

Rice *PPS1* encodes a DYW motif-containing pentatricopeptide repeat protein required for five consecutive RNA-editing sites of *nad3* in mitochondria

Haijun Xiao¹, Qiannan Zhang¹, Xiaojian Qin², Yanghong Xu¹, Chenzi Ni¹, Jishuai Huang¹, Linlin Zhu³, Feiya Zhong¹, Wei Liu¹, Guoxin Yao⁴, Yingguo Zhu¹ and Jun Hu¹ 

¹State Key Laboratory of Hybrid Rice, Engineering Research Center for Plant Biotechnology and Germplasm Utilization of Ministry of Education, College of Life Sciences, Wuhan University, Wuhan, 430072 Hubei, China; ²Chongqing Key Laboratory of Molecular Biology of Plants Environmental Adaptations, College of Life Sciences, Chongqing Normal University, Chongqing, 401331 Chongqing, China; ³No.16 Middle School of Zhengzhou, Zheng Zhou 450000, Henan, China; ⁴School of Life and Science Technology, Hubei Engineering University, Xiaogan 432000, China

Author for correspondence:

Jun Hu

Tel: +86 13707149400

Email: junhu@whu.edu.cn

Received: 26 April 2018

Accepted: 9 June 2018

New Phytologist (2018) **220**: 878–892

doi: 10.1111/nph.15347

Key words: mitochondria, *nad3*, pentatricopeptide repeat, pollen partial sterility, rice, RNA editing.

Summary

- The pentatricopeptide repeat (PPR) protein family is a large family characterized by tandem arrays of a degenerate 35-amino-acid motif whose members function as important regulators of organelle gene expression at the post-transcriptional level. Despite the roles of PPRs in RNA editing in organelles, their editing activities and the underlying mechanism remain obscure.
- Here, we show that a novel DYW motif-containing PPR protein, *PPS1*, is associated with five conserved RNA-editing sites of *nad3* located in close proximity to each other in mitochondria, all of which involve conversion from proline to leucine in rice. Both *pps1* RNAi and heterozygous plants are characterized by delayed development and partial pollen sterility at vegetative stages and reproductive stage.
- RNA electrophoresis mobility shift assays (REMSAs) and reciprocal competition assays using different versions of *nad3* probes confirm that *PPS1* can bind to cis-elements near the five affected sites, which is distinct from the existing mode of PPR-RNA binding because of the continuity of the editing sites. Loss of editing at *nad3* in *pps1* reduces the activity of several complexes in the mitochondrial electron transport chain and affects mitochondrial morphology.
- Taken together, our results indicate that *PPS1* is required for specific editing sites in *nad3* in rice.

Introduction

The mitochondrion is a unique organelle that functions as a powerhouse that generates the ATP needed for cell maintenance and growth in eukaryotes (Newmeyer & Ferguson-Miller, 2003). More than 95% of the proteins required to maintain the function of mitochondria are encoded by the nuclear genome and imported to the mitochondrion, whereas only ~40 genes are encoded by the mitochondrion itself (Heazlewood *et al.*, 2004). In eukaryotic cells, RNA editing changes the genetic information from the DNA master (Knoop, 2011). RNA editing is widespread in mitochondria and chloroplasts, particularly in plants (Moreira *et al.*, 2016; Sun *et al.*, 2016). In this process, conserved codons are converted, or start or stop codons are created, serving as a correction mechanism for otherwise defective organellar transcripts (Stern *et al.*, 2010).

The pentatricopeptide repeat (PPR) protein family are a new group of nuclear-encoded organellar processing factors defined by degenerate motifs of 35 amino acids arranged with two to 30 tandem repeats (Small & Peeters, 2000; Lurin *et al.*, 2004). PPR proteins function as important regulators of organelle gene expression at the post-transcriptional level, including playing roles in RNA editing, splicing, cleavage, RNA stability, and translation (Hashimoto *et al.*, 2003; Kotera *et al.*, 2005; Chi *et al.*, 2008; Ichinose *et al.*, 2013; Haili *et al.*, 2016; Wu *et al.*, 2016). Some PPR proteins mediate both RNA editing and splicing; for example, OsPPR6, a rice plastid localized PPR protein, was reported to be involved in RNA editing and splicing (Tang *et al.*, 2017). The first RNA-editing factors identified in chloroplasts and mitochondria were CRR4 and MEF1, respectively (Kotera *et al.*, 2005; Zehrmann *et al.*, 2009). It is reasonable to expect that higher plants can manage RNA-editing events in organelles by incorporating PPR proteins, MORFs/RIPs and

other factors (Takenaka *et al.*, 2012). We recently reported that a rice dual-localized PPR protein, OsPGL1, is required for RNA editing at *ccmFc-543* and *ndbD-878* sites in mitochondria and chloroplasts, respectively (Xiao *et al.*, 2018).

The PPR proteins can be divided into two subfamilies: the P and PLS subfamilies. The P subfamily proteins contain tandem arrays of canonical PPR motifs (P motifs, 35 amino acids in length), whereas the PLS subfamilies are characterized by triplets of P, L (35–36 amino acids) and S (31 amino acids) motifs. The latter subfamily is further classified into three subgroups: E, E+, and DYW, according to the characteristics of the C-terminal motifs. P-type PPR proteins have been shown to participate primarily in various aspects of organellar RNA processing, whereas PLS PPR proteins are almost exclusively associated with C-to-U RNA editing (Barkan & Small, 2014). Recent crystal structures show that PPR motifs adopt an anti-parallel helix-turn-helix fold, and the mechanism of sequence-specific RNA recognition by PPR motifs has been revealed; studies have shown that combinations involving amino acid 6 of the front motif and amino acid 1 of the subsequent motif correlate strongly with the identity of the RNA base that is bound (Takenaka *et al.*, 2013; Yagi *et al.*, 2013; Yin *et al.*, 2013).

Here, we characterized a DYW motif-containing PPR protein, PPS1, which is responsible for five consecutive RNA-editing sites of *nad3* in mitochondria in rice. Interestingly, all conserved editing sites involve the conversion of proline to leucine across various species, suggesting a conserved function of NAD3 in complex I. We next demonstrated the RNA-binding activities of the PPR motifs of PPS1 directly. The results showed that the five consecutive sites are edited simultaneously, which is a significant finding and is important for RNA-editing application of the PPR protein in the future.

Materials and Methods

RNAi constructs and plant materials

A 251 bp fragment (ranging from 56 to 306 bp) of *LOC_Os12g36620* cDNA was cloned into the pH7GWIWG(II) vector to construct the RNAi vector. Calli derived from ZhongHua 11 (*Oryza sativa*, L. *Japonica*) were used for *Agrobacterium*-mediated transformation. Wild-type (WT) and independent multiple RNAi lines were grown in a paddy field and glasshouse in Wuhan, China, under proper management.

pps1 mutants

pps1 mutants were generated using the CRISPR/Cas9 system; two target sites located 20 bp upstream of the protospacer-adjacent motif sequence (PAM) were designed. The combined target sequence was introduced to the gRNA-U3 and gRNA-U6 vectors, followed by two rounds of nested PCR. The PCR products were subsequently introduced to the CRISPR/Cas9 vector. The construct was verified through PCR and sequencing. Calli derived from *O. sativa japonica* Zhonghua 11 (ZH11) were used for *Agrobacterium*-mediated

transformation. WT and CRISPR/Cas9 knockout lines were grown in a paddy field and glasshouse in Wuhan, China, under proper management.

Pollen grain staining and examination of pollen grains on the stigma

Pollen grains were stained with 1% I₂-KI solution and observed under an optical microscope. The spikelets of RNAi lines and WT were harvested after flowering and fixed in FAA fixation (18:1:1 (v/v) mixture of formalin, 70% ethanol, and acetic acid) for 24 h followed by dehydrating through an ethanol series (70%, 50%, and 30%) and washed with distilled water three times. Following incubation in 10 N sodium hydroxide for 5 min at 50°C, the spikelets were washed with distilled water three times and stained with 0.1% water blue (95290; Fluka, Shanghai, China) overnight. Finally, the samples were examined and counted using a fluorescence microscope (Leica DM4000B, Wetzlar, Germany). At least 100 pollen grains were counted per spikelet, and 10 spikelets were counted for each line.

Scanning electron microscopy and transmission electron microscopy assays

Both scanning electron microscopy (SEM) and transmission electron microscopy (TEM) assays were performed as described in our recent work (Xiao *et al.*, 2018).

RNA Extraction, RT-PCR and qRT-PCR

Total RNA was extracted with 1 ml Trizol reagent according to the manufacturer's instructions (Invitrogen) and treated with RNase-free DNase I (New England Biolabs, Beijing, China). Approximately 5 µg of total RNA were reverse-transcribed using random primers (Invitrogen). Reverse transcription polymerase chain reaction (RT-PCR) analyses were performed using the SuperScript III one step RT-PCR system according to the manufacturer's instructions (Invitrogen); the primers are listed in Supporting Information Table S1.

Subcellular localization of PPS1

For transient expression in rice protoplast, full-length cDNA of the PPS1 was cloned into HBT-sGFP vector driven by the cauliflower mosaic virus 35S promoter to construct the 35S:PPS1:sGFP fusion protein. Protoplast preparation and transformation procedures were carried out as previously described (Xiao *et al.*, 2018).

Moreover, PPS1-FLAG fusion protein driven by ubiquitin (UBI) promoter was constructed for *Agrobacterium*-mediated transformation in rice. Mitochondria and chloroplasts were extracted and purified from calli and 4-wk-old seedlings, respectively, using methods described previously (Hu *et al.*, 2012). Anti-FLAG, anti-isocitrate dehydrogenase (anti-IDH) and anti-RbcL antibodies (Agriseria, Vannas, Sweden) were used in this study.

Histochemical analysis of GUS activity

A 1208 bp upstream sequence of *PPSI* from ZH11 containing the promoter cassette was amplified (Table S1) and cloned into the binary vector pCAMBIA1391 to drive the GUS reporter gene expression. GUS histochemical staining with various tissues from transgenic plants was performed as described previously (Jefferson *et al.*, 1987).

RNA electrophoresis mobility shift assays

The corresponding cDNA fragments of *PPSI* were amplified with specific primers (Table S1) and cloned into the pGEX-6p-1 vector to generate recombinant GST-PPSI⁵⁶⁻⁹¹⁹ and GST-PPSI⁵⁶⁻⁷¹⁷. RNA probes were synthesized and labelled with biotin at the 3' end by GenScript (Nanjing, China). For RNA electrophoresis mobility shift assays (REMSAs), the recombinant protein was incubated with an RNA probe in a 20 µl reaction mixture including 10 µl of 2× binding buffer (100 mM Na phosphate (pH 7.5), 10 units RNasin, 0.1 mg ml⁻¹ BSA, 10 mM dithiothreitol, 2.5 mg ml⁻¹ heparin, and 300 mM NaCl). The mixture was incubated at 25°C for 30 min, followed by separation through 5% native polyacrylamide gel electrophoresis (PAGE) in 0.5× TBE buffer and then transfer to a nylon membrane (Roche). For the competitive REMSAs, a gradually increasing concentration of unlabelled probe was added to the reaction mixture following the procedure described earlier.

