the reads coverage high enough. The merged transcriptome data were aligned against the tobacco mitochondria/plastid genome reference. The mapped reads in 'bam' format were classified into three groups, i.e. 'root', 'leaf', and 'merged'. With these 'bam' files, the variants were identified and stored in the files with 'VCF' format. For TN90 variety, the merged tissues included root, leaf, capsule and flower. For K326 and Basma varieties, the merged tissues only included leaf and root. The transcriptome alignment results showed that all the mapping depth of these groups were more than 20×, which meet the minimum requirement of identification of editing sites. However, we observed that there was a huge discrepancy in transcripts abundance in plastid roots versus leaves, demonstrating their wide variation in gene expression. Detailed mapping reads and depth information are listed in Table 1.

RNA editing sites detection and characteristic statistics

The raw RNA editing sites were detected using the REDO tool, which is an automated approach based on variants and genome annotation. To ensure the statistical reliability, only the editing sites with variant type C-to-U were selected for the analysis. Genome SNP-calling result was further used to eliminate genomic variation, and improve the accuracy of RNA editing sites filtering. Finally, there were 486 and 60 RNA editing sites involving 64 and 21 genes were identified in protein-coding region of mitochondria and plastid in TN90 variety. The editing efficiencies of these editing sites were close to the results of two previous studies (Sugiyama et al. 2005; Grimes et al. 2014). The number of RNA editing identified in K326 and Basma varieties with lower read depths was less than that of TN90 variety. The detailed information of identified RNA editing sites is listed in Table 1. The density of read depth in different groups is shown in Figs. 1, 2A and B. The results revealed that there is a positive correlation between the mapping depth of sequencing data and the number of editing sites.

Table 1 The statistics of transcriptome analysis in roots and leaves from three tobacco varieties

Organelle	Varieties	Tissue	Number of mapped reads	Average reads depth	Number of editing sites	Codon position (1,2,3)	Number of synonymous and non-synonymous editing sites
Mit	TN90	All	2,765,831	146×	486	158,300,28	39:447
		Root	108,157	25×	276	73,184,19	24:252
		Leaf	1,368,493	41×	415	124,256,35	39:376
	Basma	All	489,568	61×	338	99,218,21	23:315
		Root	193,926	43×	295	85,192,18	28:267
		Leaf	295,642	27×	196	59,128,9	20:176
	K326	All	339,767	56×	289	81,189,19	25:264
		Root	139,287	32×	233	63,155,15	22:211
		Leaf	200,480	26×	165	48,107,10	15:150
Pla	TN90	All	2,208,061	929×	60	6,51,3	4:56
		Root	20,761	23×	17	4,12,1	2:15
		Leaf	1,297,129	591×	50	3,47,0	4:46
	Basma	All	384,994	236×	35	2,33,0	3:32
		Root	43,966	29×	27	2,23,2	3:24
		Leaf	341,028	209×	36	2,34,0	4:32
	K326	All	320,708	184×	46	4,41,1	3:43
		Root	30,062	20×	18	3,13,2	1:17
		Leaf	290,646	165×	41	4,36,1	4:37

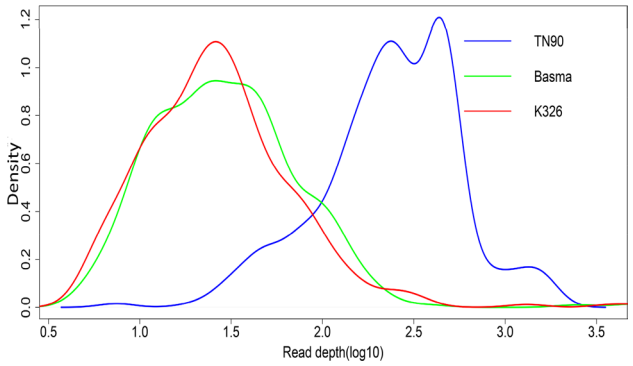

Fig. 1 The overall density of RNA-seq read depth distribution of three tobacco varieties

From the results of the number of RNA editing sites identified in Fig. 2C and D, for mitochondria, we can find that there are more editing sites identified in roots (295, 233) than that in leaves (196, 163) in Basma and K326 varieties, respectively, there's one exception, for TN90 variety, more sites were detected in leaves (415) than roots (276). Whereas for plastid, there are less editing sites in roots (17, 27 and 18) than that in leaves (50, 36 and 41) across all three varieties. The detailed information of RNA editing sites in both organelles from different tissues and varieties are listed in Tables S2 and S3, respectively. The statistics of RNA editing events also revealed that RNA editing occurred in the second codon position was mainly the largest in both organelles, accounting for more than 61% and 83% in mitochondria and plastid, respectively. This result is in accordance with the previous study (Zhang et al. 2020). Among the identified editing sites, the percentage of non-synonymous mutations reached 90%, which is significantly higher than that of synonymous mutations. Furthermore, we found that the amino acid changes tended to be hydrophobic, the proportion of changes from hydrophilic to hydrophobic accounted for about 53%. Take the TN90 mitochondria group as an example, the top amino acid changes were Ser-to-Leu (25.9%) and Pro-to-Leu (19%). These changes could increase the hydrophobic ability of proteins in critical areas and then maintain normal protein function.

