PCP
PLANT & CELL PHYSIOLOGY

PPR20 Is Required for the *cis*-Splicing of Mitochondrial *nad2* Intron 3 and Seed Development in Maize

Yan-Zhuo Yang[1], Shuo Ding[1], Yong Wang[1], Hong-Chun Wang[1], Xin-Yuan Liu[1], Feng Sun[1], Chunhui Xu[1], Baohui Liu[2] and Bao-Cai Tan[1,*]

[1]Key Lab of Plant Development and Environment Adaptation Biology, Ministry of Education, School of Life Sciences, Shandong University, Qingdao 266237, China
[2]School of Life Sciences, Guangzhou University, Guangzhou 510006, China
*Corresponding author: E-mail, bctan@sdu.edu.cn; Fax, 0532-5863-0009.
(Received May 29, 2019; Accepted October 22, 2019)

Regular Paper

Pentatricopeptide repeat (PPR) proteins are helical repeat RNA-binding proteins that function in RNA processing by conferring sequence-specific RNA-binding activity. Owing to the lethality of PPR mutants, functions of many PPR proteins remain obscure. In this study, we report the function of PPR20 in intron splicing in mitochondria and its role in maize seed development. PPR20 is a P-type PPR protein targeted to mitochondria. The *ppr20* mutants display slow embryo and endosperm development. Null mutation of PPR20 severely reduces the *cis*-splicing of mitochondrial *nad2* intron 3, resulting in reduction in the assembly and activity of mitochondrial complex I. The *ppr20-35* allele with a *Mu* insertion in the N-terminal region shows a much weaker phenotype. Molecular analyses revealed that the mutant produces a truncated transcript, coding for PPR20$^{\Delta N120}$ lacking the N-terminal 120 amino acids. Subcellular localization revealed that PPR20$^{\Delta N120}$:GFP is able to target to mitochondria as well, suggesting the sequence diversity of the mitochondrial targeting peptides. Another mutant *zm_mterf15* was also found to be impaired in the splicing of mitochondrial *nad2* intron 3. Further analyses are required to identify the exact function of PPR20 and Zm_mTERF15 in the splicing of *nad2* intron 3.

Keywords: Intron splicing • Maize (*Zea mays*) • Mitochondrion • PPR20 • Seed development • Targeting peptide.

Accession numbers: Sequence data of *PPR20* transcripts can be found in GenBank under accession numbers: **MH837624** and **MH837625**.

Introduction

Mitochondrion is derived from an α-proteobacterium-like ancestor. During evolution, many of the mitochondrial genes have been transferred to the nucleus or lost. The mitochondrial genome encodes approximately 5% of the proteins required for its normal functions (Mackenzie and McIntosh 1999). In maize, the NB (normal type in nuclear background B73) mitochondrial genome encodes 33 proteins, 22 transfer RNAs and 3 ribosomal RNAs (Clifton et al. 2004). The proteins include 9 subunits for complex I, 1 subunit for complex III, 3 subunits for complex IV, 5 subunits for complex V, 4 proteins involved in cytochrome *c* maturation, 9 ribosomal proteins, a maturase and a transporter (Clifton et al. 2004). Mitochondrial gene expression involves a number of post-transcriptional processes (Binder and Brennicke 2003), such as RNA C-to-U editing, intron splicing and 5′ or 3′ RNA maturation, which require the assistance of numerous nucleus-encoded factors. These processes are essential for the biogenesis and normal function of mitochondria (Keren et al. 2009, Kuhn et al. 2011, Zmudjak et al. 2013, Hsu et al. 2014, Cai et al. 2017, Chen et al. 2017, Qi et al. 2017, Li et al. 2018).

The mitochondrial genes contain 20–23 group II introns (Bonen 2008). Group II introns are catalytic RNAs that act as ribozymes (Pyle 2016). They are mobile genetic elements that can insert into the host genome and remove themselves by autocatalytic RNA splicing from precursor transcripts (Pyle 2016). Group II introns contain six stem–loop structures (I, II, III, IV, V and VI) that form a common secondary structure. Nucleotides from domains I, V, VI and the linker region between domains II and III form a catalytic core essential for ribozyme activity, while domain IV contains an open reading frame encoding RNA maturase required for splicing and mobility (Pyle 2016). The splicing of some group II introns is autocatalytic in vitro under high salt conditions, whereas splicing in vivo requires other proteins (Pyle 2016).

Plant mitochondrial group II introns have lost the ability of self-splicing. Their splicing requires the assistance of a multitude of nuclear proteins from different protein families. The splicing of group II introns in bacteria is mediated by maturases encoded by domain IV of the intron itself (Lambowitz and Zimmerly 2004). However, higher plant mitochondrial genome only retains a single maturase gene (*matR*) encoded by the *nad1* intron 4 (Clifton et al. 2004). MatR in *Brassicaceae* functions in the splicing of six introns, *nad1* introns 3 and 4, *nad5* intron 5, *nad7* intron 2, *rpl2* intron 1 and *rps3* intron 1 (Sultan et al. 2016). In addition to *matR*, plant nuclear genome encodes four maturases (nMATs). nMAT1 and nMAT2 function in the splicing of

Plant Cell Physiol. 61(2): 370–380 (2020) doi:10.1093/pcp/pcz204, Advance Access publication on 31 October 2019, available online at https://academic.oup.com/pcp

several group II introns in Arabidopsis mitochondria (Keren et al. 2009, Keren et al. 2012), while nMAT4 is essential for RNA processing and maturation of *nad1* introns (Cohen et al. 2014). mCSF1, a mitochondrial homolog of the chloroplast RNA splicing and ribosome maturation (CRM) protein, is required for the splicing of many introns in Arabidopsis mitochondria (Zmudjak et al. 2013). Mitochondrial transcription termination factors (mTERFs) have modular MTERF motifs characterized by tandem repeats of 30 amino acids. At_mTERF15 is involved in the splicing of *nad2* intron 3 in Arabidopsis mitochondria (Hsu et al. 2014). PMH2 (putative mitochondrial helicase 2), a DEAD-box RNA helicase that catalyzes the unwinding of duplex RNA, is required for the splicing of group II introns in Arabidopsis mitochondria (Matthes et al. 2007, Kohler et al. 2010). Plant organelle RNA recognition (PORR) proteins bind to RNA and are involved in intron splicing in plant organelles (Matthes et al. 2007; Francs-Small et al. 2012). These nucleus-encoded splicing factors are RNA-binding proteins, and most of them do not specifically recognize target RNAs.

Pentatricopeptide repeat (PPR) proteins, one of the largest protein families in higher plants, confer sequence-specific RNA-binding ability and are involved in the RNA metabolism in organelles. They contain PPR motifs (31–36 amino acids) in a tandem array, and each motif forms a helix–turn–helix structure (Yin et al. 2013), recognizing RNA in an one-nucleotide one repeat manner (Barkan et al. 2012, Takenaka et al. 2013, Yagi et al. 2013). PPR proteins are ubiquitous in eukaryotes but greatly expanded in terrestrial plants. They are classified into the following two subfamilies: P-type subfamily contains canonical PPR motifs (P, 35 amino acids) and PLS-type subfamily possesses sequential repeats of P, longer (L, 36 amino acids) and shorter (S, 31 amino acids) motifs and often carries an E1/E2 or E1/E2-DYW motifs at the C-terminus (Lurin et al. 2004, Cheng et al. 2016). Most PPR proteins are targeted to mitochondria and chloroplasts, and they function in post-transcriptional RNA processing that includes intron splicing, C-to-U editing and 5′ or 3′ RNA ends' maturation (Barkan and Small 2014). Because of its large family size, the molecular functions of many PPR proteins and their roles in plant growth and development are still unclear.

In this study, we determined the molecular function and role of a P-type PPR protein, PPR20, in maize seed development. PPR20 is required for the *cis*-splicing of mitochondrial *nad2* intron 3, complex I assembly and seed development in maize. We further provided evidence that PPR20 may function independently from Zm_mTERF15 in mediating *nad2* intron 3 splicing. In addition, we revealed that *PPR20* produces two transcripts that encode the full-length and the truncated PPR20 proteins. The latter, although lost the original signal peptide, is able to target to mitochondria and partially functional.

Results

PPR20 is a P-type PPR protein targeted to mitochondria

PPR20 (GRMZM5G818978) contains 13 PPR motifs that show a variable degree of conservation (**Fig. 1A, B**). It lacks any other domains, classifying PPR20 as a P-type PPR protein. PPR20 was predicted to be targeted to mitochondria by the TargetP algorism (Emanuelsson et al. 2000). To test this, full-length PPR20 was fused with green fluorescent protein (GFP) and the resulted PPR20:GFP fusion protein was transiently expressed in Arabidopsis protoplasts. GFP signals were found in mitochondria that were stained by MitoTracker Red (**Fig. 1C**), indicating that PPR20 is targeted to mitochondria.