Analysis of mitochondrial RNA editing

For RNA editing analysis, total RNAs were isolated from the mature anthers using the Trizol reagent and were reverse-transcribed as mentioned earlier. Primers were designed to cover all 491 mitochondrial editing sites (Table S2). The RT-PCR products were sequenced directly. This analysis was performed with three biological replicates.

Immunoblot analysis

Mitochondria were isolated and purified from rice calli as described previously (Liu *et al.*, 2012). Protein was quantified with the BCA protein assay kit (Thermo Scientific, Rockford, IL, USA). Mitochondrial proteins (10 µg) were separated by sodium dodecyl sulphate-PAGE (SDS-PAGE), transferred onto a polyvinylidene difluoride (PVDF) membrane (Bio-Rad) and incubated with various primary antibodies against NAD3, NAD7, cytochrome c₁, Cox II, Atp-alpha, IDH (Beijing Protein Innovation, Beijing, China) and alternative oxidase 1/2 (AOX1/2) antibody (Agriser). Detection was carried out by the ECL Western Blotting Detection Reagents (Bio-Rad).

Blue-native (BN)-PAGE and evaluation of mitochondrial complex activity

The equivalent of 500 µg of total mitochondrial protein was treated and then loaded into a BN-PAGE gel according to a

previous protocol (Liu *et al.*, 2012). The gel was first stained with Coomassie brilliant blue (CBB), and the activities of several complexes were assessed as described previously (Sabar *et al.*, 2005). For immunoblotting, the proteins were transferred from the gel to PVDF membranes, and antibodies against NAD3, Cox II and ATP-alpha were used for detection.

To precisely examine the activity of the mitochondrial complexes, 10 µg of total mitochondrial protein was quantified with an activity test kit (Jiancheng Bioengineering Institute, Nanjing, Jiangsu, China). The activities of complex I and complex IV and the ATP content were measured via a colorimetric assay according to the manufacturer's instructions.

Accession numbers

Sequence for rice *PPSI* and the related mitochondrial gene *nad3* can be found in the GenBank database under accession numbers XP_015618345 and YP_002000566, respectively.

Results

Phenotypic and genetic characterization of *pps1*

To identify the function of PPRs, we generated several RNAi lines corresponding to tens of PPRs using an RNA interference (RNAi) strategy in the ZhongHua 11 (ZH11, *O. sativa Japonica*) background. Some lines exhibited obvious phenotypic differences from the WT. Among the multiple independent RNAi lines, we checked the relative RNA expression level of *Os12g36620* to evaluate knockdown efficiency (Fig. S1). These multiple independent RNAi lines exhibited pleiotropic phenotypes, including growing slowly, dwarfing and delayed development in vegetative stages (Fig. 1a), smaller and shorter anthers (Fig. 1b), sterile pollen (Fig. 1c), lower germination on the stigma (Fig. 1d), shorter panicles and lower seed-setting rates (Fig. 1e). Therefore, we selected the progenies of T0–6 for subsequent experiments.

Because of the reduced seed-setting rate of the transgenic lines, we next checked the viability of the stigma and pollen after pollination. A total of 10 stigmas were assessed, and the results indicated that all stigmas exhibited normal pollination in the WT and *pps1*-RNAi plants, whereas the germination rate of the pollen was quite distinct (Table S3). The pollen germination rate on WT stigmas was *c.* 90%, whereas only *c.* 24% of the pollen was pollinated on *pps1*-RNAi stigmas. These data confirmed that the sterility of the *pps1*-RNAi plants resulted from nonfunctioning pollen, whereas the female gametes were normal.

To evaluate the morphological details of the anthers and pollen, we used SEM and TEM. The SEM results showed that the anthers were tightly packed in the WT (Fig. 2a), whereas the anthers of *pps1*-RNAi plants were curly and much shorter than those of the WT (Fig. 2g). Furthermore, reticulate anther cuticles were clearly observed on the surface of the WT (Fig. 2b), but not in *pps1*-RNAi lines (Fig. 2h). The anther cuticle is a hydrophobic layer coating the outermost surface of the anthers and plays a protective role in the development of microspores. The SEM results also showed that the pollen grains of the WT were complete,

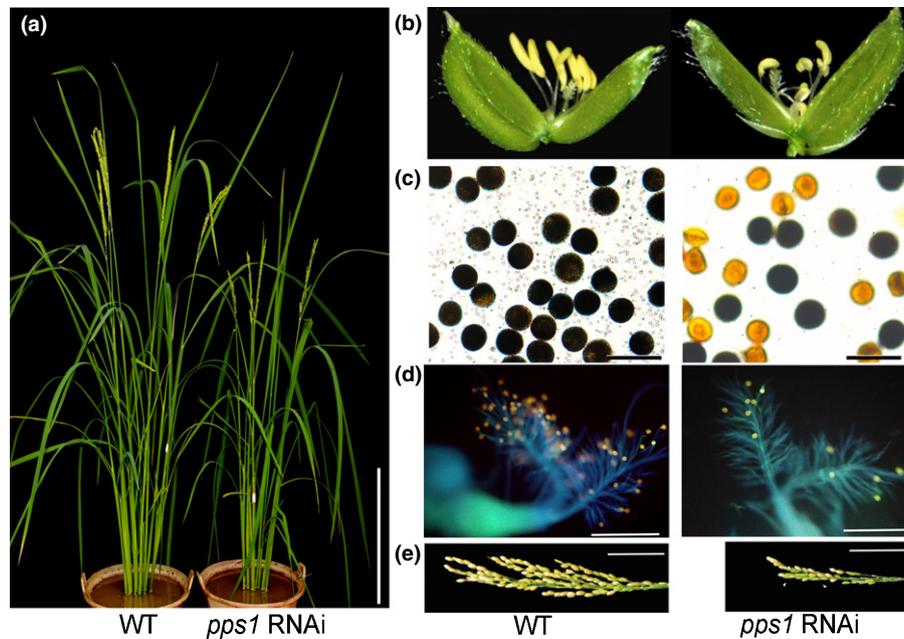


Fig. 1 Phenotypic characterization of the rice *pps1*-RNAi line. (a) The *pps1*-RNAi line (right) is slightly dwarfed relative to the wild-type (WT) plant (left) at the heading stage. (b) Comparison of a WT (left) and a *pps1*-RNAi line (right) spikelet. (c) Comparison of pollen fertility between the WT (left) and *pps1*-RNAi line (right) by 1% I₂-KI staining. Darkly stained pollen is fertile and lightly stained pollen is sterile. (d) Comparison of pollen grain germination on stigma between WT (left) and *pps1*-RNAi line (right). (e) Comparison of panicle at the harvest stage between the WT (left) and *pps1*-RNAi line (right). The spikelet fertility of the *pps1*-RNAi line is c. 52%. Arrowheads indicate the sterile spikelets. Bars: (a) 15 cm; (b) 0.5 cm; (c) 300 μ m (d) 1 mm; (e) 5 cm.

spheroidal and plump (Fig. 2c), and the germination pore was clear (Fig. 2d). However, half the pollen grains of *pps1*-RNAi plants exhibited shrunken features and were incomplete and spheroidal, and the germination pore was hard to observe (Fig. 2i,j). These results are consistent with the decreased germination rate observed on stigmas. Furthermore, the TEM assay showed that starch granules accumulated in mature pollen (Fig. 2e). However, the number of starch granules was significantly reduced in *pps1*-RNAi plants compared with that in the WT (Fig. 2k), suggesting that the pollen was sterile. Detailed observations of the pollen walls showed intact exines and intines of mature pollen in the WT (Fig. 2f), whereas the structure of the sterile pollen of the *pps1*-RNAi lines was severely impaired (Fig. 2i).

PPS1 encodes a DYW motif-containing mitochondrial protein

The results of PFAM analysis (<http://pfam.xfam.org/>) showed that PPS1 consisted of 19 PPR motifs (seven S motifs, six P motifs and six L motifs in a triple motif arrangement, P-L-S), and each PPR motif was linked by two antiparallel alpha helices (helix A and helix B) (Fig. S2). Moreover, the C-terminal region between residues 717 and 919 harboured the consensus sequences of the extension domains (E, E+, and DYW domains) (Fig. S3), indicating that PPS1 belongs to the DYW PLS subfamily. The alignment of rice PPS1 with its orthologues in various plants showed 78%, 59%, 52%, 56% and 52% similarity with *Zea mays* (GRMZM2G056996), *Nicotiana tabacum* (XP_016440080), *Brassica napus* (XP_013723128), *Glycine max* (XP_006574752)

and *Arabidopsis thaliana* (NP_193101), respectively (Fig. S3). The high similarity between PPS1 and GRMZM2G056996 also suggested that their function may be conserved in monocots.

Numerous reports have shown that PPR proteins are targeted to either plastids or mitochondria. In this study, PPS1 was also predicted to localize to mitochondria by TARGETP (<http://www.cbs.dtu.dk/services/TargetP/>). To confirm the subcellular localization of PPS1, the full-length cDNA of PPS1 was fused with green fluorescent protein (sGFP), driven by the 35S promoter, and the fusion protein was transiently expressed in rice protoplasts. The results showed that the PPS1-sGFP signals completely overlapped with the mitochondrial signals, indicating that PPS1 localizes to mitochondria (Fig. 3a). Moreover, we fused PPS1 with a FLAG tag for transgenic analysis to verify its subcellular localization. Mitochondria and chloroplasts were obtained from transgenic plants. Immunoblot analysis with an anti-FLAG antibody was then performed to detect the PPS1-FLAG fusion protein in the different extracts. A specific signal from the mitochondrial fraction confirmed that PPS1 was exclusively targeted to mitochondria (Fig. 3b). Taken together, these data indicated that *PPS1* encodes a DYW motif-containing PPR protein and is exclusively targeted to mitochondria.

