PUMPKIN, the Sole Plastid UMP Kinase, Associates with Group II Introns and Alters Their Metabolism[1]

Lisa-Marie Schmid,[a] Lisa Ohler,[b] Torsten Möhlmann,[b] Andreas Brachmann,[c] Jose M. Muiño,[d] Dario Leister,[a] Jörg Meurer,[a,2] and Nikolay Manavski[a,3,4]

[a]Plant Sciences, Faculty of Biology, Ludwig-Maximilians-University Munich, Großhaderner Street 2-4, 82152 Planegg-Martinsried, Germany
[b]Plant Physiology, Faculty of Biology, University of Kaiserslautern, Erwin Schrödinger Street, 67653 Kaiserslautern, Germany
[c]Genetics, Faculty of Biology, Ludwig-Maximilians-University Munich, Großhaderner Street 2-4, 82152 Planegg-Martinsried, Germany
[d]Humboldt University, Faculty of Life Science, Philipp Street 13, 10115 Berlin, Germany

ORCID IDs: 0000-0001-8853-3087 (L.O.); 0000-0002-5676-2042 (T.M.); 0000-0001-7980-8173 (A.B.); 0000-0002-6403-7262 (J.M.M.); 0000-0003-1897-8421 (D.L.); 0000-0003-2973-9514 (J.M.); 0000-0003-2740-5991 (N.M.).

The chloroplast hosts photosynthesis and a variety of metabolic pathways that are essential for plant viability and acclimation processes. In this study, we show that the sole plastid UMP kinase (PUMPKIN) in Arabidopsis (*Arabidopsis thaliana*) associates specifically with the introns of the plastid transcripts *trnG*-UCC, *trnV*-UAC, *petB, petD,* and *ndhA* in vivo, as revealed by RNA immunoprecipitation coupled with deep sequencing (RIP-Seq); and that PUMPKIN can bind RNA efficiently in vitro. Analyses of target transcripts showed that PUMPKIN affects their metabolism. Null alleles and knockdowns of *pumpkin* were viable but clearly affected in growth, plastid translation, and photosynthetic performance. In *pumpkin* mutants, the levels of many plastid transcripts were reduced, while the amounts of others were increased, as revealed by RNA-Seq analysis. PUMPKIN is a homomultimeric, plastid-localized protein that forms in vivo RNA-containing megadalton-sized complexes and catalyzes the ATP-dependent conversion of UMP to UDP in vitro with properties characteristic of known essential eubacterial UMP kinases. A moonlighting function of PUMPKIN combining RNA and pyrimidine metabolism is discussed.

Posttranscriptional events seem to play a larger role in regulating chloroplast gene expression than in the cyanobacterial ancestor. This regulation may be important to enable tissue-specific and developmental adaptations and responses to environmental inputs during stress and acclimation processes (Barkan, 2011a). These processes include changes of the RNA sequence, endonucleolytic cleavage, UTR trimming, RNA decay, as well as translation (Stern et al., 2010; Germain et al., 2013; Manavski et al., 2018). RNA metabolism represents a fast-evolving process. This progression depends mostly on vascular plant-specific, nuclear-encoded RNA-binding proteins and diverse conserved ribonucleases with their associated proteins (Jacobs and Kück, 2011; Manavski et al., 2012; Stoppel et al., 2012; Stoppel and Meurer, 2012, 2013).

Numerous plastid endo- and exoribonuclease-sensitive sites within genes, introns, and intergenic regions imply their complex and important role in governing transcript degradation, thus determining transcript half-lives. A variety of different nuclear encoded, newly evolved RNA-binding proteins, predominantly numerous helical-repeat proteins, bind these specific sites to safeguard transcript stability by preventing endonucleolytic attacks and progression of exonucleolytic degradation (Stern et al., 2010; Jacobs and Kück, 2011; Stoppel and Meurer, 2012, 2013; Germain et al., 2013). Depending on the number of RNA targets and secondary effects, corresponding nuclear mutants display more or less pleiotropic effects related to plastid mRNA patterns and abundance (Pfalz et al., 2009; Belcher et al., 2015; Hammani et al., 2016; Zoschke et al., 2016; Meurer et al., 2017; Teubner et al., 2017).

The acquisition and dispersal of plastid introns within the coding or tRNA regions of chloroplast genes early during the evolution of Streptophyta provide an additional platform for potential regulation of gene

[1]This research was supported by the German Science Foundation, Deutsche Forschungsgemeinschaft (ME 1794/6 to J.M.; MO 1032/4-1 to T.M.; TRR 175 D01 to J.M.M; A05 to D.L.; and A03 to J.M.).

[2]Senior author.

[3]Current address: Centre National de la Recherche Scientifique (CNRS), Institut de Biologie Moléculaire des Plantes, 12 rue du Général Zimmer, 67084 Strasbourg, France.

[4]Author for contact: nikolay.manavski@ibmp-cnrs.unistra.fr.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Nikolay Manavski (nikolay.manavski@ibmp-cnrs.unistra.fr).

L.S., T.M., J.Me., and N.M. designed the research. L.S., L.O., A.B., J.M.M., J.Me., and N.M. performed the research. L.S., T.M., D.L., J.Me., and N.M. prepared the article. J.Me. and N.M. supervised the whole study.

www.plantphysiol.org/cgi/doi/10.1104/pp.18.00687

expression (de Longevialle et al., 2010; Khrouchtchova et al., 2012; Germain et al., 2013; Schmitz-Linneweber et al., 2015). This platform seems especially relevant because splicing of individual plastid introns is often far from being complete, as is the case in the nuclear system. Members of several gene families with unrelated RNA-binding domains, among them APO, CRM, DEAD box, mTERF, PORR, PPR, and RNase III-domain and WHY proteins, act in a combinatorial manner to promote splicing of specific chloroplast introns. This action presumably works to overcome kinetically trapped structures of the fast-evolving intron sequences, allowing the structures to refold into catalytically active ribozymes (Ostersetzer et al., 2005; Germain et al., 2013; Barkan and Small, 2014). Some of the plastid splicing factors are required for additional RNA metabolism steps (Yap et al., 2015; Zhang et al., 2015; Bobik et al., 2017; Tang et al., 2017).

The impact of the metabolic and nutritional status and/or signals in chloroplasts on the plastid RNA metabolism remains largely unknown (Bohne and Nickelsen, 2017). There are just a few instances of proteins that might couple plastid gene expression with other metabolic pathways. One example of how iron regulates the expression of a nuclear gene involved in plastid gene expression was described previously (Douchi et al., 2016). The authors showed that the Half-A-Tetratricopeptide protein MAC1 stabilizes *psaC* transcripts via binding to their 5′ untranslated region in *Chlamydomonas* (*Chlamydomonas reinhardtii*) chloroplasts. MAC1 expression and phosphorylation state depends on iron availability, which in turn influences *psaC* mRNA accumulation. The involvement of the *psbA* mRNA-binding function of the dihydrolipoyl acetyltransferase subunit of the chloroplast pyruvate dehydrogenase complex was recently suggested to link photosynthetic protein and carbon biosynthesis in *Chlamydomonas reinhardtii* (Bohne et al., 2013). Under oxidizing conditions, the RNA-binding N terminus of the large Rubisco subunit becomes exposed and binds RNAs for regulation of translation, thus connecting the redox status with plastid gene expression (Yosef et al., 2004; Cohen et al., 2006). Likewise, predicted binding domains of the HCF145 protein for secondary metabolites were shown to impact RNA-binding affinity important for stabilization of the *psaA/psaB/rps14* mRNA and accumulation of PSI, suggesting a metabolic control of *psaA/psaB/rps14* gene expression (Lezhneva and Meurer, 2004; Manavski et al., 2015).

Evolutionary conserved eubacterial UMP kinases are essential homohexamers and catalyze the reversible transfer of the c-phosphoryl group from ATP to UMP to yield UDP (Serina et al., 1995; Kafer et al., 2004). UDP is further phosphorylated to UTP in an additional ATP consuming reaction by the nucleoside diphosphate kinase (Zhou et al., 2017). Cellular pyrimidine metabolism is primarily important for the synthesis of RNA as well as galacto- and sulfolipids through UDP-sugars (Zrenner et al., 2006; Okazaki et al., 2009). Here we show that the plastid UMP kinase, which we termed PUMPKIN, has a role in RNA metabolism. With characteristics comparable to eubacterial UMP kinases, PUMPKIN exhibits a high specificity toward UMP as its nucleotide substrate, and it converts UMP and ATP to UDP and ADP in vitro. With respect to PUMPKIN's role in RNA metabolism, it does not stabilize *psaA/psaB/rps14* transcripts specifically as has been proposed previously (Hein et al., 2009). But in contrast, it is associated with a subset of group-II-intron containing plastid transcripts, whose RNA metabolism it impacts.

RESULTS

PUMPKIN Is Required for Normal Growth and Photosynthetic Performance

To functionally characterize PUMPKIN in Arabidopsis (*Arabidopsis thaliana*; At3g18680), we obtained a T-DNA insertion line (GABI_154D02) from the GABI-KAT collection (Kleinboelting et al., 2012) that was generated in the Col-0 (the wild-type) background. The T-DNA was located in the first exon 134 bp downstream of the start codon as revealed by sequencing the PCR-amplified flanking regions (Fig. 1, A–C; for sequence information of oligonucleotides see Supplemental Table S1). In plants homozygous for the T-DNA insertion (*pumpkin*) obtained from the T3 generation, the full-length *pumpkin* mRNA and translational products (PUMPKIN) were not detectable (Fig. 1, D–G), indicating that *pumpkin* was a complete knockout line. The loss of PUMPKIN led to severely retarded growth and a photo-bleached phenotype (Fig. 1B). However, *pumpkin* plants were able to grow photoautotrophically, to reach plant maturity, and to produce viable homozygous seeds. The life cycle of *pumpkin* mutants—from germination to mature seeds —was much longer (~21 weeks) as compared to the wild types (~8 weeks). Introduction of the *PUMPKIN* cDNA fully complemented the *pumpkin* mutant phenotype (Fig. 1, B, D, and F), proving that the observed phenotype was entirely due to the disruption of the *PUMPKIN* gene. This proof is further supported by the fact that substantial down-regulation of Arabidopsis *PUMPKIN* in RNAi lines generated in this study led to a phenotype, which was similar but mitigated as compared to the phenotype of the *pumpkin* mutant (Fig. 1, B, D, and F). Moreover, we also generated knockdowns in tobacco (*Nicotiana tabacum*). Similar to the Arabidopsis RNAi lines, tobacco *PUMPKIN* was considerably downregulated and led to a retarded growth and chlorotic phenotype (Supplemental Fig. S1), suggesting that the function of PUMPKIN is conserved in dicots.

To investigate the effects of loss of PUMPKIN on photosynthetic performance, spectroscopic analyses were carried out with plants grown on soil in the climate chamber under identical conditions: 12-h-light/12-h-dark cycle with a photon flux density (PFD) of 100 μmol photons m^{-2} s^{-1}. While complemented lines were comparable to the wild type, the maximum PSII

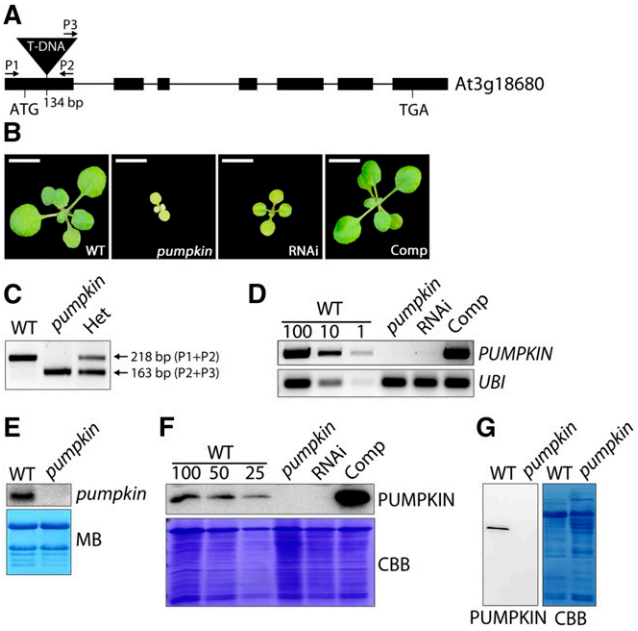

Figure 1. Phenotype and Analysis of the *pumpkin* T-DNA Insertion Line. A, Schematic representation of the Arabidopsis *PUMPKIN* gene (At3g18680). Exons are depicted as black bars, introns as black lines. The T-DNA insertion is represented by the black triangle, and its position is indicated relative to the start codon (134 bp). Primers used for genotyping in (C) are shown as arrows. Primer information can be found in Supplemental Table S1. ATG, start codon; P, Primer; TGA, stop codon. B, Images of the 3-week-old wild type, *pumpkin,* RNAi, and complemented (Comp) plants grown on soil. White scale bars correspond to 0.5 cm. C, PCR-based genotyping analysis. PCR on genomic DNA with three primers (P1, P2, P3) shown in (A) was conducted in order to distinguish between the wild-type, homozygous (*pumpkin*), and heterozygous (Het) *PUMPKIN* alleles. P1 and P2 yielded an amplicon of 218 bp corresponding to the expected size of the wild-type allele, while the combination of P2 and P3 amplified the mutated allele (163 bp). D, Analysis of the relative expression of *PUMPKIN* in the wild type, *pumpkin,* RNAi, and Comp plants by means of RT-PCR. (D) illustrates that 100%, 10%, and 1% of the wild type correspond to 100 ng, 10 ng, and 1 ng of total RNA used as template for cDNA syntheses. *UBIQUITIN* (*UBI*) served as control. E, RNA-gel–blot analysis of *pumpkin* RNA. The PCR probe was amplified with the same primers as in (D). MB, Methylene blue. F, Immunodetection of PUMPKIN in total protein extracts from the wild type, *pumpkin,* RNAi, and Comp plants using PUMPKIN-specific antibodies. In (F), 100% of the wild type corresponds to 30 μg total proteins. CBB, Coomassie Brilliant Blue. G, Immunodetection of PUMPKIN in stroma extracts (40 μg) from the wild type and *pumpkin* to prove the specificity of the PUMPKIN antibody. CBB, Coomassie Brilliant Blue; WT, wild type.

quantum yield—expressed as the ratio of the variable to the maximum fluorescence (Fv/Fm) in *pumpkin* and RNAi lines—was reduced to about 50% and 85%, respectively, indicating strong defects in PSII in mutant lines (Supplemental Fig. S2, A and B). Different from mutants in PSI and the linear electron transport, the fluorescence dropped far below the F_0 level during induction in *pumpkin* and RNAi lines. This drop was accompanied by a comparable decrease in the effective quantum yield and a severe increase in nonphotochemical quenching of excited states. After induction, the fluorescence increased again and reached the former F_0 level after about 4 min in *pumpkin*. This increase reflects an increased proton gradient in the light—thus reduced ATP synthase activity. The quantum yield of PSI was almost unchanged, although the overall signal was considerably reduced in mutant and knock down lines. In summary, *pumpkin* and RNAi lines displayed a pleiotropic phenotype with pronounced defects in the major complexes of the thylakoid membrane.