PPR20 is a conserved protein with putative orthologs in monocots and dicots. It shares a 90% identity with sorghum SbPPR20, a 78% identity with foxtail millet SiPPR20, a 72% identity with rice OsPPR20 and a 36% identity with Arabidopsis AtPPR20. Phylogenetic analysis revealed two clades that fall in monocots and dicots (**Supplementary Fig. S1**). Two monocots, date and oil palm are relatively distantly related to maize PPR20 (**Supplementary Fig. S1**).

Mutations in *PPR20* arrest maize seed development

To study the function of PPR20, we isolated two *Mutator* (*Mu*) insertional mutants from the UniformMu mutagenesis population (McCarty et al. 2005). *ppr20-35* and *ppr20-792* carry a *Mu* insertion at 35 and 792 nt downstream of the translation start codon of *PPR20*, respectively (**Fig. 2A**). The selfed progeny of *ppr20-792*/+ heterozygotes segregated defective kernels, whereas the selfed progeny of *ppr20-35*/+ heterozygotes segregated small kernels (*smk*) (**Fig. 2B, C**). The ratios of both mutants are approximately 25%, indicating monogenic recessive mutations. The *ppr20-35* kernels are smaller than the wild type (WT), but bigger than the *ppr20-792* kernels (**Fig. 2D, E**). Most *ppr20-792* mutant kernels were embryo lethal, only approximately 5% showed root growth but without shoot emergence upon germination (**Fig. 2F**). However, the *ppr20-35* kernels can germinate and the seedlings show delayed growth (**Fig. 2F**). The *ppr20-35* mutant plant can produce seeds. Cosegregation analysis on F2 populations of *PPR20-792* and *PPR20-35* showed that the seed-defective phenotype was tightly linked to the insertions (**Supplementary Fig. S2**). Crosses between *ppr20-792* and *ppr20-35* heterozygotes yielded ears segregating *smk* (**Supplementary Fig. S2**). These results indicate that the defective and small kernel phenotypes are caused by the disruption of *PPR20*, and the *ppr20-35* may be a leaky mutation.

To determine the impact on seed development, paraffin sections were performed on developing kernels of the *ppr20* mutants and WT. At 14 d after pollination (DAP), WT embryo reached the leaf stage, in which scutellum (sc), coleoptile (col), root apical meristem (RAM), shoot apical meristem (SAM) and leaf primordia (LP) were clearly observed (**Fig. 3A, D**). In contrast, the *ppr20-35* embryo only developed sc, col, RAM and SAM (**Fig. 3B, E**) and the *ppr20-792* embryo just reached the coleoptilar stage characterized by the presence of col on the periphery of the SAM (**Fig. 3C, F**). Although the endosperms were filled with starch, the size of the mutant endosperms was much smaller than that of WT (**Fig. 3A–C**). These results indicate that embryogenesis and endosperm development are delayed in the mutants, and the mutation in the *ppr20-792* causes more severe impairments in seed development than that in the *ppr20-35*.

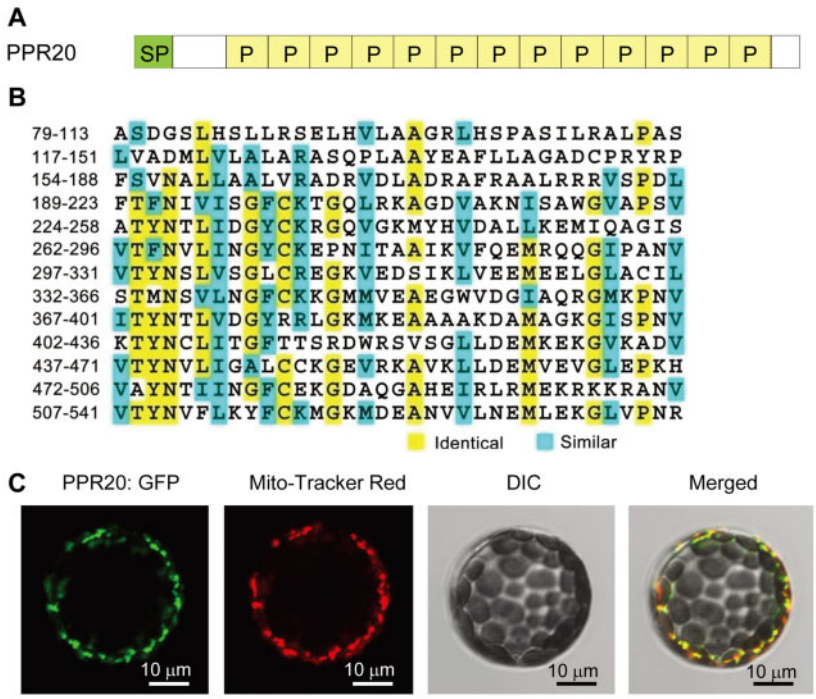

Fig. 1 *PPR20* encodes a mitochondrion-targeted P-type PPR protein. (A) Protein structures of PPR20. PPR20 contains 13 PPR motifs and an N-terminal signal peptide. (B) Alignment of 13 PPR motifs in PPR20. The motifs were identified by the TPRpred tool. Identical amino acids are highlighted in yellow and similar ones are highlighted in blue. (C) The subcellular localization of PPR20 in Arabidopsis protoplasts. The GFP fluorescence was observed under a confocal microscope. The cells were stained with MitoTracker Red before observation. DIC, differential interference contrast.

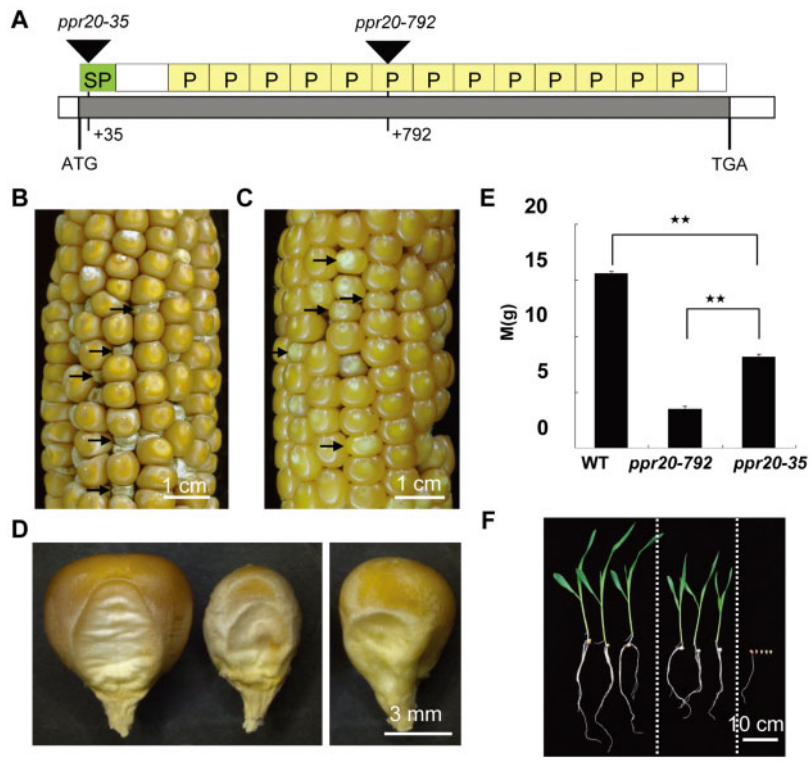

Fig. 2 Mutations in PPR20 inhibit seed development in maize. (A) Position of *Mu* insertions in *ppr20* mutants. Ears segregating *ppr20-792* (B) and *ppr20-35* (C) small kernel mutants. Arrows indicate the mutant kernels. (D) Germinal face of mature WT (left), *ppr20-792* (middle) and *ppr20-35* (right) kernels. (E) Comparison of hundred-kernel weight of mature WT, *ppr20-792* and *ppr20-35* kernels. Error bars represent ± SD of 100 kernels (**$P < 0.001$, Student's *t*-test, $n = 3$). (F) Germination of WT (left), *ppr20-35* (middle) and *ppr20-792* (right) kernels.

