



The dicot homolog of maize PPR103 carries a C-terminal DYW domain and may have a role in C-to-U editing of some chloroplast RNA transcripts

Tyra N. McCray^{1,2} · Mohammad F. Azim^{2,3} · Tessa M. Burch-Smith^{1,2,3}

Received: 10 February 2023 / Accepted: 30 January 2024
© The Author(s) 2024

Abstract

In plants, cytidine-to-uridine (C-to-U) editing is a crucial step in processing mitochondria- and chloroplast-encoded transcripts. This editing requires nuclear-encoded proteins including members of the pentatricopeptide (PPR) family, especially PLS-type proteins carrying the DYW domain. *IPI1/emb175/PPR103* is a nuclear gene encoding a PLS-type PPR protein essential for survival in *Arabidopsis thaliana* and maize. Arabidopsis IPI1 was identified as likely interacting with ISE2, a chloroplast-localized RNA helicase associated with C-to-U RNA editing in Arabidopsis and maize. Notably, while the Arabidopsis and *Nicotiana* IPI1 orthologs possess complete DYW motifs at their C-termini, the maize homolog, ZmPPR103, lacks this triplet of residues which are essential for editing. In this study we examined the function of IPI1 in chloroplast RNA processing in *N. benthamiana* to gain insight into the importance of the DYW domain to the function of the EMB175/PPR103/ IPI1 proteins. Structural predictions suggest that evolutionary loss of residues identified as critical for catalyzing C-to-U editing in other members of this class of proteins, were likely to lead to reduced or absent editing activity in the *Nicotiana* and Arabidopsis IPI1 orthologs. Virus-induced gene silencing of *NbIPI1* led to defects in chloroplast ribosomal RNA processing and changes to stability of *rpl16* transcripts, revealing conserved function with its maize ortholog. *NbIPI1*-silenced plants also had defective C-to-U RNA editing in several chloroplast transcripts, a contrast from the finding that maize PPR103 had no role in editing. The results indicate that in addition to its role in transcript stability, NbIPI1 may contribute to C-to-U editing in *N. benthamiana* chloroplasts.

Key message

The *Nicotiana benthamiana* DYW PPR protein NbIPI1 possess an intact C-terminal DYW domain and stabilizes the *rpl16-rpl14* transcript like its maize ortholog PPR103, and may also contribute to C-to-U RNA editing of some chloroplast transcripts.

Keywords Chloroplast gene expression · C-to-U editing · DYW domain · *Nicotiana benthamiana* · Pentatricopeptide repeat · PPR103 · RNA editing

Introduction

In land plant organelles, post-transcriptional processing of RNA transcripts is a crucial regulatory point for gene expression. One step of post-transcriptional RNA processing is the site-specific deamination of cytidine to uridine, called C-to-U editing. In most land plants this C-to-U RNA editing occurs within a subset of chloroplast and mitochondrial transcripts. Editing within these transcripts often produces changes in splice sites, amino acid substitutions and the addition of start and stop codons. These modifications yield

✉ Tessa M. Burch-Smith
tburch-smith@danforthcenter.org

¹ School of Genome Science and Technology, University of Tennessee, Knoxville, TN 37996, USA

² Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996, USA

³ Donald Danforth Plant Science Center, St. Louis, MO 63132, USA

transcripts that can subsequently be translated to produce proteins essential for photosynthesis and for mitochondrial function (Small et al. 2020).

C-to-U editing involves numerous nucleus-encoded proteins including several families of RNA binding proteins. The largest family of plant proteins with roles in post-transcriptional processing and C-to-U editing of organelle transcripts is the pentatricopeptide repeat (PPR) protein family. PPR proteins are site-specific RNA-binding proteins that have critical and diverse functions (Saha et al. 2007; Schmitz-Linneweber and Small 2008; Barkan and Small 2014; Small et al. 2020). Loss of a single PPR protein can result in embryonic arrest or severe developmental defects, demonstrating their fundamental importance to plant survival and development (Barkan and Small 2014). Plant PPR proteins are typically targeted to either the mitochondria or chloroplast, where they act by binding to one or several single-stranded RNA molecules via 2–30 N-terminal tandem helical repeat motifs (Shikanai 2006). Within plant mitochondria and chloroplasts, most characterized PPR proteins mediate specific events in post-transcriptional processing and maturation of RNA by influencing RNA splicing, RNA cleavage, RNA stability, translation and the site-specific sequence alteration of RNA transcripts through a process called RNA editing [reviewed in (Barkan and Small 2014)]. Other cellular processes that are affected by PPR proteins include nuclear gene expression (Ding et al. 2006; Koussevitzky et al. 2007; Liu et al. 2010) and plastid biosynthesis (Beick et al. 2008).

PPR proteins are divided into the P subfamily and the plant-specific PLS subfamily, and both classes of PPR proteins have diverse roles in RNA metabolism. In *Arabidopsis thaliana*, the P subfamily comprises approximately half of the 450 known PPR proteins. While the P family proteins contain only PPR repeats, the PLS subfamily is further subdivided according to the domains present in the extended C-terminal regions of subfamily members: (i) E proteins contain an E domain as the C-terminal region (ii) E+ proteins carry an E domain and an E+C-terminal region, and (iii) DYW proteins possess an E, E+ and an additional domain named the DYW domain due to the presence of aspartic acid, tyrosine, and tryptophan triplet of amino acids (DYW), or a variation thereof, at their extreme C-termini (Shikanai 2006). The DYW domain appears to be plant-specific and has sequence similarity to cytidine deaminases (Salone et al. 2007; Iyer et al. 2011), enzymes involved in the recognition of target cytidines in the C-to-U editing reaction (Okuda et al. 2014). The PPR DYW domain has been correlated with the occurrence of RNA editing, while the E domains are thought to recruit editing enzymes (Schmitz-Linneweber and Small 2008). PPR proteins within the PLS subfamily with known functional roles in RNA editing all belong to the E or DYW PPR subfamilies (Shikanai 2006).

The RNA editing factor interacting proteins (RIP)/Multiple organellar RNA editing factor (MORF) proteins are a small family of proteins required for C-to-U editing in both chloroplasts and mitochondria (Takenaka et al. 2012; Bentolila et al. 2013). RIP/MORF2 and 9 are required for editing of chloroplast transcripts while the other proteins are involved in mitochondrial transcript editing (Takenaka et al. 2012; Tian et al. 2019). RIP1/MORF8 is dual-targeted to chloroplasts and mitochondria (Bentolila et al. 2012). RIP/MORF proteins carry a novel conserved domain of about 100 amino acids, the MORF box, that is required for multimerization and interaction with PPR proteins (Takenaka et al. 2012; Bayer-Csaszar et al. 2017; Haag et al. 2017; Yan et al. 2017; Yang et al. 2018). MORF proteins interact with PPR proteins (Bentolila et al. 2012; Takenaka et al. 2012; Glass et al. 2015; Yan et al. 2017) to induce structural changes that increase RNA-binding and editing efficiency (Yan et al. 2017). While PPR proteins are usually involved in editing of a few specific sites, loss of some RIP/MORF proteins can cause defects in editing of all sites in chloroplasts (RIP/MORF2 and 9) (Takenaka et al. 2012) or hundreds of sites in mitochondria (Bentolila et al. 2012). Biochemical evidence has been interpreted as indicating that the organellar RNA editing machinery likely includes an RNA helicase component (Takenaka and Brennicke 2003; Hegeman et al. 2005). Editing activity in chloroplast extracts could be stimulated by ATP, CTP or dCTP (Hegeman et al. 2005) and similarly, mitochondrial extracts from pea that were used for in vitro editing assays could use any NTP or dNTP (Takenaka and Brennicke 2003). Consistent with the proposed involvement of an RNA helicase in organelle RNA editing, the chloroplast RNA helicase ISE2 was shown to be required for editing of several different *Arabidopsis* chloroplast transcripts (Bobik et al. 2017). The involvement of ISE2 in chloroplast C-to-U editing is supported by the identification of the maize ISE2 orthologue in RIP9 complexes purified from maize extracts (Sandoval et al. 2019), indicating that ISE2 is involved in editing in multiple plants. In studies to determine ISE2's involvement in chloroplast RNA processing, we identified protein partners of ISE2 including a DYW protein encoded by *At5g03800/EMB175* as a potential partner of ISE2; we subsequently named this DYW protein ISE2 PROTEIN INTERACTOR1 (IPI1) (Bobik et al. 2019; Ganusova et al. 2020). *Arabidopsis embryo defective 175 (emb175)* mutant embryos arrest at the globular-heart transition (Cushing et al. 2005). The maize ortholog of IPI1/EMB175, PPR103, functions in rRNA processing and stabilization, and loss of PPR103 resulted in seedling lethality (Hammani et al. 2016).

While ZmIPI1/PPR103 lacks the C-terminal triplet of amino acids characteristic of DYW proteins and was reported to not have a role in chloroplast C-to-U RNA editing, NbIPI1 and its dicot orthologs have a DYW motif. In

this study, sequence and structural analysis revealed that NbIPI1 and several of its orthologs lack critical residues known to be essential for C-to-U RNA editing. Our modeling predictions suggested that these proteins likely lack a complete coordination site for a critical zinc ion that is required for catalysis of editing. We investigated the function of NbIPI1 in plants where *NbIPI1* was silenced by virus-induced gene silencing (VIGS) in *Nicotiana benthamiana*, bypassing the lethality of the *emb175/ipi1* Arabidopsis mutants (Cushing et al. 2005). Analysis of chloroplast RNA transcripts revealed that PPR103/IPI1 functions in rRNA processing and stabilizing *rpl16* transcripts were partially conserved in *N. benthamiana*. In contrast to maize PPR103, *NbIPI1*-silenced plants had reduced although not completely disrupted C-to-U editing of 23 sites in 18 chloroplast transcripts. These findings suggest that NbIPI1 may have functions related to C-to-U editing although it may not directly function in deamination catalysis.

Materials and methods

Plant materials and growth conditions

Nicotiana benthamiana seedlings were grown on a light cart at 25 °C under fluorescent white light in a 16:8-h light/dark cycle. Ten-day-old seedlings were transplanted to individual pots and typically silenced using VIGS at around 2–3 weeks of age.

Transient expression and confocal microscopy

Leaves of five to six-week-old plants were agroinfiltrated with constructs for expression of AtIPI1-YFP, cTP-AtIPI1-YFP or cTP-NbIPI-YFP. Forty-eight hours later infiltrated leaf sections were vacuum infiltrated with water and mounted on slides for imaging. Confocal fluorescence microscopy was performed using a Leica SP2 or SP8X confocal laser scanning microscope (Leica Microsystems, Heidelberg GmbH). A 40× or 63× HCX PL APO objective was used for image acquisition. The samples were excited with an excitation line of 458/514 nm for YFP.

VIGS constructs and protocol

Tobacco rattle virus (TRV)-based VIGS constructs used for the non-silencing control and silencing *ISE2* and *IPI1* were described previously (Burch-Smith and Zambryski 2010; Ganusova et al. 2020). Constructs were transformed into *Agrobacterium tumefaciens* GV3101 (pMP90RK) strain. VIGS of *IPI1*, *ISE2*, *GUS* intron (negative control) and *PDS* (positive control) were performed according to (Burch-Smith and Zambryski 2010; Ganusova et al. 2020).

