J. Plant Biol. (2019) 62:351-358
DOI 10.1007/s12374-018-0452-5

ORIGINAL ARTICLE

Analysis of Tissue-specific RNA Editing Events of Genes Involved in RNA Editing in *Arabidopsis thaliana*

Umme Qulsum[1], Md. Thoufic Anam Azad[1,2] and Toshifumi Tsukahara[1,3,4,*]

[1]*School of Materials Science, Japan Advanced Institute of Science and Technology; Asahidai, Nomi City, Ishikawa 923-1292, Japan*
[2]*Rajshahi University, Rajshahi-6205, Bangladesh*
[3]*Area of Bioscience and Biotechnology, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi City, Ishikawa 923-1292, Japan*
[4]*Division of Transdisciplinary Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi City, Ishikawa 923-1292, Japan*

Received: December 20, 2018 / Accepted: July 1, 2019
© Korean Society of Plant Biologists 2019

Abstract In plants, RNA editing is a post-transcriptional modification of transcripts and commonly occurs in plastids and mitochondria. In the case of flowering plants, not only PPR but also non-PPR proteins, such as MORF/RIP, ORRM and OZ partake in the diverse RNA editing complex. In *Arabidopsis thaliana*, 12 types of RNA editing patterns have been predicted in nuclear, chloroplast, and mitochondrial transcripts. In this study, tissue-specific RNA editing events were detected in gene families involved in RNA editing. Different Arabidopsis tissues at variable developmental stages, including 4-, 8-, and 12-d-old seedlings and 16-, 21-, 27-, and 32-d-old leaves, stems, stipes and roots, were collected and used for this study. Nine types of RNA editing events, including C-to-U, U-to-C, A-to-I, A-to-C, A-to-U, G-to-A, G-to-C, U-to-A and U-to-G, were identified in target genes. Most of the editing events occurred in seedlings and leaves and a few in stem tissues. Extensive U-to-C editing (60%) and A-to-G editing (54%) was detected in 12-d-old seedlings and 21-d-old leaves, respectively. This is the first experimental report of tissue-specific nuclear RNA editing events found in plants. This study will provide important information for revealing the mechanism of RNA editing in plants.

Keywords: *Arabidopsis thaliana*, RNA editing events, Tissue-specific

*Corresponding author; Toshifumi Tsukahara
Tel : +81-761-51-1149
E-mail : tukahara@jaist.ac.jp

Introduction

RNA editing is defined as the process by which nucleotides in RNA molecules undergo modifications, such as additions or deletions. Therefore, RNA editing alters the flow of genetic information from the genome to mature RNA (Takenaka et al. 2013). Diverse types of RNA editing events have been reported in viruses, primitive eukaryotes, vertebrates, fungi, and plants. Interestingly, C-to-U and A-to-I RNA editing is a common deamination phenomenon in animals and plants, whereas U-to-C RNA editing due to amination is specific to plants (Meng et al. 2010). In plants, a C-to-U deaminase enzyme, DYW domain protein, has been described (Shikanai 2015); however, an A-to-I homologous deaminase has not yet been reported. In *Arabidopsis thaliana*, the RNA editing machinery, collectively referred to as the editosome, consists of at least four protein families. These include PPRs (pentatricopeptide repeat) protein, MORFs (multiple organellar RNA editing factors) /RIPs (RNA editing factor interacting proteins), ORRM1 (organelle RNA recognition motif) and OZ1 (organelle zinc-finger). The PPR gene family is one of the largest gene families in Arabidopsis, comprising approximately 450 genes (Lurin et al. 2004). These genes are either directly or indirectly responsible for RNA editing (Doniwa et al. 2010; Chateigner-Boutin et al. 2013; Schallenberg-Rüdinger et al. 2013; Leu et al. 2016). The MORF/RIP, ORRM and OZ proteins are also directly or indirectly involved in C-to-U RNA editing and play an important role in the editosome (Sun et al. 2016). The plant quintuple editing factor 1 (QED1), PPR protein required for *accD* RNA editing 1 (RARE1) and NUWA proteins are functional homologs of the mammalian APOBEC-1 protein and are responsible for

C-to-U RNA editing in plants (Wagoner et al. 2015; Guillaumot et al. 2017). These data indicate that RNA editing is a normal physiological phenomenon in plants.

Different theories have been proposed to describe the purpose of RNA editing. According to one theory, RNA editing is an evolutionary process, which did not occur in primitive organisms. According to another theory, RNA editing is responsible for the adaptation of organisms and repair of DNA damage caused by ultraviolet radiation or other environmental stresses (Takenaka et al. 2013; Barkan and Small 2014). Plastid and mitochondrial genome editing events have mostly been investigated using bioinformatics and experimental approaches (He et al. 2016; Chen et al. 2017; Ichinose and Sugita 2017; Rodrigues et al. 2017; Wu et al. 2017). However, RNA editing of nuclear transcripts has been studied only using bioinformatics approaches, for example, to investigate the frequency and distribution of RNA editing events in Arabidopsis (Meng et al. 2010).

The DYW domain, which shows C-to-U conversion activity, has a similar function as the cytidine deaminase protein (Shikanai 2015). The PPR proteins are a specialized type of protein that bind to their cognate RNA (Cheng et al. 2016). Like the PUF protein, the binding specificity of PPR proteins is almost reprogrammable, except for some differences (Delannoy et al. 2007; Yagi et al. 2014; Cheng et al. 2016). The majority of the PPR family proteins are responsible for RNA editing in plants and play an important role in alternative splicing (Shikanai 2015). RNA editing efficiency varies with the tissue type and developmental stage; for example, RNA editing is more common in green tissues than in non-green tissues (Tseng et al. 2013).

