Functional Characterization of PAC and PUMPKIN, Two Proteins Involved in Chloroplast RNA Metabolism

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften

an der Fakultät für Biologie

der Ludwig-Maximilians-Universität München

Lisa-Marie Schmid

München, 03/2020

Diese Dissertation wurde angefertigt unter der Leitung von PD Dr. Jörg Meurer im Bereich „Molekulare Pflanzenwissenschaften“ an der Fakultät für Biologie der Ludwig-Maximilians-Universität München

Datum der Abgabe:	31.03.2020
Datum der mündlichen Prüfung:	19.06.2020
Erstgutachter:	PD Dr. Jörg Meurer
Zweitgutachter:	Prof. Dr. Wolfgang Frank

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt wurde. Des Weiteren erkläre ich, dass ich nicht anderweitig ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen. Die folgende Dissertation liegt weder ganz noch in wesentlichen Teilen einer anderen Prüfungskommission vor.

Schmid Lisa-Marie

München, den 31.03.2020

Statutory declaration

I declare that I have authored this thesis independently and that I have not used other than the declared (re)sources. As well I declare that I have not submitted a dissertation without success and not passed the oral exam. The present dissertation (neither the entire dissertation nor parts) has not been presented to another examination board.

Schmid Lisa-Marie

München, 31.03.2020

Content

1. ABBREVIATIONS

Abbreviation	Full name
°C	degree Celsius
µg	microgram
µl	microlitre
µm	micrometre
ATP	adenosine triphosphate
bp	base pair
CFP	cyan fluorescent protein
cm	centimetre
CO_2	carbon dioxide
CTP	cytidine triphosphate
dCTP	deoxycytidine triphosphate
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
EMSA	Electrophoretic Mobility Shift Assay
FC	fold change
GTP	guanosine triphosphate
h	hour
HCO_3^-	hydrogen carbonate
Hz	hertz
IP	immunoprecipitation
LC/MS	liquid chromatography/ mass spectrometry
M	molar
MBP	maltose-binding protein
mg	milligram
min	minutes
miRNA	micro ribonucleic acid
mL	millilitre
mM	millimolar
mRNA	messenger ribonucleic acid
NEP	nuclear-encoded polymerase
nl	nanolitre
nm	nanometre

PAC	PALE CRESS
PAM	Pulse-Amplitude-Modulation
PCR	polymerase chain reaction
PEP	plastid-encoded polymerase
PUMPKIN	plastid UMP kinase
RBP	RNA-binding protein
RFP	red fluorescent protein
RIP-Seq	RNA-immunoprecipitation coupled with deep sequencing
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
TEV	Tobacco Etch Virus
tRNA	transfer ribonucleic acid
UDP	uridine diphosphate
UDPGlc	uridine diphosphate glucose
UMP	uridine monophosphate
UPP	uracil phosphoribosyltransferase
UTP	uridine triphosphate
UTR	untranslated region
w/v	weight per volume
WT	wild type
YFP	yellow fluorescent protein

2. LIST OF PUBLICATIONS WITH DECLARATION OF CONTRIBUTION

Publication I

Schmid LM, Ohler L, Möhlmann T, Brachmann A, Muiño JM, Leister D, Meurer J, Manavski N (2019) PUMPKIN, the Sole Plastid UMP Kinase, Associates with Group II Introns and Alters Their Metabolism. Plant Physiol 179: 248-264

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.

Manuscript I

Schmid LM, Möhlmann T, John A, Vicente A, Grüner F, Marino G, Lehmann M, Manavski N, Leister D, Meurer J; PUMPKIN and its Importance in the Plastid Pyrimidine Salvage Pathway

L.S., T.M., J.M., N.M. designed the research. **L.S.**, T.M., A.J., A.V., F.G., G.M., and M.L. performed the research. **L.S.**, T.M., D.L., and J.M. prepared the article with contributions from all co-authors. **L.S.** and J.M. supervised the whole study.

