

P-class pentatricopeptide repeat protein PpPPR_4 is required for splicing of the plastid tRNA^{Ile} group II intron in *Physcomitrella patens*

Seiya Goto¹, Yasuhiro Kawaguchi^{1,2}, Chieko Sugita¹, Mizuho Ichinose^{1,3}, and Mamoru Sugita^{1,*}

¹Center for Gene Research, Nagoya University, Nagoya, 464-8602 Japan

²Current address: TAIHO Chemical Industry CO., LTD. 3-1-5, Asahi, Kawaguchi-shi, Saitama, 332-0001, Japan

³Institute of Transformative Bio-Molecules, Nagoya University, Nagoya, 464-8602 Japan

*Author for correspondence

Mamoru Sugita

E-mail. sugita@gene.nagoya-u.ac.jp

Author e-mail addresses:

Seiya Goto: gotoseiya@gene.nagoya-u.ac.jp

Yasuhiro Kawaguchi: ad4dnhn@gmail.com

Chieko Sugita: chie@gene.nagoya-u.ac.jp

Mizuho Ichinose: ichinose@gene.nagoya-u.ac.jp

Mamoru Sugita: sugita@gene.nagoya-u.ac.jp

The authors declare no conflict of interest.

Running title: Moss PPR protein involved in plastid tRNA splicing

Keywords: P-class PPR protein, splicing, group II intron, tRNA^{Ile}-GAU, plastid, NEP-dependent transcription

Total word count excluding references: Count: 5549

Abstract: 248

Introduction: 718

Results: 1593

Discussion: 1150

Experimental Procedures: 1011

Acknowledgements: 56

Figure legends: 773

Numbers of Figures: 7

Supporting information: Figures S1 to S6, Tables S1 to S4.

Abstract

Pentatricopeptide repeat (PPR) proteins are widely distributed in eukaryotes and are mostly localized in mitochondria or plastids. PPR proteins play essential roles in various RNA processing steps in organelles. However, the function of the majority of PPR proteins remains unknown. To investigate the function of plastid PPR proteins, we generated and characterized *PpPPR_4* gene knockout mutants in the moss *Physcomitrella patens*. The knockout mosses displayed severe growth retardation. In addition, the quantum yield of photosystem II was reduced in these mutants. Western blot analysis showed that knockout of *PpPPR_4* resulted in a significantly reduced level of plastid-encoded proteins, such as photosystem II reaction center protein D1, the β subunit of ATP synthase, and the stromal enzyme, Rubisco. To further investigate whether knockout of the *PpPPR_4* gene affects plastid gene expression, we analyzed steady-state transcript levels of protein- and rRNA-coding genes by quantitative RT-PCR. This analysis showed that the level of many protein-coding transcripts increased in the mutants. In contrast, splicing of a spacer tRNA^{Ile} precursor encoded by the *rrn* operon was specifically impaired in the mutants while the accumulation of other plastid tRNAs and rRNAs was not largely affected. Thus, the defect in tRNA^{Ile} splicing leads to a considerable reduction of mature tRNA^{Ile}, which may be accountable for the reduced protein level. An RNA mobility shift assay showed that the recombinant PpPPR_4 bound preferentially to domain III of the tRNA^{Ile} group II intron. These results provide strong evidence that PpPPR_4 functions in RNA splicing of the tRNA^{Ile} intron.

Significance Statement: Plastid gene expression is tightly regulated at the post-transcriptional level and many of the nuclear-encoded components are involved in a transcript-specific manner. Many of these components have yet to be identified. We identified a novel P-class pentatricopeptide repeat protein PpPPR_4 that is involved in splicing of the tRNA^{Ile}-GAU precursor in moss chloroplasts. This protein binds preferentially to domain III of the tRNA^{Ile} group II intron. Its binding may facilitate splicing.

Introduction

The pentatricopeptide repeat (PPR) proteins are characterized by tandem arrays of a degenerate 35 amino acid motif, which folds into a pair of antiparallel α -helices (Small and Peeters, 2000). PPR proteins are encoded by the nucleus and constitute a large protein family in land plants, composed of 100 members in mosses to over 450 members in *Arabidopsis thaliana* and rice (*Oryza sativa* L.) (Lurin *et al.* 2004, O'Toole *et al.* 2008). The PPR proteins are structurally divided into four classes, P, PLS, E/E+, and DYW, based on their PPR motif and C-terminal domain structures. In *Arabidopsis* and rice, P-class PPR proteins represent half of all PPR proteins while the remaining half are the E/E+ and DYW-class proteins. Most plant PPR proteins are localized in either mitochondria or plastids, or both (Colcombet *et al.* 2013). In many cases, the loss of PPR proteins leads to impaired organelle-related physiological and developmental functions, i.e., photosynthesis, respiration, cytoplasmic male sterility, and early embryogenesis (Schmitz-Linneweber and Small 2008, Barkan and Small 2014).

Many studies have revealed that members of the P-class PPR proteins are implicated in RNA cleavage, splicing, stabilization, and translation. For instance, maize CRP1 is required for intergenic RNA processing of *petB-petD* pre-mRNA and translation of *petA* mRNA (Barkan *et al.* 1994, Fisk *et al.* 1999). Maize PPR4 functions in *trans*-splicing of *rps12* (Schmitz-Linneweber *et al.* 2006). *Arabidopsis* HCF152 is involved in RNA processing of *psbH-petB* pre-mRNA (Meierhoff *et al.* 2003). *Arabidopsis* PPR10 is essential for the stability of *atpH* and *psaJ* mRNAs (Pfalz *et al.* 2009). *Chlamydomonas* and *Arabidopsis* MRL1 are essential for *rbcL* mRNA stability (Johnson *et al.* 2010). OTP51 and THA8 are required for splicing of the *ycf3* intron 2 in *Arabidopsis* (de Longevialle *et al.* 2008, Khrouchtchova *et al.* 2012). P-class RPF proteins are required for efficient 5' end formation of plant mitochondrial transcripts (Binder *et al.* 2013). PPR proteins with an SMR domain have recently been shown to work for correct processing of plastid 23S-4.5S rRNA precursor (Zoschke *et al.* 2016, Wu *et al.* 2016). In contrast, the E/E+ and DYW-class PPR proteins are mainly involved in RNA editing in mitochondria and plastids (Takenaka *et al.* 2013, Shikanai 2015). Nonetheless, some members of the DYW and E/E+ PPR proteins are also involved in RNA cleavage (Hashimoto *et al.* 2003) and RNA splicing (Chateigner-Boutin *et al.* 2011,

Ichinose *et al.* 2012). Thus, the P-class PPR proteins might contribute to a wide range of RNA processing steps in a gene-specific manner. Despite the large number of PPR proteins in plants, there is still only fragmented information concerning the relationship between PPR proteins and their target RNA molecules.