Difference of average RNA editing efficiencies in roots and leaves

To investigate to what extent RNA editing affects normal protein function through base substitutions, the statistical analysis of the RNA editing efficiency was performed. The result showed that some editing sites were edited with high efficiency of more than 50% even completely edited whereas some were partially edited at low efficiency even less than

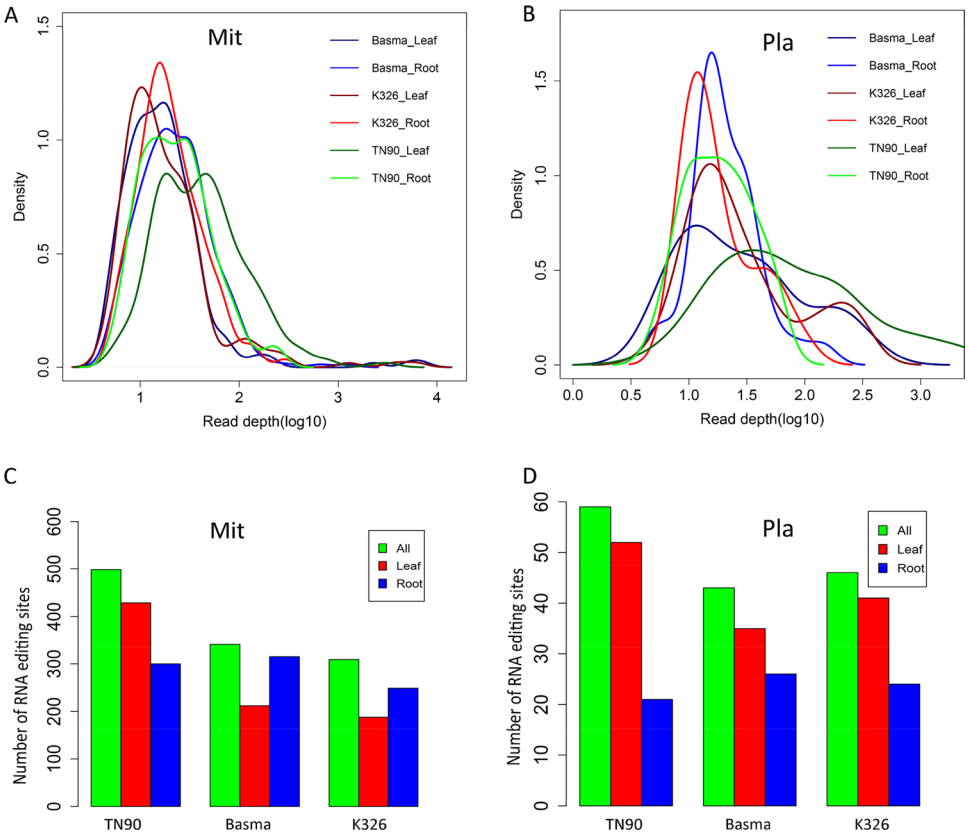

Fig. 2 Density distribution of read depth and the number of RNA editing sites identified in different tissues and varieties. **A** Density distribution of read depth of roots and leaves in mitochondria (Mit). **B** Density distribution of read depth of roots and leaves in plastid (Pla). **C** Bar plot of the numbers of RNA editing sites in mitochondria. **D** Bar plot of the numbers of RNA editing sites in plastid

5%. To clarify the relationships between varieties and tissues, clustering and heatmap plotting were performed based on normalized matrix of RNA editing efficiency. The result in Fig. 3 showed that samples with the same tissue were clustered together in spite of different varieties, reflecting the similarity of RNA editing pattern from the same tissue in tobacco.

To investigate RNA editing efficiency in a global view, the overall average editing efficiencies in mitochondria and plastid were computed. For each group, the average RNA editing efficiency was computed by summing all the values of sites and then being divided by the actual detected number of editing sites. So the comparison of average RNA editing efficiencies makes it convincible to speculate the RNA editing level in different tissues. A remarkable discrepancy (p value < 0.05) was detected in the comparison of RNA editing efficiency distribution between roots and leaves. The result in Fig. 4 showed that both RNA editing efficiencies of roots and leaves in mitochondria were around 0.9, while RNA editing efficiencies of root and leaf in plastid were about 0.7 and 0.9. In the mitochondria group, the average RNA editing efficiency in leaves of TN90 variety was 0.83, whereas it dropped to 0.81 in roots. The RNA editing efficiencies in leaves of varieties Basma and K326 were 0.89 and 0.92, whereas it dropped to 0.84 and 0.83 in roots. In the plastid group, the phenomenon of a discrepancy between roots and leaves was more apparent. The average RNA editing efficiencies in leaves of the three varieties Basma, K326