In this study, we investigated RNA editing in various

Table 1. Details of RNA editing events and the corresponding amino acid changes detected in Arabidopsis tissues at different developmental stages

Gene name-editing site in transcript	Editing site in genome and edited sequence	RNA editing in tissues (editing %) and edited sequence			Editing location (codon/ nucleotide change)	Amino acid change
PPRIIa-2969	3267 TGAGA	16-d-old leaf (–)	21-d-old leaf (14%) TGGGA	27-d-old leaf (–)	3' UTR, A>G	NA
PPRIIb-2720	2983 AGACA	21-d-old stem (–)	27-d-old stem (37%) AGGCA	32-d-old stem (–)	3' UTR, A>G	NA
PPRIIb-2694	2957 AGATT	16-d-old stem (–)	21-d-old stem (–)	27-d-old stem (28%) AGGTT	3' UTR, A>G	NA
PPRIIIb-2833	3089 CTAGT	16-d-old leaf (–)	21-d-old leaf (54%) CTGGT	32-d-old leaf (–)	3' UTR, A>G	NA
PPRIVb-1008	1333 CCAAA	8-d-old seedling (–)	16-d-old stem (23%) CCGAA	NA	3' UTR, A>G	NA
PPRVa-2267	2999 ACGAC	4-d-old seedling (–)	8-d-old seedling (–)	12-d-old seedling (38%) ACAAC	GAC>AAC,G>A	D702N
PPRVa-2293	2999 GCAGG	4-d-old seedling (–)	8-d-old seedling (–)	12-d-old seedling (28%) GCTGG	GCA>GCU,A>U	A710A
PPRVIa-134	134 TTTTC	4-d-old seedling (–)	8-d-old seedling (–)	12-d-old seedling (40%) TTCTC	UUC>CUC, U>C	F7L
PPRVIIa-1035	1144 CAAGA	16-d-old leaf (–)	21-d-old leaf (34%) CAGGA	27-d-old leaf (–)	AAG>AGG, A>G	K336R
PPRVIIIa-1260	1260 ATCAC	4-d-old seedling (–)	8-d-old seedling (–)	12-d-old seedling (30%) ATTAC	AUC>AUU, C>U	I420I
PPRIXa-2749	2648 GAGCC	4-d-old seedling (–)	8-d-old seedling (–)	12-d-old seedling (26%) GACCC	AGC>ACC, G>C	S883T
PPRXa-355	355 ATGAG	16-d-old leaf (–)	21-d-old leaf (–)	27-d-old leaf (10%) ATAAG	AUG>AUA, G>A	M111I
PPRXa-512	512 TTCAT	16-d-old leaf (–)	21-d-old leaf (–)	27-d-old leaf (9%) TTTAT	CAU>UAU, C>U	H164Y
PPRXIb-2852	3224 GGAGA	4-d-old seedling (–)	8-d-old seedling (23%) GGCGA	12-d-old seedling (–)	GAG>GCG, A>C	E831A
DAG1a-615	1296 CCTCT	16-d-old leaf (–)	21-d-old leaf (28%) CCACT	27-d-old leaf (–)	UCU>ACU, U>A	S130T
DAG1a-794	1559 AGAGG	16-d-old leaf (–)	21-d-old leaf (–)	27-d-old leaf (16%) AGTGG	AGA>AGU, A>U	R189S

Arabic numbers represent editing site in transcript and genome. Underline base indicates the edited base, "–" indicates the absence of RNA editing. NA, not applicable; 3' UTR, 3' untranslated region.

Arabidopsis tissues at different developmental stages of growth, with a special focus on PPR proteins. Our data revealed tissue-specific RNA editing events in selected genes. The information gained in this study will advance our understanding of RNA editing in plants.

Results

In this study, we mainly focused on RNA editing sites in alternatively spliced isoforms. A total of thirty (30) differentially expressed genes were identified based on their functional importance and variable expression patterns in different Arabidopsis tissues. Thirty candidate genes belonging to different families were investigated, including *PPR* (16), *ZnF* (9) *MORF/RIP* (1), *RRM* (2), *PP01* (1), and *GSDA* (1) genes (Table S1). PCR products of all genes were sequenced three times, and transcript positions were determined from the http://atgenie.org database. Overall, we identified RNA editing events in *PPR* (11) genes and *ZnF* (1) gene; no editing events were detected in any of the remaining genes.

Tissue-specific RNA Editing Events

In this study, we identified usual/transition and unusual/transversion tissue-specific RNA editing events in gene transcripts (Table 1; Fig. S1; Fig. S2). Usual types consist of amination and deamination type RNA editing. In the course of study, we identified 21 RNA editing sites in total 12 genes which are expressed in different tissues. Among these we observed usual type (13) and unusual type (8). In case of *PPR* gene family we found 19 editing sites among these usual type (13) and unusual type (6). But in case of *ZnF* gene we only detected unusual type (2) editing site.

Another thing we observed is RNA editing sites expressed in different tissues. In our analysis, we identified 10 editing sites which were expressed in seedling tissues including usual type (4) and unusual (6), 8 editing sites were expressed in leaf tissues including usual type (6) and unusual (2) and only usual type (3) was expressed in stem tissues.