Approximately 3-week-old *N. benthamiana* plants were infiltrated with a mixture of *Agrobacterium* strains containing TRV RNA1 (pYL192) and TRV RNA2 (pYL56) containing the silencing constructs, and then grown under standard growth conditions until downstream assays were performed 14 days later.

Transmission electron microscopy (TEM)

Samples were prepared for TEM as described previously (Burch-Smith and Zambryski 2010; Burch-Smith et al. 2011). Briefly, samples from young emerging leaves from control or *IPI1*-silenced plants were fixed by high-pressure freezing (HPF) and quick freeze substituted (QFS) in 1% osmium tetroxide plus 0.1% uranyl acetate in acetone. Subsequently, samples were embedded in epoxy resin Embed 12, (Ted Pella, Inc., Redding, CA), sliced into ultrathin 65–70-nm sections, and visualized on a Libra 200 M TEM/STEM (Carl Zeiss Microscopy, White Plains, NY) at 200 kilovolts.

Structural modeling

We carried out structure prediction and modeling of the C-terminal portions of NbPPR103/IPI1 using the ColabFold implementation of AlphaFold2 (ColabFold v1.5.2, AlphaFold2_mmseqs2; (Mirdita et al. 2022)). We selected residues 757–890 of NbIPI1 sequence as the input sequence. AlphaFold2 was run with the “use PDB templates” option set to “true” and instructed to output its top 5 models after Amber relaxation. Most other interface options were kept at their default values. The top 5 output models were of uniformly high predicted confidence (average plddt were extremely similar to each other [RMSd < 0.15 Å]). We therefore used the top-ranked model by plddt score for subsequent inspection and further modeling.

We inspected the structures and generated molecular graphics using PyMol v.2.50 (The PyMol Molecular Graphics System, Version 2.0, Schrödinger, LLC). We specifically carried out structural comparisons with a number of template homologous domain structures as references to identify and manually adjusted putative Zn²⁺ binding sites within the modeled NbIPI1 C-terminal domains. The homologous experimental structures included fragments and domains of *Arabidopsis thaliana* OTP86 (PDB 7O4E, 7O4F; Takenaka et al. 2021), *A. thaliana* DYW1 [PDB 7W86; (Toma-Fukai et al. 2023)], and AlphaFold pre-predicted structures of *A. thaliana* and *A. lyrata* EMB175/IPI1 (alphafold.ebi.uk entries Q9FFN1 and D7LWU0). Where appropriate, we added and manually adjusted Zn²⁺ ions and water molecules to the NbIPI1 model within PyMol and adjusted nearby side-chains using backbone-dependent rotamers.

Prediction of NbIPI1 binding sites

Potential binding sites for NbPPR103/IPI1 were predicted with the FIMO program in the MEME suite (Grant et al. 2011). The nucleotide-binding probabilities for NbPPR103/IPI1 were generated based on the amino acids found at the 6 and 1' position (first amino acid of the subsequent C-terminal PPR motif) of each PPR motif to assign a nucleotide preference according to the weighting scheme in (Takenaka et al. 2013). These nucleotide preferences scores were used to predict NbPPR103/IPI1 RNA binding sites within the *Nicotiana benthamiana* chloroplast genome using the FIMO program. The ten top-predicted binding sites were ranked by *P*-values calculated by FIMO (Grant et al. 2011).

Northern blotting

Total RNA was run on a denaturing formaldehyde gel, transferred to positively charged Roche nylon membranes (MilliporeSigma, Burlington, MA), and hybridized with DIG-labeled 5S, 23S rRNA and *rpl16* probes (Table S1) according to manufacturer's instructions (PCR DIG Probe Synthesis Kit, Roche). The amounts of RNA we used for each blot are indicated in the figure legends. Bands corresponding to ribosomal RNA species were detected using the Roche DIG High Prime DNA Labeling and Detection Starter Kit II (MilliporeSigma, Burlington, MA). The same RNA that was used to measure C-to-U editing efficiencies was used for the Northern Blot analysis.

Chloroplast isolation and RNAseq library preparation

We extracted chloroplasts according to "Extraction of Chloroplast Proteins from Transiently Transformed *Nicotiana benthamiana* Leaves" bio protocol (Klinkenberg 2014; Klinkenberg et al. 2014). Briefly, fresh leaf tissue was ground, filtered and centrifuged through a Percoll gradient and visualized on an inverted microscope. Chloroplasts were then shock-frozen and total RNA was isolated from purified chloroplasts using Trizol (Thermo Fisher Scientific, Waltham, MA) or RNeasy Plant Mini kit (Qiagen, Germantown, MD) as per manufacturers' instructions. For each plant, we ground approximately 100 mg of tissue from each leaf to isolate chloroplast RNA. Leaves from individual plants were pooled. Removal of chloroplast DNA was done by treating the samples with Ambion rDNase1 (Thermo Fisher Scientific, Waltham, MA). Because rRNA typically constitutes over 75% of total RNA and its depletion can result in very low yields of RNA for cDNA preparation, rRNA depletion was not performed. The RNA integrity of the isolated RNA was examined on a Bioanalyzer machine and quantitated on a NanoDrop 1000 spectrophotometer

(Thermo Scientific, Waltham, MA) prior to library preparation. For cDNA synthesis, about one microgram of non rRNA-depleted RNA was used to make double strand cDNA (ds-cDNA) and dsDNA was produced using the Invitrogen SuperScript II Double Stranded cDNA Synthesis kit (Thermo Fisher Scientific, Waltham, MA) with random hexamers primers for first-strand synthesis. The cleaned ds-cDNA was then used to construct a library using the Illumina Next Tera Library prep kit with no adaptations (Illumina, Inc, San Diego, CA). After examination of the library quality using the Bioanalyzer (Agilent, Santa Clara, CA), multiplexed libraries were sequenced using the Illumina MiSeq sequencing platform per standard MiSeq run parameters (Illumina protocol manuals).

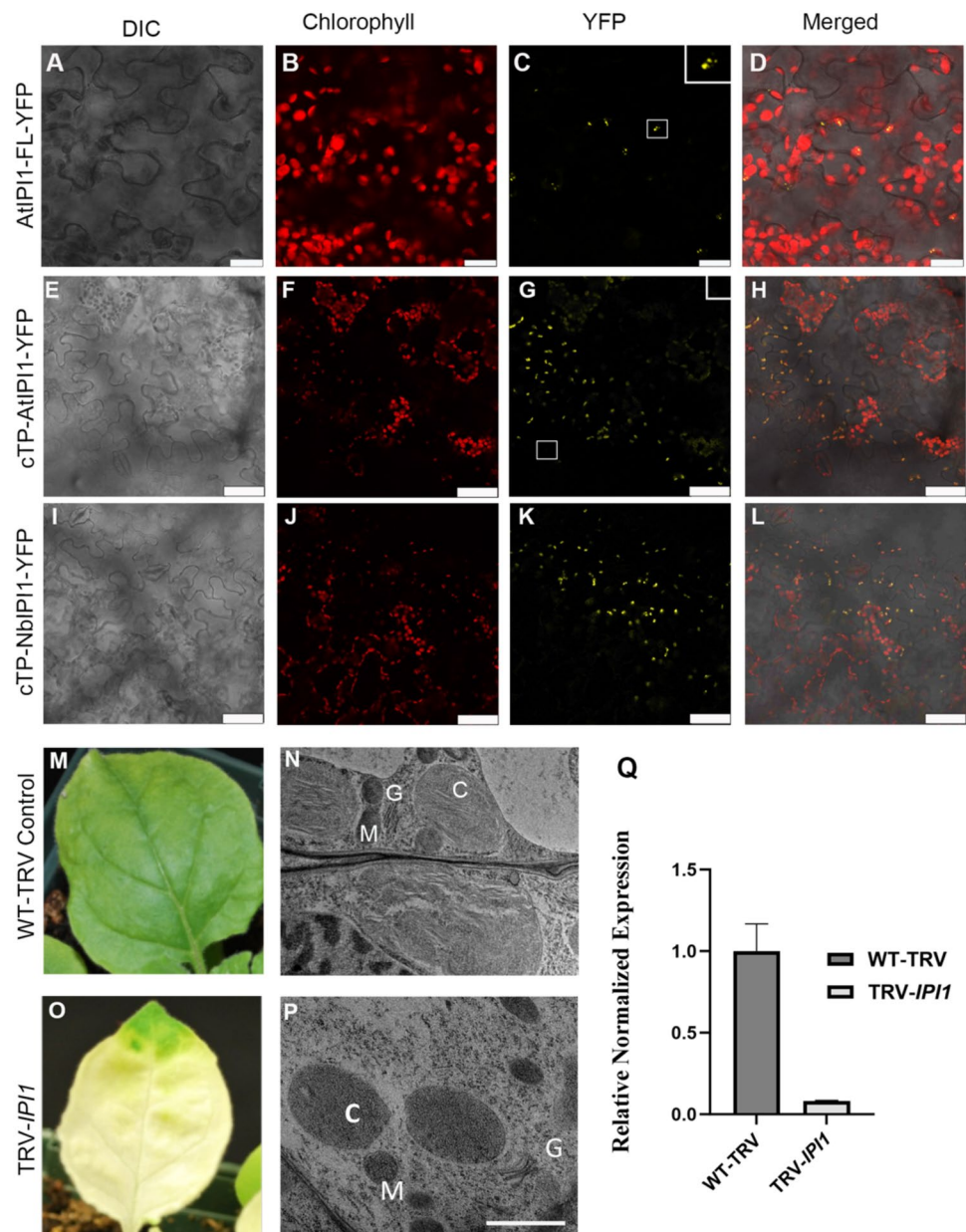
Mapping and data statistical analysis

We examined sequence reads for sequence quality, trimmed using the base space graphical user app (Base Space, Illumina, Inc) and mapped to the *N. benthamiana* genome (NC) using DNA Array Star Next Gen Seq software (version 12) permitting multiple mismatches to detect multiple SNPs. We used the mapping parameters kmer size of 21 and low SNP filter stringency (to avoid missing highly edited transcripts). Paired-end mapped contigs were visualized in Seq Man Pro software or the Integrative Genomics Viewer (IGV) software. The SeqMan NGen-mapped contigs (Supplementary Table S2) were used for subsequent analysis. The uniquely mapped reads were used to detect coverage information for each sample. The coverage summary additionally reveals the depth of sequenced reads that were mapped at each locus in the *Nicotiana* genome. Overall, a similar number of reads were mapped to the *N. benthamiana* genome in all samples.