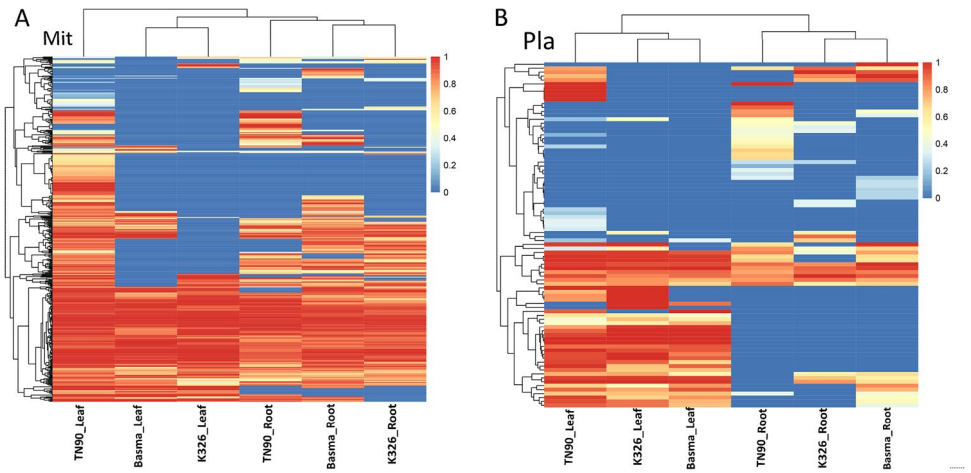

Fig. 3 Heatmaps of RNA editing efficiency of roots and leaves from three tobacco varieties. **A** RNA editing efficiency of roots and leaves in mitochondria; **B** RNA editing efficiency of roots and leaves in plastid. In the heatmap, the x-axis represents different samples, the y-axis represents RNA editing sites

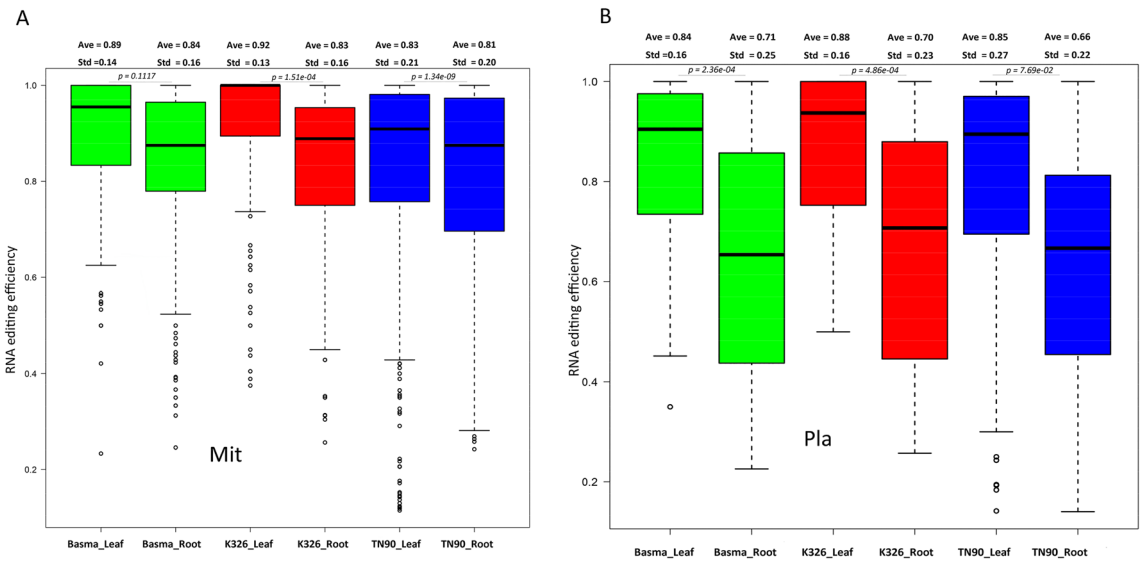

Fig. 4 Boxplots of RNA editing efficiency distribution of roots and leaves for each tobacco variety. **A** RNA editing efficiency distribution of roots and leaves in mitochondria (Mit); **B** RNA editing efficiency distribution of roots and leaves in plastid (Pla). For each tissue, the values of average and standard deviation were shown above the box. Two-tailed Wilcoxon rank-sum test was used to perform the pairwise comparison, the p value was shown on the top of the box

and TN90 were 0.84, 0.88 and 0.85, whereas it dropped to 0.71, 0.70 and 0.66, respectively, in roots. So the average RNA editing efficiency in roots was significantly lower than that of leaves in both organelles (p value < 0.05). The above results demonstrated that the overall RNA editing level was reduced in roots, and the degree of decline in plastid was higher than that in mitochondria, revealing the function of RNA editing in restoring normal protein function is both weakened in roots.