In *PPRVIIIa* gene, 30% C-to-U RNA editing was detected at 1,260 nt in 12-d-old seedlings but not in 4- or 8-d-old seedlings (Fig. 1A). In *PPRIIa* gene, 39% A-to-G another deaminase-type editing was detected at 2,799 nt in 21-d-old leaves but not in 16- or 27-d-old leaves (Fig. 1B). In *PPRIIIa* gene, 60% U-to-C an amination-type editing was detected at 2972 nt in 12-d-old seedlings but not in 4- or 8-d-old seedlings (Fig. 1C).

In *PPRIXa* gene, 31% A-to-C an unusual type of RNA editing was observed at 2,720 nt in 12-d-old seedlings but not in 4- or 8-d-old seedlings (Fig. 1D). In *PPRIb* gene, 25% A-to-U RNA editing was identified at 250 nt in 21-d-old leaves, although it was not detected in 16- and 27-d-old leaves (Fig. 1E). In *PPRXIb* gene, 41% U-to-G RNA editing was identified at 2,809 in 12-d-old seedlings but not in 4- or 8-d-old seedlings (Fig. 1F). RNA editing sites were more frequent in seedling and leaf tissues and less common in stem tissues but not in root and stipe tissues (Table 1).

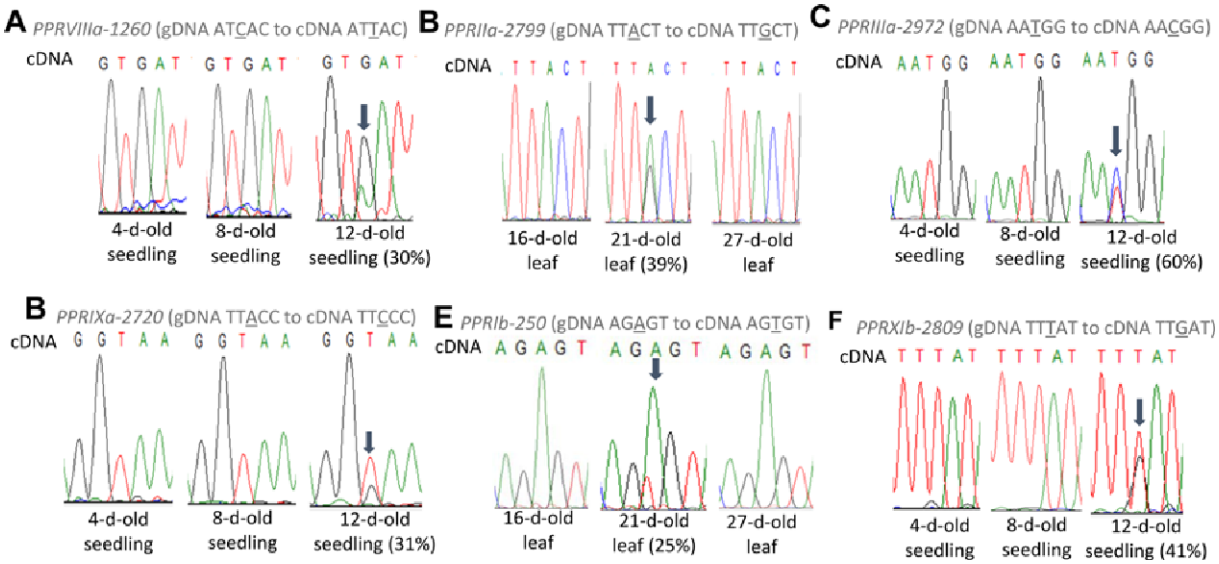

Fig. 1. Validation of usual (deaminase/aminase) and unusual types of RNA editing events in Arabidopsis tissues at different developmental stages via cDNA sequencing. Sequencing of C-to-U editing events in 12-d-old seedlings using reverse primers (A), A-to-G editing in 21-d-old leaves using forward primers (B), and U-to-C editing in 12-d-old seedlings using reverse primers (C). Sequencing of A-to-C editing events in 12-d-old seedlings using reverse primers, amino acid change (L to F) (D), A-to-U editing in 21-d-old leaves using forward primers, amino acid change (E to V) (E), and U-to-G editing in 12-d-old seedlings using forward primer2 (F). Arrow indicates the edited base. gDNA, genomic DNA.

Additionally, RNA editing was more predominant in the protein-coding regions of genes than in untranslated regions (Table 1).

Analysis of Amino Acid Substitutions Induced by RNA Editing

RNA editing resulted in several amino acid substitutions in target genes (Table 1). In *PPRIb-250*, *PPRVa-2267*, *PPRVIa-134*, *PPRVIIa-1035*, *PPRIXa-2720*, *PPRIXa-2749*, *PPRXa-355*, *PPRXa-512*, *PPRXIb-2852*, *DAG1a-615*, and *DAG1a-794*, RNA editing changed glutamic acid to valine (E-to-V), aspartic acid to asparagine (D-to-N), phenylalanine to leucine (F-to-L), lysine to arginine (K-to-R), leucine to phenylalanine (L-to-F), serine to threonine (S-to-T), methionine to isoleucine (M-to-I), histidine to tyrosine (H-to-Y), glutamic acid to alanine (E-to-A), serine to threonine (S-to-T) and arginine to serine (R-to-S), respectively, as determined using the ExPASy translation tool. Additionally, two silent mutations were detected in *PPRVa-2293* and *PPRVIIIa-1260* (Table 1).