Lack of PUMPKIN Affects Accumulation of Photosynthetic Subunits and Translation

Because *pumpkin* exhibits photosynthetic defects, we next investigated the steady-state levels of constituent subunits of the major thylakoid membrane complexes by immunoblot analysis. In *pumpkin*, the assemblies PSI (PsaA, PsaF), PSII (D1, CP43, CP47), Cytochrome b_6f (Cytb_6), ATP synthase (AtpC), LHCs (LHCa3, LHCb2), and NDH (NdhH) accumulated to roughly 25%, 20% to 40%, 25%, 25%, 75% to 120%, and 25% of the wild type, respectively (Fig. 2A). As expected, the RNAi line performed slightly better than *pumpkin*, with all tested proteins present to roughly 50% of the wild type (except for LHCs that accumulated to normal levels), while the complemented line behaved essentially like the wild type.

To analyze the formation of photosynthetic complexes, blue native (BN)-PAGE with solubilized thylakoid complexes was performed. All complexes were detectable indicating that assembly was not affected in *pumpkin* (Stoppel et al., 2011). Similar to the outcome of the immunoblot analysis, levels of PSI, PSII, cytochrome b_6f, ATP synthase, and NDH complexes were uniformly reduced in the *pumpkin* mutant, while LHCbs over-accumulated as compared to the wild type (Supplemental Fig. S3). Overall, these results suggest that the observed decreases in the levels of all thylakoid membrane complexes account for the photosynthetic deficiencies in the *pumpkin* mutants presented previously.

To investigate whether the reduced levels of proteins are due to translational impairments, de novo synthesis of chloroplast proteins was analyzed. It appeared that the incorporation of [^{35}S]-Met into newly synthesized thylakoid proteins was, in general, reduced to roughly 50% in *pumpkin* and RNAi lines, while levels of labeled proteins in the complemented line were comparable to the wild type, as revealed by repeated in vivo labeling experiments (Fig. 2B). Labeling of soluble proteins, with Rubisco's large subunit as the main representative, was reduced to an even larger extent (less than 25% of the wild type), indicating general translational deficiencies in *pumpkin* (Fig. 2C).

Besides levels of photosynthetic subunits, reduced translation efficiency usually impacts the levels of the plastid-encoded RNA polymerase (PEP), which in turn leads to reduced transcription rates of PEP-dependent

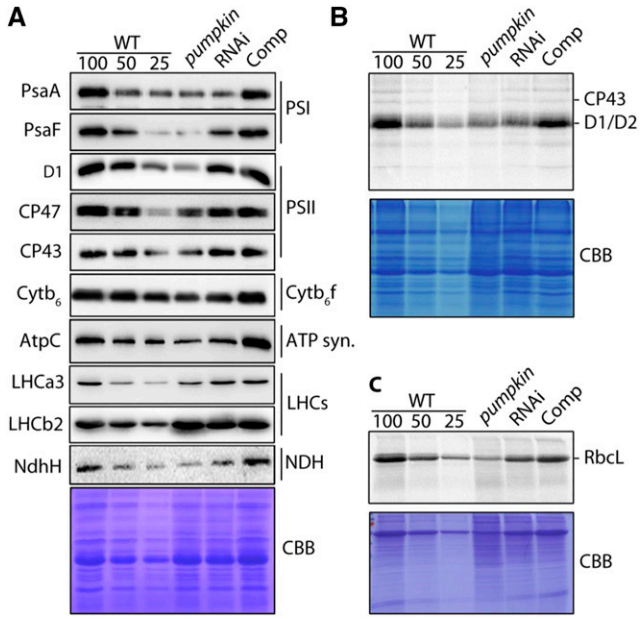

Figure 2. Protein Accumulation and Translation Rates in *pumpkin*. A, Immunotitration of thylakoid proteins in *pumpkin*. Thylakoids were extracted from the wild-type, *pumpkin*, RNAi, and Comp plants, separated on SDS gels and immunodecorated with indicated antibodies. In (A), 100, 50, and 25 of the wild type corresponds to 5 μg, 2.5 μg, and 1.25 μg chlorophyll, respectively. ATP syn., ATP synthase; CBB, Coomassie Brilliant Blue; CP43, PsbC; CP47, PsbB; Cytb_6f, cytochrome b_6f complex; D1, PsbA; LHC, light harvesting complex; NDH, NADH dehydrogenase-like complex; PSI, photosystem I; PSII, photosystem II. B, [^{35}S]-Met labeling of newly synthesized thylakoid proteins in the wild type, *pumpkin*, RNAi, and Comp plants. Insoluble protein fractions were extracted and separated on SDS-gels. Equal loading was achieved according to the determined counts/min (cpm) of the protein fractions (100 corresponds to 100,000 cpm). CBB, Coomassie Brilliant Blue; CP43, PsbC; D1, PsbA; D2, PsbD. C, In vivo labeling of newly synthesized soluble proteins in the wild type, *pumpkin*, RNAi, and Comp plants. Refer to (B). C, 100 corresponds to 1,000,000 cpm. RbcL, RUBISCO large subunit; WT, wild type.

genes, while transcription of nucleus-encoded RNA polymerase (NEP)-dependent genes increases (Meurer et al., 2017). To check whether this notion is also true for *pumpkin*, we performed transcriptome-wide expression analysis by RNA deep sequencing (-Seq) of rRNA-depleted total leaf RNA obtained from the *pumpkin* mutant and the wild type. Indeed, our RNA-Seq analysis revealed that levels of transcripts of chloroplast genes known to be preferentially transcribed by the PEP were reduced, whereas NEP-dependent transcripts accumulated (Supplemental Fig. S4).

PUMPKIN Localizes in Chloroplasts and Forms RNA-Containing Megadalton Complexes

Several public programs predicted a chloroplast subcellular location of PUMPKIN based on the presence of an N-terminal transit peptide (http://aramemnon.uni-koeln.de). To verify this prediction experimentally, the full-length PUMPKIN protein was fused to YFP and used to monitor the fluorescence in transiently transformed protoplasts of *Nicotiana benthamiana*. The PUMPKIN-YFP signals colocalized exclusively with chlorophyll fluorescence validating chloroplast localization (Fig. 3A). Interestingly, the signals within chloroplasts appeared spotty and prompted us to check for chloroplast nucleoid localization using the nucleoid marker protein RAP (Kleinknecht et al., 2014; Meurer et al., 2017). Merging the fluorescence of the PUMPKIN-YFP and RAP-CFP in cotransformed protoplasts of *N. benthamiana* led to only partial overlap of the signals, making the localization of PUMPKIN in chloroplast nucleoids ambiguous (Fig. 3B).

We next investigated the in vivo formation of PUMPKIN complexes in Arabidopsis chloroplasts. To that end, stroma extracts were fractionated by size-exclusion chromatography followed by SDS PAGE and immunodecoration. PUMPKIN was detected in all fractions ranging from 30 kD to 5 MDa (Fig. 3C). The main protein peak of PUMPKIN was found in the high M_r fractions 12 to 18, which also contained Rubisco (approximately 550 kD). In addition, PUMPKIN appeared to form MDa complexes distributed in fractions 1 to 10. Interestingly, the PUMPKIN protein in fraction one, which is known to contain nucleoids (Olinares et al., 2010), disappeared upon RNase treatment; while the signals in the remaining MDa fractions (5 to 10), which include RNase-sensitive ribosomes (Fig. 3C), weakened considerably. This result indicates that most PUMPKIN MDa complexes contain RNA. Fractions 2 to 4 contained MDa PUMPKIN complexes that comprised either RNase-resistant RNAs or were free of RNA.

PUMPKIN Is Associated with Several Chloroplast Introns in Vivo

The translation deficiency in *pumpkin* (along with the fact that PUMPKIN associates with RNA in vivo) suggests that PUMPKIN might be involved in posttranscriptional RNA metabolism processes. Moreover, pale mutants like *pumpkin* often exhibit pleiotropic effects resulting from translation deficiency associated with altered transcription, which makes it difficult to identify the primary defects (Meurer et al., 2017). We next performed a plastid genome-wide in vivo RNA association assay (RNA immunoprecipitation followed by deep sequencing [RIP-Seq]). The aims were to identify the RNA targets of PUMPKIN to differentiate between primary and secondary defects and thus to pinpoint the cause of the aberrant translation in *pumpkin*. For that purpose, wild-type stroma extracts were incubated with PUMPKIN-specific antibodies or with the corresponding preimmune serum as the negative control. Coimmunoprecipitated (Co-IP) RNAs were subjected to RNA deep sequencing. Resulting reads were mapped to the chloroplast genome and normalized to the gene length and sequencing depth. In two biological

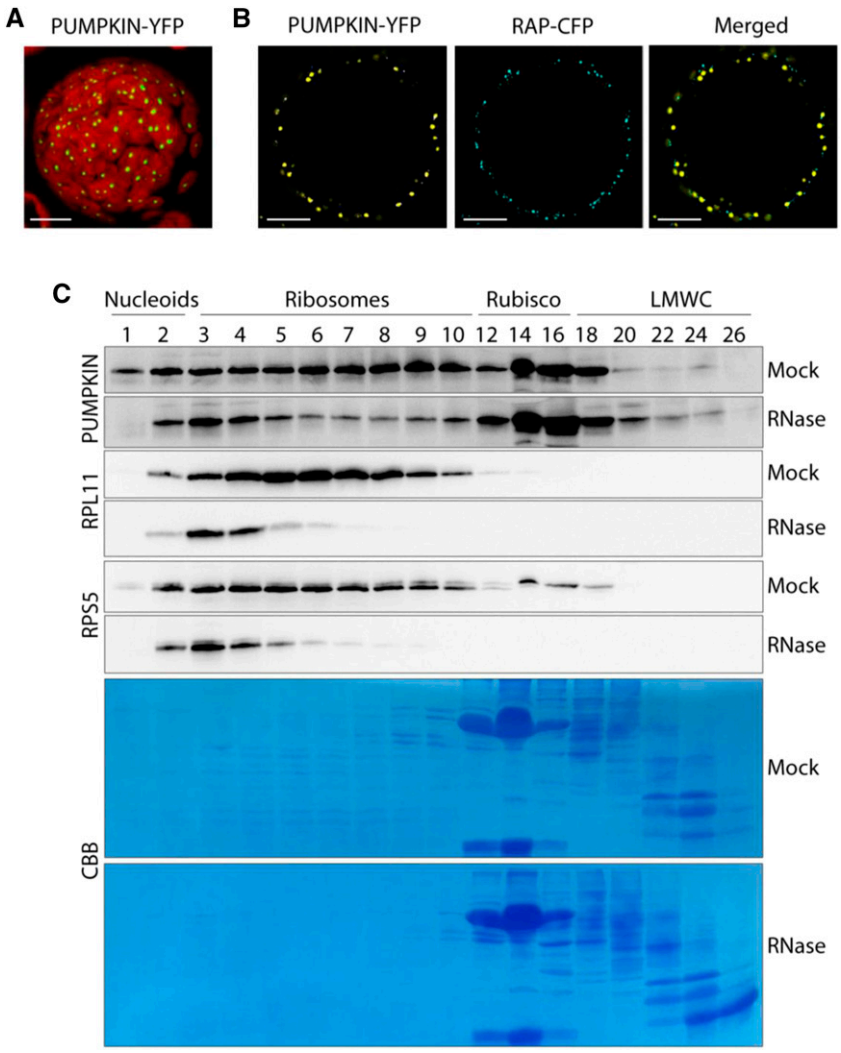

Figure 3. Subcellular Localization and in Vivo Complex Formation of PUMPKIN. A, Confocal image of a PUMPKIN-YFP-transfected *N. benthamiana* protoplast. Merged signals of chloroplast autofluorescence (red) and YFP fluorescence (yellow) are shown. The white scale bar corresponds to 10 μm. B, Macrographs of a PUMPKIN-YFP/RAP-CFP cotransfected *N. benthamiana* protoplast. Single fluorescence images of PUMPKIN-YFP (yellow) and RAP-CFP (blue) as well as a merged image of both are shown. White scale bars correspond to 10 μm. C, Size exclusion chromatography analysis of the wild-type stroma. RNase A-treated and untreated (Mock) stroma was fractionated by gel filtration using Superose 6, 10/300 GL column. Fractions (1 to 26) were separated by SDS PAGE, blotted, and immunodecorated with the indicated antibodies. Probing with antibodies against proteins of the small and large ribosomal subunits served as a control to demonstrate successful RNase-treatment of RNase-sensitive preassembled ribosomes. Indicated complexes were deduced from Olinares et al. (2010). Coomassie Brilliant Blue staining (CBB) of the membranes is shown. LMWC, low molecular weight complexes; RPS5, ribosomal protein S5 (chloroplastic); RPL11, ribosomal protein L11 (chloroplastic).