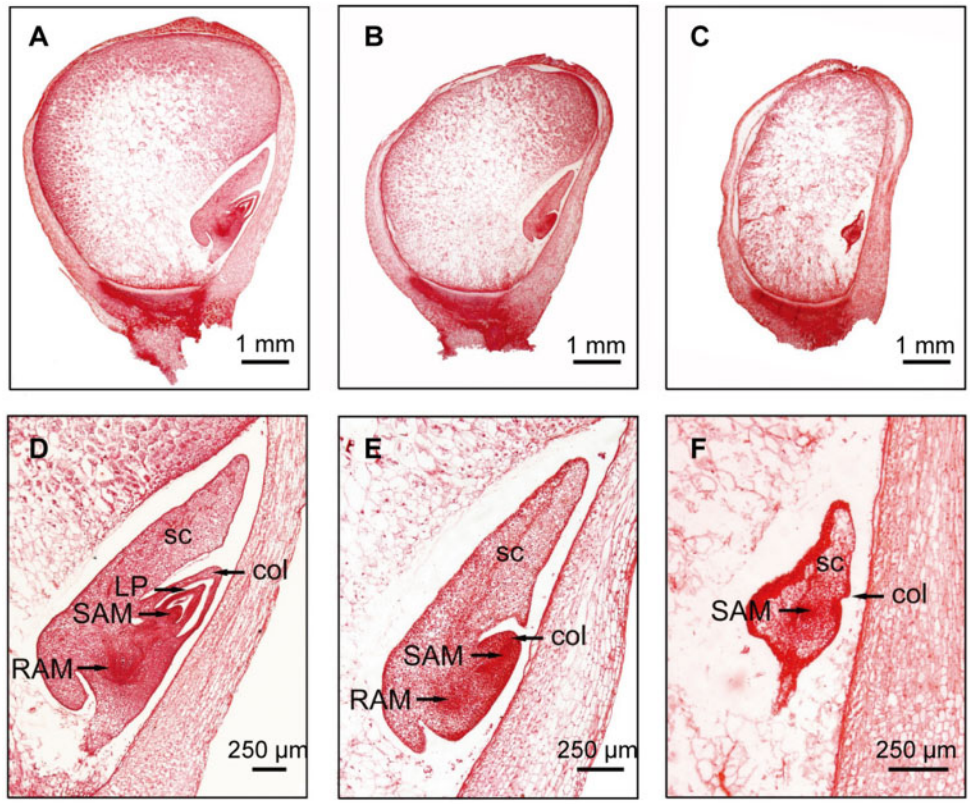

Fig. 3 Seed development is delayed in the *ppr20* mutants. Longitudinal section of WT (A and D), *ppr20-35* (B and E) and *ppr20-792* (C and F) kernels at 14 DAP.

The *cis*-splicing of *nad2* intron 3 is affected in the *ppr20* mutants

As most P-type PPR proteins are involved in organelle RNA metabolism such as RNA splicing, stabilization and translation regulation (Barkan and Small 2014), we investigated the PPR20 function on mitochondrial RNA metabolism by compraring the transcript profile between the two *ppr20* mutants and the WT. The W22 (Wisconsin 22) inbred contains an NB mitochondrial genome that encodes 33 protein-coding genes (Clifton et al. 2004). The transcripts were amplified with primers designed to capture nearly complete coding region of each gene. As shown in **Fig. 4A**, most genes showed a comparable level between the mutants and WT, some with increased levels in the mutants (e.g. *nad1*, *nad4* and *nad7*), but *nad2* showed a dramatically deceased transcript level in the two *ppr20* mutants, with more decrease in *ppr20-792* than in *ppr20-35*. The level of *nad2* transcript is consistent with the phenotypic display that *ppr20-792* showed a stronger seed-defective phenotype than *ppr20-35*.

Then, we investigated whether the reduced *nad2* transcript is the result of defects in intron splicing. *nad2* pre-RNA contains one *trans*-splicing intron and three *cis*-splicing introns (**Fig. 4B**). Primers are anchored on the exons across each intron (**Fig. 4B**). As shown in **Fig. 4C**, the *cis*-splicing of intron 3 was dramatically decreased in the *ppr20* mutants compared with WT, whereas the splicing of other three introns was normal (**Fig. 4C**). Again, the impact on splicing is consistent with the severity of the mutant phenotypes. The pre-RNA containing unspliced intron 3 was accumulated in the mutants (**Fig. 4C**). The accumulation of unspliced pre-RNA was confirmed by RT-PCR with two pairs of primers (3F/6R and 5F/4R) amplifying exon 3–intron 3 and intron 3–exon 4 junctions, respectively (**Fig. 4D**). Quantitative RT-PCR (qRT-PCR) results showed that the spliced *nad2* exon 3–exon 4 was decreased approximately 16- and 3-fold in *ppr20-792* and *ppr20-35* mutants, respectively and the unspliced *nad2* exon 3–intron 3–exon 4 was increased approximately 2-fold in both mutants (**Fig. 4E**). These results indicate that PPR20 is required for the *cis*-splicing of *nad2* intron 3 in maize.

The W22 mitochondrial genome contains 22 group II introns. To further verify the splicing specificity of *PPR20*, the splicing efficiency of all 22 mitochondrial introns was detected by qPCR. The splicing efficiency is calculated as a ratio of spliced transcript to unspliced transcript in the mutant normalized to the same ratio in the WT. Both *ppr20* mutants showed defects in the splicing efficiency of the *nad2* intron 3, whereas the splicing efficiency of other introns was not significantly affected (**Fig. 5**). These results indicate that PPR20 is a splicing factor specifically required for the *cis*-splicing of *nad2* intron 3 in maize.

The mutation in *PPR20* affects mitochondrial complex I assembly and activity

Nad2 is a subunit of mitochondrial complex I, which is an L-shaped NADH dehydrogenase composed of a membrane arm

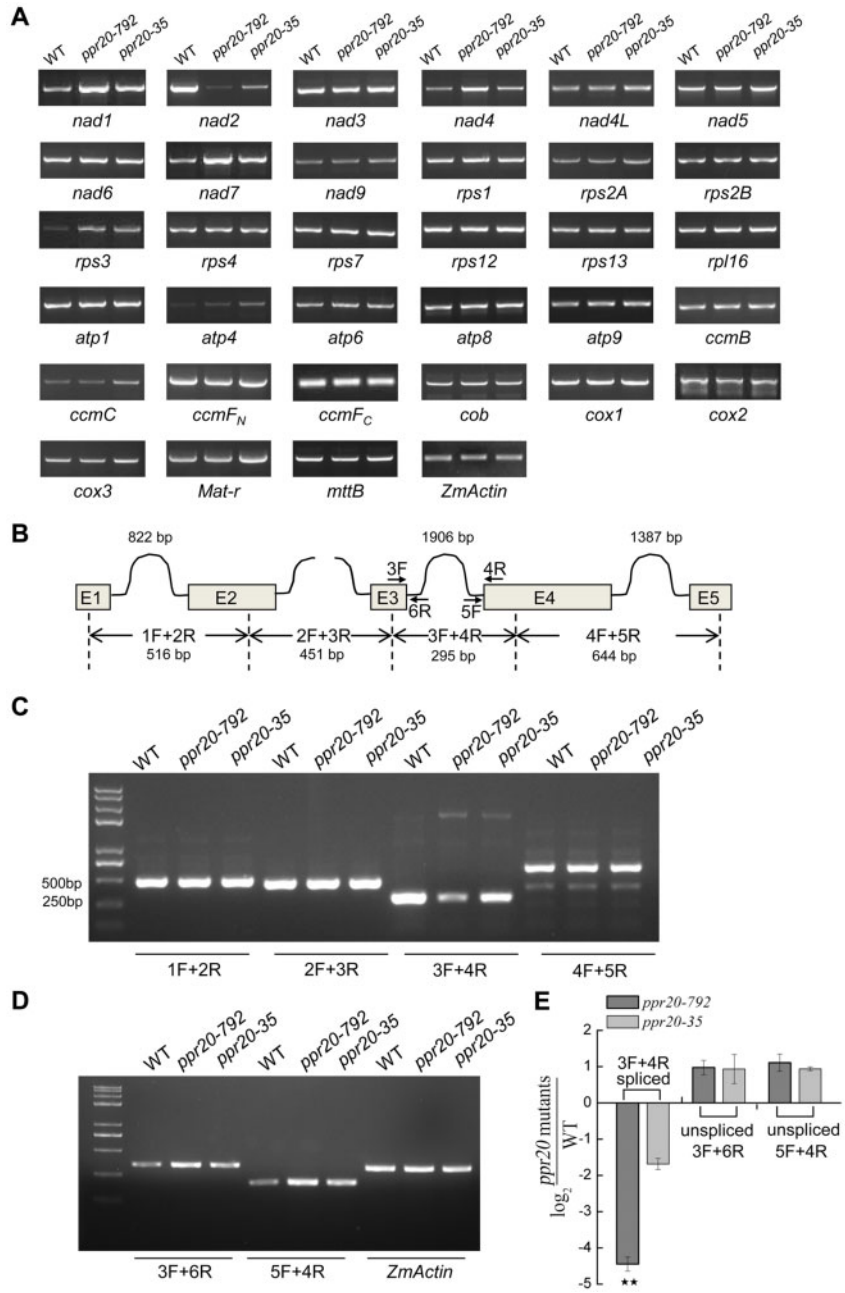

Fig. 4 The splicing of *nad2* intron 3 is impaired in the *ppr20* mutants. (A) RT-PCR analysis of 33 mitochondrion-encoded transcripts in WT, *ppr20-792* and *ppr20-35* kernels. The primers used in the analysis were designed to amplify entire coding region of each transcript. (B) Schematic representation of maize mitochondrial *nad2* gene. F and R present the primers used to detect splicing events. (C) RT-PCR analysis of the splicing of four introns in *nad2* pre-mRNA in WT and *ppr20* mutants. (D) Detection of the unspliced *nad2* intron 3 in WT and *ppr20* mutants by RT-PCR. (E) Detection of the spliced *nad2* and unspliced by qRT-PCR. Values are the mean ± SE of three biological repeats (**$P < 0.01$, Student's *t*-test).