RNA editing by sanger sequencing

For editing in the non-silenced control, *ISE2*-, *IPI1*- or *PDS*-silenced leaves, RNA was isolated from leaf number 11 of approximately six-week-old plants using Trizol (Thermo Fisher Scientific, Waltham, MA). The RNA was treated with DNase (30 min with 0.5 µL rDNase, 15 min of a 2nd 0.5 µL rDNase at 37 °C) at least once. RT-PCR was conducted according to manufactures instructions in the M-MLV RT (Promega, Madison, WI,) manual using random primer hexamers. A typical reaction consisted of PCR: 1 µg RNA, 1.2 µL random hexamer, 0.8 µL reverse transcriptase. The same reaction without the reverse transcriptase was performed in parallel with the experimental cDNA synthesis reaction to ensure the absence of genomic DNA contamination. Second-strand PCR synthesis was performed according to standard Taq polymerase protocol using primers in Table S1. PCR was conducted with an annealing temperature of 53 for almost all primer pairs. PCR products

Fig. 1 Subcellular localization of IPI1 and effects of *NbIPI1* knockdown on leaves and chloroplasts. **a–d** AtIPI1-YFP was transiently expressed in *N. benthamiana* leaves and YFP signal overlapped with chloroplast autofluorescence. Scale bars represent 20 μ m. Inset shows enlarged image of boxed region. **e–f** The predicted cTP of AtIPI1 was fused to YFP and transiently expressed in *N. benthamiana* leaves. Inset shows enlarged image of boxed region. Scale bars represent 50 μ m. **g–h** The predicted cTP of NbIPI1 was fused to YFP and transiently expressed in *N. benthamiana* leaves. Scale bars represent 50 μ m. **m–n** TRV-infected, non-silenced control leaves, and TEM image showing chloroplasts in young sink leaves with forming thylakoids and grana. C, M, and G indicate chloroplast, mitochondria, and Golgi, respectively. **o–p** *IPI1*-silenced leaves presented a severe chlorotic phenotype and TEM analysis reveals defective chloroplasts. Scale bar represents 1 μ m. **q** Silencing efficiency was measured by quantitative PCR. Statistical significance was determined by Student t test. *** $p < 0.001$



were gel purified using the Gel Extraction Wizard kit (Promega, Madison, WI). Purified amplicons were sequenced by the UTK Genomics Core Facility.

Measuring editing efficiency

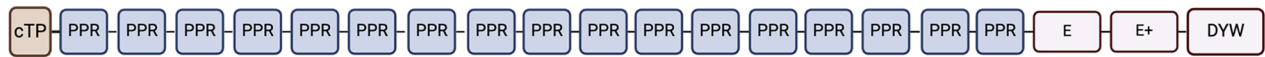
RNA was isolated using RNEasy Plant Mini Kit (QIAGEN, Germantown, MD). QPCR was conducted using SYBR Select Master Mix (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA) and primers listed in Table S2, following our previously published protocol (Ganusova et al. 2020) and using the CFX384 Touch Real Time PCR Detection System (BIO-RAD, Hercules, CA).

Results

The PPR protein, NbIPI1, is needed for chloroplast development in *N. benthamiana*

We previously identified the PPR protein encoded by *At5g03800/emb175/IPI1* as interacting with the chloroplast RNA helicase ISE2 in a yeast two-hybrid screen (Bobik et al. 2019). To determine the subcellular localization of IPI1, we cloned the full-length Arabidopsis *IPI1* coding sequence upstream of the yellow fluorescent protein (YFP). The resulting fusion protein (AtIPI1-YFP) was transiently expressed in *N. benthamiana* leaves and visualized by confocal laser scanning fluorescence microscopy. Fluorescence

A



B

		DYW motif	
<i>Physcomitrium patens</i> (moss)	-PLTPIVISKNLRV	CDCHTATKFIISKITKRQIIARDSNRF	HHFKDGVCSGDFW--- 986
<i>Salvinia cucullata</i> (fern)	-PGTPLRLVKNLRT	CPDCHKAMKPIISKIVGREIILRDANRM	HHFKDGTICSCNDYW--- 851
<i>Azolla filiculoides</i> (fern)	-KGATLRITKNLRV	CMDCHNAAKFISKLEEREILLRDVARY	HRFKNGVCSGDIYW--- 840
<i>Pinus sylvestris</i> (conifer)	-PWTSITVTKNLRV	CGDCHSRATKFIISKVVGREIIMRDANRF	HHFKNGLCSGDIYW--- 156
<i>Oryza sativa</i> (rice)	-HGETIRVVKNVVM	CGDCHSFLEYTSATGKEILVRDSAGF	HFICGGKCSGCRG--- 872
<i>Brachypodium distachyon</i> (purple false brome)	-PGQSVNVVKNIRM	CGDCHSFMKHASAATGKEISVRDSSGF	HIFRGGVCSGCRE--- 866
<i>Hordeum vulgare</i> (Barley)	-SGKIIIRVVKNIRM	CGDCHSFLEHASAATGKEISVRDSSGF	HIFRAGICSGCRE--- 861
<i>Triticum aestivum</i> (wheat)	-SGKIIIRVVKNIRM	CGDCHSFLEHASAATGKEISVRDSSGF	HVFRAGICSGCRD--- 833
<i>Triticum aestivum 2</i> (wheat)	-SGKIIIRVVKNICM	CGDCHSFLEHASAATGKEISVRDSSGF	HIFRAGICSGCRE--- 740
<i>Setaria italica</i> (foxtail millet)	-TGRAVVRVKNIRM	CGDCHSFLEHASAATGKVISVRDSSGF	HIFRGGKCTCRE--- 867
<i>Zea mays</i> (corn)	-PGQTIIRVVKNIRM	CGDCHSFLEHASAATGKVISVRDSSGF	HIFRGGKCSGCSQ--- 872
<i>Sorghum bicolor</i> (sorghum)	-PGQTVRVVKNIRM	CGDCHSFLEHASAATGSLHFLGGLSYQ	--GHGEVAMGQAFPKPF 858
<i>Zostera marina</i> (eelgrass)	IKGSRAVRVMKNVRL	CGDCHSFMKFFSVVTARKIFFRDTSGF	HHFIDGRCSCEVQ--- 921
<i>Arabidopsis thaliana</i>	'-RGKPVVRVMKNVML	CGDCHSFLKYSVSVVTKREIVLRDSSGF	HHFVNGKCSGCDLW--- 896
<i>Brassica napus</i> (rapeseed)	I-QREPVRVVKNVVM	CGDCHSFLKYSVSVVTKREIVLRDSSGF	HRFVNGKCSGCDLW--- 877
<i>Brassica rapa</i> (chinese cabbage)	I-QREPVRVVKNVVM	CGDCHSFLKYSVSVVTKREIVLRDSSGF	HRFVNGKCSGCDLW--- 877
<i>Glycine max</i> (soybean)	-PGKPIRIVKNILL	CGDCHAFKLYASIVTKRDIIFLRDSSGF	HCFNSNGQCSGDCW--- 874
<i>Citrus sinensis</i> (orange)	-AGQPVRIKVNILT	CGDCHSFLKYSVSVVTRREIFLRDASGF	HHFNLNGQCSGCDYW--- 812
<i>Cucumis sativus</i> (cucumber)	-AGQPVRIKVNILT	CGDCHSFLKYSVSVVTRREIFLRDASGF	HHFNLNGQCSGCDYW--- 812
<i>Manihot esculenta</i> (cassava)	-PGEPRIKVNILL	CGDCHSFLKYSVSVVTRREIFVRDASGF	HCFNSNGQCSGCDYW--- 912
<i>Erythranthe guttata</i> (yellow monkey flower)	-VGKPIRVVKNIHM	CGDCHTFKLYVTVTKREIHRDASGF	HCFADGECSCRDNW--- 722
<i>Nicotiana tabacum</i> (tobacco)	-PGKPVVRVMKNVHL	CGDCHTFKLYVSVVTKRDIHRDASGF	HHFVNGKCSGCRDNW--- 890
<i>Nicotiana tabacum 2</i> (tobacco)	-PGKPVVRVMKNVHL	CGDCHTFKLYVSVVTKRDIHRDASGF	HHFVNGKCSGCRDNW--- 890
<i>Nicotiana benthamiana</i>	-PGKPVVRVMKNVHL	CGDCHTFKLYVSVVTKRDIHRDASGF	HHFVNGKCSGCRDNW--- 890
<i>Capsicum annuum</i> (pepper)	-PGRPVVRVMKNVLL	CGDCHTFKLYVSVVTKRDIHRDASGF	HHFVNGKCSGCRDTCW--- 895
<i>Solanum lycopersicum</i> (tomato)	-PGKPVVRVMKNVLL	CGDCHTFKLYVSVVTKRDIHRDASGF	HHFVNGKCSGCDNWC--- 812
<i>Solanum tuberosum</i> (potato)	-PGKPVVRVMKNVLL	CGDCHTFKLYVSVVTKRDIHRDASGF	HHFVNGKCSGCDNWC--- 894

C

AtDYW1	STKVRGDKPEISGGEKKAIVDRSKAYVKLSLGE-VRDAGYVPETKYVLHDI DEEAKEK	166
AtOPT86	SWIEMDGKVHKFTARDKSHPEKSEIYEKLSVTRKLEREVGYVADTKFVLHNVEGE-KVQ	887
NtIPI1	SWIILGDKVHSFFGRDKLHSQSKDIYSGLQILPE-CLKAGYVPDTSFVLHEVEEHQKKD	817
NtIPI1-2	SWIVLGDVHSFFARDKLHSQSKDIYSGLQILPE-CLKAGYVPDTSFVLHEVEEHQKKD	817
NbIPI1	SWIILGDKVHSFFARDKLHSQSKDIYSGLQILILE-CLKAGYVPDTSFVLHEVEEHQKKD	817
AtIPI1	SWIIHENKIHSFHARDTSHPQEKDIYRGLEILIME-CLKVGYEPNTEYVLQEVDFEMKKS	822
AlIPI1	SWIIHEKKVHSFHARDTSHPQEKDIYSGLEILIME-CLKSGYEPNTEFVLQEVDFEMKKS	822
	* * : . . * * . : : . ** :*:~::~: * .	
AtDYW1	ALMHHSERLAIAGFIINTP-PGTTIRVMKNLRCGDCHNFILKILSSIEDREIIVRDNKR	225
AtOPT86	MLHGHSEIRIAIAYGLLRTPD-RACLRITKNLRVCRDCHTFCKLVSKLFRDRIVMRDANRF	946
NtIPI1	FLFYHSSKLAVTFGLLMT-RPGKPVVRVMKNVHLCGDCHTFKLYVSVVTKRDIHRDASGF	876
NtIPI1-2	FLFYHSSKLAVTFGLLMT-RPGKPVVRVMKNVHLCGDCHTFKLYVSVVTKRDIHRDASGF	876
NbIPI1	FLFYHSSKLAVTFGLLMT-RPGKPVVRVMKNVHLCGDCHTFKLYVSVVTKKDIHRDASGF	876
AtIPI1	FLFHHSAKLAVTYGILSSNTRGKPVVRVMKNVMLCGDCHSFLKYSVSVVTKREIVLRDSSGF	882
AlIPI1	FLFHHSAKLAVTYGILSSNTRGKPVVRVKNVMLCGDCHSFLKYSVSVVTKREIVLRDSSGF	882
	* ** :~::~: : : : : : : : : : : : : : : : *	
AtDYW1	HHFRDGNCSGDIW	239
AtOPT86	HHFESGLCSGDSW	960
NtIPI1	HHFVNGKCSGCRDNW	890
NtIPI1-2	HHFVNGKCSGCRDNW	890
NbIPI1	HHFVNGKCSGCRDNW	890
AtIPI1	HHFVNGKCSGCDLW	896
AlIPI1	HHFVNGKCSGCDLW	896
	*** . * * * * *	