The genome of the moss *Physcomitrella patens* encodes at least 105 PPR proteins, composed of 89 P-, 6 PLS-, and 10 DYW-class proteins but no E-class protein (Sugita *et al.* 2013). Nine PPR-DYW proteins have been identified as RNA editing factors (Tasaki *et al.* 2010, Ichinose *et al.* 2013, 2014) and one DYW-class protein as an RNA splicing factor (Ichinose *et al.* 2012). Plastid-localized P-class PpPPR_38 (formerly PpPPR531-11) is required for RNA processing of *clpP-5'-rps12* pre-mRNA (Hattori *et al.* 2007). P-class PpPPR_67 and 104 are dually-localized in plastids and mitochondria and involved in tRNA maturation (Sugita *et al.* 2014). Approximately half of the moss P-class PPR proteins show high amino acid identities (30–50%) with *Arabidopsis* PPR proteins, including PPR4, MRL1, GUN1, pTAC2, PRORP1 to 3, and THA8 (Sugita *et al.* 2013). The remaining half of P-class proteins do not correspond to any P-class proteins in *Arabidopsis* and rice and their functions are totally unknown.

To investigate the function of P-class PPR proteins, we generated and characterized *PpPPR_4* gene knockout (KO) mutants. PpPPR_4 is not a homolog of the maize and *Arabidopsis* PPR4 that is involved in *trans*-splicing of chloroplast *rps12* pre-mRNA (Schmitz-Linneweber *et al.* 2006). The obtained KO mutants displayed abnormal phenotypic characteristics, such as retarded growth of the protonemata and lower photosystem II activity. In addition, plastid-encoded protein levels in the cells of mutants were reduced to 10–20% relative to wild type (WT) levels. Analyses of the plastid transcriptome revealed that disruption of *PpPPR_4* resulted in a significant reduction of mature tRNA^{Ile}-GAU and aberrant accumulation of intron-containing tRNA^{Ile} precursor. Finally, we show the binding of PpPPR_4 to domain III of the tRNA^{Ile} group II intron.

Results

PpPPR_4 localizes in chloroplasts

The gene for PpPPR_4 (Phpat.017G039100/Pp1s105_37V6.1) comprises seven exons and six

introns (**Fig. 1A**), and encodes a polypeptide of 812 amino acids with 11 PPR motifs (**Fig. S1**). PpPPR₄ was predicted to be localized in plastids by the subcellular location prediction program TargetP (Emanuelsson *et al.* 2000) and its predicted transit peptide length was 100 amino acids. To confirm this prediction, we performed confocal fluorescence microscopic observation of the protonemal cells transiently expressing the N-terminal 128 amino acids of the PpPPR₄-green fluorescent protein (4N-GFP) fusion. GFP fluorescence was observed in the chloroplasts (Fig. S2), indicating that PpPPR₄ is a chloroplast PPR protein.

Knockout of *PpPPR_4* gene results in severely reduced growth and photosynthesis

To investigate the function of PpPPR₄, we constructed and characterized *PpPPR_4* knockout (KO) mosses, whose cognate gene transcript was not detected by RT-PCR analysis (Figs. S3 and 1B). The growth of protonemata of KO mosses ($\Delta 4-70$, $\Delta 4-72$) was severely retarded (Fig. 1C). Such small colonies might result in poor photosynthetic activity. To investigate the functional status of the photosynthetic apparatus of the KO mutants, we measured photosynthetic activity by pulse amplitude modulation. The maximum (F_v/F_m) and the effective (F_v'/F_m') quantum yield of PSII, designated as ϕ PSII, were reduced to 70% and 60%, respectively, in the KO mutants (Table S1). This implies a defect in energy transfer within PSII.

To confirm that *PpPPR_4* is responsible for the mutant phenotypes, we made stable transformants with *PpPPR_4* full-length cDNA in the KO mutant $\Delta 4-70$. The complemented moss plants thus obtained displayed WT phenotypes (Fig. 1B and C, Table S1). This confirmed that the mutant phenotypes were caused by a loss-of-function in the *PpPPR_4* gene.

Plastid-encoded protein levels are drastically reduced in the *PpPPR_4* KO mutants

Total cellular proteins were subjected to western blot analysis after detecting reduced growth and lower PSII activity in the mutant moss plants. As shown in Fig. 2, PSII reaction center protein D1 (PsbA) and the β subunit of ATP synthase (AtpB) were reduced to 10 and 20%, respectively, of the WT level. Similarly, the reduced amount (10% of the WT level) of the large subunit of Rubisco (LS) in the KO mutants was observed in the Ponceau-S-stained gel of the mutants. In contrast, the levels of the nuclear-encoded plastid proteins, such as PSII 33 kDa extrinsic protein PsbO, elongation factor Tu (EF-Tu), and light-harvesting chlorophyll

binding protein (LHCII), were slightly reduced to 80% of the WT level. The nuclear-encoded chloroplast ribonucleoprotein (cpRNP) accumulated at higher levels in the mutants than the WT. These results suggest that plastid translation is perturbed by the loss of function of PpPPR_4. A significant reduction of plastid-encoded proteins without a large alteration in the level of nuclear-encoded plastid proteins has also been observed in other PPR mutants, such as PpPPR_38 KO mutants (Hattori *et al.* 2007), Arabidopsis *ys1* mutants (Zhou *et al.* 2008) and *atecb2* mutants (Yu *et al.* 2009).

Many protein-coding transcripts accumulate at high levels in PpPPR_4 KO mutants

Many studies have shown that a loss of plastid-localized PPR proteins resulted in abnormal RNA processing or accumulation of specific plastid gene transcripts (Schmitz-Linneweber and Small 2008, Barkan and Small 2014). We therefore expected that disruption of the PpPPR_4 gene would give rise to abnormal accumulation (e.g. reduced accumulation) of plastid transcripts because of severely reduced growth, lower photosynthesis and a reduced amount of plastid-encoded proteins in the KO mutants.

To investigate this possibility, we performed quantitative RT-PCR analysis to measure plastid mRNA accumulation ratios of the KO mosses relative to WT moss. This assay covered all 85 protein-coding and two ribosomal RNA (rRNA)-coding genes of the moss plastid genome. To our surprise, this analysis showed higher accumulation of steady-state levels of many plastid protein-coding transcripts in the mutants (Fig. S4). In particular, transcripts encoding ribosomal proteins, RNA polymerase subunits and YCFs accumulated at substantial levels in the KO mutants. To confirm this result, we performed northern blot analysis of 12 selected protein-coding genes with intron(s) (Fig. 3A) and 6 intronless genes (Fig. 3B). The steady-state transcript level of most genes, except for *psbA*, *rbcL*, *petB* and *petD*, increased in the KO mutants. In particular, *ndhB*, *ycf3*, *rpl2*, *rpl16*, *rpoB*, *rpoC1*, and *atpB* transcript levels increased considerably in the mutants. In contrast, the level of transcript of photosynthesis genes accumulated to similar levels as the mutants and WT. Thus, the levels of transcript accumulation observed by northern analysis are almost coincidental (but not all) with the qRT-PCR results (Fig. S4). To assess whether splicing efficiency of protein-coding transcripts were affected in the mutants, we carried out RT-PCR analysis using exon-specific forward and reverse primers (Fig. S5). Spliced transcripts of all 12 protein-coding genes with intron(s) accumulated at similar levels in the WT and KO mutants.