Publication II

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

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

Review Article

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

* First shared

For reasons of space, the review article can be accessed via this link:

https://portlandpress.com/essaysbiochem/article-lookup/doi/10.1042/EBC20170061

München, 31.03.2020

PD Dr. Jörg Meurer Lisa-Marie Schmid

3. SUMMARY

The chloroplast – as a product of an endosymbiotic event during which a cyanobacterium was incorporated into a eukaryotic cell – displays a combination of eubacterial and eukaryotic features in terms of gene expression and RNA metabolism. In line with the increased complexity of regulation of gene expression, especially at the posttranscriptional level, numerous factors evolved to fulfil various tasks ranging from RNA processing, splicing, editing or stabilization or subsequent processes such as ribosome biogenesis. This thesis aimed to functionally characterize two of such factors involved in plastid RNA metabolism, namely PAC (PALE CRESS) and the plastid UMP kinase PUMPKIN.

The *pac* mutation in *Arabidopsis thaliana* leads to albinotic plants with arrested chloroplast development, impaired photosynthesis and seedling lethality. On the molecular level, aberrant transcript abundance and patterns of plastid-encoded genes were observed, pointing to a role of PAC in chloroplast RNA metabolism. Closer analysis finally revealed similarities to lincomycin-treated plants with inhibited plastid translation, thus PAC was assumed to be involved in plastid translation. Its nucleoid-localization in combination with its occurrence in proteinaceous, RNA-containing high molecular weight complexes and co-migration with 50S ribosomal assembly states further argued for a function in ribosome assembly. Furthermore, PAC was found to directly interact with a subset of RNA targets, including the 23S rRNA. Hence, PAC acts as an RNA chaperone and assists in ribogenesis.

Knockouts and knockdowns of *pumpkin* in *Arabidopsis thaliana* exhibit massive growth defects, accompanied by reduced photosynthetic performance and plastid translation which in turn lead to pleiotropic defects in the abundance and pattern of plastid transcripts. In contrast to its previously proposed role in transcript stabilization of *psaA*, PUMPKIN was found to associate with five other intron-containing transcripts and influences their stability. *In vitro* RNA binding studies initially suggested that PUMPKIN might directly bind to its RNA targets, but additional experiments showed that the used RNA probe unspecifically bound to several – also unrelated – proteins. Using a co-immunoprecipitation approach, we could identify a large number of potential protein interaction partners, which predominantly were annotated to function as RNA-binding proteins. Thus, the observed RNA association of PUMPKIN might rather be mediated in concert with these factors. Additionally, PUMPKIN possesses an important function in nucleotide biosynthesis as it catalyses the enzymatic phosphorylation of UMP to UDP which is part of the plastid pyrimidine salvage pathway. *In vitro* enzymatic assays confirmed this property, which was comparable in activity to eubacterial types of UMP kinases

to which PUMPKIN is closely related. Additionally, further similarities between eubacterial UMP kinases and PUMPKIN could be detected such as their homomultimeric assembly. To finally clarify which of its two functions – RNA association or enzymatic activity – is more important and to which extent these impaired functions contribute to the observed mutant phenotype, we complemented the *pumpkin* mutant either with enzymatically impaired PUMPKIN versions retaining RNA-association function or with the *E. coli* homolog possessing only the enzymatic function. We could hereby confirm a major role for PUMPKIN as a functional plastid UMP kinase, while the RNA-association function might be largely dispensable under normal conditions. Furthermore, localization studies of PUMPKIN showed a dotted localization pattern within the chloroplast, which could be attributed to co-localization and clustering of PUMPKIN with other enzymes of the plastid pyrimidine salvage pathway to perform substrate channelling, providing a first hint for the existence of a plastidial "pyrimidinosome" similar to the already described "purinosome".

.

4. ZUSAMMENFASSUNG

Als Produkt der Endosymbiose – bei dem ein Cyanobakterium von einer eukaryotischen Zelle aufgenommen wurde – besitzen Chloroplasten eine Kombination aus eubakteriellen und eukaryotischen Eigenschaften, insbesondere im Hinblick auf ihre Genexpression und ihren RNA-Metabolismus. Vor allem ihre post-transkriptionelle Genregulation ist im Laufe der Evolution deutlich komplexer geworden. Daher wurden zahlreiche Faktoren rekrutiert, die diverse Funktionen bei der post-transkriptionellen Genregulation erfüllen. Dies umfasst: das Prozessieren, Spleißen, Editieren oder Stabilisieren von RNA oder aber auch eine Rolle in der Ribosomenbiogenese. Ziel dieser Arbeit war die funktionelle Charakterisierung von zwei solcher Faktoren, die eine Rolle im plastidären RNA-Metabolismus besitzen: PAC (PALE CRESS) und PUMPKIN (plastdiäre UMP-Kinase).