Differentially RNA-edited sites and involved genes

To investigate the tissue-specific RNA editing, ttest2 was used to identify differentially edited sites between roots and leaves. As a result, a total of 27 and 14 RNA editing sites involving 12 and 10 genes with a significant efficiency change (p value < 0.05) were detected in plastid and mitochondria, respectively. The detailed information of genes with related RNA editing sites is listed in Table S4. The result of clustering and heatmap plotting based on normalized matrix of differential RNA editing efficiency is provided in Fig. 5. From the heatmap, we can find that some sites were edited in leaves but completely lost in roots across three varieties, such as *ndhB*-467, *ndhB*-737, *ndhF*-290, *ndhD*-674, *ndhD*-599 and *nad3-61*. There were also many sites that demonstrated a trend of descending, such as *ndhA*_1073, *atpF*-92, *rps14*-94, *psbE*-214, *petB*-611, *rps10*-2,*rps4*-176 and *nad9*-398, these sites were highly edited in leaves but declined in roots. For example, RNA editing site *ndhA*_1073 (genome position: 121,715) occurred in both tissues, its editing efficiencies correspond to 0.93, 0.97 and 0.96 in leaves across three varieties, but reduced to 0.79, 0.43 and 0.85 in roots, respectively. The analysis of RNA editing involved genes showed that the differentially edited sites tend to occur in NDH complex subunits, a total of 6 and 5 differentially edited sites were detected in plastid genes *ndhD* and *ndhB*, respectively. In addition, two editing sites *rps14*_149 and *rps14*_80 located in gene *rps14* were detected to have a significant change (p value < 0.05) that highly edited in leaves and at low level in roots. This result was confirmed by a previous study of RNA editing in gene *rps14* with editing efficiency 0.8 in the young protonemata and decreased to approximately 0.2 in old protonemata of moss (Miyata and Sugita 2004). Compared with plastid, less differential RNA editing sites were detected in mitochondria, four types of gene families, including *rps*, *orf*, *nad*, and *cox*, were differentially edited. Take *rps4*_176 site as an example, the editing efficiencies were more than 0.9 in leaf samples but less than 0.8 in root samples across all three varieties. The moderate discrepancy between roots and leaves in mitochondria further agreed with the comparison analysis of average RNA editing efficiency.

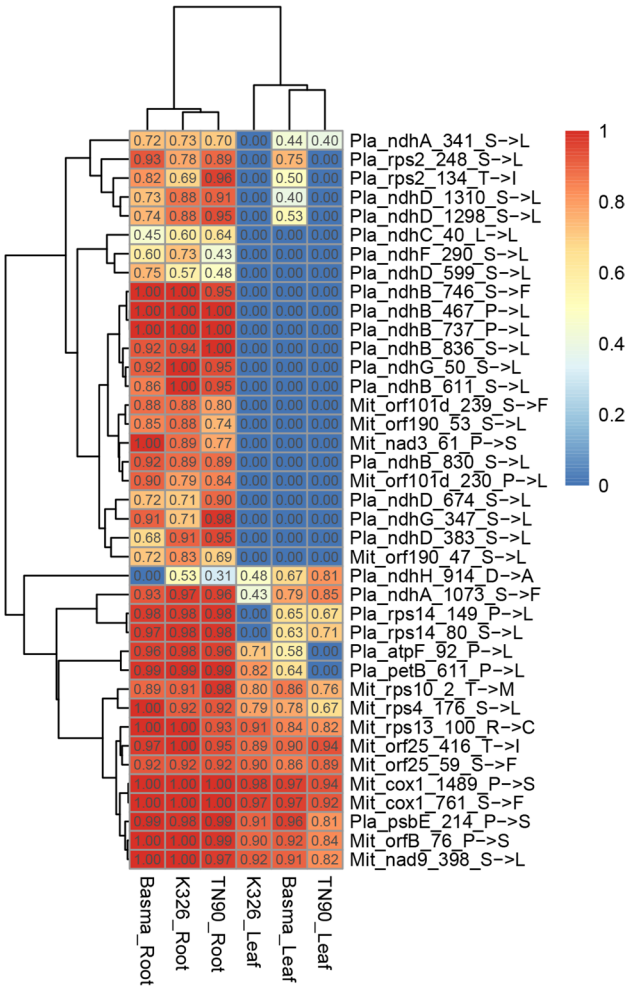

Fig. 5 Heatmap of editing efficiencies of 41 differentially edited sites. In the map, the row represents the editing sites with the notation of organelle, gene symbol, position and editing type, and the column represents the different tissues of three tobacco varieties. The values embedded in the box are editing efficiencies of the editing sites

Down-regulation of RNA editing genes and factors

To investigate the reason for the extensive loss of editing sites in roots, the expression of 22 differentially edited genes was analyzed. All these RNA-edited genes were expressed in both roots and leaves, which confirmed the reliability of the analysis. Gene expression analysis showed that there were 13 RNA-edited genes were down-regulated in roots, no genes were up-regulated in roots (p value < 0.01), see Table 2, among these genes, five genes that located in plastid were detected to have differentially edited sites, including *atpF*, *ndhB*, *ndhC*, *ndhA,* and *rps2*. Based on these observations, we suggest that both expression and editing of transcripts are tissue specifically regulated in different tissues, different types of tissues may generate different kinds of signals to affect RNA editing and the expression of organelles genes.

Table 2 The summary of differentially expressed RNA edited genes

Organelle	Genes	Fold change	*p* value
Pla	***rpl23***	− 2.29	1.22E − 06
	rps2	− 5.56	0.00014
	ndhC	− 4.86	0.000378
	rpoA	− 2.97	0.000455
	ndhB	− 29.3	0.000698
	rpl20	− 3.21	0.000911
	atpF	− 64.50	0.003726
	atpA	− 12.67	0.005696
	ndhA	− 3.21	0.007138
	psbL	− 154.76	0.008442
Mit	***rpl2***	− 59.79	7.35E − 06
	rpl16	− 45.72	0.000212
	rps3	− 3.23	0.005134