Characterization of RNA Editing Events

In this study, we identified nine types of RNA editing events. These events incorporated all possible intra-base substitutions: C-to-U, U-to-C, A-to-G, A-to-C, A-to-U, G-to-A, G-to-C, U-to-A, and U-to-G. The A-to-G substitution was the most frequent type in the UTR regions (75%), whereas A-to-U substitution was the most frequent type in CDSs (23%) (Fig. 2A). A group of substitutions, including C-to-U, A-to-C, and G-to-A, were the second most frequent editing type; each of these nucleotide changes occurred at a 15% in the CDS of genes (Fig. 2A). Depending on the editing percentage, U-to-C conversion was the most common (60%) and U-to-G was the second most common type (41%), among all of the UTR-specific RNA editing events (Fig. 2B). In the case of A-to-G editing, the editing frequency was lower than the editing percentage in CDSs of genes, whereas the reverse was true in the UTR regions (Fig. 2A, B). Additionally, the frequency of usual types of RNA editing showed the highest percentage in the CDS regions and lowest in UTR regions (Fig. 2C).

Distribution and Effects of RNA Editing

The distribution of RNA editing events showed that the frequency and type of RNA editing events were notably variable among the different Arabidopsis gene families. We identified 21 RNA editing sites in a total of 12 genes (Table 2). Of these 21 editing sites, 13 were located in the CDS of genes and 8 in the UTR regions (Table 2). RNA editing altered the secondary structure of three gene transcripts, each of which was edited in the UTR region (Table 2). The tertiary structure of only one encoded protein was affected by RNA editing (Table 2).

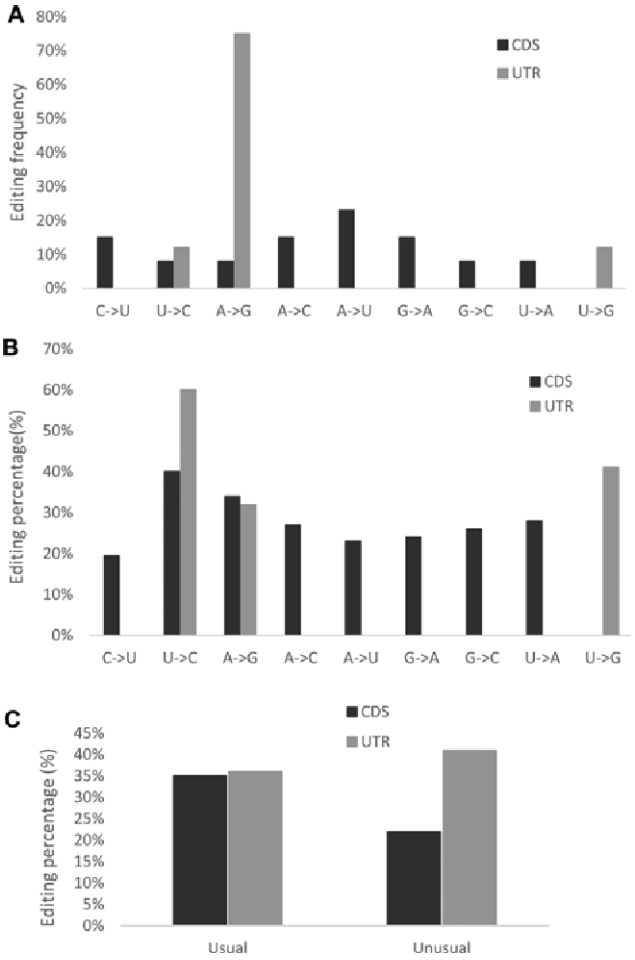

Fig. 2. Schematic representation of RNA editing events in Arabidopsis tissues. (A) Percentage of RNA editing events in the coding sequence (CDS) and untranslated region (UTR) of genes. (B) Percentage of RNA editing in the CDS and UTR. (C) Percentage of usual and unusual types of RNA editing events in the CDS and UTR. Black bar represents CDS, and gray bar indicates the UTR.

Structural Effects of RNA Editing

To determine the potential functional effects of RNA editing, we predicted the secondary structure of mRNA using RNAfold version 2.1.6 and the tertiary structure of protein using RaptorX and Swiss model. The mRNA secondary structure was affected by A-to-G, U-to-C, and U-to-G RNA editing events, whereas the protein structure was affected only by G-to-A RNA editing. Alterations observed in the mRNA secondary structure and protein tertiary structure were significant. In *PPRIIb-2720* mRNA, 'A' was located in the bulge loop before editing, whereas the edited nucleotide 'G' was located in the stem after editing (Fig. 3A). In the

Table 2. Distribution of RNA editing sites in genic regions and their effect on mRNA and protein structure.

Gene family	Given gene name	Number of RNA editing sites		Effect on	
		CDS	UTR	Secondary structure of mRNA	Tertiary structure of protein
PPR	PPRIb	1	0	-	-
	PPRIIa/PPRIIb	0	4	+	NA
	PPRIIIa/PPRIIIb	0	2	+	NA
	PPRIVb	0	1	-	NA
	PPRVa	2	0	-	+
	PPRVIa	1	0	-	-
	PPRVIIa	1	0	-	-
	PPRVIIIa	1	0	-	-
	PPRIXa	2	0	-	-
	PPRXa	2	0	-	-
	PPRXIb	1	1	+	NA
ZnF	DAG1a	2	0	-	-

Arabic number indicates numbers of RNA editing site in CDS/UTR; "-" indicates no structural effect, and "+" indicates that RNA editing altered the mRNA or protein structure. NA, not applicable; CDS, coding sequence; UTR, untranslated region.