Mutationen des *PAC* Gens in *Arabidopsis thaliana* erzeugen Pflanzen mit albinotischem Phänotyp, gehemmter Chloroplastenentwicklung sowie Defekten bei der Photosynthese und führen letztlich zur Letalität im Keimlingsstadium. Auf molekularer Ebene zeigen *pac* Mutanten eine anormale Abundanz und Prozessierung plastidärer Transkripte. Daher scheint PAC eine Rolle im RNA-Metabolismus zu erfüllen. Weitere Analysen zeigten eine Ähnlichkeit der Transkriptmuster und -vorkommen zu Pflanzen, die mit dem plastidären Translationsinhibitor Lincomycin behandelt wurden, was auf eine Funktion von PAC in der chloroplastidären Translation schließen lässt. Letztlich konnte gezeigt werden, dass PAC Teil von RNA-sensitiven hochmolekularen Proteinkomplexen ist, in den Nukleoiden lokalisiert ist und mit Assemblierungsintermediaten der 50S ribosomalen Untereinheit ko-migriert, was für eine Funktion bei der Ribosomenbiogenese spricht. Außerdem konnte eine direkte Interaktion mit einigen RNAs, unter anderem der 23S rRNA, nachgewiesen werden. Folglich fungiert PAC als RNA-Chaperon und assistiert bei der Ribosomenassemblierung.

Auch *pumpkin*-Knockouts bzw. Knockdowns zeigten massive Wachstumsstörungen, eine beeinträchtigte Photosyntheseleistung sowie eine reduzierte plastidäre Translation, die zu weiteren pleiotropen Defekten in plastidären Transkriptmustern und -vorkommen führten. Anders als bisher angenommen ist PUMPKIN nicht an der Stabilisierung der *psaA* RNA beteiligt, sondern assoziiert mit fünf anderen intronhaltigen plastidären Transkripten und beeinflusst deren Stabilität. Erste *in vitro* RNA-Bindestudien ließen vermuten, dass PUMPKIN direkt mit den plastidären RNAs interagiert. Weitere Experimente zeigten jedoch, dass die verwendete RNA-Sonde unspezifisch an mehrere andere Proteine binden konnte. Schließlich konnten mithilfe einer Co-Immunopräzipitation zahlreiche potenzielle

Proteininteraktionspartner ermittelt werden, wobei den meisten Kandidaten eine Funktion im plastidären RNA-Metabolismus zugeordnet werden konnte – folglich könnte die RNA-Assoziation von PUMPKIN auch im Zusammenspiel mit diesen Faktoren vermittelt werden. Zusätzlich besitzt PUMPKIN eine zweite wichtige Rolle in der Nukleotidbiosynthese, in der es die enzymatische Reaktion von UMP zu UDP im plastidären Pyrimidin-Recycling-Weg katalysiert. *In vitro* Experimente konnten diese enzymatische Aktivität nachweisen, wobei die ermittelten Affinitäten mit denen von anderen verwandten eubakteriellen UMP-Kinasen vergleichbar waren. Weitere Ähnlichkeiten zeigten sich auch bei der beobachteten Homomultimerisierung von PUMPKIN. Um letztlich zu klären, welche der beiden Funktionen – die RNA-Assoziation oder die Enzymaktivität – die größere Bedeutung hat und wie sehr diese beeinträchtigten Funktionen zum Phänotyp der Mutante beitragen, wurde die *pumpkin* Mutante mit verschiedenen Konstrukten komplementiert. Einerseits wurden Punktmutationen von PUMPKIN mit beeinträchtigter Enzymaktivität und andererseits das *E. coli* Homolog PYRH verwendet, das ausschließlich enzymatische Aktivität aber keine RNA-Assoziation aufweist. Damit konnte bestätigt werden, dass PUMPKIN eine bedeutende Rolle als Enzym im Pyrimidin-Recycling zukommt, während die RNA-Assoziation unter normalen Bedingungen größtenteils verzichtbar ist. Zusätzlich zeigte PUMPKIN eine punktuelle Lokalisierung innerhalb der Chloroplasten, wobei weiterführende Lokalisationsstudien eine Clusterbildung von PUMPKIN mit anderen Enzymen des Pyrimidin-Recycling-Weges zeigte, vermutlich um eine effektivere Substratkanalisierung zu erreichen. Somit gibt es erste Hinweise für ein plastidäres „Pyrimidinosome“ in Anlehnung an das bereits beschriebene „Purinosome“.