Spatial and temporal expression pattern analysis of *PPS1*

To dissect the expression pattern of *PPS1*, we first used RT-PCR and quantitative RT-PCR (qRT-PCR) to analyse various developmental stages in WT plants. The results showed that *PPS1* was constitutively expressed in both vegetative and reproductive tissues, including roots, stems, leaves, anthers and inflorescences

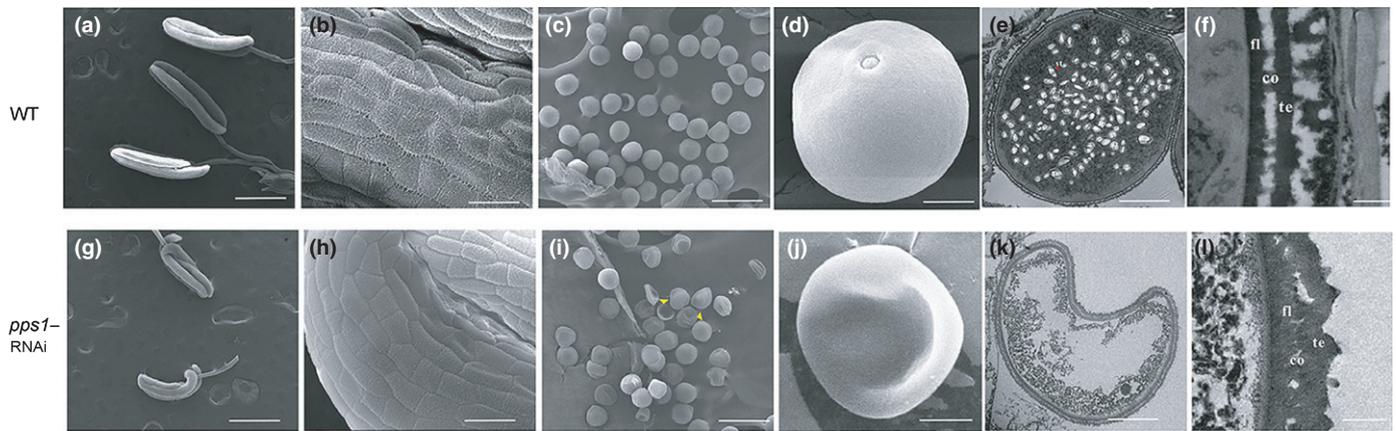


Fig. 2 Scanning electron microscopy (SEM) and transmission electron microscopy (TEM) examination of wild-type (WT) and *pps1*-RNAi anther and pollen. (a) SEM image of mature WT anther. (b) A higher-magnification SEM image of mature WT anther surface. (c) SEM image of mature WT pollen grains. (d) A higher-magnification SEM image of a mature WT pollen grain from (c). (e) TEM image of a mature pollen inner structure showing a large number of starch granules in a WT pollen grain (arrowhead indicates normal starch granules). (f) A higher-magnification TEM image of the normal-shaped pollen grain from (e). fl, foot layer; c, columella; te, tectum. (g) SEM image of mature *pps1*-RNAi anther. (h) A higher-magnification SEM image of mature *pps1*-RNAi anther surface. (i) SEM image of mature *pps1*-RNAi pollen grains (arrowheads indicate abnormal pollen grains). (j) A higher-magnification SEM image of an abnormal-shaped *pps1*-RNAi pollen grain from (i). (k) TEM image showing that there is little accumulation of starch granules in the abnormal-shaped *pps1*-RNAi pollen grain. (l) A higher-magnification TEM image of the normal-shaped pollen grain from (k) showing abnormal intine and exine. fl, foot layer; co, columella; te, tectum. Bars: (a, g) 1 mm; (b, h) 30 μ m; (c, i) 100 μ m; (d, e, j, k) 10 μ m; (f, l) 1 μ m.

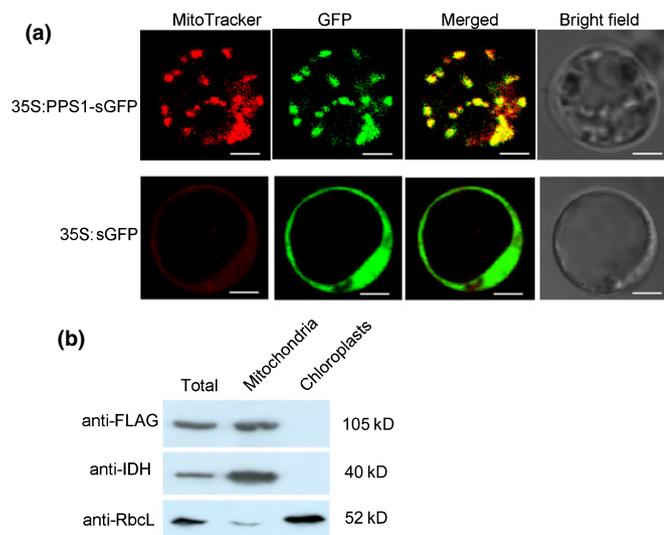


Fig. 3 Subcellular localization of PPS1. (a) Transient expression of 35S: PPS1-sGFP (top) and 35S:sGFP (bottom) in rice protoplast. MitoTracker Red was used as a mitochondria indicator. Bars, 5 μ m. (b) Total proteins, mitochondrial proteins and chloroplast proteins of transgenic plants were isolated for immunoblot assays. The large subunits of ribulose-1,5-bisphosphate carboxylase/oxygenase (RbcL) and isocitrate dehydrogenase (IDH) were used as controls for the chloroplast and mitochondrial fractions, respectively.

(Fig. 4a,b). Because the expression level in floral organs was relatively high, we further examined the expression profile of *PPS1* at various panicle development stages in detail, based on a previous study (Nonomura *et al.*, 2004). During the reproductive phase, the panicles were classified into five groups according to their development: panicles of *c.* 5 mm in length (P1); *c.* 10 mm (P2); *c.* 30 mm (P3); *c.* 60 mm (P4); and *c.* 100 mm (P5). P1 panicles have generally developed a flower primordia; P2 and P3 panicles

have developed anthers but have not entered meiosis; P4 panicles exhibit some meiocytes entering meiosis; and in P5, all meiocytes are entering meiosis. The anthers subsequently entered the post-meiotic (PM) and pollen maturation (PMA) stages. Interestingly, the expression level of *PPS1* decreased gradually from stages P1 to P5, whereas *PPS1* transcripts progressively accumulated and peaked at the stage of pollen maturation (Fig. 4c). Furthermore, we compared the expression levels of *PPS1* isolated from mature anthers between the WT and *pps1* RNAi lines to assess knock-down efficiency. The results showed that the expression level of *PPS1* was decreased to nearly half of that in the WT (Fig. 4d).

To further elucidate the spatial and temporal expression patterns of *PPS1*, we generated transgenic plants expressing β -glucuronidase (GUS) driven by the native *PPS1* promoter. GUS signals were observed in various tissues, and extremely strong signals were observed in the anthers (Fig. 4e), consistent with the results of RT-PCR and qRT-PCR. These data suggested that *PPS1* is ubiquitously expressed and accumulates to particularly high levels in the anthers. Taken together, these data supported the phenotype of the *pps1*-RNAi lines, which exhibited pollen sterility and pleiotropic effects on vegetative development.

PPS1 is involved in C-to-U RNA editing of *nad3* transcripts

Guided by the previous hypothesis that most PPR proteins of the DYW subfamily play a role in RNA editing (Salone *et al.*, 2007), we checked all 491 known editing sites in mitochondrial RNAs to dissect the function of PPS1 in rice using PCR sequencing (the primers used in these assays are listed in Table S2). The sequencing results for the RT-PCR products revealed a significantly decreased C-to-U editing efficiency at five editing sites of *nad3* transcripts in *pps1*-RNAi lines compared with that in the WT. These five sites, *nad3*-155, *nad3*-172, *nad3*-173, *nad3*-190 and

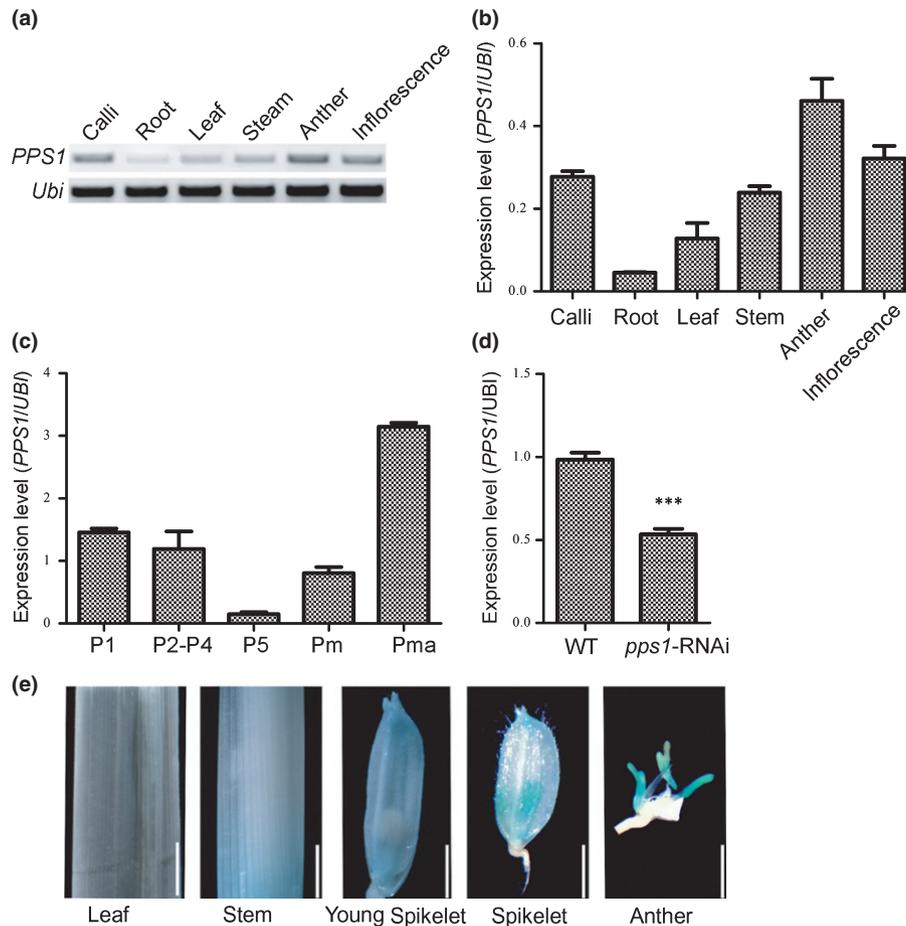


Fig. 4 Expression analysis of *PPS1*. (a) Reverse transcription polymerase chain reaction (RT-PCR) analysis of the *PPS1* in various wild-type (WT) tissues. (b) Real-time quantitative RT-PCR analysis of the *PPS1* in various WT tissues. (c) Real-time quantitative RT-PCR analysis of the *PPS1* at various panicle development stages in WT. (d) Real-time quantitative RT-PCR analysis of the *PPS1* in mature anthers between WT and *pps1*-RNAi. Data are means \pm SD from three independent biological replicates. (Student's *t*-test: ***, $P < 0.001$). (e) Histochemical staining analysis of *PPS1* promoter-GUS reporter gene in various tissues. Bars: (leaf and stem) 5 mm; (young spikelet and spikelet) 4 mm; (anther) 2 mm.

nad3-191, were edited at an efficiency of 73%, 75%, 80%, 79% and 83%, respectively, in the WT, but the efficiency was only 12%, 10%, 15%, 8% and 10% in *pps1*-RNAi plants (Fig. 5a; Table S4). We also checked the RNA-editing efficiency of *nad3* in the other independent RNAi lines, and the data showed the same decreased RNA-editing efficiency (Fig. S4). The failure of editing at the five sites resulted in the failure of amino acid codon conversion from CCG to CUG at NAD-52 and CCA to UUA at NAD3-58 and NAD3-64. Interestingly, all editing sites changed the corresponding amino acid from proline to leucine. Proline is a heterocyclic amino acid and could disrupt the secondary structure of NAD3 when it is assembled into complex I.