replicates, PUMPKIN was found to associate with group II introns of five plastid transcripts: *trnG*-UCC, *trnV*-UAC, *petB*, *petD*, and *ndhA* (Fig. 4A; Supplemental Figs. S5 and S6), where the *trnG*-UCC locus represented the most prominent peak in the RIP-Seq analysis. To verify the RIP-Seq results, we performed RIP and hybridized the recovered RNAs with probes specific for the RNA targets in slot-blot experiments (Fig. 4, B and C). Enrichment of the intron regions of *trnG*-UCC, *trnV*-UAC, *petB*, *petD*, and *ndhA* was observed in the pellet fractions, confirming the in vivo association of PUMPKIN with those RNAs. Hybridization with probes recognizing the ribosomal 23S RNA and *psaA* mRNA were used as negative controls to demonstrate that other RNAs were not enriched in the pellet fractions. In addition, the resulting RIP-Seq reads were displayed as a ratio of read coverage (*pumpkin* versus the control) across the entire genome (Supplemental Fig. S6A), confirming the identified targets. Additional peaks (*trnE*, *trnT*, *rbcL*, *accD*) were reanalyzed by RIP followed by slot-blot hybridization, clearly proving that they were false positives because no enrichment in the pellet fraction was observed (Supplemental Fig. S6B). *RbcL/accD* loci were also found as false positive targets in other RIP-Seq analysis (Meurer et al., 2017). Furthermore, like the *rbcL/accD*, the *trnE*, and *trnT* peaks, "enrichment" appeared restricted to a narrow region (<100 nt) in contrast to the peaks of the true targets, which spanned several hundred nucleotides. RNA-gel–blot analyses of the wild-type, *pumpkin*, RNAi, and complemented lines using probes specific for *trnE*, *trnT*, *accD*, and *rbcL* (Supplemental Fig. S6C) did not reveal any differences in the mutant line except for *rbcL*, which is known to be generally reduced in pale mutants (Roy and Barkan, 1998; Bollenbach et al., 2009; Meurer et al., 2017). Taken together, these experiments point to the *trnG*-UCC, *trnV*-UAC, *petB*, *petD*, and *ndhA* intron loci as the main RNA targets of PUMPKIN.

PUMPKIN Impacts the RNA Metabolism of Its Targets

Next, we addressed the question whether the RNA metabolism of the verified target RNAs *trnG*-UCC,

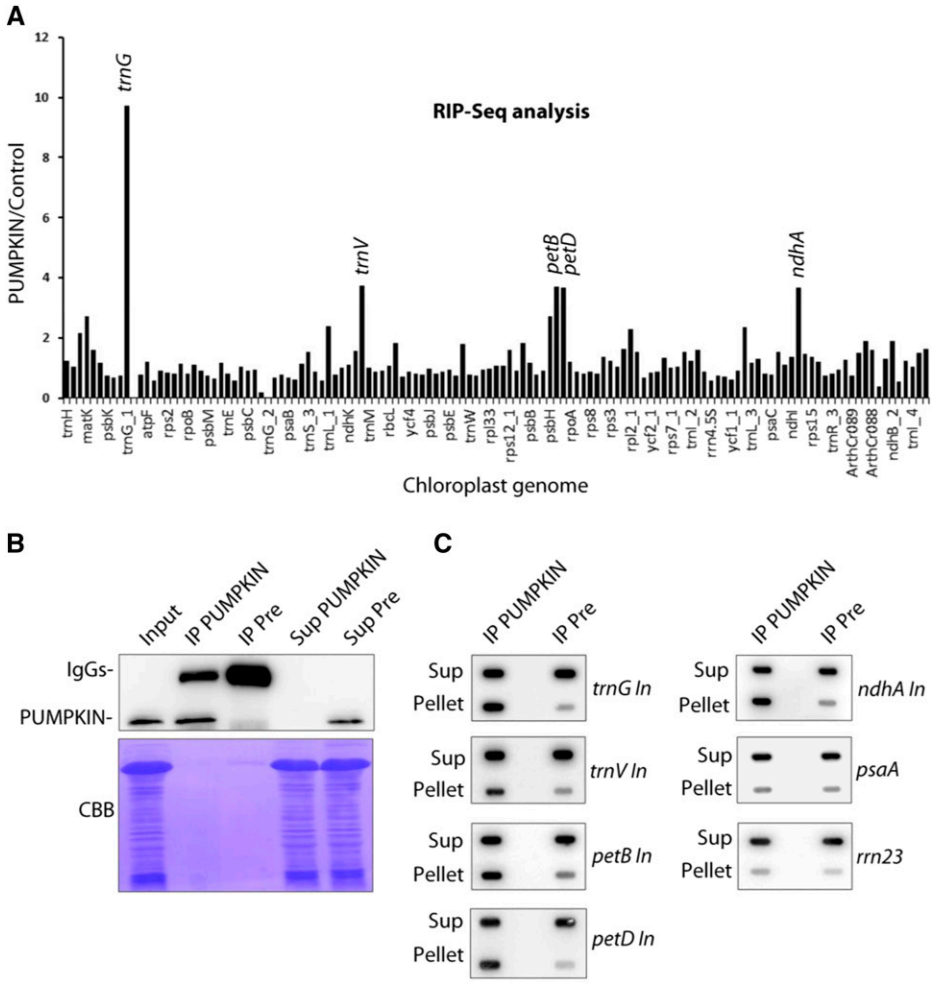

Figure 4. In Vivo Association of PUMPKIN with Chloroplast RNAs. A, RIP-Seq analysis of PUMPKIN. The RPKM (Reads Per Kilobase per Million) matrix normalizes reads for sequencing depth and gene length. The ratio of RPKM-normalized reads from PUMPKIN- and Control-IP (performed with the corresponding preimmune serum) for each chloroplast gene was calculated and plotted according to the chloroplast genome annotation (Supplementary Data Set S1). The most prominent peaks are indicated. In, intron. B, Immunodetection of immunoprecipitated PUMPKIN protein. PUMPKIN was completely depleted from the stroma extract (Input) when using PUMPKIN-specific antibodies. CBB, Coomassie Brilliant Blue; IgG, Immunglobulin G; IP, immunoprecipitation; Pre, preimmune serum; Sup, supernatant. C, Validation of RIP-Seq results via slot-blot hybridization. Immunoprecipitated (IP) RNAs were hybridized with the indicated probes. Probes *rrn23* and *psaA* served as negative controls. In, intron; IP, immunoprecipitation; Sup, supernatant.

trnV-UAC, *petB*, *petD*, and *ndhA* was affected in *pumpkin* mutants. RNA-gel blots were hybridized with probes specific for the exon and intron regions of the most prominent target, *trnG*-UCC. The analysis showed a considerable reduction of levels of both mature and precursor *trnG*-UCC RNA forms but did not reveal any obvious splicing defects, e.g. relative accumulation of unspliced molecules (Fig. 5A). This finding was also reflected by the poisoned primer extension (PPE) assay (Fig. 5B), arguing for the involvement of PUMPKIN in RNA processes other than splicing. Transcript levels of the second *trnG* gene (*trnG*-GCC) were unaffected, suggesting that tRNAs are not generally affected and that the effects on *trnG*-UCC are *pumpkin*-specific (Fig. 5A). Like *pumpkin*, levels of spliced and unspliced *trnG*-UCC RNAs were also greatly decreased in the maize *ppr5* mutant (Beick et al., 2008). Consequently, the corresponding PPR5 protein, which also associated with the intron region of *trnG*-UCC RNA in vivo, was reported to play a primary role in the stabilization rather than in the splicing of the *trnG*-UCC precursor molecules (Beick et al., 2008). Thus, given that *pumpkin* shows similar RNA defects, it is likely that PUMPKIN has a comparable function in the stabilization of *trnG*-UCC precursor RNAs. To check for stability defects of *trnG*-UCC, we conducted repeated RNA-gel–blot analyses such as the ones described for PPR5 (Beick et al., 2008) using probes specific for *trnG*-UCC exon 1 and 2 as well as the 5′ and 3′ ends of the intron (Fig. 5C). A truncated intron-containing transcript similar to that found in *ppr5* was also detected in *pumpkin* and the wild type when using 3′ intron and exon 2 probes. However, in contrast to the *ppr5* mutant, the ratio of the truncated transcript to the unspliced *trnG*-UCC RNA was not considerably higher in *pumpkin* than in the wild type. Interestingly, the *trnG* 5′ intron probe recognized a small fragment in *pumpkin* that was slightly larger than the mature *trnG*-UCC RNA. The fragment was completely absent in the wild type and in two other pale and pleiotropic mutants (*rhon1* and *pac*) that were used as controls because they exhibit decreased PEP-mediated transcripts similar to *pumpkin* (Meurer et al., 1998, 2017; Stoppel et al., 2012; Chi et al., 2014). Thus, the small fragment is likely to represent a degradation product that specifically accumulates in the absence of PUMPKIN and suggests a protective role of PUMPKIN with respect to the 5′ part of the *trnG*-UCC intron. This assumption is consistent

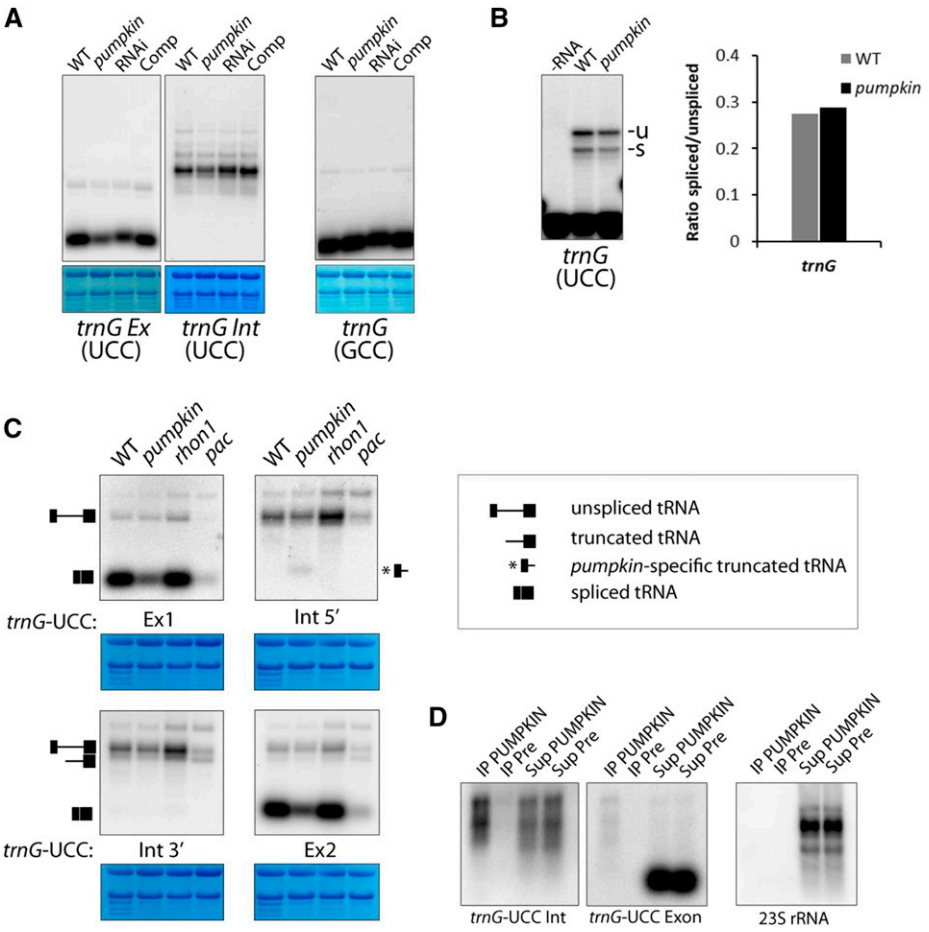

Figure 5. PUMPKIN Associates with the Intron of *trnG*-UCC and Influences Its Stability. A, RNA-gel–blot analysis of *trnG*-UCC and *trnG*-GCC in the wild-type, *pumpkin,* RNAi, and complemented (Comp) lines using 5 μg of leaf RNA. Ex, Exon; Int, Intron. Methylene blue staining is shown as loading controls (lower panels). B, Poisoned primer extension assay of *trnG*-UCC. Signals were quantified; the ratio of spliced (s) to unspliced (u) transcripts was calculated and is shown in the graph (right). C, RNA-gel–blot analysis of the wild type, *pumpkin, rhon1,* and *pac*. Probes specific for *trnG*-UCC exon 1 (Ex1), exon 2 (Ex2), and for 5′ (Int 5′) and 3′ (Int 3′) intron ends were used. The diagrams shown on the side of the blots indicate the resulting transcripts specified in the key (right panel): unspliced tRNA, truncated tRNA, *pumpkin*-specific truncated tRNA (*), and spliced tRNA. Methylene blue staining is shown as loading controls. D, RNA-gel–blot analysis of immunoprecipitated RNAs using intron- and exon-specific *trnG*-UCC probes. The 23S rRNA probe was used as a negative control. IP, immunoprecipitation; Pre, preimmune serum; Sup, supernatant. WT, wild type.

with the fact that in the RIP-Seq analysis, the highest enrichment was found in the 5′ end of *trnG*-UCC, while the coverage decreased toward the 3′ end (Supplemental Fig. S5). Probing the immunoprecipitated RNAs with intron- and exon-specific *trnG*-UCC probes confirmed that PUMPKIN associates exclusively with the intron of *trnG*-UCC (Fig. 5D).