On the other hand, unspliced transcripts were detected at higher levels in the KO mutants (Fig. S5). This result suggests that splicing of protein-coding transcripts was slightly perturbed but not defected in the KO mutants. However, an apparent splicing defect phenotype was not observed.

***PpPPR_4* KO mutants have a significantly reduced level of tRNA^{Ile}-GAU**

As shown in Fig. S4, plastid 16S and 23S rRNAs accumulated at similar levels in the WT and KO mutants. To confirm the level of accumulation of rRNAs, we performed northern blot analysis of the plastid *rrn* operon encoding 16S rRNA–tRNA^{Ile}–tRNA^{Ala}–23S rRNA–4.5S rRNA–5S rRNA. The steady-state level and processing pattern between the WT and KO mosses were not very different (Fig. 4).

Our qRT-PCR analysis was not designed to detect plastid tRNAs. Therefore, we investigated the steady-state level of plastid tRNAs. The moss plastid DNA contains 31 different tRNA genes, five of which contain a group II intron, while one gene encoding tRNA^{Leu}-UAA possesses a group I intron (Sugiura *et al.* 2003). First, we performed northern blot analysis of six intron-containing tRNA genes using oligonucleotide DNA probes encompassing the anticodon loop. This analysis showed that the tRNA^{Ile}-GAU level was significantly reduced to less than 10% of the WT level while the tRNA^{Ala}-UGC level increased slightly to 2.5-fold higher in the KO mutants than the WT level (Fig. 5). In contrast, the remaining four tRNA levels suffered minor alterations in KO and WT mosses. Furthermore, we investigated the accumulation levels of the remaining 25 plastid tRNAs in WT and a KO mutant ($\Delta 4-70$). Half of the tRNAs examined accumulated at slightly higher levels in the KO mutant while the accumulation levels of 25 tRNAs did not decrease much (Fig. S6). It is likely that a slight increase in the levels of tRNA^{Ala}-UGC and other tRNAs may be a secondary effect caused by a significant reduction in the level of tRNA^{Ile}-GAU.

Unspliced tRNA^{Ile} specifically accumulates in the KO mutants

The significant reduction of tRNA^{Ile}-GAU level can be explained by either a low transcription rate, increased turnover of mature tRNA or impaired RNA splicing of pre-tRNA^{Ile}. The first possibility can be eliminated because the tRNA^{Ile}-GAU gene is cotranscribed with the upstream 16S rRNA gene (*rrn16*). The promoter of the *rrn* operon lies upstream of the 16S rRNA gene. As shown in Fig. 4, rRNA levels were not altered in the KO

mutants and WT mosses. This implies that the transcriptional activity of the *rrn* operon is not affected in the KO mutants.

To investigate the possibility of increased turnover of mature tRNA or impaired RNA splicing, we performed northern blot analysis with two different probes. The 3' exon-specific probe detected mature tRNA and a large RNA band (0.8 kb) in the mutants. Mature tRNA accumulated at less than 10% of the WT level in the KO mutants (Fig. 6). In contrast, the large RNA was detected at a very low level in the WT and complemented mosses. Probing with the intron sequence, mature tRNA was not detected while the large RNA band was detected as a much stronger signal, representing unspliced tRNA^{Ile}-GAU. Thus, unspliced tRNA^{Ile} accumulated to substantial levels but mature tRNA^{Ile} appeared at a much lower level in the KO mosses. Similarly, in its neighboring spacer tRNA^{Ala}-UGC, unspliced tRNA^{Ala} was also observed but as a very weak signal in the KO mutants. However, mature tRNA^{Ala} accumulated at substantial levels in both WT and KO mutants. These observations clearly indicate that disruption of *PpPPR_4* resulted in significantly reduced efficiency of tRNA^{Ile} splicing and a subsequently considerable decrease in the level of mature tRNA^{Ile}. These results indicate that *PpPPR_4* is required for tRNA^{Ile} splicing.

PpPPR_4 binds to the specific region in the tRNA^{Ile} group II intron

How does *PpPPR_4* function in tRNA^{Ile} splicing? PPR proteins can bind to their target RNA molecules via a one PPR motif-one nucleotide recognition mode. Based on a combinatorial amino acid code for RNA recognition by P-class PPR proteins (Barkan *et al.* 2012), we predicted a target sequence, 5'-GUxRxxURx-3' from the 10 sequential PPR motifs out of 11 *PpPPR_4* PPR motifs (Fig. 7A). We then found five RNA sequences homologous to the target sequence in the 785-nt group II intron of pre-tRNA^{Ile} (Fig. 7A, B). To test whether *PpPPR_4* binds to this group II intron and to which part within the intron it is bound, we performed an RNA electrophoresis mobility shift assay (REMSA). For this assay, we prepared 100 kDa recombinant *PpPPR_4* (rPPR4) protein fused at its N-terminus to thioredoxin (Trx). rPPR4 was expressed in *E. coli* and purified using Ni-NTA agarose (Fig. 7C). To detect binding of rPPR4 to target RNA, synthetic RNA probes, RNA1 to RNA5, containing the predicted 8-nt sequence, were used. As shown in Fig. 7D, rPPR4 bound preferentially to RNA4 but weakly to RNA3.

Discussion

In this study, we showed that disruption of the *PpPPR_4* gene specifically impaired splicing of tRNA^{Ile}-GAU while tRNA^{Ala}-UGC splicing was not affected. This is distinct from maize and *Arabidopsis tha8* mutants, where tRNA^{Ala} splicing was strongly inhibited but tRNA^{Ile} splicing was not affected (Khrouchtchova *et al.* 2012). THA8 protein is a short P-type PPR protein with four PPR motifs that are required for the splicing of the tRNA^{Ala} intron also the *ycf3* intron 2 (Khrouchtchova *et al.* 2012). In contrast, severe splicing defects of protein-coding transcripts were not observed in the *PpPPR_4* KO mutants. Related to the involvement of pre-tRNA processing, the maize P-class PPR5 with 10 PPR motifs is required for the stabilization of pre-tRNA^{Gly} and its mature form (Beick *et al.* 2008). To date, there are only a few examples of such involvement of P-class PPR proteins in tRNA species-specific RNA splicing or stability. Thus, this study provides evidence that PpPPR_4 can be categorized as a tRNA^{Ile} specific-splicing factor. This is supported by the REMSA result in which recombinant PpPPR_4 preferentially bound to the 25-nt region forming a nanoloop in domain 3 (DIII) of the group II intron. DIII functions as a catalytic effector in group II intron splicing (Fedorova *et al.* 2003). The crystal structure of the second self-spliced group IIB intron from the green alga *Pylainella littoralis* pre-rRNA revealed interactions of DIII with a DV stem, the junction between DI and DII (J1/J2), and a DII stem (Robart *et al.* 2014). These interactions are essential for catalytically efficient splicing. Thus, it is assumed that PpPPR_4 may facilitate the interaction of DIII with other domains within the tertiary structure of tRNA^{Ile} group II intron. The maize PPR5 binding site encompasses a 40-nt region from EBS1 to α' within DI of the tRNA^{Gly} group II intron (William-Carrier *et al.* 2008), distinct from the binding site of PpPPR_4.