and a peripheral arm. As Nad2 is a central component of the membrane arm, the decrease in *nad2* transcript level might affect the assembly and activity of mitochondrial complex I (Subrahmanian et al. 2016). To assess complex I and other complexes, crude mitochondria were isolated from WT and *ppr20-792* mutant kernels and the mitochondrial complexes were analyzed by blue native polyacrylamide gel electrophoresis (BN-PAGE). As shown in **Fig.6A**, complex I and supercomplex I + III_2 were dramatically decreased in the *ppr20-792* mutant compared with WT. In-gel staining of the NADH dehydrogenase activity confirmed that the bands corresponding to complex I and supercomplex I + III_2 were almost undetectable in the *ppr20-792* mutant (**Fig. 6B**). In addition, the level of alternative oxidase (AOX), a central component of the alternative respiratory pathway activated by mitochondrial respiration deficiency, was strongly induced in the *ppr20-792* mutant (**Fig. 6C**). These results indicate that the defect in the splicing of *nad2* intron 3 causes the dysfunction of mitochondrial complex I in the *ppr20-792* mutant.

Furthermore, we detected the accumulation of mitochondrial proteins forming respiratory chain complexes by Western

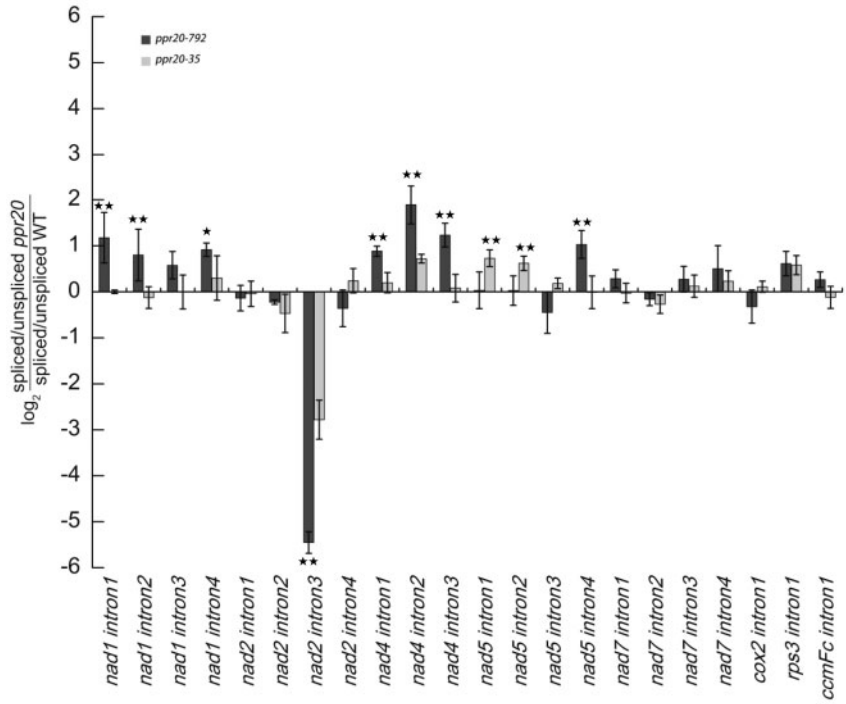

Fig. 5 Splicing efficiency of *nad2* intron 3 is decreased in the *ppr20* mutants. Quantitative RT-PCR analysis of splicing efficiency of 22 intron events of maize mitochondrial genes in the WT and *ppr20* mutants. *ZmActin* was used as an internel reference. Values are the mean ± SE of three biological repeats (**$P < 0.01$, *$P < 0.05$, Student's *t*-test).

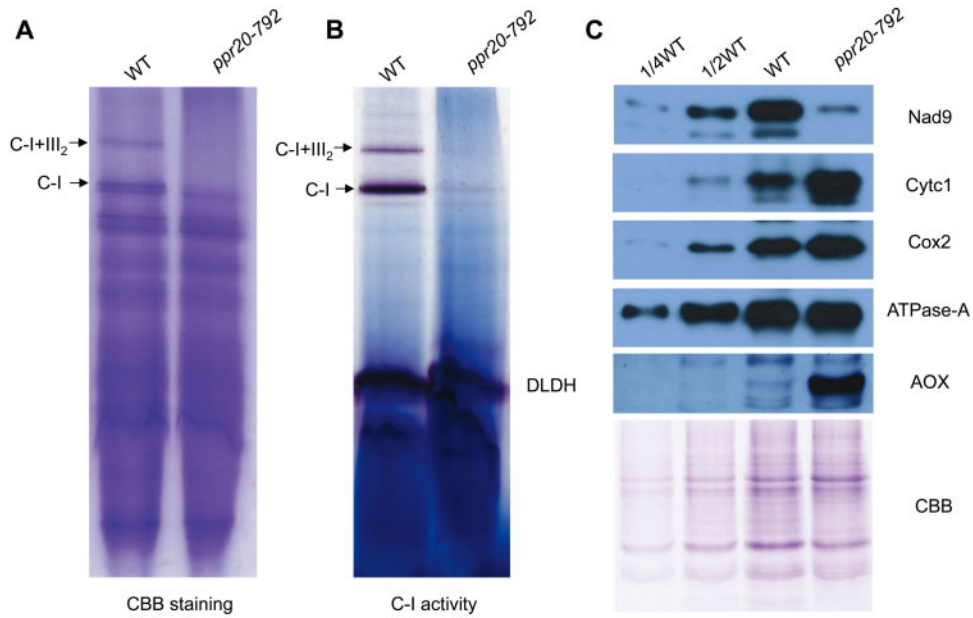

Fig. 6 Assembly and activity of mitochondrial complex I are reduced in the *ppr20-792* mutant. (A) BN-PAGE analysis of mitochondrial complex assembly. (B) In-gel analysis of mitochondrial complex I activity. The activity of dihydrolipoamide dehydrogenase (DLDH) is used as a loading control. C-I, complex I; C-I + III_2, supercomplex I + III_2. (C) Accumulation of mitochondrial proteins in the *ppr20-792* mutant. Crude mitochondrial proteins were subjected to Western blotting with specific antibodies against Nad9, Cytc1, Cox2, ATPase-A and AOX.

blot with antibodies against Nad9 (complex I), Ctyc1 (complex III), Cox2 (complex IV) and ATPase-A (complex V). The level of Nad9 was dramatically decreased in the *ppr20-792* mutant, whereas the levels of Cox2 and Ctyc1 were increased (**Fig. 6C**). There was no significant difference in the level of ATPase-A between the WT and *ppr20-792* mutant (**Fig. 6C**). These results indicate that the deficiency of Nad2 may negatively affect the accumulation of the complex I component but positively regulate the accumulation of the components from complexes III and IV.