RNA-gel–blot analyses were also performed with the remaining four targets: *trnV*-UAC, *petB*, *petD*, and *ndhA*. Like *trnG*-UCC, both spliced and unspliced *trnV*-UAC RNA forms were reduced, PPE analysis revealed no prominent splicing defects, and PUMPKIN was found to specifically associate with the intron region of *trnV*-UAC (Fig. 6, A–C). Hybridization with exon-specific probes showed reduced levels of exon-containing *ndhA*, *petB*, and *petD* transcripts in the *pumpkin* mutant and RNAi line as compared to the wild type (Fig. 6D). Interestingly, when using probes specific for the intron regions of *ndhA* and *petB*, intron-containing transcripts slightly larger than the spliced wild-type intron appeared in the *pumpkin* and RNAi line but were absent in the wild type and complemented line (Fig. 6D). To check whether these banding pattern alterations are PUMPKIN-specific, *rhon1* and *pac* mutants were examined in parallel (Fig. 6E). In the case of *petB*, the occurrence of the additional intron-containing bands is due to the specific lack of PUMPKIN. It is challenging to elaborate whether aberrant splicing or processing is responsible for the occurrence of this additional intron-containing RNA. The difficulty is that *petB* is part of the *psbB* operon, which is subjected to numerous intercistronic processing and splicing events resulting in about 20 different mono, di-, and oligocistronic transcripts (Stoppel et al., 2012). While the PPE results clearly showed that the ratio of spliced to unspliced *ndhA* and *petB* transcripts was reduced in *pumpkin* as compared to the wild type (Fig. 6F), no accumulation of the larger intron-containing *ndhA* or *petB* precursors was observed. This observation contrasts with the banding pattern of mutants previously described to be affected in the splicing of *petB* and *ndhA* mRNA (Watkins et al., 2011; Hammani and Barkan, 2014), suggesting that the presence of the additional bands is unlikely to be the result of aberrant splicing. As the additional intron-containing transcript is longer than the spliced intron, it is obvious that exon sequences are attached to the *petB* intron. Indeed, we found accumulation of a fragment starting at chloroplast genome position ~74,792 and ending at ~75,780 in the *pumpkin* mutant in our RNA-Seq analysis (Supplemental Fig. S7). This fragment perfectly matches the size of the additional intron-containing

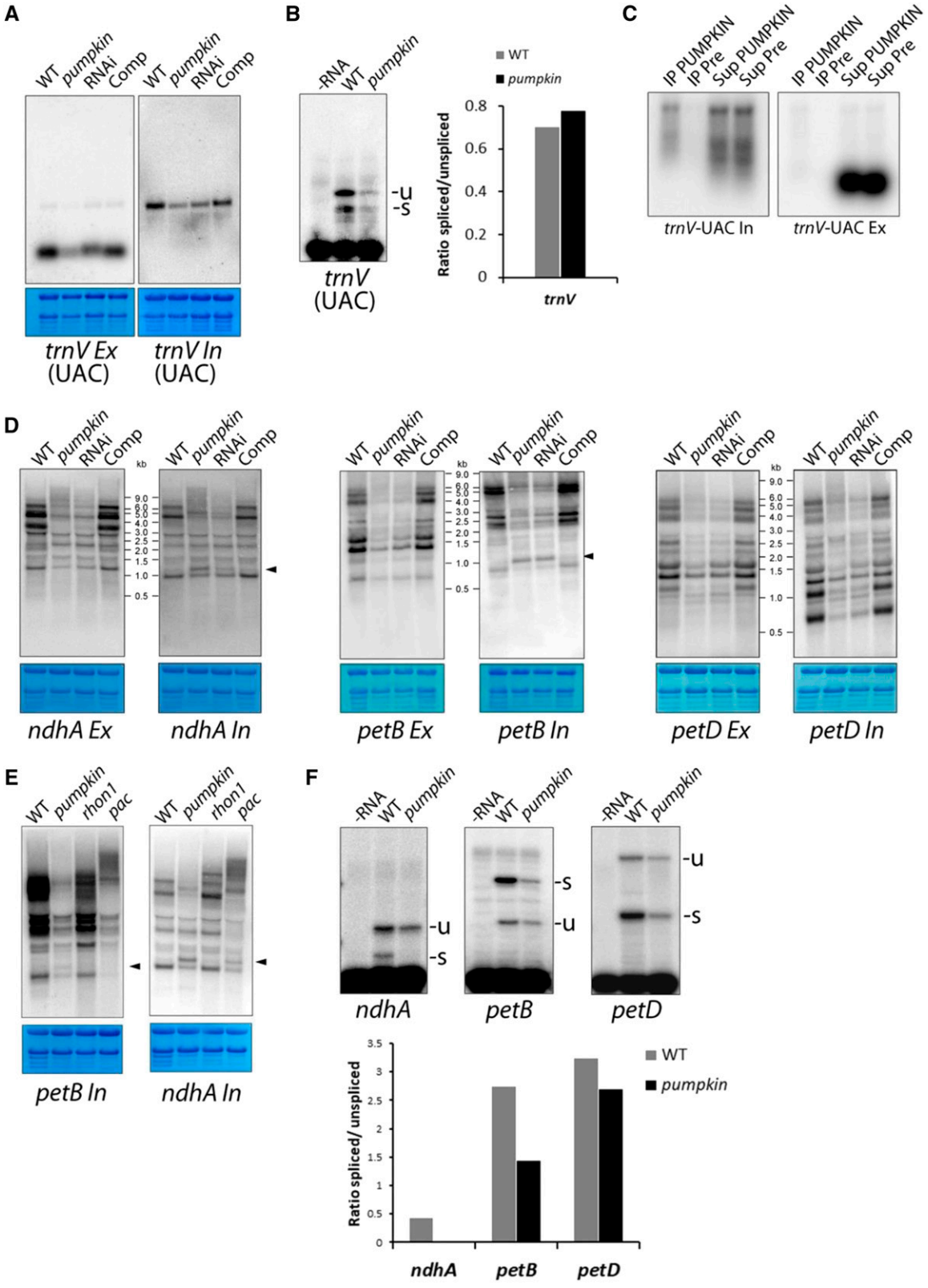

Figure 6. Analysis of the RNA Targets *trnV, ndhA, petB,* and *petD* in *pumpkin* Mutants. A, RNA-gel–blot analysis of *trnV*-UAC in the wild type, *pumpkin,* RNAi, and complemented (Comp) lines using 5 μg of leaf RNA. Ex, exon; In, intron. Methylene blue staining is shown as loading controls (lower panels). B, Poisoned primer extension assay of *trnV*-UAC using the wild-type RNA and *pumpkin* RNA (*pumpkin*). A reaction performed without RNA (–RNA) served as a negative control. Signals were

band in our RNA-gel–blot analysis (~1 kb) (Fig. 6D). The fragment contains the entire exon 1, the intron, and ~130 nt of exon 2. Thus, it seems that the absence of PUMPKIN causes aberrant processing within the exon 2 of *petB*. However, it remains elusive how exactly binding of PUMPKIN to the *petB* intron prevents the cleavage within the exon RNA. One could speculate that PUMPKIN has an impact on the RNA structure further downstream of its binding site in a way that an endonuclease sensitive site in exon 2 is masked.

Overall, these results suggest that PUMPKIN is needed for the stability of the *trnG*-UCC precursor and for preventing aberrant processing within the *petB* coding sequence. The former might also be true for *trnV*-UAC, as we found that spliced and unspliced forms of *trnV*-UAC RNA were reduced to a similar extent in *pumpkin*; and that PUMPKIN, as with *trnG*-UCC, associated exclusively with the *trnV*-UAC intron (Fig. 6C).

PUMPKIN Binds with High Affinity to the Central Region of the *petB* Intron in Vitro

Our RIP-Seq results showed that PUMPKIN is associated in vivo with the intron regions of *ndhA*, *petB*, *petD*, *trnG*-UCC, and *trnV*-UAC. To check whether PUMPKIN interacts with RNA directly, we performed an electrophoretic mobility shift assay (EMSA), using affinity- and gel filtration-purified recombinant (r)PUMPKIN protein, as well as *petB* RNA probes of similar size generated by in vitro transcription. Under these in vitro conditions, rPUMPKIN bound the probe from the central part of the *petB* intron with high affinity. But rPUMPKIN showed only a weak binding to the other two probes covering the *petB* 5′ UTR and 5′ end of the *petB* intron (Fig. 7, A and B). A protein concentration of 25 nM was sufficient to produce a distinct shift of the RNA probe, suggesting that PUMPKIN recognizes *petB* target sequences with high affinity (Fig. 7C). This result demonstrates that PUMPKIN is a bona fide RNA-binding protein that shows specificity for the central part of the *petB* intron.

PUMPKIN Forms Homomultimers and Has an Intrinsic UMP Kinase Activity in Vitro

Given that PUMPKIN and *E. coli* UMP kinase, PyrH, share 47% of amino acid identity (Supplemental Fig. S8A) and that structural predictions of both proteins revealed three-dimensional similarities (Fig. 8A), we next addressed the question of whether PUMPKIN has an UMP kinase activity.

Crystal structure analysis of the *E. coli* UMP kinase showed that the functional protein forms homohexamers (Briozzo et al., 2005). Like the *E. coli* form, PUMPKIN clearly interacted with itself when used as "prey" and "bait" in yeast two-hybrid experiments that were conducted under high stringency selection conditions (Fig. 8B). In support of the homomultimerization, when performing size-exclusion chromatography, MBP-PUMPKIN did not elute in fractions of the expected monomeric size of 72 kD (MPB, ~ 42 kD; PUMPKIN, ~ 30 kD), but much earlier with marker proteins of 669 kD in size (thyroglobulin). Cleaving off the MBP moiety reduced the size of rPUMPKIN expectedly (<440 kD, ferritin), while MBP migrated, as expected, with marker proteins of 43 kD (ovalbumin), indicating that PUMPKIN forms homomultimers (Supplemental Fig. S8, B and C).

The enzymatic function of UMP kinases is the conversion of UMP to UDP. Our aim was to check whether PUMPKIN has enzymatic activity like that of its eubacterial ancestor. We applied a coupled photometric and an HPLC-based UMP kinase assay analogous to that used to determine the UMP kinase activity of PyrH (Serina et al., 1995; Fig. 8, C–E). HPLC separation of all nucleotides in the assay was used to follow changes in their amounts during a time course. From this, consumption of UMP and ATP in a 1:1 stoichiometry to form UDP and ADP was observed, while the amount of the activator GTP was unchanged. The reaction was linear with time for at least 15 min (Fig. 8C). Further kinetic measurements were recorded by a coupled photometric assay, using the initial linear phase to determine enzyme activity. Both substrates UMP and ATP evoked a substrate-dependent increase in activity until saturation occurred (Fig. 8, D and E). Both substrates followed Michaelis-Menten kinetics and apparent affinities of 0.03 mM and 0.17 mM for UMP and ATP

Figure 6. (*Continued.*)
quantified; the ratio of spliced (s) to unspliced (u) transcripts was calculated and is shown in the graph (right). C, RNA-gel–blot analysis of immunoprecipitated RNAs using intron- and exon-specific *trnV*-UAC probes. For negative control, refer to 23S rRNA probe in Figure 5D. Ex, exon; In, intron; IP, immunoprecipitation; Pre, preimmune serum; Sup, supernatant. D, RNA-gel–blot analysis of *ndhA*, *petB*, and *petD* transcripts in the wild type, *pumpkin*, RNAi, and Comp lines using 5 μg of leaf RNA. For all tested transcripts, exon (Ex)- and intron (In)-specific probes were used as indicated. Arrowheads mark additional intron-containing transcripts detected with intron-specific *ndhA* and *petB* probes. Methylene blue staining of the membranes demonstrates equal loading (lower panels). E, RNA-gel–blot analysis of *petB* and *ndhA* in the wild type, *pumpkin*, *rhon1*, and *pac* mutants using 5 μg of leaf RNA. Methylene blue staining is shown as a loading control (lower panel). The arrowheads mark the aberrant transcripts; refer to (D). In, intron. F, Poisoned primer extension assay of PUMPKIN targets *ndhA*, *petB*, and *petD* using wild-type RNA and *pumpkin* RNA (*pumpkin*). A reaction performed without RNA (–RNA) served as negative control. Signals were quantified; the ratio of spliced (s) to unspliced (u) transcripts was calculated and is shown in the graph (lower panel). WT, wild type.

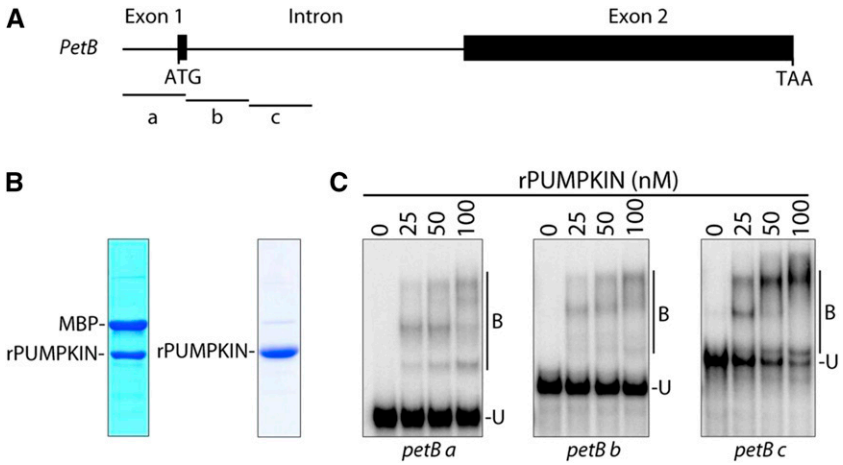

Figure 7. Recombinant PUMPKIN Binds to the *petB* Intron in Vitro. A, Schematic representation of the *petB* gene and probes (a, b, c) used in (B). ATG, start codon; TAA, stop codon. B, Recombinant MBP-PUMPKIN was purified and MBP was cleaved off (left panel). The right panel shows rPUMPKIN, obtained from an additional size exclusion chromatography purification step. Proteins were separated on an SDS gel and visualized by Coomassie Brilliant Blue staining. C, EMSA with purified rPUMPKIN and *petB* RNA probes, whose position is shown in (A). Samples were separated on a nondenaturing polyacrylamide gel. rPUMPKIN protein concentrations (nM) are indicated. B, bound; U, unbound.

were determined, respectively. Thus, PUMPKIN catalyzes the ATP-dependent conversion of UMP to UDP in vitro with properties like those of PyrH from *E. coli*, suggesting that PUMPKIN has retained its enzymatic activity over the course of evolution.