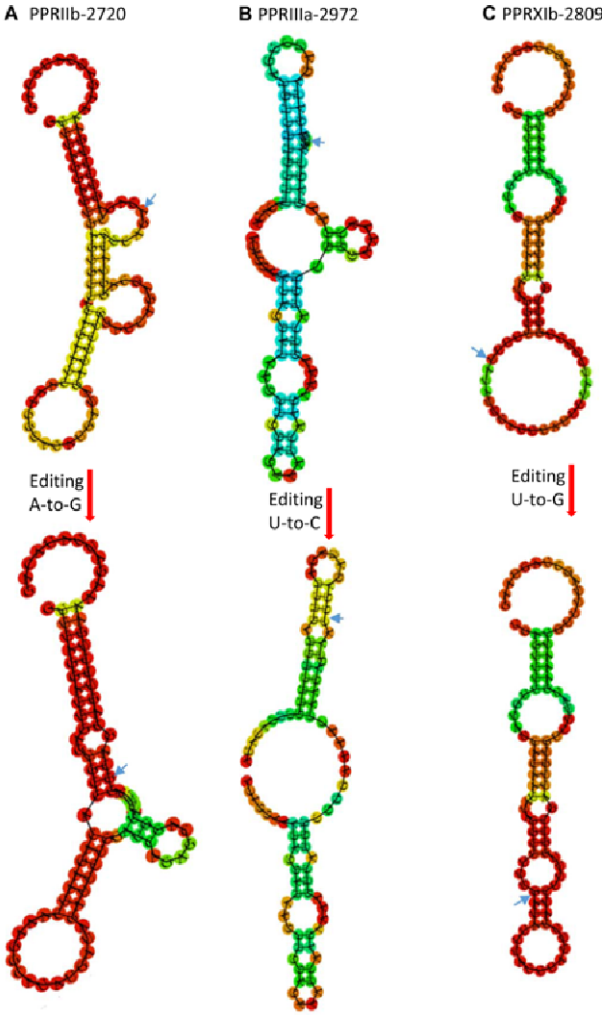

Fig. 3. Effect of RNA editing on the secondary structure of mRNA, (A) mRNA structures change due to A-to-G editing. (B) mRNA structures change due to U-to-C editing. (C) mRNA structures change due to U-to-G editing.

case of *PPRIIIa-2972* transcript, 'U' was located in the bulge loop before editing, whereas the altered nucleotide 'C' was located in the hairpin stem after editing (Fig. 3B). Similarly, in *PPRXIb-2809*, 'U' was located in the hairpin loop before editing, whereas the edited nucleotide 'G' was located in the hairpin stem after editing; additionally, a single large loop was divided into two small loops after editing (Fig. 3C). The tertiary structure of the protein encoded by the *PPRVa* gene was dramatically affected by RNA editing; Before RNA editing the protein consist of 4 domains (domain 1 to 4) but after RNA editing the protein consist of 5 domains (domain 1 to 5) (Fig. 4A, B; Fig. S5; Fig. S6). But the tertiary structure of protein of other genes are not affected by RNA editing (Fig. S4).

Nucleotide Preferences in the Editing Sites in Gene CDSs

In this study, we identified 13 RNA editing sites located in the CDS of *PPR* (11) and *ZnF* (2) gene families. The frequency of RNA editing at the first, second, and third nucleotide positions of the codon was 38%, 31%, and 31%, respectively (Fig. S3A). Among the RNA editing sites in which the first nucleotide position was altered, 50% of the sites contained 'A' (purine) at the second nucleotide position of the codon (Table 1). Sites in which the second nucleotide position of the codon was edited were of the configuration G_G (2), A_G, or A_C (Table 1; Fig 1E). The codon sequence of sites in which the third nucleotide position was edited was of the configuration GC_, AU_ (2), UU_, or AG_ (Table 1; Fig 1D). Most of the RNA editing sites were located in the gene CDS and resulted in amino acid changes. Of the 13 RNA editing sites observed in this study, 2 were silent; both of these occurred in (*PPRVa-2293* and *PPRVIIIa-1260*) and resulted in codon change but no amino acid change (Table 1).

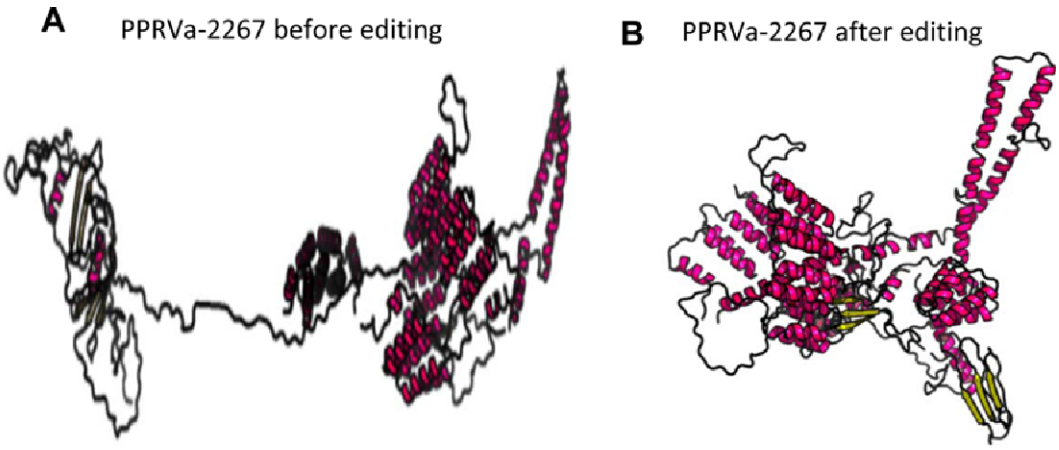

Fig. 4. Effect of RNA editing on the tertiary structure of the PPRVa protein. (A) Protein structure before editing. (B) Protein structure after G-to-A editing in *PPRVa-2267*.