To determine whether *PPS1* affects the RNA stability of *nad3*, *nad3* transcripts were quantified in the WT and *pps1*-RNAi lines. The expression level of the *nad3* transcript in the *pps1*-RNAi lines was indistinguishable from that of the WT, suggesting that *PPS1* does not alter *nad3* transcript abundance (Fig. S5). These results indicated that *PPS1* is specifically involved in the RNA editing, rather than the RNA stabilization, of target RNA.

To investigate the possible evolutionary conservation of the alteration from proline to leucine at the 52nd, 58th and 64th

positions of the NAD3 protein sequence, the five corresponding editing sites in genomic DNA (gDNA) and cDNAs from five other species (*Z. mays*, *G. max*, *B. napus*, *N. tabacum*, and *A. thaliana*) were analysed (Fig. 5b–f). The results revealed various C-to-U editing events at these editing sites in the five tested species. The three corresponding amino acids were leucine as a result of natural RNA editing only in *Z. mays* (Fig. 5b). A ‘T’ nucleotide was already present at most of these RNA-editing sites in the genomic DNA: *nad3*-44 (corresponding to *nad3*-155 in *O. sativa*) in *G. max* and *B. napus* (Fig. 5c,d); and *nad3*-61, 62 (corresponding to *nad3*-172,173 in *O. sativa*) in *G. max*. Few of these sites are edited simultaneously, and *PPS1* is the first identified RNA-editing factor for five consecutive sites in a transcript.

Alternatively, the RNA secondary structure of the *nad3* transcript may vary as a result of changes in the editing sites, thus affecting the function of its transcript. Therefore, we analysed the RNA secondary structure of *nad3* in relation to the edited and unedited states of the five affected sites. The data showed that whether the *nad3*-155 site was edited or unedited, it did not affect the *nad3* secondary structure. So we speculated that the *nad3*-155 site was less important than other sites (Fig. S6a).

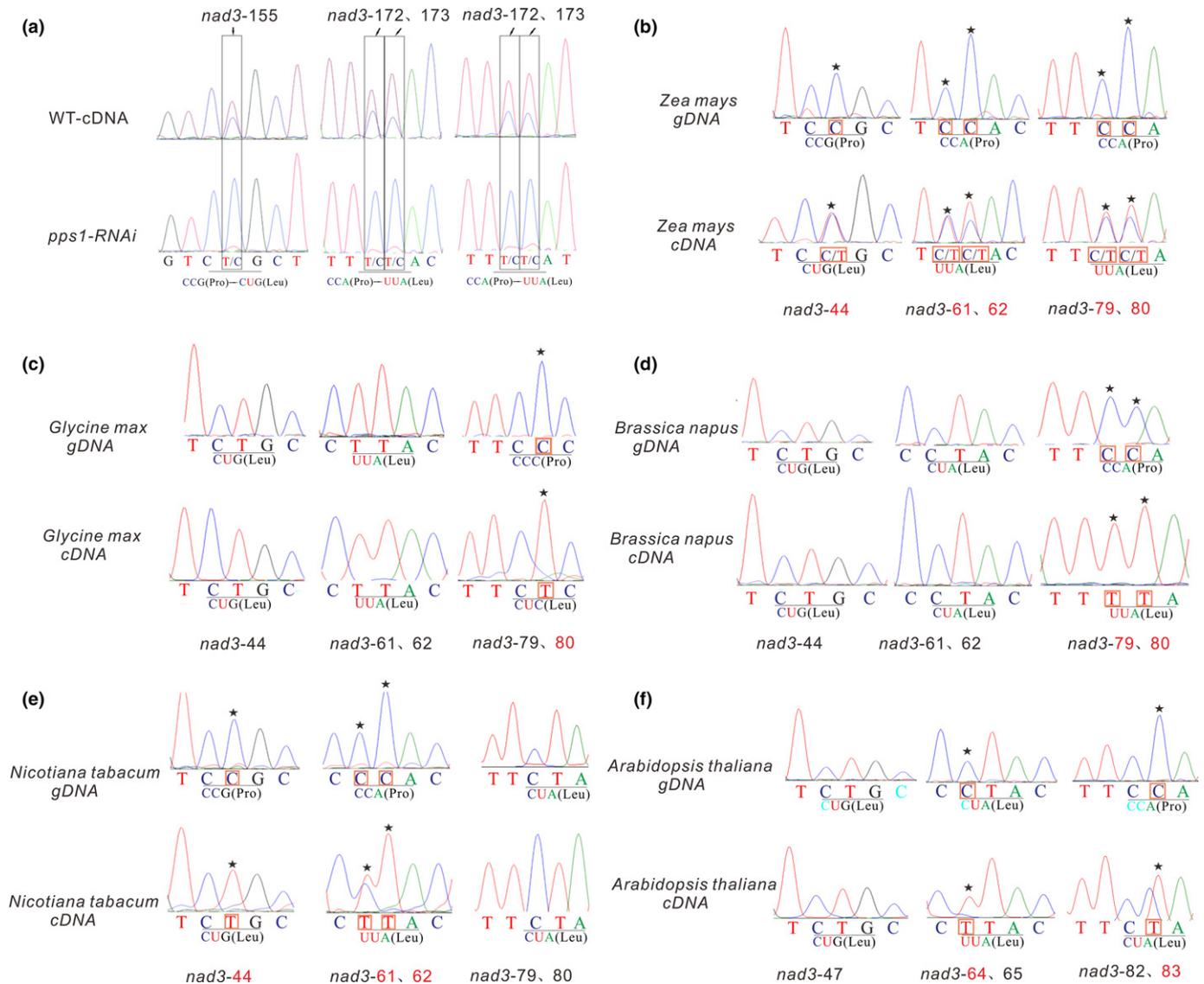


Fig. 5 Conserved RNA editing of *NAD3* in plants. (a) RNA-editing analysis of the *nad3*-155, 172, 173, 190, 191 sites in wild-type (WT) and *pps1* plants of mature anther, respectively. (b) RNA-editing analysis of the *nad3*-44, 61, 62, 79, 80 sites in *Zea mays* leaves. gDNA represents genome DNA sequence. (c) RNA-editing analysis of the *nad3*-44, 61, 62, 79, 80 sites in *Glycine max* leaves. gDNA represents genome DNA sequence. (d) RNA-editing analysis of the *nad3*-44, 61, 62, 79, 80 sites in *Brassica napus* leaves. gDNA represents genome DNA sequence. (e) RNA-editing analysis of the *nad3*-44, 61, 62, 79, 80 sites in *Nicotiana tabacum* leaves. gDNA represents genome DNA sequence. (f) RNA-editing analysis of the *nad3*-47, 64, 65, 82, 83 sites in *Arabidopsis thaliana* leaves. gDNA represents genome DNA sequence. Stars indicate the unedited 'C' in genome DNA (gDNA) and the edited 'T' in cDNA.

Interestingly, the structure was distinguishable downstream of the *nad3*-155 site for the pairs of *nad3*-172,173 and *nad3*-190,191 (Fig. S6b,c), and the structure of *nad3* was not changed when both of the consecutive sites were either edited or unedited. However, when one of the two sites was not edited, the *nad3* secondary structure was changed. Specifically, when cytidine was edited to uridine at the *nad3*-172 site while its neighbouring site, *nad3*-173, was not edited, the two sites were translocated from the loop to the stem (Fig. S6b). Analogously, when cytidine was edited to uridine at the *nad3*-191 site, while its neighbouring site, *nad3*-190, was not edited, the two sites were translocated from the stem to the loop (Fig. S6c). These data suggest that *PPS1* is a novel PPR gene that regulates the RNA editing of several consecutive sites.

Evolutionary conservation of mitochondrial protein *NAD3* and the nuclear *PPS1* proteins

To further explore the evolution of *PPS1* in plants, the full-length sequence of *PPS1* was used as a query to obtain its homologues, all of which were used to construct a phylogenetic tree. We screened the available plant sequence data and focused on species with highly similar protein data, which ultimately resulted in a widely phylogenetically distributed set comprising 68 angiosperm taxa (Fig. S7). The data showed that 14 monocots and 53 dicots were clustered together, suggesting that *PPS1* evolved from ancient plants and is conserved in monocots. Considering that the phylogenetic distribution of the DYW motif is strictly correlated with the presence of RNA editing (Salone *et al.*,

2007), we further analysed the *nad3* editing status of these 68 species; only 32 of the 68 species had available mitochondrial genome information. The results showed that the RNA-editing events *nad3*U155PL, *nad3*U172, 173PL and *nad3*U190, 191PL (corresponding to *nad3*U44PL, *nad3*U61, 62PL and *nad3*U79, 80PL, respectively, in other species) were conserved in five species, *Beta vulgaris*, *Nelumbo nucifera*, *O. sativa*, *Z. mays* and *Amborella trichopoda*. Interestingly, four of the five corresponding nucleotides in the genomic sequences of three species (*Daucus carota*, *Physcomitrella patens* and *Funaria hygrometrica*) were already ‘T’, directly encoding leucine at the corresponding sites (Fig. S7). Moreover, at least one of the five nucleotides existed in four species: *G. max*, *N. tabacum*, *A. thaliana* and *B. napus*. Conversely, loss of editing events at one of the other four editing sites occurred in 22 species, which resulted in the replacement of leucine with proline or serine (*Ziziphus jujuba*) with respect to the corresponding amino acids. The results suggested that leucine residues in the mitochondrial NAD3 protein are conserved at these corresponding editing sites and indicated that the presence of the nuclear protein PPS1 is essentially consistent with the requirements of RNA editing of the *nad3* transcript in mitochondria.

PPS1 binds directly to *nad3* transcripts

Based on the RNA-binding activities of PPR proteins, we proposed that PPS1 binds directly to the *nad3* transcript. To confirm the binding activity of PPS1, the N terminus of PPS1 (residues 1–55) was removed and fused with a glutathione S-transferase (GST) tag for expression in *Escherichia coli*. The recombinant protein (GST-PPS1^{56–919}) was analysed by western blotting using anti-GST antibodies to confirm its purity (Fig. S8a,b). The recombinant protein was further dialysed to remove contamination using RNase for REMSAs. Three *nad3* RNA probes of 35 nucleotides (from –26 bp upstream to +8 bp downstream of the *nad3*-155, 172, and 190 editing sites) were designed according to a previous study (Toda *et al.*, 2012), which were designated probe 1, probe 2 and probe 3 (Fig. 6a). Probe C (*nad3*-1580), which has been reported as a specific target of another rice PPR protein, MPR25, was used as a negative control (Fig. 6a). GST-PPS1^{56–919} and the GST tag were incubated with the biotin-labelled RNA probes. The protein–RNA complex was detected as a shifted band that migrated more slowly than the free RNA probe in the native gel; however, no shifted band was observed following incubation with GST (Fig. 6b). In addition, no shifted band was observed when probe C was incubated with GST-PPS1^{56–919} (Fig. 6b). We then performed competitor assays using unlabelled RNA probes. The binding intensity decreased as the competitor concentration increased, and the labelled RNA probes were free to undergo binding in the presence of a competitor at concentrations 10-fold greater than those of the probes (Fig. 6c–e).