DISCUSSION

The results presented here demonstrate that PUMPKIN is a chloroplast-localized enzyme acting on RNA metabolism. We show that PUMPKIN associates in vivo with group II intron-containing tRNAs and mRNAs, binds to RNA in vitro, and influences the RNA metabolism of its targets *trnG*-UCC, *trnV*-UAC, *petB*, and *ndhA*. Its function has an impact on chloroplast translation, which in turn influences chloroplast transcription, biogenesis, photosynthesis, and plant growth.

Lack and Reduction of PUMPKIN Does Not Cause a Typical PSI Phenotype

Previously, it was reported that a seedling-lethal mutant of the UMP kinase locus (At3g18680) specifically failed to accumulate tricistronic *psaA/psaB/rps14* transcripts and consequently the PSI complex (Hein et al., 2009). For that reason, the locus had been named *dpt1* for defect in *psaA/psaB/rps14* transcript accumulation 1. We decided to re-examine the T-DNA insertion line (N829192) employed by Hein et al. (2009) in their study. Our first reason was the lack of complementation experiments and RNA-binding studies as well as enzymatic tests of DPT1. Second, with HCF145, a factor required for *psaA/psaB/rps14* RNA accumulation had already been unambiguously identified (Manavski et al., 2015). Indeed, genotyping experiments in two different laboratories (Jörg Meurer and Torsten Möhlmann) failed to confirm the described T-DNA insertion into the At3g18680 (N829192) locus. Our analyses of the confirmable T-DNA insertion line of the At3g18680/PUMPKIN locus (GABI_154D02) revealed that null alleles and knockdowns of this locus in Arabidopsis and tobacco were viable and exhibited pale green leaves, a smaller size, and retarded development as compared to the wild type (Fig. 1B, Supplemental Fig. S1). In fact, many of the previously described data could not be reproduced (Hein et al., 2009). To our surprise and in contrast to Hein et al. (2009), the mutant grew photoautotrophically and produced homozygous and viable seeds under greenhouse conditions. Spectroscopic data clearly indicated a pleiotropic effect with general deficiencies in the major thylakoid membrane complexes rather than a specific PSI defect in *pumpkin*. A general decrease in the levels of membrane and soluble plastid proteins rather than a specific PSI deficiency was also confirmed by extensive immunological and molecular analyses. Almost all plastid transcripts investigated were affected when PUMPKIN was absent or decreased—not only the *psaA/psaB/rps14* mRNA, as claimed previously (Hein et al., 2009). Recently, a relatively global effect on chloroplast development was also observed in rice mutants of the orthologous plastid UMP kinase (Zhu et al., 2016; Chen et al., 2018). The comparably milder impact of the rice mutations on plant growth could be explained by allele-specific effects of the two point mutations still allowing expression and presumably residual functions of the plastid UMP kinase. For these reasons, we propose renaming the At3g18680 locus PUMPKIN.

The Pleiotropic Effects in PUMPKIN Mutants

As demonstrated previously, *pumpkin* showed clear impairments of plastid translation and consequently reduced levels of photosynthetic and housekeeping proteins causing the pleiotropic phenotype. To find the basis of these translational defects, we identified the RNA targets of PUMPKIN that included the group II introns of *trnG*-UCC, *trnV*-UAC, *petB*, *petD*, and *ndhA*. Studies of knockouts of *trnG*-UCC and *trnG*-GCC in tobacco have shown that *trnG*-GCC is dispensable,

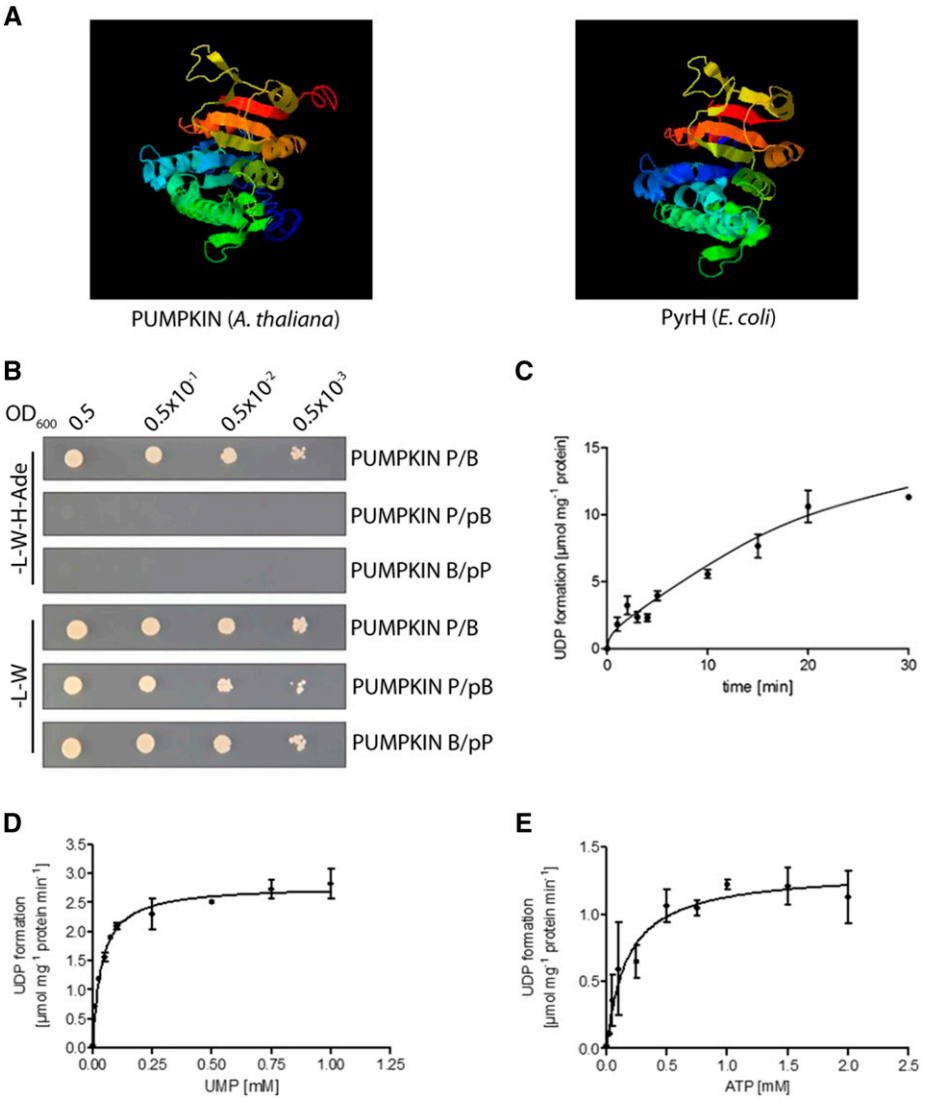

Figure 8. PUMPKIN Interacts with Itself and Exhibits Intrinsic UMP Kinase Activity in Vitro. A, Structure prediction of the UMP (Uridine monophosphate) kinases PUMPKIN (Arabidopsis) and PyrH (*E. coli*) by I-TASSER algorithm. B, Yeast-two-hybrid assay demonstrating that PUMPKIN interacts with itself. PUMPKIN was used as prey (P) and bait (B). The empty prey (pP) and bait (pB) vectors were used as negative controls. On high stringency plates (-L,-W,-H,-Ade) only true interactors are able to grow; while -L,-W plates only select for the presence of the plasmids. Culture densities were measured at 600 nm (OD_{600}) and diluted as indicated. C, Time-dependent UDP (Uridine diphosphate) formation catalyzed by rPUMPKIN was assayed by HPLC. ATP and UMP (Uridine monophosphate) were given at 1 mM concentration each. Data represent the mean of three independent experiments ± SE. D, UMP (Uridine monophosphate) concentration-dependence of UDP (Uridine diphosphate) formation catalyzed by rPUMPKIN followed Michaelis-Menten kinetics with an apparent affinity of 0.03 mM. ATP concentration was kept at 0.2 mM. Data represent the mean of three independent experiments ± SE. E, ATP concentration-dependence of UDP (Uridine diphosphate) formation catalyzed by rPUMPKIN followed Michaelis-Menten kinetics with an apparent affinity of 0.17 mM. UMP (Uridine monophosphate) concentration was kept at 0.3 mM. Data represent the mean of three independent experiments ± SE.

while *trnG*-UCC is essential for translation (Rogalski et al., 2008). Furthermore, Rogalski et al. (2008) demonstrated that *trnG*-GCC, which accumulated to the wild-type levels in *pumpkin* (Fig. 5A), was unable to compensate for a *trnG*-UCC shortage, whereas conversely *trnG*-UCC can do so by superwobbling. Another study, in maize, showed that a substantial decrease of *trnG*-UCC led to severe defects in plastid translation, causing a strong decrease in plastid-encoded proteins that in turn provokes a pronounced plant phenotype (Beick et al., 2008). Therefore, the reduced levels of the target tRNAs, *trnG*-UCC, and *trnV*-UAC, presumably resulting from decreased transcript stability, very likely represent the bottleneck for chloroplast translation in *pumpkin*. The effects in the *pumpkin* mutant are additionally enhanced by the reduced transcription of some plastid-encoded genes conditioned by insufficient translation of the PEP. At the same time, transcripts of genes preferentially transcribed by the NEP are upregulated, which is typical for PEP-deficient mutants (Pfalz et al., 2006; Cho et al., 2009; Zhou et al., 2009; Chateigner-Boutin et al., 2011; Steiner et al., 2011).

The *pumpkin* phenotype is more severe at early stages of development when high translation rates are required to allow fast establishment of the photosynthetic apparatus (Supplemental Fig. S9A). This explains an improvement of the phenotype at later stages during development implying that PUMPKIN is needed during early development. In accordance with this notion, we found that PUMPKIN levels are higher in young seedlings than in mature plants (Supplemental Fig. S9B).

Reduced levels and altered transcript metabolism of the other two targets, *petB* and *ndhA*, might further contribute to the mutant phenotype. The Arabidopsis *prfB3* mutants that failed to accumulate processed *petB* transcripts showed a specific deficiency in the cytochrome b_6f complex and were incapable of growing photoautotrophically (Stoppel et al., 2011). However,

pumpkin does not exhibit a typical cytochrome b_6f phenotype (Supplemental Fig. S2). It is not typical because processed *petB* transcripts, which seem to be essential, are not completely absent in *pumpkin* (Fig. 6D). Concerning *ndhA*, a deficiency of the NDH complex is unlikely to contribute considerably to the *pumpkin* phenotype because Arabidopsis and tobacco mutants defective in NDH complex accumulation exhibited only a weak phenotype, even under a variety of stress conditions (Barth and Krause, 2002; Muraoka et al., 2006). To what extent the reduction of *petD* might contribute to the phenotype remains elusive. In contrast to *petB* and *ndhA*, no aberrant transcripts were detected, and the ratio between spliced and unspliced transcript is not as drastically reduced when compared to *petB* and *ndhA* (Fig. 6, D and F).

The Function of PUMPKIN in RNA and Pyrimidine Metabolism

As demonstrated previously, PUMPKIN is an RNA-binding protein that has an impact on the metabolism of its RNA targets *trnG*-UCC, *trnV*-UAC, *petB,* and *ndhA* in vivo. While its involvement in the stability of *trnG*-UCC and presumably *trnV*-UAC precursors seems to be certain, the question remains whether PUMPKIN has an additional function in *petB* and *ndhA* intron splicing. Although a clear alteration in the metabolism of *petB* and *ndhA* introns is apparent (reflected by the PPE assay), the banding pattern does not resemble that of other mutants defective in plastid RNA splicing (Watkins et al., 2007; de Longevialle et al., 2008; Hammani and Barkan, 2014) because no accumulation of larger precursor molecules was observed—a common phenomenon of splicing mutants. Thus, we cannot rule out that the reduced splicing efficiencies might reflect secondary effects resulting from the altered translation and/or transcription of PEP-dependent plastid genes; PEP mutants or mutants with plastid translation deficiency were shown to exhibit pleiotropic splicing defects (Jenkins et al., 1997; Watkins et al., 2007; Chateigner-Boutin et al., 2011; Hammani and Barkan, 2014). Alternatively, PUMPKIN could additionally support the stability of *petB*- and *ndhA*-containing precursor transcripts, which would result in reduced levels of the precursor molecules when PUMPKIN is absent.

In addition to PUMPKIN's role in RNA metabolism, the conserved enzymatic characteristics imply that PUMPKIN might be also involved in plastid pyrimidine metabolism. A shortage of nucleotides could cause transcriptional changes that in turn could have effects on the RNA metabolism of some transcripts. Thus, we would like to point out that some of the RNA defects we observed might be the consequence of an altered nucleotide pool in chloroplasts. Interestingly, the PYRH homolog in *E. coli* and other eubacterial UMP kinases were found to be essential enzymes (Lee et al., 2007; Robertson et al., 2007; Meyer et al., 2008). Nevertheless, as *pumpkin* mutants are still viable, it is reasonable to assume that the enzymatic function of PUMPKIN might be at least partially compensated by other enzymatic activities. Other than PUMPKIN, Arabidopsis harbors three additional eukaryotic-type UMP/CMP kinases predicted to be localized in the cytosol; of which one (PYR6) was already functionally characterized and proven to convert UMP or CMP to UDP or CDP in vitro, respectively (Zhou et al., 1998). In addition, this UMP/CMP kinase was confirmed to be localized in the cytosol, which contains most of the uridine nucleotide pool (Dancer et al., 1990). So far only few transporters of nucleosides and other intermediates of the pathway have been identified (Möhlmann et al., 2010). Nonetheless, the absence of the sole UMP kinase in chloroplasts, PUMPKIN, does not lead to abolished transcription; thus we can conclude that uracil nucleotides, at least to a certain extent, can be transferred from the cytosol to the chloroplast.