To date, in angiosperms, splicing of pre-tRNA^{Ile} is known to require at least four nuclear-encoded factors, RNC1 (Watkins *et al.* 2007), WTF1 (Kroeger *et al.* 2009), RH3 (Asakura *et al.* 2012), mTERF4 (Hammani and Barkan 2014), and one plastid intron-encoded MatK (Zoschke *et al.* 2010). These are thought to be major splicing factors and constitute a spliceosome-like complex for splicing plastid group II introns (Schmitz-Linneweber *et al.* 2015). Since tRNA^{Ile} splicing was not completely blocked and mature tRNA^{Ile} accumulated to a low level in the *PpPPR_4* KO mutants, we suggest that PpPPR_4 may facilitate splicing

of pre-tRNA^{Ile} by interacting with some splicing factors, such as the nuclear-encoded factors identified in angiosperm chloroplasts. Indeed, the *P. patens* genome encodes an RNC1 ortholog (Phpat.005G041500), a WTF1 ortholog (Phpat.018G071800), four RH3 orthologs (Phpat.002G010700, Phpat.027G004100, Phpat.001G160200, Phpat.002G012400), and three mTERF4 orthologs (Phpat.001G100700, Phpat.022G003200, Phpat.014G046100). It will be important in the future to examine whether these orthologs are involved in splicing of tRNA^{Ile}, including in mosses.

Like angiosperms, the *P. patens* plastid genome encodes two tRNA^{Ile} isoacceptor species, tRNA^{Ile}-GAU and tRNA^{Ile}-CAU (Sugiura *et al.* 2003). The codons that specify isoleucine are AUU, AUC, and AUA. tRNA^{Ile}-GAU can recognize AUU and AUC codons and tRNA^{Ile}-CAU recognizes the AUA codon. The codon usage of protein-coding sequences in *P. patens* plastids implies that tRNA^{Ile}-GAU is more frequently utilized for translation rather than tRNA^{Ile}-CAU. We show here that the loss-of-function of PpPPR_4 leads to severe growth retardation and to a significant reduction of mature tRNA^{Ile}-GAU. We also observed a drastic decrease of plastid-encoded proteins in the PpPPR_4 KO mutants, such as the large subunit of Rubisco (encoded by *rbcL*) and PSII reaction center D1 protein (encoded by *psbA*). However, the steady-state level of mRNAs encoding these proteins accumulated at similar levels in the mutants and WT (Fig. 3). The defect in tRNA^{Ile} splicing leads to a remarkable reduction of mature tRNA^{Ile}, which may be accounted for by the reduced level of plastid-encoded proteins. This may be caused by a general decrease in translational activity in the KO mutants. A similar molecular phenotype is shown in the maize *ppr5* mutants. The absence of PPR5 leads to an albino phenotype with seedling-lethal defects and the loss of mature tRNA^{Gly} and consequently results in the absence of plastid-encoded proteins (Beick *et al.* 2008). Thus, it is likely that a certain threshold level in the accumulation of respective tRNA species, including tRNA^{Ile}-GAU, may be required for normal protein synthesis and maintenance of the photosynthetic apparatus.

In this study we observed that expression of many plastid genes, including tRNAs, was upregulated in the PpPPR_4 OK mutants. Similar upregulation of plastid genes was observed in *Arabidopsis ppr* mutants, such as *ysl* (Zhou *et al.* 2008), *atecb2* (Yu *et al.* 2009), and *vanilla cream1* (Tseng *et al.* 2010). In these *Arabidopsis* mutants, the steady-state levels of

photosynthesis-related genes decreased considerably and expression of most non-photosynthesis genes was upregulated. In contrast, transcript levels of photosynthesis-related genes such as *rbcL* and *psbA* did not decrease in the moss *PpPPR_4* mutants. Thus, the plastid transcriptomic phenotype is somewhat different between the moss *PpPPR_4* mutants and *Arabidopsis* mutants.

In angiosperm plastids, there exist at least two distinct RNA polymerases, plastid-encoded RNA polymerase (PEP) and nuclear-encoded plastid RNA polymerase (NEP). PEP and NEP primarily transcribe photosynthesis genes and non-photosynthesis/housekeeping genes, respectively (Hess and Börner 1999). PEP-dependent genes are highly expressed in the chloroplasts of green tissues where NEP-dependent transcription is regulated at a relatively low level (Legen *et al.* 2002). However, NEP also transcribes some photosynthesis-related genes in both white and green tissues. Various endogenous (e.g. photosynthesis, redox) and exogenous (e.g. light, stresses, hormones) factors are thought to stimulate and/or inhibit both PEP and NEP (Börner *et al.* 2015). Perhaps, in the *PpPPR_4* mutants, reduced photosynthetic activity and/or translational repression might promote the upregulation of NEP-dependent transcription. Although the issue whether nuclear-encoded phage-type RNA polymerase as an NEP exists in the *P. patens* chloroplasts has been debated (Richter *et al.* 2002, Kabeya and Sato 2005), our results support the existence of NEP activity in the *P. patens* chloroplasts. Further careful experiments, however, are needed to make a solid conclusion on this matter.

Plastid DNA size, gene content and gene organization are well conserved between *Arabidopsis* and *P. patens* except for a large 71-kb inversion and the existence/absence of a couple of genes, including *rpoA* and *ycf66* (Sugiura *et al.* 2003). However, there is a large difference in plastid-localized PPR proteins, with 154 in *Arabidopsis* (Colcombet *et al.* 2013) and only 54 in *P. patens* (Sugita *et al.* 2013). This may imply that posttranscriptional and translational regulation in plastids depends more tightly on PPR proteins in *Arabidopsis* rather than in *P. patens*. Intriguingly, although organization of the plastid *rrn* operon is highly conserved in land plants, we could find a *PpPPR_4* homolog in a lycophyte *Selaginella moellendorffii*, but not in *Arabidopsis* and rice (Fig. S1). Presumably, *Arabidopsis* PPR protein(s) involved in splicing of pre-tRNA^{lle} may have lost their importance during evolution while other proteins may have replaced *PpPPR_4*.

Experimental Procedures

Plant material and growth conditions

P. patens subsp. *patens* was grown at 25°C under continuous light (30 $\mu\text{E}/\text{m}^2/\text{s}$) on solidified BCD or BCDAT supplemented with ammonium tartrate (Nishiyama *et al.* 2000). For vegetative propagation, the protonemata were collected every four days and spread on fresh plates. Total cellular RNA was isolated from the four-day-old protonemata.