Genes with differentially edited sites were highlighted by black underlines

Numerous studies have reported that several protein families, including PPR proteins, MORF and ORRM proteins and so on, were involved in plant RNA editing as editing factors (Yan et al. 2018). PPR proteins recognize *cis*-elements around the editing sites and recruit editing enzymes to specific transcripts to catalyze the C-to-U conversion, other RNA editing factors form parts of the editing complex by interacting with PPR proteins (Shikanai 2015). To investigate the correlation between RNA editing pattern and editing factors or their interacting proteins, the expression of RNA editing factors were analyzed. The representative sequences of PPR, MORF and ORRM proteins of *Arabidopsis* were used as queries to blast against tobacco proteome, a total of 17 MORF proteins, 415 ORRM proteins, and 170 PPR proteins were identified. The RNA expression matrix with 6 root and 6 leaf samples from TN90 varieties was used to identify the differentially expressed RNA editing factors. As a result, a total of 149 differentially expressed RNA editing proteins were identified. Among these proteins, 6, 121 and 22 proteins belong to MORF, ORRM and PPR protein families, respectively. The detailed information on these genes can be found in Table S5. A heatmap based on normalized RNA expression matrix of differentially expressed RNA editing factors with fold change > 2 is plotted in Fig. 6. Interestingly, 107 out of 149 differentially expressed RNA editing factors were down-regulated in all root samples, the expression of 21 out of 22 PPR proteins and all 6 MORF proteins were reduced in roots. The up-regulated editing factors were only present in ORRM protein family except one is a member of PPR family. Proteins of the MORF family were once considered to be essential components of plant editosomes, one previous study demonstrated that plastid RNA editing events in rosette leaves and flowers were reduced by *morf9* mutation (Tian et al. 2019). From the results, we suggested that the editing sites that suffered loss or reduction in roots may be edited by these differential expressed PPR and MORF proteins. In other words, this factor-mediated RNA editing is subject to tissue-dependent regulation, and the resultant organelle proteomes may be pertinent to the specific tissue functions.

Discussion

As an important epigenetic mechanism that modified genome-encoded transcripts, RNA editing expands the complexity of the transcriptome. In plants, most of the editing sites are located in the coding region in the form of C-to-U conversion. Numerous studies have provided evidence that RNA editing played an important role in the restoration of conserved amino acids, which are essential for proteins' normal function in plant plastids and mitochondria (Bock et al. 1994; Hammani et al. 2009; Tillich et al. 2009). For instance, the lack of editing of tobacco *psbF* mRNA fails to restore a conserved phenylalanine codon and leads to a mutant phenotype (Bock et al. 1994). Previous studies also demonstrated that RNA editing events differed among plant species, and were relevant to developmental stages or impacted by environmental factors (Rodrigues et al. 2017; Zhang et al. 2020). Until now, discrepancy in RNA editing sites among tissues has only been detected in certain genes, such as *psbF* and *psbL* genes (Bock et al. 1993). It is still difficult to make general conclusions to what extent that discrepancy exists among different tissues. In this study, we chose two types of plant tissues, where the leaf is a green photosynthetic organ and the root is a non-green organ; their RNA editing profiles in mitochondria/plastid transcripts from three representative tobacco varieties were investigated. By comparing the numbers of editing sites across different samples, we found that the mapping depth was proportional to the numbers of editing sites, revealing that transcript abundance is an important limiting factor in identifying editing sites. The results also revealed that editing of tobacco transcripts was dynamic and site specific, the editing efficiency varies from site to site, some editing sites are highly edited even completely edited, whereas the others are partially edited at low levels. Furthermore, we found that RNA editing was differentially regulated in different tissues, the average editing efficiency in both organelles was reduced in roots. Moreover, compared with mild loss level in mitochondria, plastid suffered a more drastic loss.

In this study, a total of 41 differentially edited RNA editing sites were identified in mitochondria and plastid. There were more differentially edited sites in plastid than in mitochondria. On the one hand, a lot of sites exhibited declined editing extent, including *atpF*-92, *rps14*-94, *psbE*-214,