Fig. 2 PPR103/EMB175/IPI1 domain structure and residues important for editing. **a** The predicted domains (UniProt) of AtIPI1 including a chloroplast targeting peptide (cTP), 17 PPR domains, an E-domain (E), an E+ domain (E+), and a DYW domain (DYW). **b** Amino acid sequence alignment of the extreme C-terminal regions of select NbPPR103/IPI1 orthologs. DYW triplet is shown in blue shading. Asterisks, colons, and periods indicate identical amino acid residue, conserved substitution, and semi-conserved amino acids, respectively. Zn²⁺ coordinating residues are shaded in yellow. **c** Alignment of the putative DYW domains of *N. benthamiana*, *A. thaliana* and *A. lyrata* IPI1 sequences with the defined DYW domains of *A. thaliana* DYW1 and OPT86. Coordinating residues for the catalytic and structural Zn²⁺ binding sites are respectively outlined in blue and red. The catalytically important glutamic acid residue of canonical DYW domains is highlighted in purple

from the IPI-YFP fusion colocalized with chloroplast autofluorescence, indicating that AtIPI1-YFP localized to chloroplasts (Fig. 1a–d). AtIPI1-YFP also localized to punctae within the chloroplast, suggesting that it may localize to the chloroplast stroma. A similar pattern of fluorescence was observed when the predicted chloroplast targeting peptide (cTP) plus 20 amino acids downstream of the cTP were cloned as a translation fusion to YFP (cTP-AtIPI1-YFP; Fig. 1e–h). To determine NbIPI1's subcellular localization, a similar construct carrying the predicted cTP of NbIPI1, cTP-NbIPI1-YFP, was transiently expressed in *N. benthamiana* leaves. Fluorescence from cTP-NbIPI1-YFP colocalized with chloroplast autofluorescence, indicating that the fusion was imported into chloroplasts and that NbIPI1 localizes to the chloroplast stroma (Fig. 1i–l). Unfortunately, we were unable to clone full-length NbIPI1 to test its subcellular localization despite numerous attempts. Based on our results with the cTP-NbIPI1-YFP construct, we conclude that it is likely that NbIPI1 localizes to chloroplasts given that its chloroplast targeting peptide is functional, although localization to additional compartments cannot be ruled out.

Arabidopsis emb75 mutants fail to develop past the very early stage of embryonic development, arresting at the globular embryo stage (Cushing et al. 2005). We therefore used VIGS to silence *NbIPI1* in young *N. benthamiana* plants and examine NbIPI1's function. VIGS of *NbIPI1* in *N. benthamiana* caused severe leaf chlorosis and reduced chlorophyll content [(Fig. 1m, o and q); (Ganusova et al. 2020)]. Transmission electron microscopy on young sink leaves from silenced plants revealed profound effects of *IPI1* knockdown on chloroplast development. While thylakoids and nascent grana were observed in TRV-infected non-silenced controls (Fig. 1n), these structures were largely absent from chloroplasts in *NbIPI1*-silenced leaves (Fig. 1p). These observations are consistent with localization of NbIPI1 to chloroplasts (Fig. 1a–l) and severe chlorosis of *NbIPI1*-silenced leaves (Fig. 1o) and suggest that NbIPI1 has critical roles in chloroplast development.

IPI1 orthologues have divergent sequences hinting at differing functions

NbIPI1 is predicted to comprise a plastid-localization sequence at its N-terminus, 13 PLS PPR motifs and C-terminal E and DYW domains (Fig. 2a and Supplementary Figure S1). The DYW domain is named after the C-terminal aspartate, tyrosine and tryptophan sequence that is often found in PPR proteins associated with C-to-U RNA editing in organelles (Gutmann et al. 2020).

IPI1 orthologues from early land plants including mosses and ferns also contain a so-called DYW motif consisting of 9 C-terminal residues (Takenaka et al. 2021) (Fig. 2b). The DYW motif is also present in dicots represented in the alignment (Fig. 2B and Supplementary Figure S1). In these sequences, the tyrosine residue is most often substituted by leucine as in *Arabidopsis* (DLW) or asparagine as in *N. benthamiana* (DNW). Interestingly, the C-terminal DYW triplet is absent from the DYW motifs of the examined monocots, as has been reported for maize [Fig. 2b, (Hammani et al. 2016)]. This finding is supported by phylogenetic analysis, which shows that the monocot proteins form a distinct clade (Supplementary Figure Fig. S2). The absence of the DYW motif from the maize IPI orthologue, PPR103, likely explains previous observations that suggest that maize PPR103 is not involved in C-to-U RNA editing of maize chloroplast transcripts (Hammani et al. 2016).

Recent findings revealed that a canonical DYW domain catalyzes the C-to-U editing through a novel mechanism in which a catalytic zinc atom is regulated by its coordination state (Takenaka et al. 2021). In all orthologues we examined, almost all the residues involved in Zinc binding (highlighted in yellow) are highly conserved (Fig. 2b and c). An important exception is a key glutamic acid residue (E894 in OPT86) within a gating subdomain that is essential for catalysis (Boussardon et al. 2014; Hayes et al. 2015; Oldenkott et al. 2019) (Fig. 2c, highlighted in purple). This glutamate residue, which coordinates the catalytic zinc ion, is substituted with serine in NbIPI1 and with alanine in *Arabidopsis* EMB175/PPR103 proteins (Fig. 2c).

Structural prediction of IPI1 structures

To follow up on the amino acid sequence analysis and get a better understanding of possible NbIPI1 functions, we used the protein structure prediction package AlphaFold2 to generate structural models of the DYW domain of NbIPI1 and compare them with the protein structures of DYW1 (PDB: 7W86) or the DYW domain of OPT86 (PDB: 704F). Based on these comparisons, NbIPI1 appears to contain two Zn⁺² ion binding sites similar to those of canonical DYW domains, one of which participates in catalysis in such DYW domains (Fig. 3a and Supplementary Figures S3 [model

of whole DYW domain] and S4 [structural Zn⁺² binding domain]). However, the catalytic zinc appears to coordinate with only three residues: H822, C850 and C853 (Fig. 3b). In DYW1 and OTP86, the catalytically important glutamate residues E173 and E894, respectively, bind a water molecule that acts as the fourth coordinating ligand for the catalytic zinc ion (Takenaka et al. 2021). Based on our modeling, there is no obvious equivalent residue either at a similar position in NbIPI1 as these critical glutamate sidechains. Nor is there a different residue that could directly substitute for the coordinating water molecule observed in the OTP86 DYW domain structure. We therefore suggest that NbIPI1 and its orthologs likely do not play a direct role in catalyzing C-to-U deamination, at least using the mechanism described for OTP86 (Takenaka et al. 2021).

Identification of predicted NbIPI1 targets

PPR proteins bind their RNA targets in a combinatorial manner, using residues from adjacent PPR modules to identify specific RNA bases (Barkan et al. 2012). We used publicly available software tools for prediction of PPR-binding sequences (Barkan et al. 2012; Takenaka et al. 2013; Kobayashi et al. 2019) to identify potential NbIPI1 targets. Computationally predicted sites for Arabidopsis IPI1/EMB175 included an *rps3-rps16* intergenic region that is conserved in maize (Hammani et al. 2016). Supporting the value of computationally predictions in revealing the action of a PPR protein, maize PPR103 bound to and stabilized the 5' end of the processed maize *rps16* mRNA. A refined algorithm suggested that the mitochondrial *NAD2* transcript was the most likely target for AtIPI1 (Kobayashi et al. 2019). However, this finding is not likely relevant to EMB175/IPI1/PPR103 function since IPI1 is likely localized to chloroplasts and not in mitochondria, at least under our experimental conditions (Fig. 1). We therefore focused on identifying potential chloroplast targets for NbIPI1 (Fig. 4a) using the FIMO tool, as previously described (Takenaka et al. 2013; Grant et al. 2011). We identified several potential target sites including sequences within the *rps3-rpl16* intergenic region, like for maize PPR103 (Fig. 4b).

NbIPI1 is necessary for the accumulation of chloroplast rRNA species

Knockdown of *NbIPI1* expression by TRV-mediated VIGS led to severe leaf chlorosis (Fig. 1o) and reduced chlorophyll content and photosystem II quantum efficiency (Ganusova et al. 2020). Similarly, maize *ppr103* mutant seedlings had albino leaves and did not continue development past the seedling stage. Mutant seedlings contained drastically reduced chloroplast ribosomal RNA (rRNA) levels in the *ppr103* albino leaves (Hammani et al. 2016). To test whether

NbIPI1 may have a similar role in chloroplast rRNA biogenesis, we performed Northern blotting analysis for chloroplast rRNAs in *NbIPI1*-silenced *N. benthamiana* leaves using sequence-specific probes (Table S1). Defects in rRNA levels were severe enough to be observed on an agarose gel stained with ethidium bromide (Fig. 5a). The Northern blots revealed that silencing *NbIPI1* led to major defects in chloroplast rRNA, with drastic reductions in the 23S rRNA compared to the rRNA levels in the TRV-containing non-silenced controls (Fig. 5b) although there was no obvious change in 5S rRNAs (Fig. 5c). These results suggest that IPI1's role in rRNA processing is conserved in maize and *N. benthamiana*, although ribosome-associated transcripts were not among the top hits for NbIPI1 targets in our computational analysis (Fig. 4).

NbIPI1 is required for stability of *Nbrpl16* transcripts

The *rps3-rps16* intergenic region was one of the top 10 predicted targets for NbIPI1 (Fig. 4b), similar to maize PPR103. PPR103 binding to this intergenic region was found to be required for specification of the 5' end of *rpl16* transcripts (Hammani et al. 2016). The role of NbIPI1 in processing of *Nbrpl16* transcripts was therefore examined. RNA gel blot hybridization using a probe for the *rpl16* exons only (Fig. 6a) revealed a marked accumulation of unprocessed transcripts of higher molecular weights, 3 to > 6 kb but did not detect the expected processed transcripts of 1.4 and 0.4 kb (Fig. 6A and B). However, when probes against the *rpl16* intron were used, the 1.4-kb species representing the full-length transcript and the 1.9 kb bicistronic *rpl16-rpl14* transcript were absent in *NbIPI1*-silenced tissues compared to TRV-only infected and uninfected control (WT) plants (Fig. 6c). Importantly, silencing *NbISE2* did not change the stability of these RNA species (Fig. 6c), suggesting that NbIPI1 is required for processing the *Nbrpl16* transcript and for determining its stability. This is consistent with the function of PPR103 in processing maize *rpl16* transcripts and suggests a conserved function for NbIPI1.

NbIPI1 functions in chloroplast C-to-U RNA editing in *N. benthamiana*

Given that NbIPI1 appeared to have distinct functions from maize PPR103, we examined its role in RNA editing. RNAseq analysis was conducted on total RNA from chloroplasts from leaves of *NbIPI1*-silenced or non-silenced control *N. benthamiana* plants. To ensure that enough chloroplasts were available for RNA isolation and subsequent cDNA library preparation, we pooled all chlorotic leaves from one silenced plant to produce one biological replicate.