The mutation in *ppr20-35* is leaky

The *Mu* insertion in *ppr20-35* was expected to completely disrupt the PPR20 function because the insertion is in the N-terminal region and would eliminate the mitochondrial targeting peptide. However, the less severe seed phenotype and reduced impact on *nad2* intron 3 splicing than the *ppr20-792* mutation point to a possibility of leaky mutation. We investigated the possibility whether a functional PPR20 protein can be produced in the *ppr20-35* mutant. Extensive RT-PCR analysis failed to detect any full-length *PPR20* transcripts in *ppr20-35*, ruling out the possibility of a full-length protein (**Supplementary Fig. S3**). Then, we examined the 5′ ends of *PPR20* transcripts in the mutant by 5′-rapid amplification of cDNA ends (5′-RACE) (**Supplementary Fig. S4**). 5′-RACE mapped the 5′ end at 41 nt downstream of the translation start codon (**Fig. 7A**). An in-frame ATG was identified at 360 nt from the original ATG in the WT *PPR20* transcript (**Fig. 7A**). This transcript encodes a truncated PPR20 protein with an N-terminal truncation of 120 amino acids (PPR20$^{\Delta N120}$), eliminating a predicted mitochondrial signal peptide and two PPR motifs (**Fig. 7A**). For this protein to be functional, it has to be targeted to the mitochondrion. We tested the subcellular localization of this truncated protein by PPR20$^{\Delta N120}$:GFP fusion (the full-length PPR20 without N-terminal 120 amino acids fused to GFP) expression in Arabidopsis protoplasts. Surprisingly, GFP signals were observed in mitochondria, which were marked by MitoTracker Red staining (**Fig. 7B**), indicating the truncated protein gaining a new mitochondrial signal peptide.

This surprising finding promoted us to investigate the transcripts in the WT kernels. 5′-RACE analysis resulted a single band (**Supplementary Fig. S4**), but cloning and sequencing of this band uncovered two transcripts, one with a 5′ end at −45 nt and the other at +24 nt from the *PPR20* ATG, respectively (**Fig. 7C**). Of the 12 random clones examined, nine clones initiate at −45 nt and three clones at +24 nt. The former is predicted to encode a full-length PPR20 protein, while the latter is predicted to encode exactly the PPR20$^{\Delta N120}$ that was detected in the *ppr20-35* mutant (**Fig. 7C**). We tried to prepare an anti-PPR20 antibody to prove the presence of two PPR20 isoforms in WT, unfortunately failed.

We further analyzed the expression of *PPR20* in WT and the *ppr20-35* mutant by qPCR. Consistently, the truncated transcript is expressed at a level approximately 1.6-fold higher in *ppr20-35* than in the WT (the full-length and truncated *PPR20* combined, **Fig. 7D**). Together, these results suggest that the truncated protein PPR20$^{\Delta N120}$ with mitochondrion-targeting capability could be produced in the *ppr20-35*. The less severe seed phenotype of *ppr20-35* mutant suggests that this truncated protein is compromised but still functional.

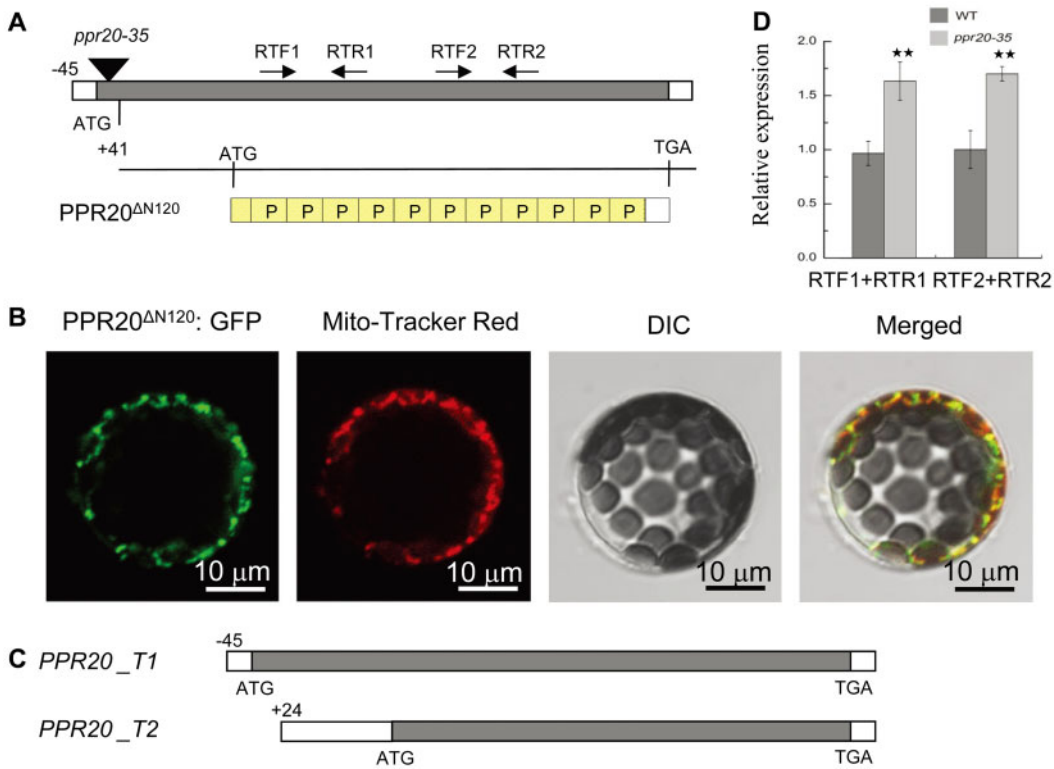

Fig. 7 The *ppr20-35* mutant may encode a truncated PPR20$^{\Delta N120}$ with partial splicing function. (A) The transcription start site of *PPR20* transcript in *ppr20-35* mutant mapped by 5′-RACE. This transcript encodes PPR20$^{\Delta N120}$ with 11 PPR motifs. (B) The subcellular localization of PPR20$^{\Delta N120}$ in Arabidopsis protoplasts. DIC, differential interference contrast. (C) Mapping of the transcription initiation site of *PPR20* transcript in WT. This approach identified two transcripts that were initiated from 45 nt upstream of start codon and 24 nt downstream of start codon. (D) The expression level of *PPR20* transcripts in the WT and *ppr20-35* mutant. Values are the mean ±SE of three biological repeats (**$P < 0.01$, Student's *t*-test).

Zm_mTERF15 is required for the *cis*-splicing of *nad2* intron 3

As Arabidopsis mTERF15 is a splicing factor that exclusively affects the splicing of *nad2* intron 3, we suppose that its ortholog in maize may has the same function. The most similar protein of At_mTERF15 in maize is GRMZM2G177019 (thereafter named Zm_mTERF15) with a 30% identity, and mTERF domains of two proteins share a 52% identity (**Supplementary Fig. S5**). A *zm_mterf15*/+*Mu* insertion line was isolated from UniformMu population, and the insertion was verified by PCR at 918 nt downstream of the translation start codon (**Fig. 8A**). Selfed heterozygotes of *zm_mterf15*/+ produced small kernel mutants (**Fig. 8B, C**). To determinate the function of Zm_mTERF15, the transcription levels of 33 mitochondrial genes were evaluated by RT-PCR. Among these genes, the *nad2* transcript was dramatically decreased in the *zm_mterf15* mutants, whereas other transcripts did not show significant difference (**Fig. 8D**). Further analysis showed that the *cis*-splicing of *nad2* intron 3 was impaired in the mutants (**Fig. 8E**). These results indicate that the orthologous proteins At_mTERF15 and Zm_mTERF15 have a similar function in the *cis*-splicing of *nad2* intron 3. Although we did not isolate a second allele of *Zm_mTERF15*, the similarity of molecular function and amino acid sequence between At_mTERF15 and Zm_mTERF15 suggests that the small kernel phenotype is caused by the disruption of *Zm_mTERF15*.

Discussion

Through the molecular analyses, we have provided compelling evidence that PPR20, a P-type PPR protein, is required for the *cis*-splicing of *nad2* intron 3 in maize mitochondria. This conclusion is supported by the deficiency of the splicing in two alleles of the *ppr20* mutants and further corroborated by the deficiency in mitochondrial complex I assembly and activity. Because of the essential role of mitochondria to cellular activity, it provides a plausible explanation for the defective seed phenotype in maize. In addition, we characterized the maize ortholog of At_mTERF15 and revealed a conserved function in the *cis*-splicing of mitochondrial *nad2* intron 3. This study raised the following issues.

The function of PPR20 in the *cis*-splicing of *nad2* intron 3

RT-PCR and qRT-PCR data showed that the spliced *nad2* exon 3–exon 4 was decreased in *ppr20* mutants, and the unspliced pre-RNA containing intron 3 was increased (**Fig. 4**), indicating that PPR20 is required for the splicing of *nad2* intron 3.