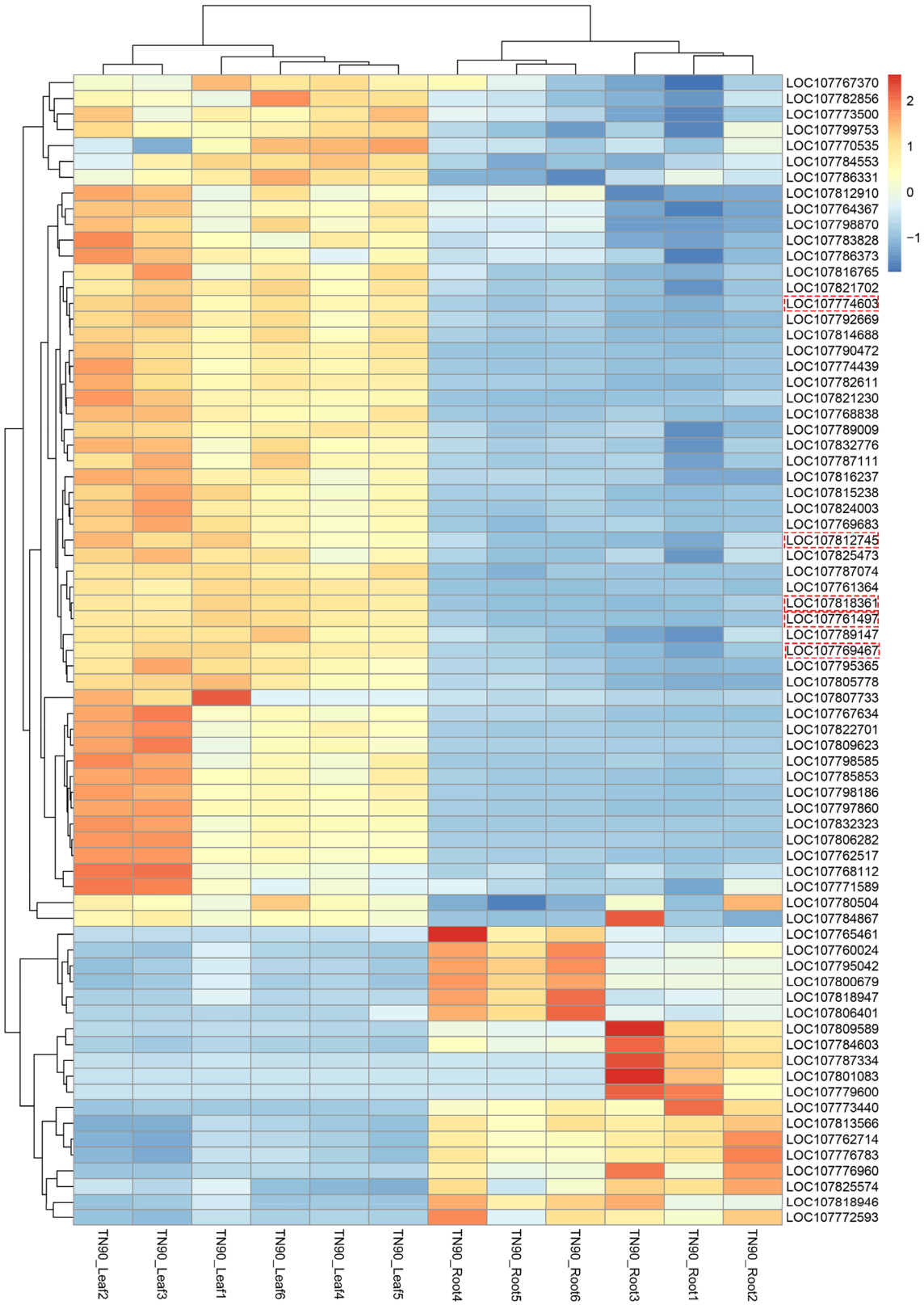

Fig. 6 Heatmap of differential expression of RNA editing factors in roots and leaves from TN90 variety. The *x*-axis represents different samples and the *y*-axis represents RNA editing factors. There are six replicates for each tissue. The *morf* genes are indicated by red dotted box

petB-611, *rps10*-2, *rps4*-176, and *nad9*-398 that were highly edited in leaves and lowly edited in roots. On the other hand, a lot of sites were completely lost in root across three varieties, for instance, the editing of *ndhB*-467, *ndhB*-737, *ndhF*-290, *ndhD*-674, *ndhD*-599 that located in plastid were completely lost in all root samples. These observations are consistent with one previous study of plastid RNA editing (Chateigner-Boutin and Hanson 2003), which also observed that transcripts of most NDH subunits are edited inefficiently in roots. In general, most of editing sites that suffered losses in our study are located in plastid transcripts encoding NDH subunits. Plant NDH is a multiple-subunit (NDH A~K) complex in the thylakoid membranes mediating cyclic electron transport (CET) and is assumed to be essential for plant growth and development in normal growth and stress periods (Ma et al. 2021). NDH complex plays a key role in the process of antioxidation under certain conditions even in wild-type plants, the oxidative stress of chloroplast can be alleviated and the over-reduction of stroma can be prevented (Peng et al. 2008; Ma et al. 2021). The loss or reduction of *ndh* genes editing in roots but not in leaves suggests that the NDH complex is not active in roots, because the root is composed of non-photosynthetic cells without photo-oxidative stress; hence, there is a relaxed selection on editing in photosynthetic genes. By contrast, leaf functions as photosynthesis site and has a much higher expression of *ndh* genes; the functional discrepancy between roots and leaves may partly explain the reduction in plastid RNA editing in roots. Whereas for mitochondria, we found that only the editing of *nad3-61* was completely lost in all root samples.

Regardless of the above dramatic discrepancy between roots and leaves, there are still many editing sites that are highly edited, up to 0.9, in both tissues, such as *rps12*-146, *cox1*-590, *atp9*-20, their stable and high editing ensures their encoded proteins are homogeneous in assembling into functional complexes that are equally important for both tissues. Compared with plastid, the reduction of editing efficiency in mitochondria tends to be mild, suggesting that respiration is essential for both roots and leaves. However, many mitochondria genes maybe only actively expressed in dividing cells but not in mature cells (Li et al. 1996); once the respiratory complexes are synthesized in young cells, they are not actively turned over in mature cells, so it is not necessary to edit RNA in mature cells. The selection pressure on RNA editing in respiratory genes is also slightly relaxed, which may partly explain the mild reduction of mitochondria RNA editing in roots. The above observations suggest that RNA editing in tobacco is subject to tissue regulation, the resultant organelle proteomes in turn might be related to the specific tissue functions. For an editing gene, if the unedited and partially edited transcripts can both be translated, its heterogeneity will be increased simultaneously. It will be interesting to further investigate how the plant accommodates these transcript variants.