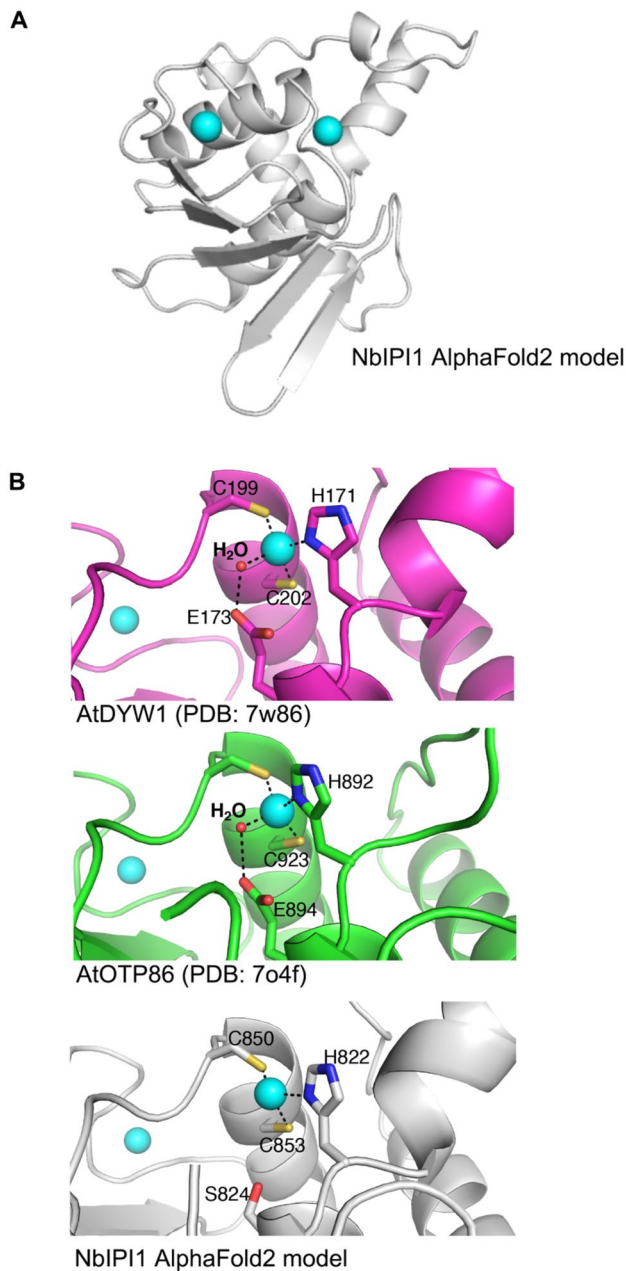


Fig. 3 Structural modeling of the NbIPI1 DYW domain. **a** AlphaFold2 model of the C-terminal sequences of NbIPI1, comprising the DYW domain and corresponding to residues 757–890 of the full-length protein. Cyan spheres indicated modeled Zn^{2+} ions. **b** Closeups of the catalytic Zn^{2+} binding site of *A. thaliana* DYW1 and OTP86 (top and middle panels; PDB 7W86, 7O4F) showing the full coordination of the zinc ion (cyan sphere), including a coordinating bound water molecule (small red sphere). The sidechain of the key catalytic glutamic acid residue is shown in each panel. The lowest panel shows the equivalent region of the NbIPI1 model. The NbIPI1 serine that occupies the site of the key glutamic acid cannot effectively substitute for the glutamate as a water-binding residue nor directly coordinate the Zn^{2+} ion

Samples from TRV-infected non-silenced control plants were similarly generated. The chloroplasts from control plants were phenotypically normal as observed by bright-field microscopy, whereas almost all chloroplasts isolated from *IPI1*-silenced mutant tissue were distinctively defective, with no thylakoid structures or starch granules apparent (Supplementary Figure S3), as expected from TEM observations (Fig. 1). Three control RNA-seq libraries (TRV-infected, non-silenced) and four test libraries (VIGS-*IPI1*) were used for Illumina MiSeq analysis (see Materials and Methods and Supplementary Figure S5). The reads had high quality mapping scores before and after the adapters were trimmed, ensuring that good quality contigs were subsequently mapped to the *N. benthamiana* chloroplast reference genome. We mapped the 250-nucleotide long sequenced reads to the *N. benthamiana* chloroplast reference genome curated by the Queensland University of Technology (<https://sefapps02.qut.edu.au/benWeb/subpages/chloroplast.php>) with no mismatch penalty to allow the detection of multiple editing events within the same transcript. Editing events were detected using an embedded SNP detection algorithm in the DNA Array Star workflow and are represented as the fraction of reads with an edited base out of the total reads (edited + unedited) for a given site. Additionally, reads that mapped to multiple locations were not excluded from subsequent analysis to allow the detection of potential SNPs on the inverted repeat strand of *ndhB*, a transcript that is heavily edited (Fig. 7). The editing efficiency at 23 edited sites in 18 chloroplast transcripts was statistically significantly reduced in *IPI1*-silenced leaves compared to control leaves (Fig. 7). This result suggested that reduced levels of NbIPI1 impacted editing of multiple chloroplast transcripts.

NbIPI1 may have roles in editing specific transcripts

Since AtIPI1/EMB175 interacts with AtISE2 which was previously shown to be required for chloroplast C-to-U editing, it is possible that some defects in editing *IPI1*-silenced chloroplasts were due to indirect effects on ISE2 activity. In addition, general stress in chlorotic leaves is also known to have deleterious effects on C-to-U editing (Kakizaki et al. 2009; Tseng et al. 2013; Zhu et al. 2014). For these reasons and to confirm the results from the deep sequencing approach, we bulk sequenced selected chloroplast transcripts from *IPI1*-, *ISE2*- or *PHYTOENE DESATURASE* (*PDS*)-silenced plants and compared the results to editing in non-silenced TRV-infected control plants. Sanger sequencing confirmed reduced editing in transcripts from *IPI1*-silenced leaves (Fig. 8). As expected, knockdown of *ISE2* or *PDS* in silenced plants also resulted in defective C-to-U editing (Fig. 8). The defective editing of the *rpoA*, *ndhB-4*, and *ndhD-1* sites is likely a secondary effect of chloroplast dysfunction since these

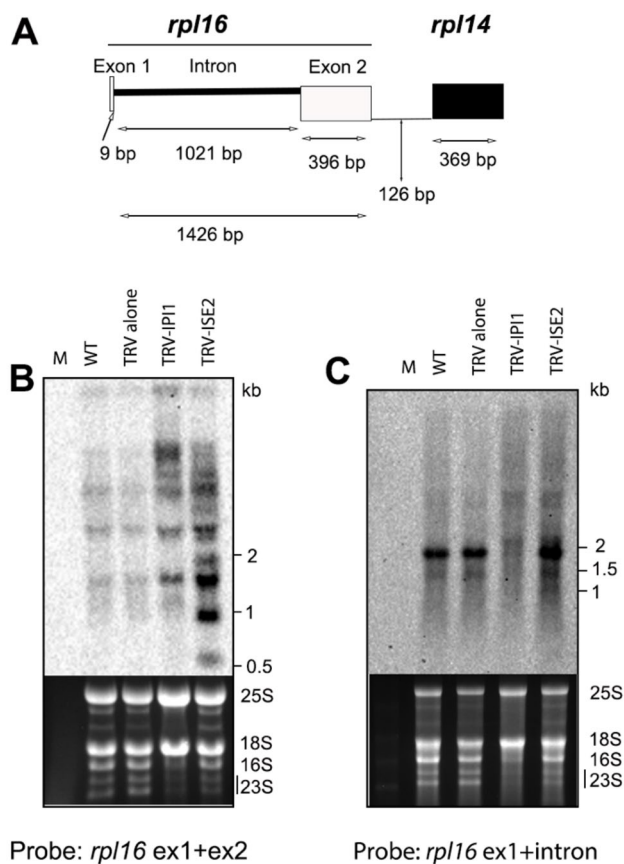


Fig. 6 Silencing of *NbIPI1* affects the *rpl16* transcripts. **a** The diagram shows the *rpl16* gene containing two exons (exon 1 is 9 bp and exon 2 is 374 bp) and an intron (1020 bp). RNA from *NbIPI1*- and *NbISE2*-silenced plants were probed with *rpl16* ex1+ex2 (exon 1 and exon 2) (**b**) and *rpl16* exon1 and intron (**c**), are compared to wild type (WT) and non-silenced control (TRV alone). 4 μ g and 2 μ g of total RNA were loaded for (**b**) and (**c**), respectively

for its maize orthologue PPR103 (Hammani et al. 2016). We find that NbIPI1 is involved in the processing of rRNA transcripts (Figs. 5 and 6). The results also extend NbIPI1's possible role to C-to-U editing of chloroplast transcripts (Figs. 7 and 8). Our analysis of RNA editing in *NbPDS*-silenced leaves suggests that NbIPI1 may contribute to editing three sites within the *ndhB* transcript (Fig. 8). Most DYW PPR proteins have a single editing target and little redundancy between editing factors has been observed, probably due to the specificity with which each protein binds its target RNA (Barkan and Small 2014; Gutmann et al. 2020). Indeed, at least six PPR proteins have been identified as required for editing the Arabidopsis *ndhB* transcript, with each PPR protein targeting a single site (Hammani et al. 2009; Okuda et al. 2009, 2010; Hayes et al. 2013; Fu et al. 2022;). In contrast, Arabidopsis QED1 edits one site in each of five Arabidopsis chloroplast transcripts, an usually large number of sites for a single PPR protein (Wagoner et al. 2015).