Among a total of 21 RNA editing events, 11 resulted in amino acid alterations. RNA editing also affected the hydrophobicity/hydrophilicity of the encoded proteins; the proportion of hydrophobic amino acids increased from 38.46% before editing to 53.85% after editing (Fig. S3B).

Discussion

In this study, it is interesting that the RNA editing components could be edited. We identified nine types of RNA editing events, including all possible nucleotide substitutions, except C-to-A, C-to-G, and G-to-U, in Arabidopsis. Previously, 12 types of RNA editing events have been reported in *Salvia miltiorrhiza* and *Arabidopsis thaliana*; in Arabidopsis, these RNA editing events have been reported in nuclear transcripts using bioinformatics approaches (Meng et al. 2010; Wu et al. 2017). To our knowledge, this is the first experimental report of tissue- and development-specific RNA editing in Arabidopsis. Although the nine types of RNA editing events identified in this study were based on a small-scale screening of target genes (approximately 21%), these events were specific to seedling, leaf, and stem tissues of different developmental ages (Table 1). The tissue-specific characteristic of these RNA editing events implies that these were post-transcriptional modifications, not genomic mutations. We observed A-to-U editing events surrounding positions with a higher number of 'A' nucleotide. Previous research indicates that a *cis*-regulatory element neighboring the editing site is required for its recognition by the PPR-associated editing enzyme in plants (Shikanai 2006). Further studies are needed to better understand the processes involved in RNA editing, including the identification of *cis*-regulatory elements, isolation of editing enzymes, and validation of editing sites. We also identified RNA editing events in the UTR regions of genes. RNA editing in introns and UTR regions affects mRNA stability or splicing due to the modification of its secondary structure (Drescher et al. 2002; Jean-Claude Farré et al. 2012; Zeng et al. 2007; Wu et al. 2017). RNA-Seq analysis in Arabidopsis has shown that more than 61% of multi-exon genes generate alternative mRNA isoforms, which are expressed in a tissue-specific manner (Marquez et al. 2012). However, RNA editing at some sites has an adverse effect on plant growth, development, and fertility (Hammani and Giegé 2014). In the Arabidopsis nuclear transcripts *AT1G29930.1* and *AT1G52400.1*, C-to-U and U-to-C RNA editing has been reported at the translation borders (Meng et al. 2010). These deamination and amination reactions occur in adjacent sites; therefore, the deamination reaction is considered as the donor of amino group for the amination reaction, although the frequency of the amination reaction is higher than that of the deamination reaction (Meng et al. 2010). It is possible that another factor acts as the amino group donor for the amination reaction. Through bioinformatic approach it is found that U-to-C editing frequency is higher than C-to-U in nuclear transcript (Meng et al. 2010). In our experimental study, we found U-to-C editing frequency is higher as well (Fig. 2A). However, in chloroplast and mitochondrial RNA editing C-to-U editing is very dominant in most of the plants (Ichinose and Sugita, 2016). Additionally, A-to-I editing events were prevalent in the present study. Such editing events have also been reported in previous studies using bioinformatics approaches (Meng et al. 2010). However, plant homologs of cytidine deaminase or adenosine deaminase of mammals responsible for this editing event have not yet been discovered. It is stated that the adenosine deaminase family of enzymes is derived from the cytidine deaminase family, as the only difference between the two enzymes is the deaminase groove, which facilitates the deamination reaction (Goodman et al. 2012). Therefore A-to-I editing activity may be performed by the unique members of the PPR protein family which can be discovered by individual screening of the most probable

PPR protein candidates.

In this study, it is found the transversion editing is lower than the transitional editing. Out of 21 editing sites 13 (7 in coding region, 6 in noncoding region) are transition whereas transversion is 8 (7 in coding region, 1 in noncoding region) (Fig. 2C). However, in animal transition editing is also reported higher (Luo et al. 2016). Transversion editing is more complicated process than transition editing which might have adverse effect (Luo et al. 2016).

Codon bias is an important consideration for RNA editing, although targeting a nucleotide mainly depends on its neighboring nucleotides (Kuttan and Bass 2012). In this study, we found that the nucleotide in the third position in the codon showed higher editing frequency than those in the first and second positions. We observed that, for codons in which the third nucleotide was edited, the neighboring nucleotides were GC_, AU_ (2), UU_, and AG_. In contrast to our data, He and colleagues reported a 3-fold higher frequency of RNA editing at the second nucleotide position of the codon than at the first and third nucleotide positions (He et al. 2016). This discrepancy may be due to differences in the preferences of editing enzymes for the adjacent nucleotides; in the case of adenosine deaminase, editing is greatly affected by neighboring nucleotides.