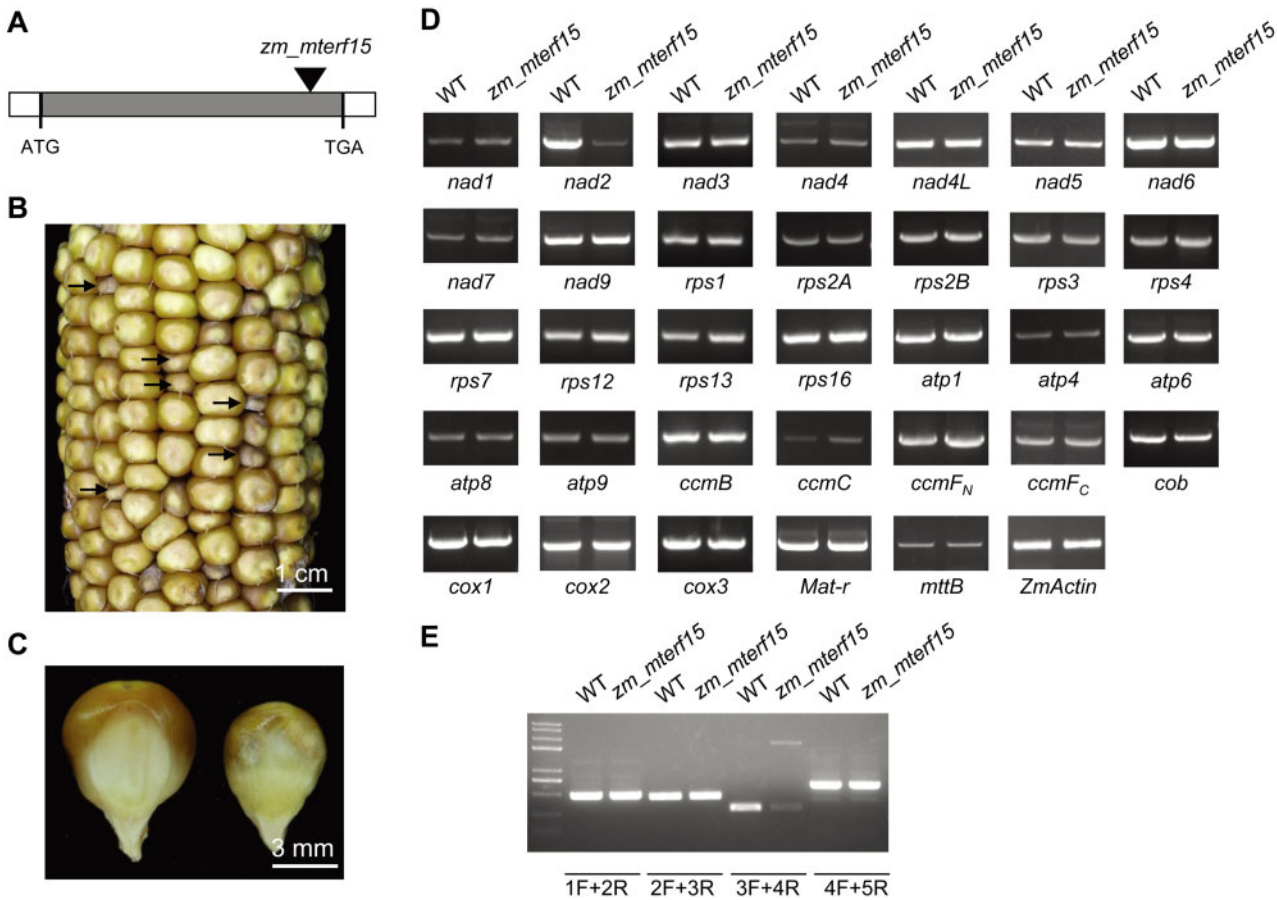

Fig. 8 *Zm_mTERF15* is required for the splicing of *nad2* intron 3 in maize. (A) Gene structure of *Zm_mTERF15* and *Mu* insertion position. (B) Ears segregating *zm_mterf15* small kernel mutants. Arrows indicate the mutant kernels. (C) Germinal face of WT (left) and *zm_mterf15* (right) kernels. Scale bar, 3 mm. (D) RT-PCR analysis of thetranscript levels of intron-containing genes in WT and *zm_mterf15*. (E) RT-PCR analysis of the splicing of four introns in *nad2* pre-mRNA in *zm_mterf15* mutants.

However, the spliced *nad2* exon 3–exon 4 was decreased approximately 16-fold in the *ppr20-792* mutant and the unspliced *nad2* exon 3–intron 3–exon 4 was increased only 2-fold. This phenomenon was also found in another *nad2* intron 3 splicing defective mutant *abo5*. In *abo5* mutant, the spliced *nad2* exon 3–exon 4 was almost undetectable, whereas the intron 3 transcripts were merely accumulated 2-fold compared with WT (Liu et al. 2010). This evidence suggests that pre-RNA containing *nad2* intron 3 may be unstable in these mutants and PPR20 may play a role in the stabilization of *nad2* intron 3 containing pre-RNA.

Multiple factors are required for the splicing of *nad2* intron 3

In addition to PPR20 and Zm_mTERF15 identified in this study, several other proteins have been reported to participate in the splicing of *nad2* intron 3. ABO5, a P-type PPR protein, is required for the *cis*-splicing of *nad2* intron 3 in Arabidopsis (Liu et al. 2010; Hsu et al. 2014). ABO5 (AT1G46210) is not the Arabidopsis ortholog of PPR20 as these two shares a mere 20% identity in protein sequences. The likely ortholog of PPR20 in Arabidopsis is AT1G09820 with a 36% identity. RUG3, an RCC motif-containing protein, is required for the splicing of both *nad2* introns 2 and 3 (Kuhn et al. 2011). The CRM protein mCSF1 and the DEAD-box proteins PMH2 and ABO6 are essential for the splicing of many introns in mitochondria, including *nad2* intron 3 (Kohler et al. 2010, He et al. 2012, Zmudjak et al. 2013). Clearly, these proteins belong to different families, but they are RNA-binding proteins with known RNA-binding domains, such as the RCC and CRM domains. The CRM and DEAD-box proteins appear to play a general role in the intron splicing as they are involved in multiple introns. Whereas PPR20, ABO5 and Zm_mTERF15 are specifically required for *nad2* intron 3 splicing, implying that these three proteins may confer the specificity of substrate recognition.

Although how these different proteins coordinate in the splicing is unclear, multiple lines of evidence suggest that the splicing of plant organellar introns appears to be mediated by protein complexes. RNA helicase and mTERF proteins are found in high molecular mass protein complexes (Matthes et al. 2007, Hammani and Barkan 2014). Similarly, PPR proteins are found in protein complexes in the molecular mass range between 120 and 1,000 kDa in mitochondria (Senkler et al. 2017). In particular, Zm_mTERF4 coimmunoprecipitates with several chloroplast splicing factors, including CRM domain proteins CAF2, CFM2, CFM3 and CRS1, PORR protein WTF1 and PPR protein THA8 (Hammani and Barkan 2014). As PPR20 and Zm_mTERF15 are specifically required for the splicing of the same intron, it is possible that these two proteins interact physically. However, we failed to detect such interaction in yeast two-hybrid analyses (**Supplementary Fig. S6**). It is possible that these proteins indeed interact during intron splicing, but rather weakly or transiently that the interaction escapes the detection by the yeast two-hybrid system. It is equally possible that these proteins might cooperate in a complex without direct interaction. It is also possible that these proteins facilitate the splicing by probably binding to *nad2* intron 3 at different sites, not through protein–protein interaction with each other. In any case, further studies are necessary to elucidate the nature of the intron-splicing machinery in organelles.

PPR20 is essential for mitochondrial complex I assembly and seed development in maize

Complex I assembly involves the sequential addition of >40 subunits and the sequential formation of 200, 400, 450, 650 and 850 kDa subcomplexes and 1,000 kDa mature complex (Subrahmanian et al. 2016). Nad2 is proposed to be added at the early stage of assembly, and the lack of Nad2 leads to the loss of the assembly and activity of complex I in higher plants (Subrahmanian et al. 2016). In *at_mterf15* mutant, an intron-splicing defect in *nad2* intron 3 leads to reduced complex I assembly and activity (Hsu et al. 2014). Arabidopsis RUG3 is required for the splicing of *nad2* introns 2 and 3 (Kuhn et al. 2011). The *rug3* mutant shows greatly reduced complex I abundance and activity (Kuhn et al. 2011). In maize *emp10* and *emp16* mutants, the defects in the splicing of *nad2* introns impair the formation and activity of complex I (Xiu et al. 2016, Cai et al. 2017). In this study, we showed that abundance and activity of complex I are dramatically decreased in *ppr20-792* (**Fig. 6A, B**). In addition, we found that the level of Nad9 is greatly decreased (**Fig. 6C**). Nad9 is located in the matrix arm of the L-shaped complex I and is added after Nad2 is assembled (Subrahmanian et al. 2016). The accumulation of Nad9 is affected in Nad2-deficient mutants, such as *emp16* (Xiu et al. 2016), *at_mterf15* (Hsu et al. 2014) and *rug3* (Kuhn et al. 2011). We propose that the deficiency in Nad2 affects the assembly of subcomplex at the early stage and hence blocks the assembly of Nad9, leading to the degradation of the unassembled Nad9, which is free in the mitochondrial matrix. As only trace amounts of active complex I can be detected in *ppr20-792* mutant (**Fig. 6**), we conclude that Nad2 is essential for mitochondrial complex I assembly.