To gain an insight into the molecular mechanism of PPS1 RNA binding, we dissected the protein and expressed only the PPR motifs (from 56 to 717 amino acids) in *Escherichia coli* (Fig. S8a,c). The recombinant protein PPS1^{56–717} was next quantified and dialysed for REMSAs. GST-PPS1^{56–717} and GST tag

were incubated with the biotin-labelled RNA probes. The protein–RNA complex was also detected as a shifted band in the native gel, but the shifted band was not observed when incubated with GST (Fig. 6f–i). In addition, the shifted band was not observed when probe C was incubated with GST-PPS1^{56–717}. At the same time, the binding intensity decreased in the presence of a competitor at concentrations 10-fold greater than those of the probes (Fig. 6f–i). Probe C was also used as a competitor to compete with three target RNAs for two recombinant PPS1 proteins. Data showed no combination between PPS1 and probe C (Fig. S9).

Moreover, to check whether one probe could displace another in the binding experiments among these three probes, we used the three RNA templates reciprocally in competition assays. The signal of PPS1 and the labelled RNA probe complex was reduced in the presence of another nonlabelled RNA probe, which indicated that these probes can be bound by PPS1 (Fig. S10). These data definitively confirmed that PPS1 can bind directly and specifically to the *nad3* transcript based on the PPR motifs.

Deficient mitochondrial electron transport chain complex activities

Given that the amino acid of NAD3^{52nd}, NAD3^{58th} and NAD3^{64th} was converted from proline to leucine in the WT, we further checked the stability of NAD3 and other mitochondrial proteins in the *pps1*-RNAi lines. Surprisingly, the NAD3 signal was much weaker in *pps1*-RNAi plants than in the WT, suggesting that defective editing in the *pps1*-RNAi lines resulted in instability of the NAD3 protein (Fig. 7a). Furthermore, NAD7, CytC1, COXII, and ATP-alpha of complex I, complex III, complex IV and complex V, respectively, were also investigated, and IDH served as an internal control (Fig. 7a). The results showed that these proteins were also reduced in *pps1*-RNAi plants, except for NAD7 and CytC1, indicating decreased activities of these complexes. Generally, a mature defence mechanism opposing blockage of the electron transfer chain is to increase in the alternative respiratory pathway characterized by AOX. Therefore, we investigated the accumulation of AOX1/2 in *pps1*-RNAi plants (Fig. 7a). The results demonstrated strong induction in the *pps1*-RNAi lines, suggesting that the AOX pathway is activated when mitochondrial electron transport chain (ETC) complexes are impaired.

To obtain an insight into the underlying mechanism, we further examined the activities of the ETC complexes. Mitochondrial complexes isolated from calli were separated in BN gels and subjected to staining to detect the activity of complex I, complex IV and complex V. CBB staining showed that several complexes displayed a distinguishable reduction in *pps1*-RNAi plants, including super complex I+III₂, and complex I, IV and V (Fig. 7b). Further assays for the detection of in-gel enzyme activities also revealed decreases in the NADH dehydrogenase activity of complex I, the cytochrome c oxidase activity of complex IV and the ATP synthase activity of complex V in *pps1*-RNAi lines (Fig. 7c–e).

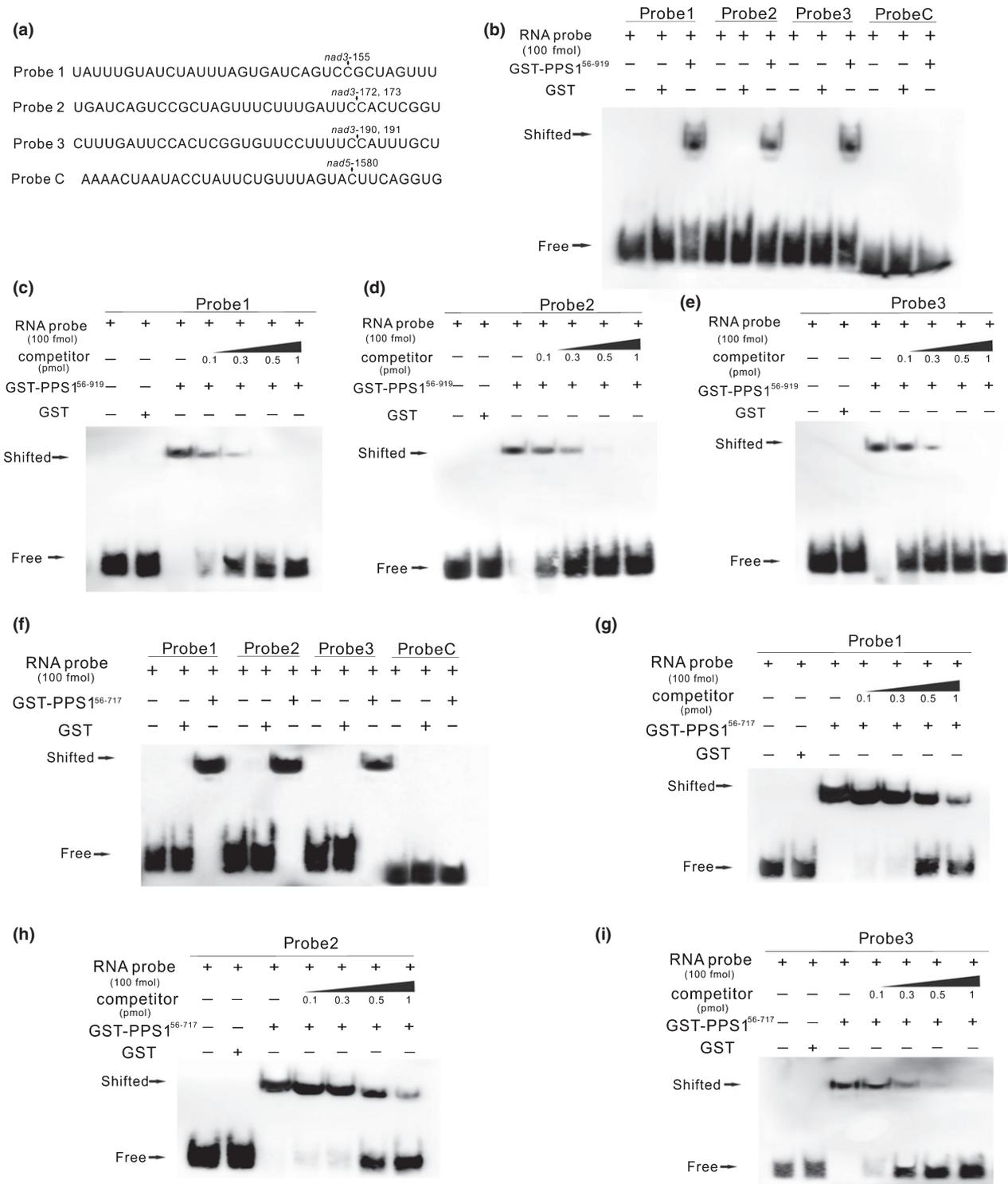


Fig. 6 PPS1 possesses RNA-binding activities with its pentatricopeptide repeat (PPR) motifs. (a) Schematic sequences of RNA probes. Edited sites are indicated and marked in red. Probe C is negative control in this study. (b) RNA electrophoresis mobility shift assays (REMSAs) of GST-PPS1⁵⁶⁻⁹¹⁹ and glutathione S transferase (GST) tag with various RNA probes. GST tag and probe C were used as negative controls. (c) Nonlabelled probe 1 was used as a competitor at a range of concentrations for competitive REMSA with recombinant GST-PPS1⁵⁶⁻⁹¹⁹. The GST tag was a negative control. (d) Nonlabelled probe 2 was used as a competitor at a range of concentrations for competitive REMSA with recombinant GST-PPS1⁵⁶⁻⁹¹⁹. The GST tag was a negative control. (e) Nonlabelled probe 3 was used as a competitor at a range of concentrations for competitive REMSA with recombinant GST-PPS1⁵⁶⁻⁹¹⁹. The GST tag was a negative control. (f) REMSA of GST-PPS1⁵⁶⁻⁷¹⁷ and GST tag with various RNA probes. GST-tag and probe C were used as negative controls. (g) Nonlabelled probe 1 was used as a competitor at a range of concentrations for competitive REMSA with recombinant GST-PPS1⁵⁶⁻⁷¹⁷. The GST tag was a negative control. (h) Nonlabelled probe 2 was used as a competitor at a range of concentrations for competitive REMSA with recombinant GST-PPS1⁵⁶⁻⁷¹⁷. The GST tag was a negative control. (i) Nonlabelled probe 3 was used as a competitor at a range of concentrations for competitive REMSA with recombinant GST-PPS1⁵⁶⁻⁷¹⁷. The GST tag was a negative control.