Subcellular localization

DNA-free total cellular RNA was reverse-transcribed with ReverTra Ace (Toyobo) to synthesize cDNA using Primer Mix (Toyobo). A cDNA fragment (384 bp) encoding the N-terminal 128 amino acids of PpPPR_4 was amplified using PrimeSTAR GXL DNA polymerase (TaKaRa) and appropriate primers (Table S2), and cloned in-frame into the *Sma*I site in pKSPGFP9 (Tasaki *et al.* 2010) to obtain p4N-GFP. Plasmids p4N-GFP and pMt-RFP (Uchida *et al.* 2011) were simultaneously introduced into the *P. patens* protonemata by particle bombardment. GFP or RFP fluorescence was monitored using a confocal fluorescent microscope FLUOVIEW FV10i (Olympus).

Construction of plasmid and moss transformation

The genomic region (3034 bp) was amplified with the primers 4KO-F and 4KO-R (Table S2, Fig. S3), and was cloned into pGEM-T Easy to generate pMI4. The *hpt* gene expression cassette was inserted into the *Nru*I site in the fifth exon of the *PpPPR_4* gene in pMI4. The resulting plasmid, p4KO, was linearized with *Not*I and introduced into the *P. patens* protonemata, and hygromycin-resistant mosses were selected as described by Ichinose *et al.* (2012). Gene disruption by homologous recombination was confirmed by PCR with the primers P1 to P4 (Table S2, Fig. S3).

Detection of *PpPPR_4* transcript

To detect *PpPPR_4* transcript, RT-PCR was carried out using cDNA, PrimeSTAR GXL DNA polymerase, and appropriate primers (Table S2). As a control, *P. patens actin1* (*PpAct1*, Pp1a337_21V6.1) cDNA was amplified using appropriate primers.

Construction of complemented moss plants

The full-length cDNA encoding PpPPR₄ was amplified using primers 4GFP-F and 4-full-R2 (Table S2) and cloned in-frame into the *Swa*I site of the modified plasmid p9WmycZ3, which was derived from the over-expression vector pOX9WZ1 (Sugita *et al.* 2012). Myc tag was generated by mutagenesis using PrimeSTAR GXL DNA polymerase and primers (myc-F and myc-R, Table S2) containing a Myc sequence and an in-frame stop codon. The resulting construct was introduced into the KO mutant $\Delta 4-70$ by particle bombardment. Several independent complementation lines were selected on BCDAT plates containing zeocin for further analysis.

Phenotypic observation and chlorophyll fluorescence analysis

To compare external phenotypes, the moss protonemata were grown on BCDAT plates without antibiotics. Chlorophyll fluorescence was measured with a kinetics multispectral fluorescence imaging FluorCam 800 MF (Photon System Instruments) according to the manufacturer's instruction. Three-week-old protonemal colonies were dark-adapted for 10 min before measurement. Minimum fluorescence (*F_o*) was determined by a weak red light and maximum fluorescence of dark-adapted state (*F_m*) was measured during a subsequent saturating pulse of white light (2000 μ E/m²/s for 0.8 sec). The moss colony was then continuously illuminated with actinic light intensity at an intensity of 400 μ E/m²/s for 1 min.

Immunoblot analysis

Total cellular protein was extracted from the moss protonemata. Ten micrograms of protein and each protein dilution were separated on 5–20% gradient polyacrylamide-SDS gels and transferred to Immobilon[®] Transfer Membranes (Millipore). Immunodetection was performed using ECL Prime (GE Healthcare UK Ltd.) protocols. Anti- α tubulin (eBioscience), anti-PsbA/D1 protein of PSII (Agrisera), anti-ATP synthase β -subunit (kindly provided by Tohru Hisabori), anti-spinach PsbO (kindly provided by Fumio Sato), and anti-tobacco chloroplast ribonucleoprotein cp28 (Nakamura *et al.* 1999) were used. The antibody against tobacco chloroplast elongation factor Tu (Murayama *et al.* 1993) was supplied by Medical & Biological Laboratories Co., Ltd.

qRT-PCR

To quantify the plastid transcripts, qRT-PCR was performed using specific primers designed for exon regions of each gene listed in Table S3. For qRT-PCR analysis, the first strand of cDNA was synthesized using total RNA prepared from 4-day old protonemata, ReverTra Ace qPCR RT Kit (Toyobo), and primer mix (random primer and Oligo(dT) primer). qRT-PCR was performed using Fast SYBR Green Master Mix (Applied Biosystems) and StepOnePlus real-time PCR system (Applied Biosystems) with the following thermal cycling program: 95°C for 20 sec, followed by 40 cycles of 95°C for 3 sec, 60°C for 30 sec. The data were analyzed using the StepOne software version 2.2.2 (Applied Biosystems). All of the quantifications were normalized to the nuclear α -tubulin gene, *PpTUA1* (AB096718). qRT-PCR was performed in duplicate for each sample in three independent experiments.

Northern blot analysis

For northern blot analysis of tRNAs encoded by intron-containing genes, total cellular RNA (5 μ g) was loaded onto a 10% polyacrylamide-7 M urea gel. To analyze all tRNAs, rRNAs, and mRNAs, total cellular RNA was loaded onto a 1% (for mRNA) or 1.2% (for tRNA, rRNA) agarose gel containing formaldehyde. RNAs in the gel were transferred to a nylon membrane that was hybridized with gene-specific DNA probes. Plastid tRNA gene-specific oligonucleotide probes (Table S2) were labeled with DIG-ddUTP and terminal deoxynucleotidyl transferase (Roche). Gene-specific probes were amplified by PCR using appropriate primers (Table S2).

Production of recombinant protein

The DNA sequence (2271 bp) encoding PpPPR_4 protein excluding a putative transit peptide was amplified by PCR from cDNA using appropriate primers (Table S2). The amplified fragment was inserted in-frame into the pBAD/Thio-TOPO vector (Invitrogen), allowing the protein to be expressed as an N-terminal thioredoxin fusion protein with six histidine residues at the C-terminus. The recombinant protein was expressed in *E. coli* XL1-blue and purified using Ni-NTA agarose (Qiagen) as described previously (Tasaki *et al.* 2010).

REMSA

Synthetic RNA probes were 5'-end labeled with T4 polynucleotide kinase and [γ - ^{32}P]ATP, and then extracted by ethanol precipitation. REMSA was performed as described previously (Ichinose *et al.* 2013). The recombinant protein (ranging from 0 to 200 nM) was incubated at 25°C for 10 min in a reaction mixture (100 mM NaCl, 40 mM HEPES-KOH (pH 7.7), 4 mM DTT, 0.1 mg/mL bovine serum albumin, and 10% glycerol), then ^{32}P -labeled RNA probe (50 pM) was added and incubated for a further 15 min. The aliquot of the reaction mixture was subjected to native 6% polyacrylamide gel electrophoresis. Free-RNA and protein-bound RNA in the gel were detected by using a STORM 820 Phosphorimager (GE Healthcare).