Given the function of mitochondria as the powerhouse of the cell, it is reasonable to conclude that the defect in seed development in *ppr20* mutants is caused by the failed mitochondrial respiratory chain. In maize, mutants blocking the assembly of complex I arrest seed development. Loss of *Emp10* and *Emp16* functions arrests embryogenesis and endosperm development (Xiu et al. 2016, Cai et al. 2017). Maize *dek2* mutant shows reduced splicing efficiency of mitochondria *nad1* intron 1 and small kernel phenotype (Qi et al. 2017). The decrease in the splicing efficiency of *nad4* intron 1 in maize *dek35* mutant causes a lethal seed (Chen et al. 2017). Our results highlight the function of PPR20 in *nad2* intron 3 splicing, mitochondrial complex I assembly and seed development in maize.

The *ppr20-35* mutant is a leaky mutation

Surprisingly, the *Mu* insertion in the *ppr20-35* mutant does not abolish the PPR20 function, but only impairs partial function. Molecular analysis indicates that the *ppr20-35* mutant allows the expression of a truncated PPR20$^{\Delta N120}$ with the deletion of the N-terminal signal peptide and two PPR motifs (**Fig. 7A**). It is possible that removal of signal peptide and two PPR motifs does not completely eliminate mitochondrial target ability and splicing function of PPR20$^{\Delta N120}$. Although no typical signal peptide was predictable, transient expression experiment showed that

PPR20$^{\Delta N120}$ can be translocated into mitochondria (**Fig. 7B**). Mitochondrial proteins are mainly imported by a post-translational pathway in which mitochondrial signal sequences are recognized by receptors on the mitochondrial surface (Schleiff and Becker 2011). The signal sequences are enriched in positively charged and hydrophobic amino acid residues, forming an amphipathic α-helix (Schleiff and Becker 2011). Analysis of N-terminus of PPR20$^{\Delta N120}$ indicated that the sequence is enriched in those residues (**Supplementary Fig. S7**). In addition, each PPR motif forms a hairpin of α-helices (Yin et al. 2013). Therefore, it is likely that the N-terminal sequence of PPR20$^{\Delta N120}$ can be recognized and imported into mitochondria. However, it is difficult to evaluate the translocation efficiency by transient expression experiment.

The N-terminal motifs of PPR protein may not essential for its function. In our previous study, we found that the deletion of N-terminal two PPR motifs does not completely eliminate the editing function of an E+ class PPR protein, EMP9 (Yang et al. 2017). In this study, the accumulation of PPR20$^{\Delta N120}$ protein in the *ppr20-35* mutant was likely higher than that of PPR20 in WT (**Fig. 7**), whereas the *ppr20-35* mutant showed defects in molecular function and physiologic phenotype, suggesting that the lack of N-terminal two PPR motifs only impairs partial function of PPR20. Thus, the leaky mutation in the *ppr20-35* mutant is likely caused by the lower efficiency of PPR20$^{\Delta N120}$ translocation or/and the deletion of two N-terminal PPR motifs.

Materials and Methods

Plant materials

The *ppr20-792* (UFMu-02628) and *ppr20-35* (UFMu-09213) mutant lines were obtained from the Maize Genetics Cooperation Stock Center. The *Mu* insertion in each line was verified by PCR with *Mu*-specific primer TIR8 and *PPR20*-specific primers PPR20-R4 (for *ppr20-35*) and PPR20-F3 (for *ppr20-792*). The insertion sites were identified by the sequencing of specific PCR products. All plants were grown under natural conditions in the Shandong University experimental field.

Subcellular localization

To generate a PPR20:GFP fusion (the full-length PPR20 fused to GFP), the full-length coding region of PPR20 without stop codon was amplified by primers PPR20-F2 and PPR20-R2 on cDNA from maize inbred line B73 and then cloned into the pBI221 vector. To generate a PPR20$^{\Delta N120}$:GPF fusion, the sequence encoding C-terminal 454 amino acids of PPR20 was amplified by primers PPR20-CACC-F1 and PPR20-R2 and then cloned into the pBI221 vector. The resulting constructs were introduced into Arabidopsis protoplasts by polyethylene glycol-mediated transformation. Arabidopsis protoplasts were isolated from leaves as described previously (Yoo et al. 2007). The signals were observed under a confocal laser scanning microscope after 20 h incubation at 24°C. To label mitochondria, the protoplasts were stained with MitoTracker Red (Thermo Fisher Scientific, Waltham, MA, USA) for 30 min at 37°C.

5′-RACE

5′-RACE was performed on WT and *ppr20-35* mRNA using the SMART RACE cDNA amplification kit (Clontech, Mountain View, CA, USA). The 5′-RACE cDNA products were cloned into the pGEM-T vector (Promega, Mannheim, Germany) and sequenced. The primers used in 5′-RACE are shown in **Supplementary Table S1**.

RNA extraction and qRT-PCR

Total RNA was extracted from fresh kernels using the PureLink RNA Mini Kit (Ambion Carlsbad, CA, USA) following manufacturer's instruction. Prior to reverse transcription, RNA was treated with DNase I (New England Biolabs, Ipswich, MA, USA) to eliminate DNA contamination. First-strand cDNA was synthesized from 2 μg of total RNA with random hexamers. RT-PCR analysis of 33 mitochondrial transcripts was performed as described preciously (Xiu et al. 2016). qRT-PCR was performed using the Light Cycler Fast Start DNA Master SYBR Green I Kit (Roche, Mannheim, Germany). The relative quantification of gene expression was calculated with the $2^{-\Delta\Delta Ct}$ method. *ZmActin* was used as a control for RNA normalization in RT-PCR and qRT-PCR analysis. Primer sequences were listed in **Supplementary Table S1** or published as descried preciously (Yang et al. 2017).

BN-PAGE and complex I in-gel assay

Crude mitochondria were isolated from maize kernels as described (Xiu et al. 2016). Mitochondrial proteins were solubilized with 1% dodecyl maltoside (β-DM), and the native mitochondrial complexes were subjected to a gradient NativePAGE Bis-Tris gel (Thermo Fisher Scientific). To observe mitochondrial complexes, the gel was stained with Coomassie Brilliant Blue and then destained with 25% ethanol and 8% acetic acid. To analyze complex I in-gel activity, the gel was immersed in the assay buffer (0.1 M Tris/HCl, pH 7.4, 0.14 mM NADH and 1.22 mM nitrotetrazolium blue). The reaction was stopped with 40% methanol and 10% acetic acid.

Western blotting

Mitochondrial proteins were extracted from crude mitochondria and separated by SDS–PAGE. The separated proteins were transferred to a nitrocellulose membrane and then incubated with specific antibodies against Nad9, Cox2, ATPase-A, Cytc1 and AOX, followed by a Horseradish Peroxidase-conjugated goat anti-rabbit IgG secondary antibody. Protein concentration was quantified by the Protein Assay reagent (Bio-Rad, CA, USA). Each antibody was used at a 1:5,000 dilution.

Cytological sections

Immature *ppr20-792* mutant kernels and their WT siblings were isolated from the self-pollinated ear of a heterozygous plant at 13 DAP. The kernels were fixed in a 4% paraformaldehyde solution, dehydrated in a 50–100% ethanol gradient, embedded in paraffin, sectioned at 8 μm with a microtome, stained with Johansen's Safranin O and observed under a Zeiss microscope.

Supplementary Data

Supplementary data are available at PCP online.

Acknowledgments

We thank Dr. Tsuyoshi Nakagawa (Shimane University, Japan) for the pGWB vectors and the Maize Genetics Cooperation Stock Center for the seed stocks.

Funding

National Natural Science Foundation of China [Project Nos. 91735301, 31630053 and 91435201] to B.-C.T.; Natural Science Foundation of Shandong Province [Project No. ZR2019MC005] to Y.-Z.Y.; and Fundamental Research Funds of Shandong University [Project No. 2017HW0013] to Y.-Z.Y.

Disclosures

No conflicts of interest declared.

References

Barkan, A., Rojas, M., Fujii, S., Yap, A., Chong, Y.S., Bond, C.S., et al. (2012) A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet.* 8: e1002910.