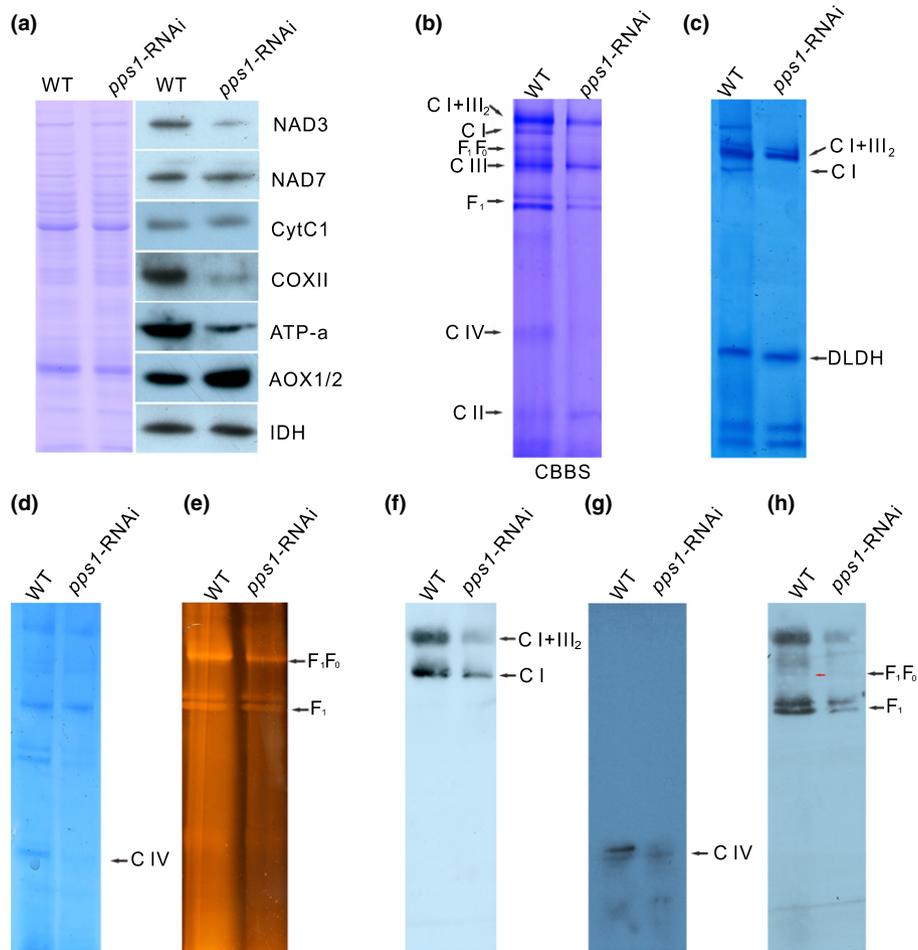


Fig. 7 Immunoblot detection of mitochondrial proteins. (a) Coomassie brilliant blue (CBB) staining (left) showing equal loading of the wild-type (WT) and *pps1*-RNAi total mitochondrial proteins separated by 10% sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Immunoblot detection of NAD3 and NAD7 of complex I, CytC₁ of complex III, COX II of complex IV, Atp-a of complex IV, AOX1/2 with specific antibodies. Isocitrate dehydrogenase (IDH) used as a mitochondrial reference antibody. (b) CBB staining (left) showing equal loading of the WT and *pps1*-RNAi total mitochondrial proteins separated by blue native PAGE (BN-PAGE). Complexes are indicated by arrows. (c) NADH dehydrogenase activity of complex I in the presence of NADH and nitroblue tetrazolium (NBT). Dihydropyridine dehydrogenase (DLDH) is used as a loading control. Positions of the complexes are indicated by arrows. (d) Succinate dehydrogenase and cytochrome c oxidase activity of complex IV detection with its own substrate, NBT. Positions of the complexes are indicated by arrows. (e) ATP content determination with ATP and inorganic phosphate Pb(NO₃)₂. Positions of the complexes are indicated by arrows. (f) Anti-NAD3 immunoblot to detect complex I by BN-PAGE. (g) Anti-COXII immunoblot to detect complex IV by BN-PAGE. (h) Anti-ATP-a immunoblot to detect complex V by BN-PAGE.

Furthermore, protein gel blotting with anti-NAD3, anti-ATP-alpha and anti-COXII antibodies was performed to investigate the number of corresponding complexes. A similar tendency to that described as above occurred, suggesting that reduced enzyme activity is accompanied by deficient mitochondrial respiratory complex abundance (Fig. 7f–h).

Furthermore, mitochondria were measured in solution to quantify the activity of complex I, complex IV and complex V. The results also showed that the activities of these three complexes were significantly lower in *pps1*-RNAi plants than in the WT (Fig. S11a–c). These results indicated that aberrant NAD3 derived from unedited *nad3* could be unstable and cause deficiency of complex I, resulting in a lower electron delivery efficiency. Therefore, the activities of the mitochondrial respiratory complexes are affected by other unknown retrograde signals.

Altered mitochondrial morphology in *pps1*-RNAi plants

To determine whether the reduced complex activities observed in *pps1*-RNAi also compromise mitochondrial morphology, we investigated the mitochondrial ultrastructure via TEM. In the WT, sausage-shaped mitochondria with normal compact cristae were clearly observed, whereas most of the mitochondria in the *pps1*-RNAi lines displayed a spherical morphology (Fig. 8a). Chloroplast ultrastructure was also assessed in this study. No obvious structural difference between WT and *pps1*-RNAi plants was observed (Fig. 8b). These results suggested that the deficiency of mitochondrial complex activities may only compromise mitochondrial morphology, leading to pleiotropic phenotypes and pollen sterility in *pps1*-RNAi plants.

pps1 knockout plants generated using the CRISPR/Cas9 system exhibit similar problematic results.

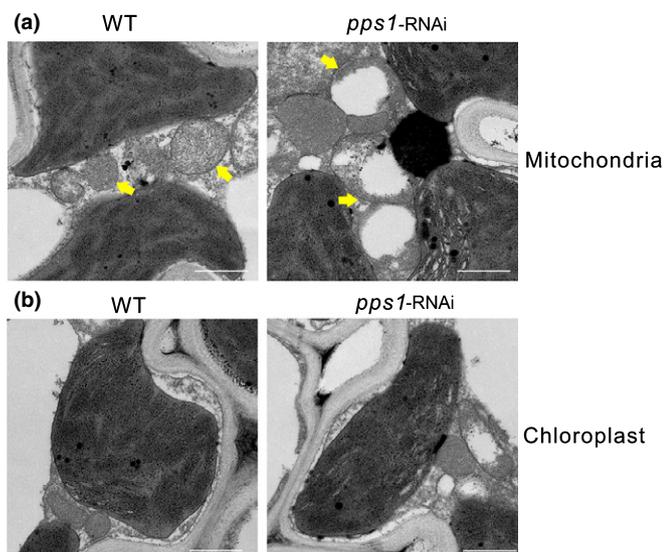


Fig. 8 Organelle ultrastructure in wild-type (WT) and *pps1*-RNAi leaves. (a) Mitochondrial ultrastructure observation by TEM of mesophyll cells from WT (left) and *pps1*-RNAi (right). Mitochondria are indicated with arrows. (b) Chloroplast ultrastructure observation by transmission electron micrographs of mesophyll cells from WT (left) and *pps1*-RNAi (right). Bars, 1 μ m.

To further verify whether the defects were indeed caused by the *PPS1* gene, the CRISPR/Cas9 system was used to generate *pps1* mutants. Two different regions in the first exon of *PPS1* were selected (Fig. S12a). No homozygous plants were obtained; we propose that the pollen sterility may belong to gametophytic sterility, and the male gametes cannot be transmitted normally. Two independent heterozygous plants were used for further analysis; *pps1*^{+/+} #1, a single-base insertion at target site 1, and *pps1*^{+/+} #2, a four-base deletion at target site 2 (Fig. S12b). Both plants exhibited similar pleiotropic phenotypes to the RNAi lines, including slow growth, dwarfing and delayed development during the vegetative stage (Fig. 9a), sterile pollen (Fig. 9b), lower germination on stigmas (Fig. 9c), and low seed-setting rates (Fig. 9d). *pps1*^{+/+} #1 was used for dissecting the morphologic details of pollen using SEM and TEM. The results showed that the pollen grains of the WT were completely spheroidal and plump; the germination pore was clear; and starch granules were fully accumulated in mature pollen. However, in the *pps1* plant, half of the pollen grains exhibited shrunken features; the germination pores were difficult to observe; and the number of starch granules was significantly reduced (Fig. 9e,f).

The editing efficiency at the five editing sites of the *nad3* transcript was also investigated in both knockout plants. The *nad3*-155, *nad3*-172, *nad3*-173, *nad3*-190 and *nad3*-191 sites were edited at efficiencies of 73%, 75%, 80%, 79% and 83%, respectively, in the WT, but the editing efficiency decreased to *c.* 0% in both knock-out plants, with the exception of a 10% efficiency at *nad3*-155 and a 15% efficiency at both *nad3*-172 and *nad3*-173 in *pps1*^{+/+} #2 (Fig. S13).

The activities of the three complexes (I, IV, V) in the ETC were further validated in the *pps1* knockout plants. The data

showed that these complexes (I, I+III₂, IV and V) in the *pps1* knockout plants displayed distinct reductions (Fig. S14a–d). Taken together, these data suggest that the defects, including phenotypes, loss of editing efficiency and the activities of the three complexes in the ETC, were indeed caused by the loss of function of *PPS1*.

Discussion

The conserved leucine is essential for the function of complex I

Although > 30 edited genes in mitochondria or chloroplasts have been reported, few reports have focused on the editing factors of *nad3* transcripts. In *A. thaliana*, MEF22 is involved in the editing of site 149 of *nad3*, at which serine is converted to phenylalanine (Takenaka *et al.*, 2010). SLG1 is associated with a single editing site, *nad3*-250, in *A. thaliana* and controls conversion from CCT to TCT, resulting in a change from proline to serine (Yuan & Liu, 2012). PpPPR₅₆ is involved in RNA editing at the *nad3*-230 and *nad4*-272 sites in moss (Ohtani *et al.*, 2010). Despite the availability of a few reports on rice PPR proteins, PPS1 is the first reported editing factor for *nad3* in rice.

We recently reported a dual-localized PPR protein, OsPGL1, that is required for both chloroplast RNA editing of *ndhD*-878 and mitochondrial RNA editing of *ccmF*c-543 in rice (Xiao *et al.*, 2018). Despite the synonymous editing of *ccmF*c-543, the loss of *ndhD*-878 editing causes failure of serine to leucine conversion. Conversion from proline to leucine is a common phenomenon in RNA editing (Kim *et al.*, 2009; Liu *et al.*, 2013; Li *et al.*, 2014; Yap *et al.*, 2015). One interesting finding regarding these editing sites is that they all contribute to converting proline to leucine, suggesting that the function of leucine in NAD3 could be evolutionarily conserved in plants. The highly conserved leucine of NAD3 may also be related to the structure and function of this protein in plants. Given that PPS1 shares 80% similarity with GRMZM2G056996 (*Z. mays*), the coincident editing patterns of *nad3* in rice and maize suggest a conserved function in monocots. These cases suggest that leucine plays a crucial role in mitochondrial function and that conversion from proline to leucine is critical for maintaining the structure or function of NAD3.

The PPR protein family has been recognized as a large family for almost two decades and it is well known for its essential functions in post-transcriptional processes, including RNA editing, RNA splicing, RNA cleavage, RNA stability and translation within mitochondria and chloroplasts (Schmitz-Linneweber & Small, 2008; Barkan & Small, 2014). In this study, we further demonstrated that PPS1 is involved in the RNA editing of *nad3*, rather than other post-transcriptional processes. The results of the evaluation of proteins suggested that abnormal NAD3 either cannot be correctly loaded into complex I or impairs the activity of complex I when it is loaded, which activates complicated retrograde signals in cells. Consequently, the accumulation of NAD3 in *pps1* plants is significantly decreased, and other components of the ETC in mitochondria are also impaired.