Our data showed that different kinds of RNA editing events occur not only in chloroplast and mitochondrial transcripts but also in nuclear transcripts. These events were very frequent in the *PPR* than *ZnF* gene families but not in the *MORF/RIP*, *RRM*, *PP01*, and *GSDA* genes, indicating that RNA editing is gene-specific. The *PPR* genes were highly affected by RNA editing. Additionally, RNA editing events are more frequent in seedlings than in any other plant tissues. Seedlings play a very important role in plant physiology. Early chloroplast development 1 (ECD1), which belongs to the PLS subfamily of the PPR protein family, is responsible for early chloroplast development in seedlings (Jiang et al. 2018). In this study, we observed that single nucleotide conversion adversely affected mRNA secondary structure and protein domain structure in some case. Therefore, it can be stated that the RNA editing in the nuclear transcript of such important family genes might have big impact on the plant physiology. Further investigation is needed to determine the algorithm via which PPR proteins recognize and bind to their cognate mRNAs and to understand the precise effect of RNA editing events on the highly programmable PPR protein family and other protein families.

Materials and Methods

Growth Conditions and Sample Collection

Arabidopsis ecotype Columbia (Col-0) were sown in paper pots filled with a mixture of horticultural perlite, peat moss, and vermiculite in a 1:2:1 ratio. Pots were covered with plastic, stored in the dark for 3–4 days, and then transferred to a growth room in green farm U.ING (Made in Japan) maintained at 22°C and 45% relative humidity, and with a 16 h/8 h light/dark cycle. Plants were watered twice daily, in the morning and evening and fertilizer was applied twice a week. Pools of Arabidopsis plant was used for collection of samples. Different tissues were harvested at different developmental stages, including from 4-, 8-, and 12-d-old seedlings and 16-, 21-, 27-, and 32-d-old leaves, stipes, stems and roots.

Extraction of RNA and cDNA Synthesis

Total RNA was extracted from the Arabidopsis tissues using QIAGEN Plant Mini Kit (Hilden city, Germany; catalog no. 74904), according to the manufacturer's instructions, and treated with RQ1 RNase-free DNase (Promega, Madison, WI, USA) to remove genomic DNA contamination. The RNA samples were then purified using phenol-chloroform extraction and ethanol precipitation and quantified using a NanoDrop Spectrophotometer (Thermo Scientific). Purified RNA was subjected to cDNA synthesis using reverse transcriptase (Superscript III, Invitrogen) and oligo dT primers. The Arabidopsis *GAPDH* gene was used as a housekeeping gene and was amplified using the forward primer GTTGTCATCTCTGCCCCAAG and reverse primer TGCAACTAGCGTTGGAAACA.

Selection of Genes Involved in RNA Editing

To identify the best candidate genes involved in RNA editing, accession numbers of *PPR*, *ZnF*, *MORF/RIP*, *RRM* and other genes were obtained from the National Center for Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov/). These accession numbers were entered in the www.plantgdb.org/AtGDB/ database to determine the genomic map of each gene. The genomic map location of each gene was then used to identify alternatively spliced genes from the Arabidopsis TAIR10 whole genome reference assembly (https://www.arabidopsis.org/). The full-length genomic DNA, mRNA, and cDNA sequences of each gene obtained from TAIR10 were validated against the sequence information (gene, coding sequence [CDS], intron-exon boundaries, transcript, and amino acid) available at http://atgenie.org.

Primer Design

Primers were designed using Primer3 (bioinfo.ut.ee/primer3-0.4.0/primer3/) and verified using the NCBI Primer-BLAST tool. In the case of failure or inappropriate outcome with the first set of primers, new primer sets were designed. Primers were purchased from Eurofins (Japan) in TE buffer at a concentration of 50 pmol/μl in a salt-free condition. Each primer was diluted to a working concentration of 10 pmol/μl using TE buffer.

Sequencing of PCR Products

To identify RNA editing sites, PCR products were subjected to three times of direct sequencing on a 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Austin, TX, USA; catalog no. 4336935). Sequences of the reverse strand were reverse complemented using online software (http://www.bioinformatics.org/sms2/reference.html). All sequences were aligned with the Arabidopsis genome sequence using BLAST (Kent 2002).

Quantification of Different RNA Editing Events

RNA editing events were quantified from the sequencing results using the peak height ratio method. The peak height of the corresponding dual

peaks was calculated using ImageJ software. RNA editing (%) was quantified based on the maximum peak height, according to the following equation (Eggington et al. 2011): RNA editing (%) = [C height / (T height + C height)] × 100.

Prediction of mRNA and Protein Structures

The secondary structures of mRNAs were predicted using RNA fold version 2.1.6, with default parameters (Lorenz et al. 2011). Protein structures were predicted using RaptorX (http://raptorx.uchicago.edu/) and Swiss model (https://swissmodel.expasy.org/), using default parameters (Guex et al. 2009; Morten et al. 2012).

Acknowledgements

Part of this work was supported by Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (26670167, 17H02204 and 18K19288).

Author's Contributions

TT, UQ and MTAA designed research; UQ and MTAA performed research, analyzed the data and wrote the paper; TT, UQ and MTAA revised the manuscript. All authors have agreed on the contents of the manuscript and declare no conflict of interests.