Acknowledgements

We thank Tohru Hisabori (Tokyo Institute of Technology) and Fumihiko Sato (Kyoto University) for providing the antibodies, and Toshiharu Shikanai (Kyoto University) for discussing on physiological phenotype of the KO mutants. This work was supported by JSPS KAKENHI Grant Numbers 2529105, 15K14917 (to MS), by JSPS Research Fellowship for Young Scientists to MI (25 • 3052).

Supplementary information

Figure S1. Alignment of PpPPR_4 and its homolog from *Selaginella moellendorffii*.

Figure S2. Subcellular localization of PpPPR_4 protein.

Figure S3. Generation of *PpPPR_4* gene knockout mutants and verification of the targeted disruption.

Figure S4. Abundance of plastid transcripts in the KO mutants.

Figure S5. RT-PCR analysis of protein-coding transcripts with introns.

Figure S6. Northern blot analysis of plastid tRNAs.

Table S1. Fluorescence parameters for wild type, complemented moss, and KO mutants.

Table S2. Primers and oligonucleotides used for cloning, preparation of DNA probes, verification of KO mutants, detection of transcripts.

Table S3. Primers used for qRT-PCR analysis of plastid transcripts.

Table S4. Primers used for RT-PCR for detection of spliced and unspliced transcripts.

References

- Asakura, Y., Galarneau, E., Watkins, K. P., Barkan, A. and van Wijk, K.J.** (2012) Chloroplast RH3 DEAD box RNA helicases in maize and Arabidopsis function in splicing of specific group II introns and affect chloroplast ribosome biogenesis. *Plant Physiol.* **159**, 961-974.
- Barkan, A. and Small, I.** (2014) Pentatricopeptide repeat proteins in plants. *Annu. Rev. Plant Biol.* **65**, 415-442.
- Barkan, A., Rojas, M., Fujii, S., Yap, A., Chong, Y.S., Bond, C.S. and Small, I.** (2012) A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet.* **8**, e1002910.
- Barkan, A., Walker, M., Nolasco, M. and Johnson, D.** (1994) A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. *EMBO J.* **13**, 3170-3181.
- Beick, S., Schmitz-Linneweber, C., Williams-Carrier, R., Jensen, B. and Barkan, A.** (2008) The pentatricopeptide repeat protein PPR5 stabilizes a specific tRNA precursor in maize chloroplasts. *Mol. Cell. Biol.* **28**, 5337-5347.
- Binder, S., Stoll, K. and Stoll, B.** (2013) P-class pentatricopeptide repeat proteins are required for efficient 5' end formation of plant mitochondrial transcripts. *RNA Biol.* **10**, 1511-1519.
- Börner, T., Aleynikova, A.Y., Zubo, Y.O. and Kusnetsov, V.V.** (2015) Chloroplast RNA polymerases: Role in chloroplast biogenesis. *Biochim. Biophys. Acta* **1847**, 761-769.
- Chateigner-Boutin, A.L., des Francs-Small, C.C., Delannoy, E., Kahlau, S., Tanz, S.K., Falcon de Longevialle, A., Fujii, S. and Small, I.** (2011) OTP70 is a pentatricopeptide repeat protein of the E subgroup involved in splicing of the plastid transcript *rpoCl*. *Plant J.* **65**, 532-542.
- Colcombet, J., Lopez-Obando, M., Heurtevin, L., Bernard, C., Martin, K., Berthomé, R., and Lurin, C.** (2013) Systematic study of subcellular localization of Arabidopsis PPR proteins confirms a massive targeting to organelles. *RNA Biol.* **10**, 1557-1575.
- de Longevialle, A.F., Hendrickson, L., Taylor, N. L., Delannoy, E., Lurin, C., Badger, M., Millar, A.H. and Small, I.** (2008) The pentatricopeptide repeat gene *OTP51* with two LAGLIDADG motifs is required for the *cis*-splicing of plastid *ycf3* intron 2 in

- Arabidopsis thaliana*. *Plant J.* **56**, 157-168.
- Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G.** (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol. Biol.* **300**, 1005-1016.
- Fisk, D.G., Walker, M.B. and Barkan, A.** (1999) Molecular cloning of the maize gene *crp1* reveals similarity between regulators of mitochondrial and chloroplast gene expression. *EMBO J.* **18**, 2621-2630.
- Fedorova, O., Mitros, T. and Pyle, A.M.** (2003) Domains 2 and 3 interact to form critical elements of the group II intron active site. *J. Mol. Biol.* **330**, 197-209.
- Hammani, K. and Barkan, A.** (2014) An mTERF domain protein functions in group II intron splicing in maize chloroplasts. *Nucleic Acids Res.* **42**, 5033-5042.
- Hashimoto, M., Endo, T., Peltier, G., Tasaka, M. and Shikanai, T.** (2003) A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast *ndhB* in *Arabidopsis*. *Plant J.* **36**, 541-549.
- Hattori, M., Miyake, H. and Sugita, M.** (2007) A Pentatricopeptide repeat protein is required for RNA processing of *clpP* pre-mRNA in moss chloroplasts. *J. Biol. Chem.* **282**, 10773-10782.
- Hess, W.R. and Börner, T.** (1999) Organellar RNA polymerases of higher plants. *Int. Rev. Cytol.* **190**, 1-59.
- Ichinose, M., Tasaki, E., Sugita, C. and Sugita, M.** (2012) A PPR-DYW protein is required for splicing of a group II intron of *cox1* pre-mRNA in *Physcomitrella patens*. *Plant J.* **70**, 271-278.
- Ichinose, M., Sugita, C., Yagi, Y., Nakamura, T. and Sugita, M.** (2013) Two DYW subclass PPR proteins are involved in RNA editing of *ccmF_c* and *atp9* transcripts in the moss *Physcomitrella patens*: First complete set of PPR editing factors in plant mitochondria. *Plant Cell Physiol.* **54**, 1907-1916.
- Ichinose, M., Uchida, M. and Sugita, M.** (2014) Identification of a pentatricopeptide repeat RNA editing factor in *Physcomitrella patens* chloroplasts. *FEBS Lett.* **588**, 4060-4064.
- Johnson, X., Wostrikoff, K., Finazzi, G., Kuras, R., Schwarz, C., Bujaldon, S., Nickelsen, J., Stern, D.B., Wollman, F.A. and Vallon, O.** (2010) MRL1, a conserved pentatricopeptide repeat protein, is required for stabilization of *rbcL* mRNA in *Chlamydomonas* and *Arabidopsis*. *Plant Cell* **22**, 234-248.