In the plant and cyanobacterial PUMPKIN protein sequence we found conserved residues of the *E. coli* UMP kinase (PyrH) with dedicated functions, including Asp-146 as part of the catalytic center and binding sites Arg-62 and Asp-77 for the allosteric effectors GTP (activator) and UTP (inhibitor), respectively (Bucurenci et al., 1998; Supplemental Fig. S8A). The same applies to residues supporting UMP binding, including Gly-57, Gly-58, Arg-62, Gly-63, Asp-77, Thr-138, Asn-140, Asp-174, and Thr-145 (Briozzo et al., 2005). Interestingly, residues Asn-72 and Asp-93, which are important for dimer formation, are not conserved in PUMPKIN (Briozzo et al., 2005), indicating a diversified assembly process. According to the conserved residues and the overall similarity, PUMPKIN exhibited enzymatic properties with characteristics similar to eubacterial UMP kinases (Fig. 8, C–E) (Serina et al., 1995). When compared directly, *E. coli* PyrH exhibits apparent affinities of 0.12 mM (ATP) and 0.05 mM (UMP), which are similar to the 0.17 mM (ATP) and 0.03 mM (UMP) observed for PUMPKIN (Serina et al., 1995).

A moonlighting function has been reported for PyrH also, the PUMPKIN homolog in *E. coli* (Bucurenci et al., 1998; Kholti et al., 1998). Apart from PyrH's catalytic UMP kinase activity, it was found to be part of a large protein complex consisting of the homohexameric PepA (aminopeptidase A), IHE (integration host factor), PurR (purine repressor), and RutR (HTH-type transcriptional regulator). The complex associates with promoter elements of the *carAB* operon and controls transcriptional activity, presumably by modulating the DNA fold of adjacent regions (Kholti et al., 1998; Minh et al., 2009). Thus, in *E. coli,* PyrH incorporates two functions: pyrimidine metabolism and control of gene expression. As such, PUMPKIN seems to have retained the enzymatic function of its bacterial ancestor in pyrimidine nucleotide metabolism and the moonlighting control of gene expression. Although it did switch regulation from transcriptional to post-transcriptional levels, which generally predominate in chloroplasts.

So far, the predicted structure and domain searches have not provided hints of any known RNA-binding motif that could explain the RNA-binding activity of PUMPKIN demonstrated in this study. Yet the nucleotide-binding properties, together with the fact that PUMPKIN forms large homomultimers, might account for the intrinsic RNA-binding capacity. The hexameric ring-shaped protein Hfq from *E. coli*, for example, is an RNA-binding protein that controls bacterial gene expression on different levels. It has a so-called proximal RNA-binding site, where it binds six nucleotides of the RNA in a one-nucleotide-per-monomer manner (Schumacher et al., 2002; Sauer et al., 2012). A similar RNA-binding mechanism might also exist for PUMPKIN, whereby the nucleotide-binding pockets of PUMPKIN are occupied by nucleotides of the RNA molecule, one by each monomer. This scenario implies that free ribonucleotides compete with RNA binding, so that one could speculate that PUMPKIN-dependent RNA metabolism is likely to be under free ribonucleotide control and/or vice versa. A similar mechanism has already been observed for other mammalian RNA-binding enzymes, such as iron regulatory protein (IRP) 1, thymidylate synthase, dihydrofolate reductase, and glyceraldehyde-3-phosphate, where RNA-binding and enzymatic activity are mutually exclusive (Hentze, 1994; Castello et al., 2015).

In other organisms, such as human and yeast, several dual-functioning enzymes that participate in metabolic pathways and are capable of binding RNA were identified from RNA interactome studies, thereby supporting the so-called RNA-enzyme-metabolite (REM) hypothesis. This concept proposes a connection between intermediary metabolism and posttranscriptional regulation of gene expression, mediated by RNA-binding enzymes (Hentze and Preiss, 2010; Castello et al., 2015; Hentze et al., 2018). Similar interactome studies were also performed in Arabidopsis, leading to the identification of unique RNA-binding proteins (RBPs) that (1) are lacking classical RNA-binding motifs (e.g. the RRM, KH, or DEAD box helicase domains) and (2) were assigned to exhibit enzymatic activity or participate in metabolic pathways. The identification opened a regulatory network between gene expression and metabolic processes (Köster et al., 2017).

In conclusion, PUMPKIN adds another layer of complexity to the already complex chloroplast RNA metabolism as it seems to link plastid gene expression with processes of primary nucleotide metabolism.

MATERIALS AND METHODS

Plant Material and Cultivation

The *pumpkin* Arabidopsis (*Arabidopsis thaliana*) T-DNA line (accession D02154) was obtained from the GABI-Kat collection (Kleinboelting et al., 2012). All Arabidopsis plants were grown on soil under the following controlled conditions: 12-h-light (20°C)/12-h-dark (18°C) at a PFD of 100 μM photons m^{-2} sec^{-1}. For RNA-gel blots, chlorophyll fluorescence measurements, immunodetection, and BN PAGE plants of a comparable developmental stage were used (wild type [Col-0] and complemented line: 3 weeks; *pumpkin* and RNAi lines: around 6 weeks). Chloroplast-soluble extracts for size exclusion chromatography and IPs were prepared from 2-week-old Col-0 plants. For comparison of *pumpkin* with *rhon1* and *pac* mutants, plants were grown on 1x MS medium with 1.5% w/v Suc under following conditions: 16-h-light /8-h-dark (20°C) at a PFD of 30 μM photons m^{-2} sec^{-1}.

Nicotiona tabacum plants were grown on soil under standard greenhouse conditions (16-h-light/8-h-dark cycle; 22°C; 200 μM photons m^{-2} sec^{-1}).

Generation of *pumpkin* RNAi and Complemented Lines

For generation of *pumpkin* RNAi lines, a 240 bp-fragment of the coding sequence of *PUMPKIN* was amplified with primers Fw_CACC_Pumpkin_RNAi/Rev_Pumpkin_RNAi. For the mutant complementation, the complete CDS of *PUMPKIN* (*AT3G68180*) was RT-PCR-amplified using primers Fw_Pumpkin_CACC/Rev_Pumpkin_Stop to obtain a construct without additional tags. The respective PCR products were cloned into pENTR/SD/D-TOPO (Invitrogen, Carlsbad, CA) and transferred into GATEWAY pB7GWIWG2 (I) for RNAi line generation or pB7YWG2 for complementation via the LR reaction (LR clonase II, Thermo Fisher Scientific, Waltham, MA), according to the manufacturer's protocols. *Agrobacterium* strain GV3101 was transfected with the respective construct and used for floral dip transformation of either the Col-0 plants (for RNAi line generation) or heterozygous D02154 plants (for complementation). Seeds obtained from T_0-plants were sown on soil and seedlings were treated with BASTA (Bayer AG, Leverkusen, Germany) to select for resistant plants.

For generation of RNAi lines in tobacco, a 269 bp-fragment of the coding sequence of *PUMPKIN* from *N. tabacum* (GeneRef.: XM_016635857.1) was amplified via RT-PCR with primers Fw_CACC_NT_RNAi_Pumpkin/Rev_NT_RNAi_Pumpkin and cloned into pB7GWIWG2 (I) as described previously. *Agrobacterium*-mediated (strain GV3101) transformation of surface-sterilized leaf disks was carried out as described (Horsch et al., 1985). Selection of antibiotic-resistant cells and shoot regeneration was as follows: After 2 d on 1/2 MS plates with 1 mg/L 6-benzylaminopurine (BAP) in the dark at 25°C, leaf discs were washed in sterile water and transferred to 1/2 MS plates containing 2 mg/L BAP, 500 mg/L carbenicillin and 400 mg/L cefotaxime and cultivated under long-day conditions (16-h-light/8-h-dark) at 25°C at 100 μM photons m^{-2} sec^{-1}. For selection of transformants, callus/shoots were regularly transferred on identical MS plates with increasing concentrations of phosphinothricin (1 to 5 mg/L). After induction of root growth by substituting BAP with 0.1 mg/L indole-3-butyric acid, plants were transferred to soil for seed production.

RNA-Gel–Blot Hybridization, RT-PCR and PPE

RNA-gel–blot analyses, total RNA isolation, electrophoresis, blotting, and hybridization with radioactive labeled probes were carried out as described by Manavski et al. (2015). Oligonucleotides and primers for PCR-probes are listed in Supplemental Table S1. For RT-PCR, DNA contaminations were removed by the DNA-free DNA Removal Kit (Thermo Fisher Scientific, Waltham, MA). Reverse transcription was performed with BIORAD's (Hercules, CA) iScript cDNA Synthesis Kit. RT-PCR (25 cycles; Fig. 1D) was performed with *PUMPKIN*-specific primers (Fw_Pumpkin_CACC/ Rev_Pumpkin_stop), while primers for *UBIQUITIN (UBI)* (Fw_Ubi/ Rev_Ubi) served as control. PPE was conducted as described previously by Barkan (2011b). Signals were quantified by ImageJ.

RIP-Seq, RNA-Seq, and Slot-Blot Analysis

RIP-Seq analysis was essentially performed as described by Meurer et al. (2017). Briefly, stroma extracts were isolated from the 2-week-old wild-type plants and incubated with either PUMPKIN-specific antibodies or with the preimmune serum. IgGs were captured with SiMAG-Protein G beads (Chemicell, Berlin, Germany). Recovered RNA was used for generation of libraries with the ScriptSeq v2 RNA-seq Library Preparation Kit (Epicentre, Madison, WI), according to manufacturer's instructions. Deep sequencing (2x 150 bp, v2 chemistry) was performed on a MiSeq sequencer (Illumina, San Diego, CA), yielding 6.6 Mio and 10.3 Mio paired reads for PUMPKIN antibodies and the preimmune serum control, respectively. To obtain the depths of coverage (reads/nucleotides), primary reads were aligned to the Arabidopsis chloroplast

genome (accession no. NC_000932.1) using CLC Genomics Workbench 6.5.1 (Qiagen, Valencia, CA) with the following parameters: *mismatch cost = 2; insertion cost = 3; deletion cost = 3; length fraction = 0.5; similarity fraction = 0.8; global alignment = no; auto-detect paired distances = yes*. Aligned reads were extracted as BAM files and sorted in Galaxy using the SortSam tool (version 2.7.1.1). Sorted BAM files were converted into RPKM-normalized bigwig files using the bamCoverage tool (version 3.0.2) and displayed in IGB. The ratio of PUMPKIN/control of the two replicates was displayed across the entire chloroplast genome.

For RPKM calculation of the RIP-Seq data, primary reads were aligned to NC_000932.1 using the RNA-Seq Analysis tool with the following settings: *Reference type = Genome annotated with genes only; Reference sequence = Genome (Genome); Gene track = Genome (Gene); Mapping type = Map to gene regions only; Mismatch cost = 2; Insertion cost = 3; Deletion cost = 3; Length fraction = 0.8; Similarity fraction = 0.8; Global alignment = no; Auto-detect paired distances = yes; Strand specific = both; Maximum hits for a read = 3; Count paired reads as two = no; Expression value = total counts; Calculate RPKM for genes without transcripts = yes; Use EM estimation = no; Minimum read count fusion gene table = 5*. Numeric data are provided in Supplemental Data Set S1.

For RNA-Seq analysis, isolated RNA of the wild-type and *pumpkin* mutants was treated with RiboMinusTM Plant Kit (Thermo Fisher Scientific, Waltham, MA) following the manufacturer's protocol and used for library preparation followed by deep sequencing as described previously. Primary reads (the wild-type samples 11.2 Mio reads each, *pumpkin* samples 13 and 13.7 Mio reads) were aligned to the complete Arabidopsis genome using the CLC RNA-Seq analysis tool with the following parameters: *Reference type = Genome annotated with genes and transcripts; Reference sequence = Arabidopsis thaliana sequence; Gene track = Arabidopsis thaliana_Gene; mRNA track = Arabidopsis thaliana_mRNA; Mapping type = Map to gene regions only; Mismatch cost = 2; Insertion cost = 3; Deletion cost = 3; Length fraction = 0.8; Similarity fraction = 0.8; Global alignment = no; Auto-detect paired distances = yes; Strand specific = both; Maximum hits for a read = 3; Count paired reads as two = no; Expression value = Total counts; Calculate RPKM for genes without transcripts = yes; Use EM estimation (recommended) = no*. Genome references were imported using the Download Reference Genome Data tool; references were created from ftp.ensemblgenomes.org/pub/current/plants/fasta/Arabidopsis_thaliana/dna/Arabidopsis_thaliana.TAIR10.dna.toplevel.fa.gz. Mapping results were sorted for chloroplast genes; the mean value of the two independent replicates and the fold changes were calculated. Data was provided in Supplemental Data Set S2. In order to characterize the 5′ and 3′ ends of *petB* transcripts, fastq files were subjected to illumination adapter removal and sequence trimming using Trimmomatic (v 0.36) with parameters: *CROP:100 TruSeq3-PE-2.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:20*. Remaining pair-end reads were mapped to the whole Arabidopsis genome (TAIR10) using STAR (v 2.5.2a). The only mapped pair-end reads retained were those with a mapping quality bigger than 100 and mapping length bigger than 50 bp. To identify the 5′-ends of the intron-containing *petB* transcript: The mapping position of the 5′-end of the first pair-end reads *which* second read-pair map inside of the *petB* intron was retained. To identify the 3′-ends of the intron-containing *petB* transcript: The mapping position of the 3′-end of the second pair-end reads *which* first read-pair map inside of the *petB* intron was retained.

Slot-blot analysis was conducted as described by Meurer et al. (2017).

Immunoblot Analysis, BN PAGE, and Antibody Production

Separation of soluble and thylakoid protein complexes and total leaf proteins were performed as described by Torabi et al. (2014). Thylakoid membranes for BN PAGE were isolated, solubilized, and run in the first dimension as described by Schwenkert et al. (2006). Second dimension SDS gels were either stained with Roti-Blue quick (Carl Roth, Karlsruhe, Germany) or blotted and used for immunodetection.