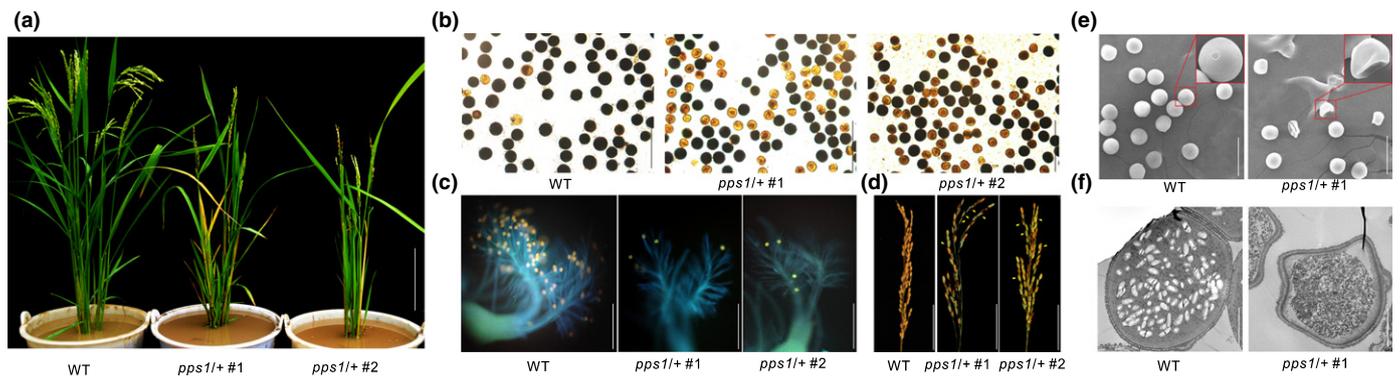


Fig. 9 Phenotypic characterization of the *pps1* knockout plants. (a) The two *pps1* knockout plants (right) are slightly dwarfed relative to the wild-type (WT) plant (left) at the heading stage. (b) Comparison of pollen fertility of a WT (left) and two *pps1* knockout plants (right) by 1% I₂-KI staining. Darkly stained pollen is fertile and lightly stained pollen is sterile. (c) Comparison of pollen grain germination on stigma between a WT (left) and two knockout plants (right). (d) Comparison of panicle of a WT (left) and two knockout plants (right) at the harvest stage. Arrowheads indicate the sterile spikelets. (e) Scanning electron microscopy (SEM) image of pollen grains of a mature WT (left) and *pps1* knockout plants (right). Red rectangle shows a higher-magnification image of a single mature pollen grain's inner structure in WT (left) and *pps1* knockout plants (right), Bars: (a) 15 cm; (b) 300 μ m; (c) 1 mm; (d) 5 cm; (e) 100 μ m; (f) 10 μ m.

PPS1 regulates consecutive editing sites of *nad3* transcripts

As organellar RNA-editing factors, most PPR proteins are associated with one site in RNA editing. Few scenarios exist in which one PPR protein acts on multiple editing sites of the same RNA transcript simultaneously. Several continuous pairs of RNA-editing sites have been reported in Arabidopsis. SLO2 is an RNA-editing factor associated with two continuous sites, *mttB*-144 and *mttB*-145 (Zhu *et al.*, 2012). In the rice *ogr1* mutant, among its seven RNA-editing sites, *nad4*-401, 416 and 433 are consecutively distributed in the *nad4* transcript (Kim *et al.*, 2009). Nevertheless, RNA-binding activity between these editing factors and the cis-elements near the editing sites has not been detected, and the recognition of these nearby sites and PPRs remains unclear. In *Physcomitrella patens*, two RNA-editing sites, *ccmFc*-103 and *ccmFc*-122, are separated by 18 nucleotides. Two studies have indicated that PpPPR₇₁ specifically binds to *ccmFc* RNA containing the *ccmFc*-103 editing site and is required for *ccmFc*-122 RNA editing (Tasaki *et al.*, 2010). Additionally, PpPPR₆₅ was confirmed to be required for RNA editing of both sites but binds specifically only to the upstream target site, *ccmFc*-103 (Schallenberg-Rudinger *et al.*, 2013).

In our study, we solidly confirmed the binding capacity between all editing sites and recombinant PPS1 proteins, suggesting that RNA editing of five *nad3* sites occurs simultaneously. However, there does not seem to be any biological or phenotypic significance caused by the five or one RNA-editing sites. We think consecutive is important to further study, which is significant for designed RNA editing on gene therapy. We also speculate that the defect of *nad3* RNA editing is unlikely to be a unique reason for the abnormal phenotype. Therefore, the PPR proteins should have another function to induce the phenotype. Another interesting finding is that the interval spaces between *nad3*-155 and *nad3*-172,173 and between *nad3*-172,173 and *nad3*-190,191 are both 16 nucleotides. To our knowledge, few studies have shown that an RNA-editing factor can directly

possess several consecutive RNA-editing sites in a transcript. PPS1 is required for five consecutive editing sites in *nad3*, suggesting a complex mechanism of RNA binding and deamination.

Prediction of the modular recognition mechanism of target *nad3* editing sites for PPS1

Recently, some research groups have decoded the recognition mechanism of the binding of PPR proteins to RNAs (Barkan *et al.*, 2012; Takenaka *et al.*, 2013; Yagi *et al.*, 2013; Shen *et al.*, 2016). Bioinformatics and structural analyses have indicated that three amino acid positions distributed in two adjacent PPR repeats are of great importance in recognizing the target RNA base (Takenaka *et al.*, 2013). Position 6 in the first helix motif is predicted to be more important for distinguishing between purines (A and G) and pyrimidines (U and C), whereas position 1 in the second helix motif directly recognizes amino (A and C) or keto (G and U) groups (Yagi *et al.*, 2013). Position 3 is less important and imposes an additional constraint on the binding nucleotides (Takenaka *et al.*, 2013). However, this mode is suitable only for P and S motifs and is supported by experimental determination between a P-subfamily PPR protein and its target RNA (Yin *et al.*, 2013). Recently, researchers proposed that PPR proteins generally bind to their target nucleotides from the -4 nucleotide position of the target to the 5' RNA region over up to as many nucleotides as the number of repeats in the related PPR protein (Barkan *et al.*, 2012).

To evaluate the matching of PPS1 with its five target sites, we performed computational prediction based on recent reports. PPS1 has 19 PPR repeats, and the putative binding nucleotide sequences could be located in the -22 to -4 region of the sequence upstream each of the targets sites. Alignment of the target sites of PPS1 showed that these editing site-containing regions exhibit comparatively high PPR-nucleotide matches among the motifs of PPS1. Notably, because two pairs of the five *nad3* editing sites are distributed consecutively, we propose that

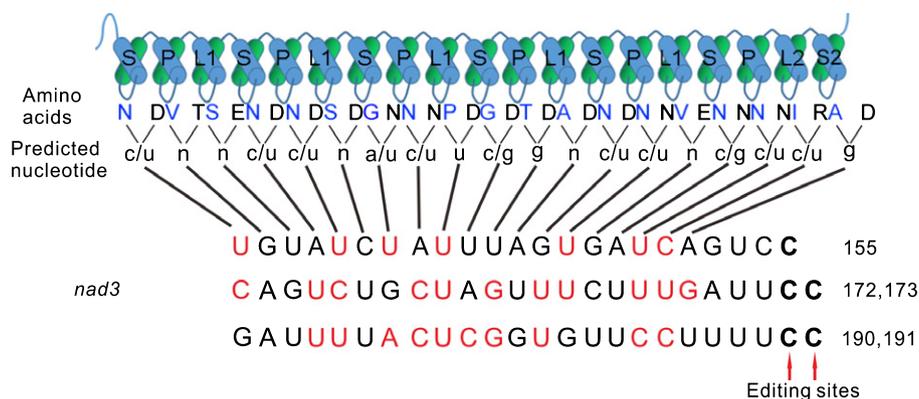


Fig. 10 Alignment of pentatricopeptide repeat (PPR) motifs of PPS1 with the cis-elements. The editing sites are indicated by the arrows. The blue amino acids are at the 6 position in the PPR motif, while black amino acids are at the 1 position of next PPR motif. The nucleotide identical to the predicted nucleotide recognized by PPS1 is in red. 'n' in the predicted nucleotide means no specific nucleotide for recognition.

these five sites may be located in three cis-elements (Fig. 10). Based on the continuous character of these editing sites, we further analysed the matching scores (predicted matches/all matches) between the PPR motifs and cis-elements near each site. The results showed that the scores for the *nad3*-172 and *nad3*-190 sites were relatively higher than those for the other three sites (Fig. S15), suggesting that each site makes a different contribution to the PPR recognition of the corresponding cis-element, which also supports the conclusion that the five editing sites may indeed be distributed in three cis-elements. An artificial PPR scaffold was generated to illustrate a structurally robust and a programmed recognition mechanism for sequence-specific RNA binding between 23,916 PPR sequences and the predefined RNA targets (Coquille *et al.*, 2014). The results suggested that each PPR motif shows varying nucleotide-binding intensities, even for the same amino acid combinations at positions 6 and 1', suggesting limitations of the current prediction tool, which weighs all possible contact sites equally, and further suggesting that combinations between PPR motifs and targeted nucleotides are very flexible. Very few PPR proteins regulate more than five RNA-editing sites. Here, PPS1 is reported, for the first time, to possess five consecutive RNA-editing sites, which further implies the complex mechanism of RNA recognition, an important area for future detailed research.

Acknowledgements

We thank Yaoguang Liu laboratory for providing us with pYLCRISPR/Cas9 vectors. We declare there is no conflict of interest. This work was supported by funds from the National Key Research and Development Program of China (2016YFD0100804) and the National Natural Science Foundation of China (31371698 and 31670310), and the Suzhou Science and Technology Project (SNG2017061).

Author contributions

H.X., J. Hu. and Y.Z. designed the study. H.X., Q.Z. and X.Q. contributed to the construction of the *pps1*-RNAi lines and *pps1*

knockout plants. H.X. and Q.Z. carried out most experiments, H.X., Q.Z. and C.N. conducted SEM and TEM. H.X., F.Z., W.L. and C.N. performed expression and purification of recombinant proteins, H.X., L.Z. performed the subcellular localization of PPS1. J. Huang., Y.X. and G.Y. contributed to field management. H.X. and J. Hu. wrote the manuscript with feedback from all authors.

ORCID

Jun Hu  <http://orcid.org/0000-0001-9107-8118>

References

- Barkan A, Rojas M, Fujii S, Yap A, Chong YS, Bond CS, Small I. 2012. A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genetics* 8: e1002910.
- Barkan A, Small I. 2014. Pentatricopeptide repeat proteins in plants. *Annual Review of Plant Biology* 65: 415–442.
- Chi W, Ma J, Zhang D, Guo J, Chen F, Lu C, Zhang L. 2008. The pentatricopeptide repeat protein DELAYED GREENING1 is involved in the regulation of early chloroplast development and chloroplast gene expression in *Arabidopsis*. *Plant Physiology* 147: 573–584.
- Coquille S, Filipovska A, Chia T, Rajappa L, Lingford JP, Razif MF, Thore S, Rackham O. 2014. An artificial PPR scaffold for programmable RNA recognition. *Nature Communications* 5: 5729.
- Haili N, Planchard N, Arnal N, Quadrado M, Vrielynck N, Dahan J, des Francs-Small CC, Mireau H. 2016. The MTL1 pentatricopeptide repeat protein is required for both translation and splicing of the mitochondrial NADH DEHYDROGENASE SUBUNIT7 mRNA in *Arabidopsis*. *Plant Physiology* 170: 354–366.
- Hashimoto M, Endo T, Peltier G, Tasaka M, Shikanai T. 2003. A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast *ndhB* in *Arabidopsis*. *Plant Journal* 36: 541–549.
- Heazlewood JL, Tonti-Filippini JS, Gout AM, Day DA, Whelan J, Millar AH. 2004. Experimental analysis of the *Arabidopsis* mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. *Plant Cell* 16: 241–256.
- Hu J, Wang K, Huang W, Liu G, Gao Y, Wang J, Huang Q, Ji Y, Qin X, Wan L *et al.* 2012. The rice pentatricopeptide repeat protein RF5 restores fertility in Hong-Lian cytoplasmic male-sterile lines via a complex with the glycine-rich protein GRP162. *Plant Cell* 24: 109–122.