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

Binder, S. and Brennicke, A. (2003) Gene expression in plant mitochondria: transcriptional and post-transcriptional control. *Phil. Trans. R. Soc. Lond. B* 358: 181–189.

Bonen, L. (2008) *Cis*- and *trans*-splicing of group II introns in plant mitochondria. *Mitochondrion* 8: 26–34.

Cai, M., Li, S., Sun, F., Sun, Q., Zhao, H., Ren, X., et al. (2017) *Emp10* encodes a mitochondrial PPR protein that affects the *cis*-splicing of *nad2* intron 1 and seed development in maize. *Plant J.* 91: 132–144.

Chen, X., Feng, F., Qi, W., Xu, L., Yao, D., Wang, Q., et al. (2017) *Dek35* encodes a PPR protein that affects *cis*-splicing of mitochondrial *nad4* intron 1 and seed development in maize. *Mol. Plant* 10: 427–441.

Cheng, S., Gutmann, B., Zhong, X., Ye, Y., Fisher, M.F., Bai, F., et al. (2016) Redefining the structural motifs that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants. *Plant J.* 85: 532–547.

Clifton, S.W., Minx, P., Fauron, C.M.R., Gibson, M., Allen, J.O., Sun, H., et al. (2004) Sequence and comparative analysis of the maize NB mitochondrial genome. *Plant Physiol.* 136: 3486–3503.

Cohen, S., Zmudjak, M., Colas, D. F.-S C., Malik, S., Shaya, F., Keren, I., et al. (2014) nMAT4, a maturase factor required for *nad1* pre-mRNA processing and maturation, is essential for holocomplex I biogenesis in Arabidopsis mitochondria. *Plant J.* 78: 253–268.

Duchene, A.M., Pujol, C. and Marechal-Drouard, L. (2009) Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. *Curr. Genet.* 55: 1–18.

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.

Francs-Small, C.C., Kroeger, T., Zmudjak, M., Ostersetzer-Biran, O., Rahimi, N., Small, I., et al. (2012) A PORR domain protein required for *rpl2* and *ccmFC* intron splicing and for the biogenesis of *c*-type cytochromes in Arabidopsis mitochondria. *Plant J.* 69: 996–1005.

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.

He, J., Duan, Y., Hua, D., Fan, G., Wang, L., Liu, Y., et al. (2012) DEXH box RNA helicase-mediated mitochondrial reactive oxygen species production in Arabidopsis mediates crosstalk between abscisic acid and auxin signaling. *Plant Cell* 24: 1815–1833.

Hsieh, W.Y., Liao, J.C., Chang, C.Y., Harrison, T., Boucher, C. and Hsieh, M.H. (2015) The SLOW GROWTH3 pentatricopeptide repeat protein is required for the splicing of mitochondrial NADH dehydrogenase subunit7 intron 2 in Arabidopsis. *Plant Physiol.* 168: 490–501.

Hsu, Y.W., Wang, H.J., Hsieh, M.H., Hsieh, H.L. and Jauh, G.,Y. (2014) Arabidopsis mTERF15 is required for mitochondrial *nad2* intron 3 splicing and functional complex I activity. *PLoS One* 9: e112360.

Keren, I., Bezawork-Geleta, A., Kolton, M., Maayan, I., Belausov, E., Levy, M., et al. (2009) AtnMat2, a nuclear-encoded maturase required for splicing of group-II introns in Arabidopsis mitochondria. *RNA* 15: 2299–2311.

Keren, I., Tal, L., Des Francs-Small, C.C., Araujo, W.L., Shevtsov, S., Shaya, F., et al. (2012) nMAT1, a nuclear-encoded maturase involved in the trans-splicing of *nad1* intron 1, is essential for mitochondrial complex I assembly and function. *Plant J.* 71: 413–426.

Kohler, D., Schmidt-Gattung, S. and Binder, S. (2010) The DEAD-box protein PMH2 is required for efficient group II intron splicing in mitochondria of *Arabidopsis thaliana*. *Plant Mol. Biol.* 72: 459–467.

Kuhn, K., Carrie, C., Giraud, E., Wang, Y., Meyer, E.H., Narsai, R., et al. (2011) The RCC1 family protein RUG3 is required for splicing of *nad2* and complex I biogenesis in mitochondria of *Arabidopsis thaliana*. *Plant J.* 67: 1067–1080.

Lambowitz, A.M. and Zimmerly, S. (2004) Mobile group II introns. *Annu. Rev. Genet.* 38: 1–35.

Li, X., Gu, W., Sun, S., Chen, Z., Chen, J., Song, W., et al. (2018) Defective Kernel 39 encodes a PPR protein required for seed development in maize. *J. Integr. Plant Biol.* 60: 45–64.

Liu, Y., He, J., Chen, Z., Ren, X., Hong, X. and Gong, Z. (2010) *ABA overly-sensitive 5* (*ABO5*), encoding a pentatricopeptide repeat protein required for *cis*-splicing of mitochondrial *nad2* intron 3, is involved in the abscisic acid response in Arabidopsis. *Plant J.* 63: 749–765.

Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., et al. (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* 16: 2089–2103.

Mackenzie, S. and McIntosh, L. (1999) Higher plant mitochondria. *Plant Cell* 11: 571–586.

Matthes, A., Schmidt-Gattung, S., Kohler, D., Forner, J., Wildum, S., Raabe, M., et al. (2007) Two DEAD-box proteins may be part of RNA-dependent high-molecular-mass protein complexes in Arabidopsis mitochondria. *Plant Physiol.* 145: 1637–1646.

McCarty, D.R., Settles, A.M., Suzuki, M., Tan, B.C., Latshaw, S., Porch, T., et al. (2005) Steady-state transposon mutagenesis in inbred maize. *Plant J.* 44: 52–61.

Pyle, A.M. (2016) Group II intron self-splicing. *Annu. Rev. Biophys.* 45: 183–205.

Qi, W., Yang, Y., Feng, X., Zhang, M. and Song, R. (2017) Mitochondrial function and maize kernel development requires *Dek2*, a pentatricopeptide repeat protein involved in *nad1* mRNA splicing. *Genetics* 205: 239–249.

Schleiff, E. and Becker, T. (2011) Common ground for protein translocation: access control for mitochondria and chloroplasts. *Nat. Rev. Mol. Cell Biol.* 12: 48–59.

Senkler, J., Senkler, M., Eubel, H., Hildebrandt, T., Lengwenus, C., Schertl, P., et al. (2017) The mitochondrial complexome of *Arabidopsis thaliana*. *Plant J.* 89: 1079–1092.

Subrahmanian, N., Remacle, C. and Hamel, P.P. (2016) Plant mitochondrial complex I composition and assembly: a review. *BBA Bioenergetics* 1857: 1001–1014.

Sultan, L.D., Mileshina, D., Grewe, F., Rolle, K., Abudraham, S., Głodowicz, P., et al. (2016) The reverse transcriptase/RNA maturase protein MatR is required for the splicing of various group II introns in Brassicaceae mitochondria. *Plant Cell* 28: 2805–2829.

Takenaka, M., Zehrmann, A., Brennicke, A. and Graichen, K. (2013) Improved computational target site prediction for pentatricopeptide repeat RNA editing factors. *PLoS One* 8: e65343.

Xiu, Z., Sun, F., Shen, Y., Zhang, X., Jiang, R., Bonnard, G., et al. (2016) EMPTY PERICARP16 is required for mitochondrial *nad2* intron 4 *cis*-splicing, complex I assembly and seed development in maize. *Plant J.* 85: 507–519.

Yagi, Y., Hayashi, S., Kobayashi, K., Hirayama, T. and Nakamura, T. (2013) Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. *PLoS One* 8: e57286.

Yang, Y.Z., Ding, S., Wang, H.C., Sun, F., Huang, W.L., Song, S., et al. (2017) The pentatricopeptide repeat protein EMP9 is required for mitochondrial *ccmB* and *rps4* transcript editing, mitochondrial complex biogenesis and seed development in maize. *New Phytol.* 214: 782–795.

Yin, P., Li, Q., Yan, C., Liu, Y., Liu, J., Yu, F., et al. (2013) Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* 504: 168–171.

Yoo, S.D., Cho, Y.H. and Sheen, J. (2007) Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* 2: 1565–1572.

Zmudjak, M., Colas, D. F.-S C., Keren, I., Shaya, F., Belausov, E., Small, I., et al. (2013) mCSF1, a nucleus-encoded CRM protein required for the processing of many mitochondrial introns, is involved in the biogenesis of respiratory complexes I and IV in Arabidopsis. *New Phytol.* 199: 379–394.