PUMPKIN antibodies were raised against the synthetic peptide VDGVFDDDPKRNPN and used in 1:1,000 dilutions. Peptide synthesis, immunization, and IgG purification were done by BioGenes GmbH (Berlin, Germany). For the RIP-Seq and slot-blot experiments, the respective preimmune serum was used as negative control. The NdhH antibody was kindly provided by Tsuyoshi Endo (Kyoto University). The RPL11 antibody was generated recently (Meurer et al., 2017). All other antisera (PsaA, PsaF, PsbA, PsbB, PsbC, PetB, AtpC, LHCb2, LHCa1, RPS5) were obtained from Agrisera (Vännäs, Sweden). Signals were quantified by ImageJ.

Preparation and Size Exclusion Chromatography of Stroma Extracts

Size exclusion chromatography analysis of chloroplast stroma extracts obtained from the 2-week-old wild-type plants was performed as described by Meurer et al. (2017). Fractions were separated on 15% SDS-PAGE.

Size exclusion chromatography experiments with recombinant MBP-PUMPKIN and AcTEV-digested MBP-PUMPKIN were conducted with a Superdex 200 10/300 GL column (GE Healthcare, Little Chalfont, UK). Reaction conditions and determination of the complex sizes were specified previously (Meurer et al., 2017).

EMSA

For in vitro RNA-protein–binding studies, EMSA experiments were carried out. Trace amounts of radioactively labeled RNA probes generated by in vitro transcription were incubated with increasing concentrations of recombinant PUMPKIN protein and separated on native PAA gels. Reaction conditions and experimental setup were described elsewhere (Meurer et al., 2017). Supplemental Table S1 offers sequence information of the probes.

Yeast-Two Hybrid Screening

For analysis of protein-protein interactions, we used the Matchmaker Two-Hybrid System Kit (Clontech, Mountain View, CA). The coding region of *PUMPKIN* lacking the transit peptide (aa 1-53) was PCR amplified (Fw_Pumpkin_*Nde*I/ Rev_Pumpkin_stop_*Bam*HI) and cloned into the *Nde*I/*Bam*HI restriction sites of the pGADT7 and pGBKT7 vectors to obtain bait and prey constructs. The bait and prey vectors were cotransformed into yeast strain AH109 and selected on SD plates without Leu and Trp (-L-W). Protein-protein interactions were analyzed by growing the double transformants on highly stringent SD plates without Leu, Trp, His, and adenine (-L-W-H-Ade) as suggested by the Matchmaker Two-Hybrid System Kit manufacturer. As a negative control, yeast cells containing the empty vector were transformed with the respective bait/prey *PUMPKIN* construct.

In Vivo Protein Labeling

In vivo labeling was performed as described by Meurer et al. (2017), with minor changes. Leaf discs from plants of the same developmental stage (wild type/complemented: 2 weeks; *pumpkin*/RNAi: around 4 weeks) were used. Samples were incubated at RT in ambient light for 20 min. Proteins were isolated and separated into soluble and thylakoid fractions as described for immunoblot analysis. After determination of counts per minute (cpm), equal amounts of sample corresponding to 1,000,000 cpm for soluble proteins and 100,000 cpm for thylakoid proteins were separated on a 10% denaturing polyacrylamide gel. Proteins were visualized by coomassie staining. Gels were dried and signals were detected using a Typhoon TRIO.

Localization Studies

Transient expression of PUMPKIN fused to YFP was performed as described previously (Witte et al., 2004). *A. tumefaciens* strain GV310 transformed with PUMPKIN C-terminal YFP constructs [pHygII_UT_mVenusC (Kunz et al., 2014)] was infiltrated through the lower epidermis of 6-week-old *N. benthamiana* leaves. After three days, protoplasts were isolated by incubation in 0.45 M sorbitol, 10 mM $CaCl_2$, 1% w/v Cellulase R-10, 0.25% Macerozyme R-10, and 20 mM MES-KOH pH 5.7 for 3 h in the dark and filtered through a stainless steel gauze (125 μm). The protoplasts were pelleted (70 g, 1 min), resuspended in W5 (145 mM NaCl, 125 mM $CaCl_2$, 5 M KCl, 5 mM Glc, pH 5.5) and analyzed for the presence of fluorescence signals with a Leica TCS SP5II microscope (514-nm excitation and 525- to 582-nm detection of emission through a HCX PL APO 63x/1.2 W water immersion objective). Chlorophyll autofluorescence was detected with 514-nm excitation and a 651- to 704-nm emission wavelength.

For colocalization analysis, the full-length CDS of RAP (*AT2G31890*) was amplified with primers Fw_RAP_CACC/ Rev_RAP, cloned into pENTR/SD/D-TOPO (Invitrogen, Carlsbad, CA) and transferred into pH7CWG2. Transient coexpression of PUMPKIN-YFP and RAP-CFP was performed as just described.

Generation and Purification of Recombinant Protein

The coding sequence of PUMPKIN lacking the transit peptide (aa 1-53) was PCR-amplified (Fw_Pumpkin_*Bam*HI/ Rev_Pumpkin_stop_*Sal*I) and cloned into pMAL-TEV (kindly provided by Alice Barkan, Institute of Molecular Biology, University of OR, Eugene, OR 97403) via the *Bam*HI/*Sal*I restriction sites. Overexpression and affinity purification were performed as described by Chi et al. (2014). To remove the MBP-tag, AcTEV protease (Thermo Fisher Scientific, Waltham, MA) was added. Proteins were separated by size exclusion chromatography on a Superose 6 10/300 GL column (GE Healthcare, Little Chalfont, UK), with buffer containing 0.1 M Tris HCl pH 8.0; 0.15 M NaCl; and 1 mM EDTA. Fractions containing rPUMPKIN were concentrated using Amicon Ultra-15 centrifugal filter units (Ultracel 3K; Merck Millipore, Darmstadt, Germany).

UMP Kinase Assay

Initially, rPUMPKIN was tested for UMP kinase activity in an HPLC-based assay as follows: 2 to 5 μg rPUMPKIN was incubated in Tris/HCl (50 mM, pH 7.4); KCl (50 mM), $MgCl_2$ (2 mM); UMP (1 mM); ATP (1 mM), and GTP (0.5 mM, required for full activity) at 30°C for the given time. Reaction products were separated as specified previously (Daumann et al., 2015). UV-Vis detection was at 265 nm, and data analysis was performed with Chromeleon software (Dionex, Idstein, Germany). For further analysis, a coupled photometric assay as described by Blondin et al. (1994) was exploited. Samples were measured in 96 well plates in an Infinite M200 reader (Tecan, Crailsheim, Germany), determining linearity with time and protein amount (0.2 to 3 μg per assay). Substrate dependence for ATP and UMP were determined holding the nontested substrate constant (0.2 mM for ATP and 0.3 mM for UTP). Substrate affinities were calculated by the Best-Fit method for Michaelis-Menten kinetics as part of the GraphPad Prism software (GraphPad software, Inc.).

Accession Numbers

Sequence data from this article can be found in the EMBL/GenBank libraries under the following accession numbers: PUMPKIN in Arabidopsis (At3g18680), PUMPKIN in *N. tabacum* (XM_016635857.1), Arabidopsis chloroplast genome (NC_000932.1), PyrH in *E. coli* (NC_000913.3), SYN_UMPK in *Synechocystis sp.* PCC6803 (NC_000911.1), and RAP in Arabidopsis (AT2G31890). RNA-Seq and RIP-Seq data are available at NCBI GEO accession GSE121778.

SUPPLEMENTAL DATA

The following supplementary materials are available:

Supplemental Figure S1. RNAi Lines in Tobacco.

Supplemental Figure S2. Spectroscopic Analysis of *pumpkin* Mutants.

Supplemental Figure S3. Formation of Photosynthetic Complexes Demonstrated by Blue Native PAGE in the First and SDS PAGE in the Second Dimension.

Supplemental Figure S4. Up- and Down-Regulation of Plastid Genes Preferentially Transcribed by the NEP and the PEP, Respectively.

Supplemental Figure S5. Enrichment of PUMPKIN RNA Targets Identified by RIP-Seq.

Supplemental Figure S6. Identification and Validation of PUMPKIN RNA-Targets.

Supplemental Figure S7. In-Depth Analysis of the Aberrant Intron-Containing Transcript of *petB* in the *pumpkin* Mutant Using RNA-Seq.

Supplemental Figure S8. rPUMPKIN Forms Homomultimers in Vitro.

Supplemental Figure S9. Development-Dependent Expression of PUMPKIN.

Supplemental Table S1. Sequence Information of Oligonucleotides

Supplemental Data Set S1. Mapped Reads to the Arabidopsis Chloroplast Genome Resulting from RIP-Seq.

Supplemental Data Set S2. RPKM Values of Chloroplast Transcripts Obtained by RNA-Seq.

ACKNOWLEDGMENTS

We wish to thank Christian Schmitz-Linneweber and Julia Legen for performing preliminary RIP-Chip experiments. We thank Tsuyoshi Endo for providing NdhH antisera and Alice Barkan for the pMAL-TEV vector.

Received October 8, 2018; accepted October 29, 2018; published November 8, 2018.

LITERATURE CITED

Barkan A (2011a) Expression of plastid genes: Organelle-specific elaborations on a prokaryotic scaffold. Plant Physiol **155:** 1520–1532

Barkan A (2011b) Studying the structure and processing of chloroplast transcripts. Methods Mol Biol **774:** 183–197

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

Barth C, Krause GH (2002) Study of tobacco transformants to assess the role of chloroplastic NAD(P)H dehydrogenase in photoprotection of photosystems I and II. Planta **216:** 273–279

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:** 5337–5347

Belcher S, Williams-Carrier R, Stiffler N, Barkan A (2015) Large-scale genetic analysis of chloroplast biogenesis in maize. Biochim Biophys Acta **1847:** 1004–1016

Blondin C, Serina L, Wiesmüller L, Gilles AM, Bârzu O (1994) Improved spectrophotometric assay of nucleoside monophosphate kinase activity using the pyruvate kinase/lactate dehydrogenase coupling system. Anal Biochem **220:** 219–221

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:** 114–131

Bohne AV, Nickelsen J (2017) Metabolic control of chloroplast gene expression: An emerging theme. Mol Plant **10:** 1–3

Bohne AV, Schwarz C, Schottkowski M, Lidschreiber M, Piotrowski M, Zerges W, Nickelsen J (2013) Reciprocal regulation of protein synthesis and carbon metabolism for thylakoid membrane biogenesis. PLoS Biol **11:** e1001482

Bollenbach TJ, Sharwood RE, Gutierrez R, Lerbs-Mache S, Stern DB (2009) The RNA-binding proteins CSP41a and CSP41b may regulate transcription and translation of chloroplast-encoded RNAs in Arabidopsis. Plant Mol Biol **69:** 541–552

Briozzo P, Evrin C, Meyer P, Assairi L, Joly N, Barzu O, Gilles AM (2005) Structure of *Escherichia coli* UMP kinase differs from that of other nucleoside monophosphate kinases and sheds new light on enzyme regulation. J Biol Chem **280:** 25533–25540

Bucurenci N, Serina L, Zaharia C, Landais S, Danchin A, Bârzu O (1998) Mutational analysis of UMP kinase from *Escherichia coli*. J Bacteriol **180:** 473–477

Castello A, Hentze MW, Preiss T (2015) Metabolic enzymes enjoying new partnerships as RNA-binding proteins. Trends Endocrinol Metab **26:** 746–757

Chateigner-Boutin AL, des Francs-Small CC, Delannoy E, Kahlau S, Tanz SK, de Longevialle AF, Fujii S, Small I (2011) OTP70 is a pentatricopeptide repeat protein of the E subgroup involved in splicing of the plastid transcript *rpoC1*. Plant J **65:** 532–542

Chen F, Dong G, Ma X, Wang F, Zhang Y, Xiong E, Wu J, Wang H, Qian Q, Wu L, et al (2018) UMP kinase activity is involved in proper chloroplast development in rice. Photosynth Res **137:** 53–67

Chi W, He B, Manavski N, Mao J, Ji D, Lu C, Rochaix JD, Meurer J, Zhang L (2014) RHON1 mediates a Rho-like activity for transcription termination in plastids of *Arabidopsis thaliana*. Plant Cell **26:** 4918–4932

Cho WK, Geimer S, Meurer J (2009) Cluster analysis and comparison of various chloroplast transcriptomes and genes in *Arabidopsis thaliana*. DNA Res **16:** 31–44

Cohen I, Sapir Y, Shapira M (2006) A conserved mechanism controls translation of Rubisco large subunit in different photosynthetic organisms. Plant Physiol **141:** 1089–1097

Dancer J, Neuhaus HE, Stitt M (1990) Subcellular compartmentation of uridine nucleotides and nucleoside-5′ -diphosphate kinase in leaves. Plant Physiol **92:** 637–641 16667327

Daumann M, Fischer M, Niopek-Witz S, Girke C, Möhlmann T (2015) Apoplastic nucleoside accumulation in Arabidopsis leads to reduced photosynthetic performance and increased susceptibility against *Botrytis cinerea.* Front Plant Sci **6:** 1158

de Longevialle AF, Hendrickson L, Taylor NL, Delannoy E, Lurin C, Badger M, Millar AH, Small I (2008) The pentatricopeptide repeat gene *OTP51* with two LAGLIDADG motifs is required for the cis-splicing of plastid *ycf3* intron 2 in *Arabidopsis thaliana.* Plant J **56:** 157–168

de Longevialle AF, Small ID, Lurin C (2010) Nuclearly encoded splicing factors implicated in RNA splicing in higher plant organelles. Mol Plant **3:** 691–705