- Ichinose M, Sugita C, Yagi Y, Nakamura T, Sugita M. 2013. Two DYW subclass PPR proteins are involved in RNA editing of *ccmFc* and *atp9* transcripts in the moss *Physcomitrella patens*: first complete set of PPR editing factors in plant mitochondria. *Plant and Cell Physiology* 54: 1907–1916.
- Jefferson RA, Kavanagh TA, Bevan MW. 1987. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO Journal* 6: 3901–3907.
- Kim SR, Yang JI, Moon S, Ryu CH, An K, Kim KM, Yim J, An G. 2009. Rice OGR1 encodes a pentatricopeptide repeat-DYW protein and is essential for RNA editing in mitochondria. *Plant Journal* 59: 738–749.
- Knoop V. 2011. When you can't trust the DNA: RNA editing changes transcript sequences. *Cellular and Molecular Life Sciences* 68: 567–586.
- Kotera E, Tasaka M, Shikanai T. 2005. A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature* 433: 326–330.
- Li XJ, Zhang YF, Hou M, Sun F, Shen Y, Xiu ZH, Wang X, Chen ZL, Sun SS, Small I *et al.* 2014. Small kernel 1 encodes a pentatricopeptide repeat protein required for mitochondrial *nad7* transcript editing and seed development in maize (*Zea mays*) and rice (*Oryza sativa*). *Plant Journal* 79: 797–809.
- Liu G, Tian H, Huang YQ, Hu J, Ji YX, Li SQ, Feng YQ, Guo L, Zhu YG. 2012. Alterations of mitochondrial protein assembly and jasmonic acid biosynthesis pathway in Honglian (HL)-type cytoplasmic male sterility rice. *Journal of Biological Chemistry* 287: 40051–40060.
- Liu YJ, Xiu ZH, Meeley R, Tan BC. 2013. Empty pericarp5 encodes a pentatricopeptide repeat protein that is required for mitochondrial RNA editing and seed development in maize. *Plant Cell* 25: 868–883.
- Lurin C, Andres C, Aubourg S, Bellaoui M, Bitton F, Bruyere C, Caboche M, Debast C, Gualberto J, Hoffmann B *et al.* 2004. Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* 16: 2089–2103.
- Moreira S, Valach M, Aoulad-Aissa M, Otto C, Burger G. 2016. Novel modes of RNA editing in mitochondria. *Nucleic Acids Research* 44: 4907–4919.
- Newmeyer DD, Ferguson-Miller S. 2003. Mitochondria: releasing power for life and unleashing the machineries of death. *Cell* 112: 481–490.
- Nonomura K, Nakano M, Fukuda T, Eiguchi M, Miyao A, Hirochika H, Kurata N. 2004. The novel gene HOMOLOGOUS PAIRING ABERRATION IN RICE MEIOSIS1 of rice encodes a putative coiled-coil protein required for homologous chromosome pairing in meiosis. *Plant Cell* 16: 1008–1020.
- Ohtani S, Ichinose M, Tasaki E, Aoki Y, Komura Y, Sugita M. 2010. Targeted gene disruption identifies three PPR-DYW proteins involved in RNA editing for five editing sites of the moss mitochondrial transcripts. *Plant and Cell Physiology* 51: 1942–1949.
- Sabar M, Balk J, Leaver CJ. 2005. Histochemical staining and quantification of plant mitochondrial respiratory chain complexes using blue-native polyacrylamide gel electrophoresis. *Plant Journal* 44: 893–901.
- Salone V, Rudinger M, Polsakiewicz M, Hoffmann B, Groth-Malonek M, Szurek B, Small I, Knoop V, Lurin C. 2007. A hypothesis on the identification of the editing enzyme in plant organelles. *FEBS Letters* 581: 4132–4138.
- Schallenberg-Rudinger M, Kindgren P, Zehrmann A, Small I, Knoop V. 2013. A DYW-protein knockout in *Physcomitrella* affects two closely spaced mitochondrial editing sites and causes a severe developmental phenotype. *Plant Journal* 76: 420–432.
- Schmitz-Linneweber C, Small I. 2008. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends in Plant Science* 13: 663–670.
- Shen C, Zhang D, Guan Z, Liu Y, Yang Z, Yang Y, Wang X, Wang Q, Zhang Q, Fan S *et al.* 2016. Structural basis for specific single-stranded RNA recognition by designer pentatricopeptide repeat proteins. *Nature Communications* 7: 11285.
- Small ID, Peeters N. 2000. The PPR motif – a TPR-related motif prevalent in plant organellar proteins. *Trends in Biochemical Sciences* 25: 46–47.
- Stern DB, Goldschmidt-Clermont M, Hanson MR. 2010. Chloroplast RNA metabolism. *Annual Review of Plant Biology* 61: 125–155.
- Sun T, Bentolila S, Hanson MR. 2016. The unexpected diversity of plant organelle RNA editosomes. *Trends in Plant Science* 21: 962–973.
- Takenaka M, Verbitskiy D, Zehrmann A, Brennicke A. 2010. Reverse genetic screening identifies five E-class PPR proteins involved in RNA editing in mitochondria of *Arabidopsis thaliana*. *Journal of Biological Chemistry* 285: 27122–27129.
- Takenaka M, Zehrmann A, Brennicke A, Graichen K. 2013. Improved computational target site prediction for pentatricopeptide repeat RNA editing factors. *PLoS ONE* 8: e65343.
- Takenaka M, Zehrmann A, Verbitskiy D, Kugelmann M, Hartel B, Brennicke A. 2012. Multiple organellar RNA editing factor (MOREF) family proteins are required for RNA editing in mitochondria and plastids of plants. *Proceedings of the National Academy of Sciences, USA* 109: 5104–5109.
- Tang J, Zhang W, Wen K, Chen G, Sun J, Tian Y, Tang W, Yu J, An H, Wu T *et al.* 2017. OsPPR6, a pentatricopeptide repeat protein involved in editing and splicing chloroplast RNA, is required for chloroplast biogenesis in rice. *Plant Molecular Biology* 95: 345–357.
- Tasaki E, Hattori M, Sugita M. 2010. The moss pentatricopeptide repeat protein with a DYW domain is responsible for RNA editing of mitochondrial *ccmFc* transcript. *Plant Journal* 62: 560–570.
- Toda T, Fujii S, Noguchi K, Kazama T, Toriyama K. 2012. Rice MPR25 encodes a pentatricopeptide repeat protein and is essential for RNA editing of *nad5* transcripts in mitochondria. *Plant Journal* 72: 450–460.
- Wu W, Liu S, Ruwe H, Zhang D, Melonek J, Zhu Y, Hu X, Gusewski S, Yin P, Small ID *et al.* 2016. SOT1, a pentatricopeptide repeat protein with a small MutS-related domain, is required for correct processing of plastid 23S-4.5S rRNA precursors in *Arabidopsis thaliana*. *Plant Journal* 85: 607–621.
- Xiao H, Xu Y, Ni C, Zhang Q, Zhong F, Huang J, Liu W, Peng L, Zhu Y, Hu J. 2018. A rice dual-localized pentatricopeptide repeat protein is involved in organellar RNA editing together with OsMORFs. *Journal of Experimental Botany* 69: 2923–2936.
- Yagi Y, Hayashi S, Kobayashi K, Hirayama T, Nakamura T. 2013. Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. *PLoS ONE* 8: e57286.
- Yap A, Kindgren P, Colas des Francs-Small C, Kazama T, Tanz SK, Toriyama K, Small I. 2015. AEF1/MPR25 is implicated in RNA editing of plastid *atpF* and mitochondrial *nad5*, and also promotes *atpF* splicing in Arabidopsis and rice. *Plant Journal* 81: 661–669.
- Yin P, Li Q, Yan C, Liu Y, Liu J, Yu F, Wang Z, Long J, He J, Wang HW *et al.* 2013. Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* 504: 168–171.
- Yuan H, Liu D. 2012. Functional disruption of the pentatricopeptide protein SLG1 affects mitochondrial RNA editing, plant development, and responses to abiotic stresses in Arabidopsis. *Plant Journal* 70: 432–444.
- Zehrmann A, Verbitskiy D, van der Merwe JA, Brennicke A, Takenaka M. 2009. A DYW domain-containing pentatricopeptide repeat protein is required for RNA editing at multiple sites in mitochondria of *Arabidopsis thaliana*. *Plant Cell* 21: 558–567.
- Zhu Q, Dugardeyn J, Zhang C, Takenaka M, Kuhn K, Craddock C, Smalle J, Karampelias M, Denecke J, Peters J *et al.* 2012. SLO2, a mitochondrial pentatricopeptide repeat protein affecting several RNA editing sites, is required for energy metabolism. *Plant Journal* 71: 836–849.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 The relative expression level of independent *pps1*-RNAi lines.

Fig. S2 Schematic structural sequence of PPS1.

Fig. S3 Sequence alignment of PPS1 and its orthologues.

Fig. S4 RNA editing analysis of *nad3* in three independent RNAi lines and the WT.

Fig. S5 Expression of *nad3*.

Fig. S6 Predicted RNA secondary structure of *nad3* transcript in the region of the five editing sites.

Fig. S7 A phylogenetic tree constructed from protein sequences in various species.

Fig. S8 Expression and purification of PPS1⁵⁶⁻⁹¹⁹ and PPS1⁵⁶⁻⁷¹⁷.

Fig. S9 Competitive binding assays of two recombinant PPS1 proteins with probe C.

Fig. S10 Reciprocally competitive RNA EMSA of PPS1 with RNA probes.

Fig. S11 In-solution investigation of the activity of ETC complexes.

Fig. S12 Mutation schematic structures in two *pps1* knockout plants.

Fig. S13 RNA editing efficiency examination of the five *nad3* sites in WT and two *pps1* knock-out plants.

Fig. S14 The abundance and enzyme activities of complexes I, IV, SN V in the *pps1* knockout plants, *pps1/+* #1, compared with the WT.

Fig. S15 Prediction of RNA target sequences of PPS1 on *nad3* transcript.

Table S1 Primers used in this study

Table S2 Primers used for mitochondrial RNA editing assay

Table S3 Viability of stigma and pollen after pollination

Table S4 RNA editing efficiency between wild-type (WT) and *pps1*-RNAi

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.



About New Phytologist

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews and Tansley insights.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <26 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@lancaster.ac.uk)
- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**