- Kabeya, Y. and Sato, N.** (2005) Unique translation initiation at the second AUG codon determines mitochondrial localization of the phage-type RNA polymerases in the moss *Physcomitrella patens*. *Plant Physiol.* **138**, 369-382.
- Khrouchtchova, A., Monde, R.A. and Barkan, A.** (2012) A short PPR protein required for the splicing of specific group II introns in angiosperm chloroplasts. *RNA* **18**, 1197-1209.
- Kroeger, T.S., Watkins, K.P., Friso, G., van Wijk, K.J. and Barkan, A.** (2009) A plant-specific RNA-binding domain revealed through analysis of chloroplast group II intron splicing. *Proc. Natl. Acad. Sci. USA* **106**, 4537-4542.
- Legen, J., Kemp, S., Krause, K., Profanter, B., Herrmann, R.G. and Maier, R.M.** (2002) Comparative analysis of plastid transcription profiles of entire plastid chromosomes from tobacco attributed to wild-type and PEP-deficient transcription machineries. *Plant J.* **31**, 171-188.
- 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, M-L., Mireau, H., Peeters, N., Renou, J-P., Szurek, B., Taconnat, L. and Small, I.** (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* **16**, 2089-2103.
- Meierhoff, K., Felder, S., Nakamura, T., Bechtold, N. and Schuster, G.** (2003) HCF152, an Arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast *psbB-psbT-psbH-petB-petD* RNAs. *Plant Cell* **15**, 1480-1495.
- Michel, F., Umesono, K. and Ozeki, H.** (1989) Comparative and functional anatomy of group II catalytic introns – a review. *Gene* **82**, 5-30.
- Murayama, Y., Matsubayashi, T., Sugita, M. and Sugiura, M.** (1993) Purification of chloroplast elongation factor Tu and cDNA analysis in tobacco: the existence of two chloroplast elongation factor Tu species. *Plant Mol. Biol.* **22**, 767-774.
- Nakamura, T., Ohta, M., Sugiura, M. and Sugita, M.** (1999) Chloroplast ribonucleoproteins are associated with both mRNAs and intron-containing precursor tRNAs. *FEBS Lett.* **460**, 437-441.
- Nishiyama, T., Hiwatashi, Y., Sakakibara, K., Kato, M., and Hasebe, M.** (2000) Tagged mutagenesis and gene-trap in the moss, *Physcomitrella patens* by shuttle mutagenesis. *DNA Res.* **7**, 9-17.
- O'Toole, N., Hattori, M., Andres, C., Iida, K., Lurin, C., Schmitz-Linneweber, C.,**

- Sugita, M. and Small, I.** (2008) On the expansion of the pentatricopeptide repeat gene family in plants. *Mol. Biol. Evol.* **25**, 1120-1128.
- Pfalz, J., Bayraktar, O.A., Prikryl, J. and Barkan, A.** (2009) Site-specific binding of a PPR protein defines and stabilizes 5' and 3' mRNA termini in chloroplasts. *EMBO J.* **28**, 2042-2052.
- Richter, U., Kiessling, J., Hedtke, B., Decker, E., Reski, R., Borner, T. and Weihe, A.** (2002) Two *RpoT* genes of *Physcomitrella patens* encode phage-type RNA polymerases with dual targeting to mitochondria and plastids. *Gene* **290**, 95-105.
- Robart, A.R., Chan, R.T., Peters, J.K., Rajashankar, K.R. and Toor, N.** (2014) Crystal structure of a eukaryotic group II intron lariat. *Nature* **514**, 193-197.
- Schmitz-Linneweber, C., Williams-Carrier, R.E, Williams-Voelker, P.M., Kroeger, T.S., Vichas, A. and Barkan, A.** (2006) A pentatricopeptide repeat protein facilitates the *trans*-splicing of the maize chloroplast *rps12* pre-mRNA. *Plant Cell* **18**, 2650-2663.
- Schmitz-Linneweber, C., Lampe, M-K., Sultan, L.D. and Ostersetzer-Biran, O.** (2015) Organellar maturases: A window into the evolution of the spliceosome. *Biochim. Biophys. Acta* **1847**, 798-808.
- Schmitz-Linneweber, C. and Small, I.** (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci.* **13**, 663-670.
- Shikanai, T.** (2015) RNA editing in plants: Machinery and flexibility of site recognition. *Biochim. Biophys. Acta* **1847**, 779-785.
- Small, I.D. and Peeters, N.** (2000) The PPR motif – a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem. Sci.* **25**, 46-47.
- Sugita, C., Kato, Y., Yoshioka, Y., Tsurumi, N., Iida, Y., Machida, Y. and Sugita, M.** (2012) *CRUMPLED LEAF (CRL)* homologs of *Physcomitrella patens* are involved in the complete separation of dividing plastids. *Plant Cell Physiol.* **53**, 1124-1133.
- Sugita, C., Komura, Y., Tanaka, K., Kometani, K., Satoh, H. and Sugita, M.** (2014) Molecular characterization of three PRORP proteins in the moss *Physcomitrella patens*: nuclear PRORP protein is not essential for moss viability. *PLoS ONE* **9**, e108962.
- Sugita, M., Ichinose, M., Ide, M. and Sugita, C.** (2013) Architecture of the PPR gene family in the moss *Physcomitrella patens*. *RNA Biol.* **10**, 1439-1445.

- Sugiura, C., Kobayashi, Y., Aoki, S., Sugita, C. and Sugita, M.** (2003) Complete chloroplast DNA sequence of the moss *Physcomitrella patens*: evidence for the loss and relocation of *rpoA* from the chloroplast to the nucleus. *Nucleic Acids Res.* **31**, 5324-5331.
- Takenaka, M., Zehrmann, A., Verbitskiy, D., Härtel, B. and Brennicke, A.** (2013) RNA editing in plants and its evolution. *Annu. Rev. Genet.* **47**, 335-352.
- Tasaki, E., Hattori, M. and Sugita, M.** (2010) The moss pentatricopeptide repeat protein with a DYW domain is responsible for RNA editing of mitochondrial *ccmF* transcript. *Plant J.* **62**, 560-570.
- Tseng, C.C., Sing, T.Y., Li, Y.C., Hsu, S.J., Lin, C.L. and Hsieh, M.H.** (2010) Editing of *accD* and *ndhF* chloroplast transcripts is partially affected in the *Arabidopsis thaliana* *cream1* mutant. *Plant Mol. Biol.* **73**, 309-323.
- Uchida, M., Ohtani, S., Ichinose, M., Sugita, C. and Sugita, M.** (2011) The PPR-DYW proteins are required for RNA editing of *rps14*, *cox1* and *nad5* transcripts in *Physcomitrella patens* mitochondria. *FEBS Lett.* **585**, 2367-2371.
- Watkins, K.P., Kroeger, T.S., Cooke, A.M., Williams-Carrier, R.E., Friso, G., Belcher, S.E., van Wijk, K.J. and Barkan, A.** (2007) A ribonuclease III domain protein functions in group II intron splicing in maize chloroplasts. *Plant Cell* **19**, 2606-2623.
- Williams-Carrier, R., Kroeger, T. and Barkan, A.** (2008) Sequence-specific binding of a chloroplast pentatricopeptide repeat protein to its native group II intron ligand. *RNA* **14**, 1930-1941.
- Wu, W., Liu, S., Ruwe, H., Zhang, D., Melonek, J., Zhu, Y., Hu, X., Gusewski, S., Yin, P., Small, I., Katharine, H. and Huang, J.** (2016) SOT1, a pentatricopeptide repeat protein with a small MutS-related domain, is required for correct processing of plastid 23S-4.5S rRNA precursor in *Arabidopsis thaliana*. *Plant J.* in press.
- Yu, Q.B., Jinag, Y., Chong, K. and Yang, Z.N.** (2009) AtECB2, a pentatricopeptide repeat protein, is required for chloroplast transcript *accD* RNA editing and early chloroplast biogenesis in *Arabidopsis thaliana*. *Plant J.* **59**, 1011-1023.
- Zhou, W., Cheng, Y., Yap, A., Chateigner-Boutin, A.L., Delannoy, E., Hammani, K., Small, I. and Huang, J.** (2008) The Arabidopsis gene *YS1* encoding a DYW protein is required for editing of *rpoB* transcripts and the rapid development of chloroplasts during early growth. *Plant J.* **58**, 82-96.