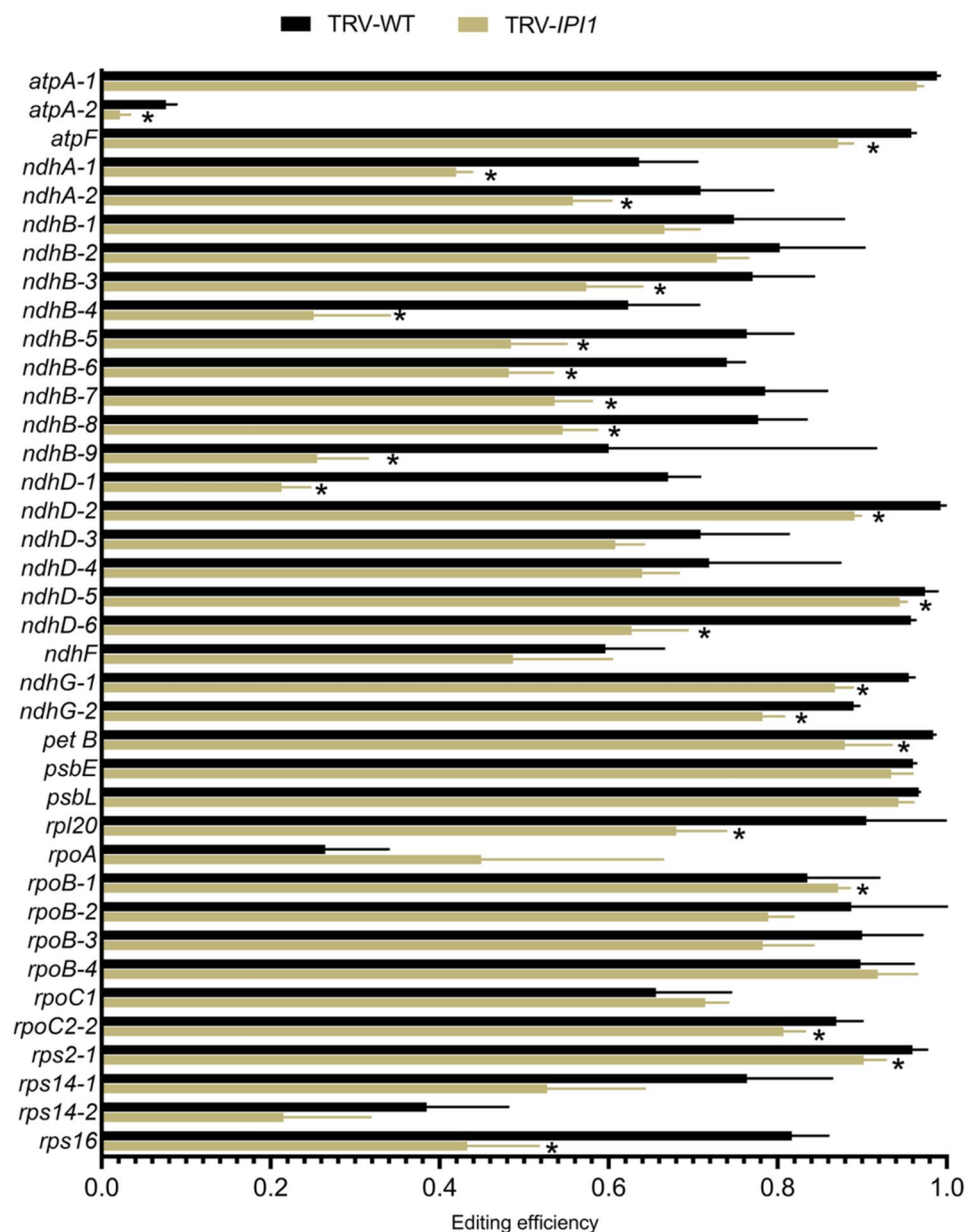
NbIPI1 is potentially involved in editing at least two sites in the *ndhB* transcript from *N. benthamiana* chloroplasts; however, further analysis is required to determine whether these sites and others in edited transcripts (Fig. 7) are specific NbIPI1 targets. Indeed, the computational prediction of targets suggests that editing of other transcripts could potentially involve NbIPI1 (Fig. 4).

The role of the DYW domain in editing

Nucleus-encoded RNA processing factors that are organelle-targeted are responsible for RNA editing and consistent with this, defects in chloroplast translation do not affect RNA editing (Zeltz et al. 1993). Thus far, nuclear-encoded PPR proteins have been classified as site-specific trans-factors involved in RNA editing (Mach 2009). The DYW domain's predicted structure resembles that of cytidine deaminases which bind zinc as part of their catalytic activity (Salone et al. 2007; Iyer et al. 2011). Mutations within the DYW domain of DYW1, a PPR protein that is similar to IPI1, greatly impair both zinc-binding and RNA editing, indicating that the DYW domain may confer cytidine deaminase activity (Okuda et al. 2009). Interestingly, the maize IPI1 orthologue PPR103 does not have the specific C-terminal amino acid residues DYW or any variation thereof are not at its C-terminus (Supplementary Fig. S2), and no editing events attributed to PPR103 were disrupted in the maize *ppr103* mutant (Hammani et al. 2016). This observation suggests that the specific DYW triad motif or its variants contribute to the editing reaction at several chloroplast editing sites. Several recent studies have clarified the importance of the DYW domain in RNA editing. Heterologous expression of a moss (*Physcomitrium patens*) PPR protein, PPR65, in *E. coli* was sufficient to edit the co-expressed transcript of the corresponding mitochondrial site, *ccmFCeU103PS*, with 70–100% efficiency (Oldenkott et al. 2019). A second moss PPR protein, PPR56, was also able to edit both its targets, *nad4eU272SL* and *nad3eU230SL* in the *E. coli* system with similar efficiencies as those observed in *planta*. In vitro assays demonstrated that purified recombinant moss PPR65 could successfully perform C-to-U editing of its synthetically generated target RNA, and that the editing activity required zinc and was enhanced by ATP or nonhydrolyzable nucleotide analogs (Hayes and Santibanez 2020). Together these studies advanced the idea that DYW proteins can independently carry out C-to-U editing.

The mechanism of the DYW domain has been more fully revealed by details gleaned from recent crystal structures of the DYW domain of *Arabidopsis thaliana* OTP86 protein, which specifically edits a site in the *rps14* transcript (Takenaka et al. 2021). These studies revealed a cytidine deaminase fold and a DYW domain containing zinc atoms that were critical to editing activity through

Fig. 7 C-to-U editing in *IP11*-silenced plants. RNA-seq analysis of *N. benthamiana* chloroplast transcripts confirms editing of all sites identified by Sanger sequencing and reveals reduced editing of several sites in *NbIP11*-silenced chloroplasts. Editing efficiency was calculated by number of reads carrying the edited base as a proportion of the total number of reads for a given site. Error bars represent standard deviation of reads from individual libraries, with four replicates used for *NbIP11*-silenced chloroplasts and three for the TRV-infected non-silenced controls. Asterisks denote P -value < 0.05 as determined using a one-tailed t -test assuming unequal variance

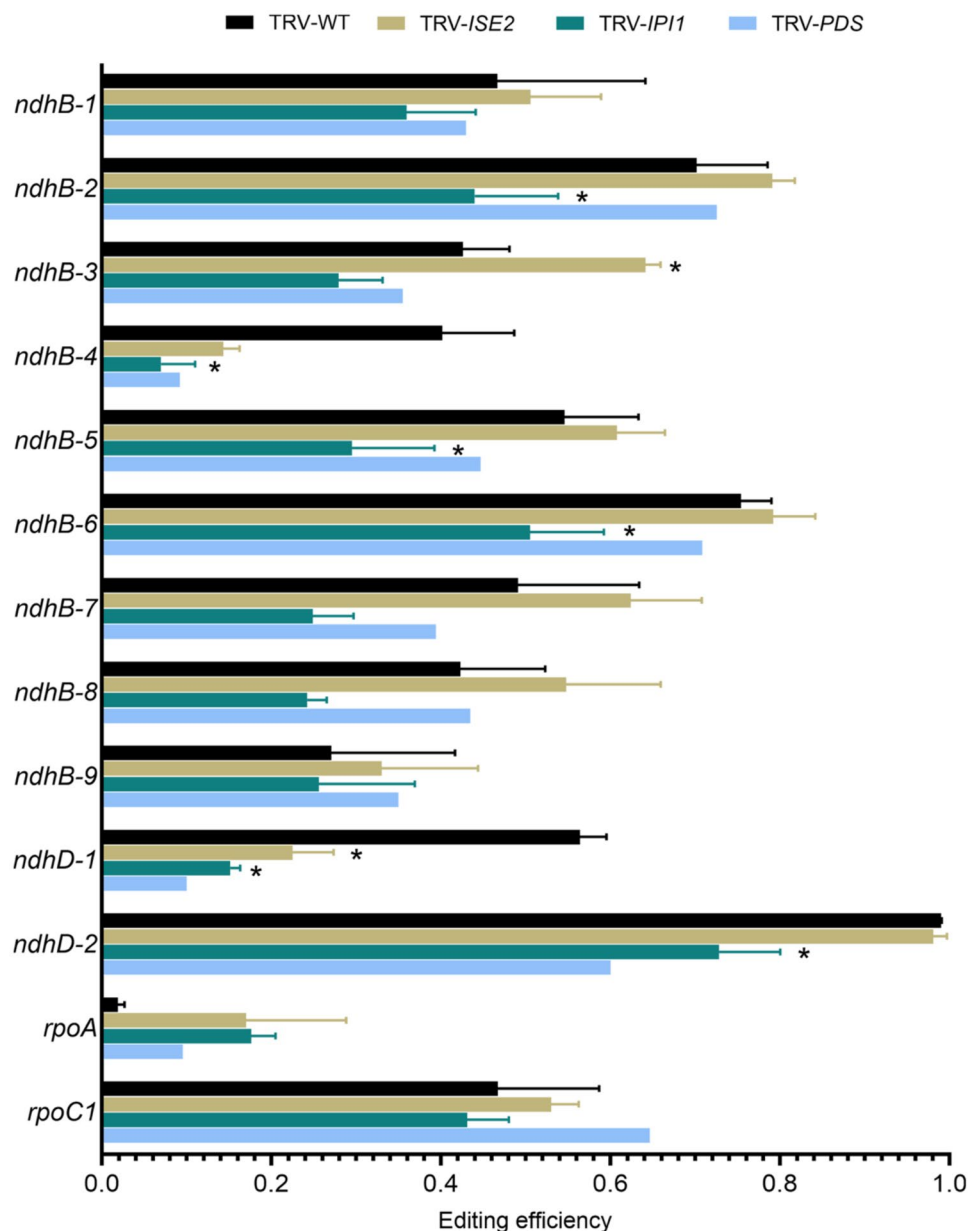


the regulation of a “gated zinc shutter”. In vitro RNA editing assays confirmed the importance of highly conserved residues to catalysis and highlighted the importance of the coordinating Zn ions for catalysis (Zn1) or stability of the DYW motif (Zn2) (Takenaka et al. 2021), supporting findings from previous mutational studies with other DYW proteins (Hayes et al. 2013, 2015; Boussardon et al. 2014; Oldenkott et al. 2019; Hayes and Santibanez 2020). Our data support the importance of the DYW domain for editing as they demonstrate that a PPR103 family member carrying the DYW triad motif may contribute to editing (Figs. 7 and 8). However, As *NbIP11* and *Arabidopsis* EMB175 family members are missing a critical glutamic

acid residue that is intimately involved in the proposed “gated zinc shutter” mechanism, it is not currently clear whether *NbIP11* can edit its target RNAs by itself. We suggest that PPR/variant DYW proteins such as *NbIP11* may a) be catalytically active through a mechanism that is distinct from that of canonical DYW proteins such as OTP86; or b) may not be catalytically active themselves but may function as a specific scaffold or adaptor for a catalytically active partner. The latter could include other PPR proteins within editosome complexes.

Our analysis does not rule out the possibility that the observed reductions in editing in *NbIP11*-silenced plants are a secondary effect of defective ribosome assembly and the

Fig. 8 Effects of knockdown of *NbIP1* or *NbISE2* on editing efficiency. RNA editing compared between *N. benthamiana* plants displaying chlorosis resulting from silencing *IP11*, *ISE2* or *PDS* and TRV-infected, non-silenced controls by Sanger sequencing of transcripts. Results represent two non-silencing control biological replicates, three biological replicates for silencing *IP11* or *ISE2*, and one biological replicate for silencing *PDS*. Error bars represent standard deviation for each replicate. Asterisks denote P -value < 0.05 as determined using a one-tailed Student's t -test assuming unequal variance, pairwise comparisons were made between controls and silenced plants



resulting chlorosis. However, editing at some sites, *ndhB-5* and *-6* for example, was only affected in the *NbIP11*-silenced plants (Figs. 7 and 8). It is typical that loss of catalytic RNA editing factors will result in a complete abrogation of editing of their target editing sites, but there was only about a 30–50% reduction in editing efficiency of *ndhB-5* and *-6* in the silenced plants. It is possible that this reduction (and not loss) was due to the VIGS assay causing a reduction of *NbIP11* levels and not total loss of the protein combined with the heterogeneity of VIGS. Further, if *NbIP11* is not the catalytic protein involved in editing a transcript but instead is playing a supporting role in the context of a larger complex, it is possible that there would be remnant editing activity when *NbIP11* levels are reduced.