The current hypothesis is that PPR, MORF and ORRM proteins participate together in the plant RNA editing apparatus, where PPR proteins recognize *cis*-elements around the editing sites, and then recruit the editing enzymes specific transcripts to catalyze the C-to-U conversion (Schmitz-Linneweber and Small 2008). Therefore, plant editing status of transcripts in different tissues might be regulated by the availability of editing factors and their interacting proteins. We characterized four types of RNA editing factors and quantified their expression level, an extraordinarily high percentage of down-regulated RNA editing factor in roots was detected, particularly for PPR and MORF proteins. Hence, we suggest that RNA editing and their editing factors were both modulated in a tissue-specific manner, the editing sites that suffered loss or reduction in roots may be edited by these differentially regulated PPR and MORF proteins. Further studies are still required to identify components of the editing machinery and clarify the regulation relationships between editing factors and editing sites.

Conclusion

In this study, we implemented an integrative analysis of RNA-seq and genome re-sequencing data of roots and leaves from three tobacco varieties (TN90, Basma and K326) to identify RNA editing events and investigate the tissue specificity of RNA editing in plants. We found that the average RNA editing level was obviously reduced in roots versus leaves; however, the reduction of RNA editing efficiency in mitochondria was mild compared with plastid. The gene expression analysis of RNA editing factors further revealed that down-regulated RNA editing factors PPR and MORF proteins might limit the editing of sites that suffered loss or reduction. Our analysis provided references and insights into understanding the tissue specificity of plant RNA editing and the roles of RNA editing in tissue differentiation.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11816-021-00692-3.

Acknowledgements This work was supported by grants from the CAS Pioneer Hundred Talents Program, the National Natural Sciences Foundation of China (Nos. 61402457, 32070682), and the National Science and Technology Innovation Zone Project (Nos. 1716315XJ00200303, 1816315XJ00100216).

Author contributions A-DZ, JF, and X-JZ conceived and designed the research, JF performed data analysis, A-DZ, JF wrote the manuscript. X-HJ, T-FW and X-JZ provided many critical suggestions. All authors reviewed the manuscript.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Bock R, Hagemann R, Kossel H, Kudla J (1993) Tissue- and stage-specific modulation of RNA editing of the psbF and psbL transcript from spinach plastids–a new regulatory mechanism? Mol Gen Genet 240(2):238–244

Bock R, Kossel H, Maliga P (1994) Introduction of a heterologous editing site into the tobacco plastid genome: the lack of RNA editing leads to a mutant phenotype. EMBO J 13(19):4623–4628

Brown J, Pirrung M, McCue LA (2017) FQC dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. Bioinformatics 33(19):3137–3139

Chateigner-Boutin AL, Hanson MR (2003) Developmental co-variation of RNA editing extent of plastid editing sites exhibiting similar cis-elements. Nucleic Acids Res 31(10):2586–2594

Dai Z, Sheridan JM, Gearing LJ, Moore DL, Su S, Wormald S, Wilcox S, O'Connor L, Dickins RA, Blewitt ME, Ritchie ME (2014) edgeR: a versatile tool for the analysis of shRNA-seq and CRISPR-Cas9 genetic screens. F1000Res 3:95

Danecek P, McCarthy SA (2017) BCFtools/csq: haplotype-aware variant consequences. Bioinformatics 33(13):2037–2039

Edera AA, Gandini CL, Sanchez-Puerta MV (2018) Towards a comprehensive picture of C-to-U RNA editing sites in angiosperm mitochondria. Plant Mol Biol 97:215–231

Grewe F, Herres S, Viehover P, Polsakiewicz M, Weisshaar B, Knoop V (2011) A unique transcriptome: 1782 positions of RNA editing alter 1406 codon identities in mitochondrial mRNAs of the lycophyte *Isoetes engelmannii*. Nucleic Acids Res 39(7):2890–2902

Grimes BT, Sisay AK, Carroll HD, Cahoon AB (2014) Deep sequencing of the tobacco mitochondrial transcriptome reveals expressed ORFs and numerous editing sites outside coding regions. BMC Genomics 15:31

Grosskopf D, Mulligan RM (1996) Developmental- and tissue-specificity of RNA editing in mitochondria of suspension-cultured maize cells and seedlings. Curr Genet 29(6):556–563

Hammani K, Okuda K, Tanz SK, Chateigner-Boutin AL, Shikanai T, Small I (2009) A study of new Arabidopsis chloroplast RNA editing mutants reveals general features of editing factors and their target sites. Plant Cell 21(11):3686–3699

Jo H, Koh G (2015) Faster single-end alignment generation utilizing multi-thread for BWA. Biomed Mater Eng 26(Suppl 1):S1791-1796

Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12(4):357–360

Knie N, Grewe F, Fischer S, Knoop V (2016) Reverse U-to-C editing exceeds C-to-U RNA editing in some ferns - a monilophyte-wide comparison of chloroplast and mitochondrial RNA editing suggests independent evolution of the two processes in both organelles. BMC Evol Biol 16:1–12

Li XQ, Zhang MD, Brown GG (1996) Cell-specific expression of mitochondrial transcripts in maize seedlings. Plant Cell 8(11):1961–1975

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, S. Genome Project Data Processing (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25(16):2078–2079