Zoschke, R., Nakamura, M., Liere, K., Sugiura, M., Börner, T. and Schmitz-Linneweber, C. (2010) An organellar maturase associates with multiple group II introns. *Proc. Natl. Acad. Sci. USA* **107**, 3245-3250.

Zoschke, R., Watkins, K.P., Miranda, R.G. and Barkan, A. (2016) The PPR-SMR protein PPR53 enhances the stability and translation of specific chloroplast RNAs in maize. *Plant J.* in press. DOI: 10.1111/tpj.13093

Figure legends

Figure 1. Structure of *PpPPR_4* gene and external phenotype of the KO mutants.

(A) Exons of the *PpPPR_4* gene are represented by boxes. The *hpt* gene cassette was inserted into the *NruI* site of exon 5. (B) RT-PCR to detect cognate transcripts was performed using primers 4cds-F and 4KO-R (Table S2) in the wild type (WT), the *PpPPR_4* KO mutants ($\Delta 4-70$ and $\Delta 4-72$), complemented moss (Comp), and the *PpPPR_38* KO mutant ($\Delta 38$) as a control (Hattori *et al.* 2007). RT-PCR for *PpActin1* transcript was used as the control. (C) Protonemata colony of the wild type (WT), *PpPPR_4* KO mutants, and the complemented moss (Comp). The mosses were grown for three weeks on BCDAT medium plates without hygromycin B. Scale bar = 10 mm.

Figure 2. Immunoblot analysis of chloroplast proteins.

Total proteins (10 μ g or the indicated dilution of the WT sample) were analyzed by immunoblots with antibodies for photosystem II subunits (PsbA, PsbO), ATP synthase β -subunit (AtpB), chloroplast elongation factor Tu (EF-Tu), chloroplast ribonucleoprotein (cpRNP), and α -tubulin (α -tub). The blotted membrane was stained with Ponceau S to visualize the loaded proteins (bottom). The large subunit of Rubisco (RbcL) and light-harvesting chlorophyll binding protein (LHCII) are indicated. Size markers (kDa) are indicated on the right (Precision Plus Protein Standards; BIO-RAD).

Figure 3. Northern blot analysis of plastid protein-coding genes.

Total cellular RNA (10 μ g) from wild type (WT), KO mutants $\Delta 4-70$ (lanes 1) and $\Delta 4-72$ (lanes 2) was separated on a 1% agarose gel subjected to northern blot analysis using plastid gene-specific DNA probes of intron-containing genes (A) and intronless genes (B). The

positions of 25S and 18S cytoplasmic rRNAs are marked on the right hand side of each panel.

Figure 4. RNA gel blot analysis of plastid rRNA.

Total cellular RNA (2.5 µg) from the wild type (WT) and KO mutants ($\Delta 4-70$ and $\Delta 4-72$) was analyzed. Map of the plastid rRNA gene (*rrn*) region. The hybridization probes 1 to 5 and their locations are indicated below the map. RNA gel blot analysis with probes 1 to 5. The left panel is the same gel stained with ethidium bromide (EtBr).

Figure 5. Steady-state transcript levels of plastid intron-containing tRNA genes.

Total cellular RNA (5 µg/lane) extracted from the wild type (WT) and KO mutants ($\Delta 4-70$, $\Delta 4-72$) was separated on a 10% polyacrylamide-7 M urea gel, and subjected to northern blot analysis. Ethidium bromide (EtBr)-stained gel served as the loading control. The positions of RNA size markers (80 and 150 nt, Low Range ssRNA Ladder, New England BioLab Inc.) are indicated on the left.

Figure 6. RNA gel blot analysis of plastid tRNA^{Ile}-GAU and tRNA^{Ala}-UGC.

Total RNA (10 µg) from the wild type (WT), KO mutants ($\Delta 4-70$, $\Delta 4-72$), and complementation line (Comp) was separated on 1.2% agarose gels and subjected to northern blot analysis using the 3' exon or intron-specific probes (*trnI* 3' exon, *trnA* 3' exon, and *trnI* intron). Shown on the right are the same gels stained with ethidium bromide. RNA size markers 200 to 2,000 nt (Prestain Marker for RNA High, BioDynamics Laboratory Inc.) are indicated on the left.

Figure 7. Detection of binding of the recombinant PpPPR₄ to the predicted RNA binding site.

(A) Prediction of PpPPR₄ binding site. The amino acids at positions 6 and 1' in PPR motifs 1 to 10 are indicated. Position 1' is the first amino acid of the respective C-terminally adjacent PPR repeat. The 6/1' amino acid combinations for each PPR motif 1 to 10 predict a binding site GUXRxxURx (R = A or G, x = A, G, C, or U). RNA1 to 5 contain a 9-mer sequence homologous to the predicted binding site and their nucleotide positions (pos.) in the 785-bp tRNA^{Ile} intron are also indicated. DI, II, III and V indicate domain I, II, III and V, respectively. EBS2 (exon binding site 2) and J2/3 (the junction between DII and DIII) are

indicated. **(B)** Schematic secondary structure of the tRNA^{Ile} group II intron and the positions of RNA probes RNA1 to 5 are indicated. EBS1 and IBS1, EBS2 and IBS2, α and α' refer to three-dimensional pairings (Michel *et al.* 1989). DI to VI indicate domains I to VI. **(C)** Coomassie brilliant blue-stained recombinant PpPPR_4 protein (rPPR4, 1 μ g) and thioredoxin (rTrx, 1 μ g) separated on a 5–20% gradient polyacrylamide-SDS gel. The size marker is the Precision Plus Protein Standards (BIO-RAD). **(D)** REMSA was performed with recombinant proteins (rTrx, rPPR4) and ³²P-labeled RNA1 to 5. The concentration of recombinant proteins is indicated above each lane. B indicates bound probe, U indicates unbound probe. All RNA probes were used at 50 pM with protein concentrations ranging from 0 to 200 nM.