NbISE2 and NbIP1 function in RNA editing in *N. benthamiana*

ISE2 is an evolutionarily conserved chloroplast-localized RNA helicase that has many roles in RNA processing (Carlotto et al. 2016; Bobik et al. 2017). ISE2 is required for C-to-U RNA editing at multiple sites in Arabidopsis (Bobik et al. 2017), and the maize ortholog was identified in multi-protein complexes that edited the C473 site in maize *ndhA* transcripts (Sandoval et al. 2019). The ISE2-containing editing complex from maize also contained the non-PPR editing factors RIP1/MORF8 and RIP9 as well as ORRM, containing an RNA-Recognition Motif (RRM) and OZ1, a RanBP2-type Zn finger protein. Several PPR proteins including six P-type PPR proteins previously

not known to be involved with RNA editing were also identified (Sandoval et al. 2019). The interaction of ISE2 with other editing factors and PPR proteins is consistent with the identification of several RNA binding proteins including IPI1 as interacting with ISE2 in a yeast two-hybrid screen (Bobik et al. 2019). Reduced ISE2 levels in Arabidopsis led to significantly reduced editing specifically at three sites (*rpoB-338*, *rpoB-551* and *rps14-149*) (Bobik et al. 2017). Interestingly, none of these were among the top predicted targets for NbIPI1 binding (Fig. 4) and the editing of the equivalent sites was unaffected in *NbIPI1*-silenced leaves although there was a small but statistically significant increase for the *rpoB-1* (Arabidopsis *rpoB-338*) site (Fig. 7). Perhaps IPI1 and ISE2 may function as part of an editing complex for only some sites, although this remains to be tested experimentally. It is also possible that IPI1 and ISE2 may function as part of a dynamic editing complex for some sites dependent upon plant developmental stage or tissue type. The composition of the possible NbIPI1-containing editosome should be the focus of future investigations as such information will reveal more about chloroplast RNA editing and the molecular machineries involved.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11103-024-01424-1>.

Acknowledgements We thank Dr Saurav Misra for his assistance with structural analysis and use of AlphaFold2, and for helpful comments on the manuscript. We thank George Butler for excellent assistance with plant growth and VIGS, Drs. Joe May and Veronica Brown of the University of Tennessee, Knoxville Molecular Biology Research Facility for assistance with Sanger sequencing and Next-gen sequencing, respectively.

Author's contribution TNM and TBS conceived the work and designed the experiments, TNM, MFA and TBS acquired, analyzed, and interpreted the data. MFA contributed significantly to figure production. TNM and TBS wrote and edited the manuscript. All authors read and approved the final manuscript.

Funding This research was supported by NSF awards IOS 1456761 and MCB 2210127 to TBS. TNM was supported by an NIH PEER Fellowship, NIH-2R25GM086761.

Data Availability The dataset supporting the conclusions of this article is available in the NCBI Sequence Read Archive (SRA) repository, PRJNA826083 and <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA826083>.

Declarations

Conflict of interest The author(s) declare(s) that they have no competing financial or nonfinancial interests to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated

otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Asakura Y, Hirohashi T, Kikuchi S, Belcher S, Osborne E, Yano S, Terashima I, Barkan A, Nakai M (2004) Maize mutants lacking chloroplast FtsY exhibit pleiotropic defects in the biogenesis of thylakoid membranes. *Plant Cell* 16(1):201–214. <https://doi.org/10.1105/tpc.014787>
- Barkan A, Rojas M, Fujii S, Yap A, Chong YS, Bond CS, Small I (2012) A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet* 8(8):e1002910. <https://doi.org/10.1371/journal.pgen.1002910>
- Barkan A, Small I (2014) Pentatricopeptide repeat proteins in plants. *Annu Rev Plant Biol* 65:415–442. <https://doi.org/10.1146/annurev-arplant-050213-040159>
- Bayer-Csaszar E, Haag S, Jorg A, Glass F, Hartel B, Obata T, Meyer EH, Brennicke A, Takenaka M (2017) The conserved domain in morf proteins has distinct affinities to the PPR and E elements in PPR RNA editing factors. *Biochim Biophys Acta Gene Regul Mech* 1860(8):813–828. <https://doi.org/10.1016/j.bbagr.2017.05.004>
- Beick S, Schmitz-Linneweber C, Williams-Carrier R, Jensen B, Barkan A (2008) The Pentatricopeptide repeat protein PPR5 stabilizes a specific tRNA precursor in maize chloroplasts. *Mol Cell Biol* 28(17):5337–5347. <https://doi.org/10.1128/MCB.00563-08>
- Bentolila S, Heller WP, Sun T, Babina AM, Friso G, van Wijk KJ, Hanson MR (2012) RIP1, a member of an arabidopsis protein family, interacts with the protein RARE1 and broadly affects RNA editing. *Proc Natl Acad Sci USA* 109(22):E1453–E1461. <https://doi.org/10.1073/pnas.1121465109>
- Bentolila S, Oh J, Hanson MR, Bukowski R (2013) Comprehensive high-resolution analysis of the role of an arabidopsis gene family in RNA editing. *PLoS Genet* 9(6):e1003584. <https://doi.org/10.1371/journal.pgen.1003584>
- Bobik K, Fernandez JC, Hardin SR, Ernest B, Ganusova EE, Staton ME, Burch-Smith TM (2019) The essential chloroplast ribosomal protein uL15c interacts with the chloroplast RNA helicase ISE2 and affects intercellular trafficking through plasmodesmata. *New Phytol* 221(2):850–865. <https://doi.org/10.1111/nph.15427>
- Bobik K, McCray TN, Ernest B, Fernandez JC, Howell KA, Lane T, Staton M, Burch-Smith TM (2017) The chloroplast RNA helicase ISE2 is required for multiple chloroplast RNA processing steps in arabidopsis thaliana. *Plant J* 91(1):114–131. <https://doi.org/10.1111/tpj.13550>
- Boussard C, Avon A, Kindgren P, Bond CS, Challenor M, Lurin C, Small I (2014) The cytidine deaminase signature HxE(x)n CxxC of DYW1 binds zinc and is necessary for RNA editing of ndhD-1. *New Phytol* 203(4):1090–1095. <https://doi.org/10.1111/nph.12928>
- Bryant N, Lloyd J, Sweeney C, Myouga F, Meinke D (2011) Identification of nuclear genes encoding chloroplast-localized proteins required for embryo development in arabidopsis. *Plant Physiol* 155(4):1678–1689. <https://doi.org/10.1104/pp.110.168120>
- Burch-Smith TM, Brunkard JO, Choi YG, Zambryski PC (2011) Organelle-nucleus cross-talk regulates plant intercellular communication via plasmodesmata. *Proc Natl Acad Sci U S A* 108(51):E1451–E1460. <https://doi.org/10.1073/pnas.1117226108>
- Burch-Smith TM, Zambryski PC (2010) Loss of increased size exclusion limit (ISE)1 or ISE2 increases the formation of secondary