Douchi D, Qu Y, Longoni P, Legendre-Lefebvre L, Johnson X, Schmitz-Linneweber C, Goldschmidt-Clermont M (2016) A nucleus-encoded chloroplast phosphoprotein governs expression of the Photosystem I Subunit PsaC in *Chlamydomonas reinhardtii.* Plant Cell **28:** 1182–1199

Germain A, Hotto AM, Barkan A, Stern DB (2013) RNA processing and decay in plastids. Wiley Interdiscip Rev RNA **4:** 295–316

Hammani K, Barkan A (2014) An mTERF domain protein functions in group II intron splicing in maize chloroplasts. Nucleic Acids Res **42:** 5033–5042

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:** 4278–4288

Hein P, Stöckel J, Bennewitz S, Oelmüller R (2009) A protein related to prokaryotic UMP kinases is involved in *psaA/B* transcript accumulation in *Arabidopsis.* Plant Mol Biol **69:** 517–528

Hentze MW (1994) Enzymes as RNA-binding proteins: A role for (di)nucleotide-binding domains? Trends Biochem Sci **19:** 101–103

Hentze MW, Preiss T (2010) The REM phase of gene regulation. Trends Biochem Sci **35:** 423–426

Hentze MW, Castello A, Schwarzl T, Preiss T (2018) A brave new world of RNA-binding proteins. Nat Rev Mol Cell Biol **19:** 327–341

Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT (1985) A simple and general method for transferring genes into plants. Science **227:** 1229–1231

Jacobs J, Kück U (2011) Function of chloroplast RNA-binding proteins. Cell Mol Life Sci **68:** 735–748

Jenkins BD, Kulhanek DJ, Barkan A (1997) Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. Plant Cell **9:** 283–296

Kafer C, Zhou L, Santoso D, Guirgis A, Weers B, Park S, Thornburg R (2004) Regulation of pyrimidine metabolism in plants. Front Biosci **9:** 1611–1625

Kholti A, Charlier D, Gigot D, Huysveld N, Roovers M, Glansdorff N (1998) *pyrH*-encoded UMP-kinase directly participates in pyrimidine-specific modulation of promoter activity in *Escherichia coli.* J Mol Biol **280:** 571–582

Khrouchtchova A, Monde RA, Barkan A (2012) A short PPR protein required for the splicing of specific group II introns in angiosperm chloroplasts. RNA **18:** 1197–1209

Kleinboelting N, Huep G, Kloetgen A, Viehoever P, Weisshaar B (2012) GABI-Kat SimpleSearch: New features of the *Arabidopsis thaliana* T-DNA mutant database. Nucleic Acids Res **40:** D1211–D1215

Kleinknecht L, Wang F, Stübe R, Philippar K, Nickelsen J, Bohne AV (2014) RAP, the sole octotricopeptide repeat protein in *Arabidopsis,* is required for chloroplast *16S* rRNA maturation. Plant Cell **26:** 777–787

Köster T, Marondedze C, Meyer K, Staiger D (2017) RNA-binding proteins revisited—The Emerging Arabidopsis mRNA interactome. Trends Plant Sci **22:** 512–526

Kunz HH, Zamani-Nour S, Häusler RE, Ludewig K, Schroeder JI, Malinova I, Fettke J, Flügge UI, Gierth M (2014) Loss of cytosolic phosphoglucose isomerase affects carbohydrate metabolism in leaves and is essential for fertility of Arabidopsis. Plant Physiol **166:** 753–765

Lee SE, Kim SY, Kim CM, Kim MK, Kim YR, Jeong K, Ryu HJ, Lee YS, Chung SS, Choy HE, et al (2007) The *pyrH* gene of *Vibrio vulnificus* is an essential in vivo survival factor. Infect Immun **75:** 2795–2801

Lezhneva L, Meurer J (2004) The nuclear factor HCF145 affects chloroplast *psaA-psaB-rps14* transcript abundance in *Arabidopsis thaliana.* Plant J **38:** 740–753

Manavski N, Torabi S, Stoppel R, Meurer J (2012) Phylogenetic and ontogenetic integration of organelles into the compartmentalized genome of the eukaryotic cell. Journal of Endocytobiosis and Cell Research **23:** 25–31

Manavski N, Torabi S, Lezhneva L, Arif MA, Frank W, Meurer J (2015) HIGH CHLOROPHYLL FLUORESCENCE145 binds to and stabilizes the *psaA* 5′ UTR via a newly defined repeat motif in Embryophyta. Plant Cell **27:** 2600–2615

Manavski N, Schmid LM, Meurer J (2018) RNA-stabilization factors in chloroplasts of vascular plants. Essays Biochem **62:** 51–64

Meurer J, Grevelding C, Westhoff P, Reiss B (1998) The PAC protein affects the maturation of specific chloroplast mRNAs in Arabidopsis thaliana. Mol Gen Genet **258:** 342–351

Meurer J, Schmid LM, Stoppel R, Leister D, Brachmann A, Manavski N (2017) PALE CRESS binds to plastid RNAs and facilitates the biogenesis of the 50S ribosomal subunit. Plant J **92:** 400–413

Meyer P, Evrin C, Briozzo P, Joly N, Bârzu O, Gilles AM (2008) Structural and functional characterization of *Escherichia coli* UMP kinase in complex with its allosteric regulator GTP. J Biol Chem **283:** 36011–36018

Minh PN, Devroede N, Massant J, Maes D, Charlier D (2009) Insights into the architecture and stoichiometry of *Escherichia coli* PepA*DNA complexes involved in transcriptional control and site-specific DNA recombination by atomic force microscopy. Nucleic Acids Res **37:** 1463–1476

Möhlmann T, Bernard C, Hach S, Ekkehard Neuhaus H (2010) Nucleoside transport and associated metabolism. Plant Biol **12**(Suppl 1): 26–34

Muraoka R, Okuda K, Kobayashi Y, Shikanai T (2006) A eukaryotic factor required for accumulation of the chloroplast NAD(P)H dehydrogenase complex in Arabidopsis. Plant Physiol **142:** 1683–1689

Okazaki Y, Shimojima M, Sawada Y, Toyooka K, Narisawa T, Mochida K, Tanaka H, Matsuda F, Hirai A, Hirai MY, et al(2009) A chloroplastic UDP-glucose pyrophosphorylase from *Arabidopsis* is the committed enzyme for the first step of sulfolipid biosynthesis. Plant Cell **21:** 892–909

Olinares PD, Ponnala L, van Wijk KJ (2010) Megadalton complexes in the chloroplast stroma of *Arabidopsis thaliana* characterized by size exclusion chromatography, mass spectrometry, and hierarchical clustering. Mol Cell Proteomics **9:** 1594–1615

Ostersetzer O, Cooke AM, Watkins KP, Barkan A (2005) CRS1, a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains. Plant Cell **17:** 241–255

Pfalz J, Liere K, Kandlbinder A, Dietz KJ, Oelmüller R (2006) pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. Plant Cell **18:** 176–197

Pfalz J, Bayraktar OA, Prikryl J, Barkan A (2009) Site-specific binding of a PPR protein defines and stabilizes 5′ and 3′ mRNA termini in chloroplasts. EMBO J **28:** 2042–2052

Robertson D, Carroll P, Parish T (2007) Rapid recombination screening to test gene essentiality demonstrates that *pyrH* is essential in *Mycobacterium tuberculosis.* Tuberculosis (Edinb) **87:** 450–458

Rogalski M, Karcher D, Bock R (2008) Superwobbling facilitates translation with reduced tRNA sets. Nat Struct Mol Biol **15:** 192–198

Roy LM, Barkan A (1998) A SecY homologue is required for the elaboration of the chloroplast thylakoid membrane and for normal chloroplast gene expression. J Cell Biol **141:** 385–395

Sauer E, Schmidt S, Weichenrieder O (2012) Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition. Proc Natl Acad Sci USA **109:** 9396–9401

Schmitz-Linneweber C, Lampe MK, Sultan LD, Ostersetzer-Biran O (2015) Organellar maturases: A window into the evolution of the spliceosome. Biochim Biophys Acta **1847:** 798–808

Schumacher MA, Pearson RF, Møller T, Valentin-Hansen P, Brennan RG (2002) Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: A bacterial Sm-like protein. EMBO J **21:** 3546–3556

Schwenkert S, Umate P, Dal Bosco C, Volz S, Mlçochová L, Zoryan M, Eichacker LA, Ohad I, Herrmann RG, Meurer J (2006) PsbI affects the stability, function, and phosphorylation patterns of photosystem II assemblies in tobacco. J Biol Chem **281:** 34227–34238

Serina L, Blondin C, Krin E, Sismeiro O, Danchin A, Sakamoto H, Gilles AM, Bârzu O (1995) *Escherichia coli* UMP-kinase, a member of the aspartokinase family, is a hexamer regulated by guanine nucleotides and UTP. Biochemistry **34:** 5066–5074

Steiner S, Schröter Y, Pfalz J, Pfannschmidt T (2011) Identification of essential subunits in the plastid-encoded RNA polymerase complex reveals building blocks for proper plastid development. Plant Physiol **157:** 1043–1055

Stern DB, Goldschmidt-Clermont M, Hanson MR (2010) Chloroplast RNA metabolism. Annu Rev Plant Biol **61:** 125–155

Stoppel R, Meurer J (2012) The cutting crew—Ribonucleases are key players in the control of plastid gene expression. J Exp Bot **63:** 1663–1673

Stoppel R, Meurer J (2013) Complex RNA metabolism in the chloroplast: An update on the *psbB* operon. Planta **237:** 441–449

Stoppel R, Lezhneva L, Schwenkert S, Torabi S, Felder S, Meierhoff K, Westhoff P, Meurer J (2011) Recruitment of a ribosomal release factor for light- and stress-dependent regulation of *petB* transcript stability in *Arabidopsis* chloroplasts. Plant Cell **23:** 2680–2695

Stoppel R, Manavski N, Schein A, Schuster G, Teubner M, Schmitz-Linneweber C, Meurer J (2012) RHON1 is a novel ribonucleic acid-binding protein that supports RNase E function in the *Arabidopsis* chloroplast. Nucleic Acids Res **40:** 8593–8606

Tang J, Zhang W, Wen K, Chen G, Sun J, Tian Y, Tang W, Yu J, An H, Wu T, et al (2017) OsPPR6, a pentatricopeptide repeat protein involved in editing and splicing chloroplast RNA, is required for chloroplast biogenesis in rice. Plant Mol Biol **95:** 345–357

Teubner M, Fuß J, Kühn K, Krause K, Schmitz-Linneweber C (2017) The RNA recognition motif protein CP33A is a global ligand of chloroplast mRNAs and is essential for plastid biogenesis and plant development. Plant J **89:** 472–485

Torabi S, Umate P, Manavski N, Plöchinger M, Kleinknecht L, Bogireddi H, Herrmann RG, Wanner G, Schröder WP, Meurer J (2014) PsbN is required for assembly of the photosystem II reaction center in *Nicotiana tabacum.* Plant Cell **26:** 1183–1199

Watkins KP, Kroeger TS, Cooke AM, Williams-Carrier RE, Friso G, Belcher SE, van Wijk KJ, Barkan A (2007) A ribonuclease III domain protein functions in group II intron splicing in maize chloroplasts. Plant Cell **19:** 2606–2623

Watkins KP, Rojas M, Friso G, van Wijk KJ, Meurer J, Barkan A (2011) APO1 promotes the splicing of chloroplast group II introns and harbors a plant-specific zinc-dependent RNA binding domain. Plant Cell **23:** 1082–1092

Witte CP, Noël LD, Gielbert J, Parker JE, Romeis T (2004) Rapid one-step protein purification from plant material using the eight-amino acid StrepII epitope. Plant Mol Biol **55:** 135–147

Yap A, Kindgren P, Colas des Francs-Small C, Kazama T, Tanz SK, Toriyama K, Small I (2015) AEF1/MPR25 is implicated in RNA editing of plastid *atpF* and mitochondrial *nad5,* and also promotes *atpF* splicing in Arabidopsis and rice. Plant J **81:** 661–669

Yosef I, Irihimovitch V, Knopf JA, Cohen I, Orr-Dahan I, Nahum E, Keasar C, Shapira M (2004) RNA binding activity of the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit from *Chlamydomonas reinhardtii.* J Biol Chem **279:** 10148–10156

Zhang HD, Cui YL, Huang C, Yin QQ, Qin XM, Xu T, He XF, Zhang Y, Li ZR, Yang ZN (2015) PPR protein PDM1/SEL1 is involved in RNA editing and splicing of plastid genes in *Arabidopsis thaliana.* Photosynth Res **126:** 311–321

Zhou K, Xia J, Wang Y, Ma T, Li Z (2017) A *Young Seedling Stripe2* phenotype in rice is caused by mutation of a chloroplast-localized nucleoside diphosphate kinase 2 required for chloroplast biogenesis. Genet Mol Biol **40:** 630–642

Zhou L, Lacroute F, Thornburg R (1998) Cloning, expression in *Escherichia coli*, and characterization of *Arabidopsis thaliana* UMP/CMP kinase. Plant Physiol **117:** 245–2549576794

Zhou W, Cheng Y, Yap A, Chateigner-Boutin AL, Delannoy E, Hammani K, Small I, Huang J (2009) The Arabidopsis gene *YS1* encoding a DYW protein is required for editing of *rpoB* transcripts and the rapid development of chloroplasts during early growth. Plant J **58:** 82–96

Zhu X, Guo S, Wang Z, Du Q, Xing Y, Zhang T, Shen W, Sang X, Ling Y, He G (2016) Map-based cloning and functional analysis of *YGL8,* which controls leaf colour in rice (*Oryza sativa*). BMC Plant Biol **16:** 134

Zoschke R, Watkins KP, Miranda RG, Barkan A (2016) The PPR-SMR protein PPR53 enhances the stability and translation of specific chloroplast RNAs in maize. Plant J **85:** 594–606

Zrenner R, Stitt M, Sonnewald U, Boldt R (2006) Pyrimidine and purine biosynthesis and degradation in plants. Annu Rev Plant Biol **57:** 805–836