Ma M, Liu Y, Bai C, Yong JWH (2021) The significance of chloroplast NAD(P)H dehydrogenase complex and its dependent cyclic electron transport in photosynthesis. Front Plant Sci 12:661863

Miyata Y, Sugita M (2004) Tissue- and stage-specific RNA editing of rps 14 transcripts in moss (*Physcomitrella patens*) chloroplasts. J Plant Physiol 161(1):113–115

Nakajima Y, Mulligan RM (2001) Heat stress results in incomplete C-to-U editing of maize chloroplast mRNAs and correlates with changes in chloroplast transcription rate. Curr Genet 40(3):209–213

Peng L, Shimizu H, Shikanai T (2008) The chloroplast NAD(P)H dehydrogenase complex interacts with photosystem I in Arabidopsis. J Biol Chem 283(50):34873–34879

Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL (2016) Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc 11(9):1650–1667

Rodrigues NF, da Fonseca GC, Kulcheski FR, Margis R (2017) Salt stress affects mRNA editing in soybean chloroplasts. Genet Mol Biol 40(1):200–208

Schmitz-Linneweber C, Small I (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci 13(12):663–670

Shikanai T (2015) RNA editing in plants: machinery and flexibility of site recognition. BBA Bioenerg 1847(9):779–785

Sierro N, Battey JN, Ouadi S, Bakaher N, Bovet L, Willig A, Goepfert S, Peitsch MC, Ivanov NV (2014) The tobacco genome sequence and its comparison with those of tomato and potato. Nat Commun 5:3833

Sosso D, Mbelo S, Vernoud V, Gendrot G, Dedieu A, Chambrier P, Dauzat M, Heurtevin L, Guyon V, Takenaka M, Rogowsky PM (2012) PPR2263, a DYW-Subgroup Pentatricopeptide repeat protein, is required for mitochondrial nad5 and cob transcript editing, mitochondrion biogenesis, and maize growth. Plant Cell 24(2):676–691

Sugiyama Y, Watase Y, Nagase M, Makita N, Yagura S, Hirai A, Sugiura M (2005) The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. Mol Genet Genomics 272(6):603–615

Sun F, Wang XM, Bonnard G, Shen Y, Xiu ZH, Li XJ, Gao DH, Zhang ZH, Tan BC (2015) Empty pericarp7 encodes a mitochondrial E-subgroup pentatricopeptide repeat protein that is required for ccmF(N) editing, mitochondrial function and seed development in maize. Plant J 84(2):283–295

Takenaka M, Zehrmann A, Verbitskiy D, Kugelmann M, Hartel B, Brennicke A (2012) Multiple organellar RNA editing factor (MORF) family proteins are required for RNA editing in mitochondria and plastids of plants. Proc Natl Acad Sci USA 109(13):5104–5109

Takenaka M, Zehrmann A, Verbitskiy D, Hartel B, Brennicke A (2013) RNA editing in plants and its evolution. Annu Rev Genet 47(47):335–352

Tang JP, Zhang WW, Wen K, Chen GM, Sun J, Tian YL, Tang WJ, Yu J, An HZ, Wu TT, Kong F, Terzaghi W, Wang CM, Wan JM (2017) OsPPR6, a pentatricopeptide repeat protein involved in editing and splicing chloroplast RNA, is required for chloroplast biogenesis in rice. Plant Mol Biol 95(4–5):345–357

Tian F, Yu J, Zhang Y, Xie Y, Wu B, Miao Y (2019) MORF9 functions in plastid RNA editing with tissue specificity. Int J Mol Sci 20(18):4635

Tillich M, Hardel SL, Kupsch C, Armbruster U, Delannoy E, Gualberto JM, Lehwark P, Leister D, Small ID, Schmitz-Linneweber C (2009) Chloroplast ribonucleoprotein CP31A is required for editing and stability of specific chloroplast mRNAs. Proc Natl Acad Sci USA 106(14):6002–6007

Tseng CC, Lee CJ, Chung YT, Sung TY, Hsieh MH (2013) Differential regulation of Arabidopsis plastid gene expression and RNA editing in non-photosynthetic tissues. Plant Mol Biol 82(4–5):375–392

Wu S, Liu W, Aljohi HA, Alromaih SA, Alanazi IO, Lin Q, Yu J, Hu S (2018) REDO: RNA editing detection in plant organelles based on variant calling results. J Comput Biol 25(5):509–516

Xiong J, Tao T, Luo Z, Yan S, Liu Y, Yu X, Liu G, Xia H, Luo L (2017) RNA editing responses to oxidative stress between a wild abortive type male-sterile line and its maintainer line. Front Plant Sci 8:2023

Yan J, Zhang Q, Yin P (2018) RNA editing machinery in plant organelles. Sci China Life Sci 61(2):162–169

Zahn LM (2017) The evolution of edited RNA transcripts. Science 355(6331):1278–1279

Zehrmann A, Hartel B, Glass F, Bayer-Csaszar E, Obata T, Meyer E, Brennicke A, Takenaka M (2015) Selective homo- and heteromer interactions between the multiple organellar RNA editing factor (MORF) proteins in *Arabidopsis thaliana*. J Biol Chem 290(10):6445–6456

Zhang A, Jiang X, Zhang F, Wang T, Zhang X (2020) Dynamic response of RNA editing to temperature in grape by RNA deep sequencing. Funct Integr Genomics 20(3):421–432