- plasmodesmata. *Curr Biol* 20(11):989–993. <https://doi.org/10.1016/j.cub.2010.03.064>
- Carlotto N, Wirth S, Furman N, Ferreyra Solari N, Ariel F, Crespi M, Kobayashi K (2016) The chloroplastic DEVH-Box RNA helicase increased size exclusion limit 2 involved in plasmodesmata regulation is required for group II intron splicing. *Plant Cell Environ* 39(1):165–173. <https://doi.org/10.1111/pce.12603>
- Cushing DA, Forsthoefel NR, Gestaut DR, Vernon DM (2005) Arabidopsis Emb175 and other Ppr knockout mutants reveal essential roles for pentatricopeptide repeat (PPR) proteins in plant embryogenesis. *Planta* 221(3):424–436. <https://doi.org/10.1007/s00425-004-1452-x>
- Ding YH, Liu NY, Tang ZS, Liu J, Yang WC (2006) Arabidopsis glutamine-rich protein23 is essential for early embryogenesis and encodes a novel nuclear PPR motif protein that interacts with RNA polymerase II subunit III. *Plant Cell* 18(4):815–830. <https://doi.org/10.1105/tpc.105.039495>
- Fu M, Lin X, Zhou Y, Zhang C, Liu B, Feng D, Wang J, Wang H, Jin H (2022) OTP970 is required for RNA editing of chloroplast *ndhB* transcripts in arabidopsis thaliana. *Genes* 13(1):139. <https://doi.org/10.3390/genes13010139>
- Ganusova EE, Reagan BC, Fernandez JC, Azim MF, Sankoh AF, Freeman KM, McCray TN, Patterson K, Kim C, Burch-Smith TM (2020) Chloroplast-to-nucleus retrograde signalling controls intercellular trafficking via plasmodesmata formation. *Philos Trans R Soc Lond B Biol Sci* 375(1801):20190408. <https://doi.org/10.1098/rstb.2019.0408>
- Glass F, Härtel B, Zehrmann A, Verbitskiy D, Takenaka M (2015) MEF13 requires MORF3 and MORF8 for RNA editing at eight targets in mitochondrial mRNAs in arabidopsis thaliana. *Mol Plant* 8(10):1466–1477. <https://doi.org/10.1016/j.molp.2015.05.008>
- Grant CE, Bailey TL, Noble WS (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27(7):1017–1018. <https://doi.org/10.1093/bioinformatics/btr064>
- Gutmann B, Royan Mareike Schallenberg-Rüdinger S, Lenz H, Castleden IR, McDowell R, Vacher MA et al. (2020a) The expansion and diversification of pentatricopeptide repeat RNA-editing factors in plants. *Molecular*
- Haag S, Schindler M, Berndt L, Brennicke A, Takenaka M, Weber G (2017) Crystal structures of the Arabidopsis thaliana organellar RNA editing factors MORF1 and MORF9. *Nucleic Acids Res* 45(8):4915–4928. <https://doi.org/10.1093/nar/gkx099>
- Hammani K, Takenaka M, Miranda R, Barkan A (2016) A PPR protein in the PLS subfamily stabilizes the 5'-end of processed Rpl16 mRNAs in maize chloroplasts. *Nucleic Acids Res* 44(9):4278–4288. <https://doi.org/10.1093/nar/gkw270>
- Hammani K, Okuda K, Tanz SK, Chateigner-Boutin A-L, Shikanai T, Small I (2009) A study of new Arabidopsis chloroplast rna editing mutants reveals general features of editing factors and their target sites. *Plant Cell* 21(11):3686–3699. <https://doi.org/10.1105/tpc.109.071472>
- Hayes ML, Dang KN, Diaz MF, Michael Mulligan R (2015) A Conserved glutamate residue in the C-terminal deaminase domain of pentatricopeptide repeat proteins is required for RNA editing activity. *J Biol Chem* 290(16):10136–10142. <https://doi.org/10.1074/jbc.M114.631630>
- Hayes ML, Giang K, Berhane B, Michael Mulligan R (2013) Identification of two pentatricopeptide repeat genes required for RNA editing and zinc binding by C-terminal cytidine deaminase-like domains. *J Biol Chem* 288(51):36519–36529. <https://doi.org/10.1074/jbc.M113.485755>
- Hayes ML, Santibanez PI (2020) A Plant Pentatricopeptide Repeat protein with a DYW-deaminase domain is sufficient for catalyzing C-to-U RNA editing in vitro. *J Biol Chem* 295(11):3497–3505. <https://doi.org/10.1074/jbc.RA119.011790>
- Hegeman CE, Hayes ML, Hanson MR (2005) Substrate and cofactor requirements for RNA editing of chloroplast transcripts in arabidopsis in vitro: substrate and cofactor requirements for RNA editing in arabidopsis. *Plant J* 42(1):124–132. <https://doi.org/10.1111/j.1365-3113X.2005.02360.x>
- Hirose T, Kusumegi T, Tsudzuki T, Sugiura M (1999) RNA editing sites in tobacco chloroplast transcripts: editing as a possible regulator of chloroplast RNA polymerase activity. *Mol Gen Genet* 262(3):462–467
- Iyer LM, Zhang D, Rogozin IB, Aravind L (2011) Evolution of the deaminase fold and multiple origins of eukaryotic editing and mutagenic nucleic acid deaminases from bacterial toxin systems. *Nucleic Acids Res* 39(22):9473–9497. <https://doi.org/10.1093/nar/gkr691>
- Kakizaki T, Matsumura H, Nakayama K, Che FS, Terauchi R, Inaba T (2009) Coordination of plastid protein import and nuclear gene expression by plastid-to-nucleus retrograde signaling. *Plant Physiol* 151(3):1339–1353. <https://doi.org/10.1104/pp.109.145987>
- Klinkenberg J (2014) Extraction of chloroplast proteins from transiently transformed Nicotiana Benthamiana Leaves. *BIO-PROTOCOL* 4 (18). <https://doi.org/10.21769/BioProtoc.1238>
- Klinkenberg J, Faist H, Saupe S, Lambert S, Krischke M, Stingl N, Fekete A et al (2014) Two fatty acid desaturases, stearyl-acyl carrier protein δ^9 -desaturase6 and fatty acid desaturase3, are involved in drought and hypoxia stress signaling in arabidopsis crown galls. *Plant Physiol* 164(2):570–583. <https://doi.org/10.1104/pp.113.230326>
- Kobayashi T, Yagi Y, Nakamura T (2019) Comprehensive prediction of target RNA editing sites for PLS-class PPR proteins in Arabidopsis Thaliana. *Plant Cell Physiol* 60(4):862–874. <https://doi.org/10.1093/pcp/pcy251>
- Koussevitzky S, Nott A, Mockler TC, Hong F, Sachetto-Martins G, Surpin M, Lim J, Mittler R, Chory J (2007) Signals from chloroplasts converge to regulate nuclear gene expression. *Science* 316(5825):715–719. <https://doi.org/10.1126/science>
- Liu Y, He J, Chen Z, Ren X, Hong X, Gong Z (2010) ABA over-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(5):749–765. <https://doi.org/10.1111/j.1365-3113X.2010.04280.x>
- Mach J (2009) Chloroplast RNA editing by pentatricopeptide repeat proteins. *Plant Cell* 21(1):17. <https://doi.org/10.1105/tpc.109.210114>
- Maier UG, Bozarth A, Funk HT, Zauner S, Rensing SA, Schmitz-Linneweber C, Börner T, Tillich M (2008) Complex chloroplast RNA metabolism: Just debugging the genetic programme? *BMC Biol* 6(1):36. <https://doi.org/10.1186/1741-7007-6-36>
- Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M (2022) ColabFold: making protein folding accessible to all. *Nat Methods* 19(6):679–682. <https://doi.org/10.1038/s41592-022-01488-1>
- Okuda K, Chateigner-Boutin AL, Nakamura T, Delannoy E, Sugita M, Myouga F, Motohashi R, Shinozaki K, Small I, Shikanai T (2009) Pentatricopeptide repeat proteins with the DYW motif have distinct molecular functions in RNA Editing and RNA cleavage in arabidopsis chloroplasts. *Plant Cell* 21(1):146–156. <https://doi.org/10.1105/tpc.108.064667>
- Okuda K, Hammani K, Tanz SK, Peng L, Fukao Y, Myouga F, Motohashi R, Shinozaki K, Small I, Shikanai T (2010) The Pentatricopeptide repeat protein OTP82 is required for RNA editing of plastid *ndhB* and *ndhG* transcripts. *Plant J* 61(2):339–349. <https://doi.org/10.1111/j.1365-3113X.2009.04059.x>

- Okuda K, Shoki H, Arai M, Shikanai T, Small I, Nakamura T (2014) Quantitative analysis of motifs contributing to the interaction between PLS-subfamily members and their target RNA sequences in plastid RNA editing. *Plant J* 80(5):870–882. <https://doi.org/10.1111/tjp.12687>
- Oldenkott B, Yang Y, Lesch E, Knoop V, Schallenberg-Rüdinger M (2019) Plant-type pentatricopeptide repeat proteins with a DYW domain drive C-to-U RNA editing in *Escherichia coli*. *Commun Biol* 2(1):85. <https://doi.org/10.1038/s42003-019-0328-3>
- Saha D, Prasad AM, Srinivasan R (2007) Pentatricopeptide repeat proteins and their emerging roles in plants. *Plant Physiol Biochem* 45(8):521–534. <https://doi.org/10.1016/j.plaphy.2007.03.026>
- Salone V, Rüdinger M, Polsakiewicz M, Hoffmann B, Groth-Malonek M, Szurek B, Small I, Knoop V, Lurin C (2007) A Hypothesis on the identification of the editing enzyme in plant organelles. *FEBS Lett* 581(22):4132–4138. <https://doi.org/10.1016/j.febslet.2007.07.075>
- Sandoval R, Boyd RD, Kiszter AN, Mirzakhanyan Y, Santibañez P, Gershon PD, Hayes ML (2019) Stable native RIP 9 complexes associate with C-to-U RNA editing activity, PPR s, RIP s, OZ 1, ORRM 1 and ISE 2. *Plant J* 99(6):1116–1126. <https://doi.org/10.1111/tjp.14384>
- Sasaki T, Yukawa Y, Miyamoto T, Obokata J, Sugiura M (2003) Identification of RNA editing sites in chloroplast transcripts from the maternal and paternal progenitors of tobacco (*Nicotiana Tabacum*): comparative analysis shows the involvement of distinct trans-factors for *ndhB* editing. *Mol Biol Evol* 20(7):1028–1035. <https://doi.org/10.1093/molbev/msg098>
- Schmitz-Linneweber C, Small I (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci* 13(12):663–670. <https://doi.org/10.1016/j.tplants.2008.10.001>
- Shikanai T (2006) RNA editing in plant organelles: machinery, physiological function and evolution. *Cell Mol Life Sci* 63(6):698–708. <https://doi.org/10.1007/s00018-005-5449-9>
- Small ID, Schallenberg-Rüdinger M, Takenaka M, Mireau H, Oster-setzer-Biran O (2020) Plant organellar RNA editing: what 30 years of research has revealed. *Plant J* 101(5):1040–1056. <https://doi.org/10.1111/tjp.14578>
- Stern DB, Goldschmidt-Clermont M, Hanson MR (2010) Chloroplast RNA metabolism. *Annu Rev Plant Biol* 61:125–155. <https://doi.org/10.1146/annurev-arplant-042809-112242>
- Takenaka M, Brennicke A (2003) In vitro RNA editing in pea mitochondria requires NTP or dNTP, suggesting involvement of an RNA helicase. *J Biol Chem* 278(48):47526–47533. <https://doi.org/10.1074/jbc.M305341200>
- Takenaka M, Zehrmann A, Verbitskiy D, Kugelmann M, Hartel B, Brennicke A (2012) Multiple organellar RNA editing factor (MORF) family proteins are required for RNA editing in mitochondria and plastids of plants. *Proc Natl Acad Sci USA* 109(13):5104–5109. <https://doi.org/10.1073/pnas.1202452109>
- Takenaka M, Takenaka S, Barthel T, Frink B, Haag S, Verbitskiy D, Oldenkott B et al (2021) DYW domain structures imply an unusual regulation principle in plant organellar RNA editing catalysis. *Nat Catal* 4(6):510–522. <https://doi.org/10.1038/s41929-021-00633-x>
- Takenaka M, Zehrmann A, Brennicke A, Graichen K (2013) Improved computational target site prediction for pentatricopeptide repeat RNA editing factors, Edited by Stefan Maas. *PLoS ONE* 8(6):e65343. <https://doi.org/10.1371/journal.pone.0065343>
- Tian F, Yu J, Zhang Y, Xie Y, Wu B, Miao Y (2019) MORF9 functions in plastid RNA editing with tissue specificity. *Int J Mol Sci* 20(18). <https://doi.org/10.3390/ijms20184635>
- Toma-Fukai S, Sawada Y, Maeda A, Shimizu H, Shikanai T, Takenaka M, Shimizu T (2023) Structural insight into the activation of an arabidopsis organellar C-to-U RNA editing enzyme by active site complementation. *Plant Cell* 35(6):1888–1900. <https://doi.org/10.1093/plcell/koac318>
- Tseng CC, Lee CJ, Chung YT, Sung TY, Hsieh MH (2013) Differential regulation of arabidopsis plastid gene expression and RNA editing in non-photosynthetic tissues. *Plant Mol Biol* 82(4–5):375–392. <https://doi.org/10.1007/s11103-013-0069-5>
- Wagoner JA, Sun T, Lin L, Hanson MR (2015) Cytidine deaminase motifs within the DYW domain of two pentatricopeptide repeat-containing proteins are required for site-specific chloroplast RNA editing. *J Biol Chem* 290(5):2957–2968. <https://doi.org/10.1074/jbc.M114.622084>
- Wang X, An Y, Pan Xu, Xiao J (2021) Functioning of PPR proteins in organelle RNA metabolism and chloroplast biogenesis. *Front Plant Sci* 12(February):627501. <https://doi.org/10.3389/fpls.2021.627501>
- Yan J, Zhang Q, Guan Z, Wang Q, Li L, Ruan F, Lin R, Zou T, Yin P (2017) MORF9 Increases the RNA-binding activity of PLS-type pentatricopeptide repeat protein in plastid RNA editing. *Nat Plants* 3(April):17037. <https://doi.org/10.1038/nplants.2017.37>
- Yang J, Zhang M, Wang X (2018) Crystal structure of the chloroplast RNA editing factor MORF2. *Biochem Biophys Res Commun* 495(2):2038–2043. <https://doi.org/10.1016/j.bbrc.2017.12.044>
- Zeltz P, Hess WR, Neckermann K, Borner T, Kossel H (1993) Editing of the chloroplast *rpoB* transcript is independent of chloroplast translation and shows different patterns in barley and maize. *EMBO J* 12(11):4291–4296
- Zhu Q, Dugardeyn J, Zhang C, Muhlenbock P, Eastmond PJ, Valcke R, De Coninck B et al (2014) The arabidopsis thaliana RNA editing factor SLO2, which affects the mitochondrial electron transport chain, participates in multiple stress and hormone responses. *Mol Plant* 7(2):290–310. <https://doi.org/10.1093/mp/sst102>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.