Supporting Information

Fig. S1. Validation of tissue-specific usual (deaminase/aminase) types of RNA editing events in Arabidopsis via cDNA sequencing.
Fig. S2. Validation of tissue-specific unusual RNA editing events in Arabidopsis via cDNA sequencing.
Fig. S3. Schematic representation of RNA editing events in Arabidopsis.
Fig. S4. Effect of RNA editing on tertiary structure of protein *PPRVIa-134, PPRVIIa-1035, PPRIXa-2749, PPRXa-355, PPRXIb-2852, PPRIXa-2720, DAG1a-615, DAG1a-794, PPR1-250* (a-i).
Fig. S5. PPRVa-2267 before editing, yellow color indicate usual amino acid.
Fig. S6. PPRVa-2267 after editing, yellow color indicate edited amino acid.
Table S1. The accession number of our studied genes of RNA editing related family proteins.

References

Barkan A, Small I (2014) Pentatricopeptide repeat proteins in plants. Annu Rev Plant Biol 65:415–442

Chen TC, Liu YC, Wang X, Wu CH, Huang CH, Chang CC (2017) Whole plastid transcriptomes reveal abundant RNA editing sites and differential editing status in *Phalaenopsis aphrodite* subsp. *formosana*. Bot Stud 58:1–14

Cheng S, Gutmann B, Zhong X, Ye Y, Fisher MF, Bai F, Castleden I, Song Y, Song B, Huang J, Liu X, Xu X, Lim BL, Bond CS, Yiu SM, Small I (2016) Redefining the structural motifs that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants. Plant J 85:532–547

Delannoy E, Stanley WA, Bond CS, Small ID (2007) Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. Biochem Soc Trans 35:1643–1647

Drescher A, Hupfer H, Nickel C, Albertazzi F, Hohmann U, Herrmann R, Maier R (2002) C-to-U conversion in the intercistronic ndhI/ndhG RNA of plastids from monocot plants: Conventional editing in an unconventional small reading frame? Mol Genet Genomics 267:262–269

Eggington JM, Greene T, Bass BL (2011) Predicting sites of ADAR editing in double-stranded RNA. Nat Commun 2:1–9

Goodman RA, MacBeth MR, Beal PA (2012) ADAR Proteins: structure and catalytic mechanism. Curr Top Microbiol Immunol 353:1–33

Guillaumot D, Lopez-Obando M, Baudry K, Avon A, Rigaill G, Falcon de Longevialle A, Broche B, Takenaka M, Berthomé R, De Jaeger G, Delannoy E, Lurin C (2017) Two interacting PPR proteins are major Arabidopsis editing factors in plastid and mitochondria. Proc Natl Acad Sci USA 114:8877–8882

Hammani K, Giegé P (2014) RNA metabolism in plant mitochondria. Trends Plant Sci 19:380–389

He P, Huang S, Xiao G, Zhang Y, Yu J (2016) Abundant RNA editing sites of chloroplast protein-coding genes in Ginkgo biloba and an evolutionary pattern analysis. BMC Plant Biol 16:1–12

Ichinose M, Sugita M (2016) RNA editing and its molecular mechanism in plant organelles. Genes 8:1–15

Jean-Claude Farré, Cindy Aknin, Alejandro Araya BC (2012) RNA editing in mitochondrial for splicing Trans -Introns is required for splicing. PLoS One 7:1–10

Kent WJ (2002) BLAT -The BLAST -like alignment tool. Genome Res 12:656–664

Kuttan A, Bass BL (2012) Mechanistic insights into editing-site specificity of ADARs. Proc Natl Acad Sci USA 109:E3295–E3304

Luo GH, Li XH, Han ZJ, Zhang ZC, Yang Q, Guo HF, Fang JC (2016) Transition and transversion mutations are biased towards GC in transposons of *Chilo suppressalis*. Genes 7:1–12

Lurin C, Andrés C, Aubourg S, Bellaoui M, Bitton F, Bruyère C, Caboche M, Debast C, Gualberto J, Hoffmann B, Lecharny A, Le Ret M, Martin-Magniette ML, Mireau H, Peeters N, Renou JP, Szurek B, Taconnat L, Small I (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell 16:2089–2103

Meng Y, Chen D, Jin Y, Mao C, Wu P, Chen M (2010) RNA editing of nuclear transcripts in Arabidopsis thaliana. BMC Genomics 11:1–7

Rodrigues NF, Christoff AP, da Fonseca GC, Kulcheski FR, Margis R (2017) Unveiling chloroplast RNA editing events using next generation small RNA sequencing data. Front Plant Sci 8:1–13

Shikanai T (2015) RNA editing in plants: Machinery and flexibility of site recognition. Biochim Biophys Acta 1847:779–785

Takenaka M, Zehrmann A, Verbitskiy D, Härtel B, Brennicke A (2013) RNA Editing in Plants and Its Evolution. Annu Rev Genet 47:335–352

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 J 82:375–392

Wagoner JA, Sun T, Lin L, Hanson MR (2015) Cytidine deaminase motifs within the DYW domain of Two pentatricopeptide repeat-containing proteins are required for site-specific chloroplast RNA editing. J Biol Chem 290:2957–2968

Wu B, Chen H, Shao J, Zhang H, Wu K, Liu C (2017) Identification of symmetrical RNA editing events in the mitochondria of Salvia miltiorrhiza by strand-specific RNA sequencing. Sci Rep 7:1–11

Yagi Y, Nakamura T, Small I (2014) The potential for manipulating RNA with pentatricopeptide repeat proteins. Plant J 78:772–782

Zeng WH, Liao SC, Chang CC (2007) Identification of RNA editing sites in chloroplast transcripts of Phalaenopsis aphrodite and comparative analysis with those of other seed plants. Plant Cell Physiol 48:362–368