5. AIMS OF THE THESIS

Analysing the complex RNA metabolism in chloroplasts is of utter importance to understand how gene expression is controlled. Especially in terms of post-transcriptional gene regulation plastids developed a plethora of regulatory steps and numerous factors are involved in processes, ranging from RNA processing, maturation, stabilization, splicing to editing (Stern et al., 2010). Many factors have been characterized that mostly possess classical RNA-binding motifs, but latest research argues for many, yet unrecognized proteins, which also participate in post-transcriptional gene regulation without containing known RNA-binding domains (Köster et al., 2017).

This thesis is aimed to functionally characterize two of such plastid-localized proteins involved in chloroplast RNA metabolism: PAC (PALE CRESS) and PUMPKIN (plastid UMP kinase). Neither of them possesses classical domains or motifs for RNA-recognition but knock-out of both led to severe defects and massively impaired growth or even a seedling lethal phenotype which could be tied into post-transcriptional processes and underlines the importance of these factors.

In previous studies on *pac* mutants, it became obvious that knock-out plants of *PAC* exhibited a quite pleiotropic and severe phenotype (Reiter et al., 1994). First analysis provided evidence for a disturbance of many plastid-encoded transcripts regarding their abundance and transcript patterns suggesting a role of PAC in plastid RNA metabolism (Meurer et al., 1998). Still, for almost twenty years it remained unclear which molecular function the PAC protein precisely fulfils. Thus, the hereafter presented in-depth characterization intended to identify in which step of gene expression PAC participates, which RNA targets it may bind and how it then influences these targets.

Previous studies proposed a function of PUMPKIN as a stabilization factor of *psaA* mRNA (Hein et al., 2009), thus making PUMPKIN an interesting candidate for in-depth analysis of its RNA-binding repertoire and properties. Especially as HCF145, a previously characterized factor in our group, was shown to bind to the same target (*psaA 5'UTR)* (Manavski et al., 2015). When starting molecular analyses, we observed clear differences between the previously described *pumpkin* mutant line and our confirmed independent line, which prompted us to completely re-analyse the molecular defects in our mutant and thus elucidating the real function of the corresponding protein. Besides characterizing its role in RNA metabolism, we further focused on analysing its second function as a eubacterial-like UMP kinase converting UMP to

UDP in the plastid pyrimidine salvage pathway, which was not examined in detail before. To gain even deeper insight in its dual-functionality a complementation approach with either enzymatically impaired PUMPKIN versions or the *E. coli* homolog PYRH only possessing enzymatic activity but no RNA-association capability was conducted, finally helping to decipher which of the two functions is causative for the observed phenotype and thus of more relevance.

6. INTRODUCTION

6.1. The evolution of the chloroplast

The acquisition of organelles also referred to as the "endosymbiotic theory" represents the most crucial evolutionary step towards eukaryotic cells and organisms (Figure 1). More than one billion years ago a cyanobacterial ancestor was engulfed by a heterotrophic host which previously acquired mitochondria by the uptake of an α-proteobacterium (Maréchal, 2018). Even though there is still an ongoing debate about the sequence of events that generated this heterotrophic eukaryotic host – as it is assumed that evolution of eukarya could only be accomplished by a mitochondria-containing cell due to bioenergetic means (Martin et al., 2015) – the next primary endosymbiotic event by which a cyanobacterium was engulfed and finally developed into the semiautonomous chloroplast, is widely accepted. Moreover, it seems as if horizontal gene transfer provided additional genetic material, including the potential contribution of Chlamydia and other bacteria (Huang and Gogarten, 2007; Eme et al., 2017).

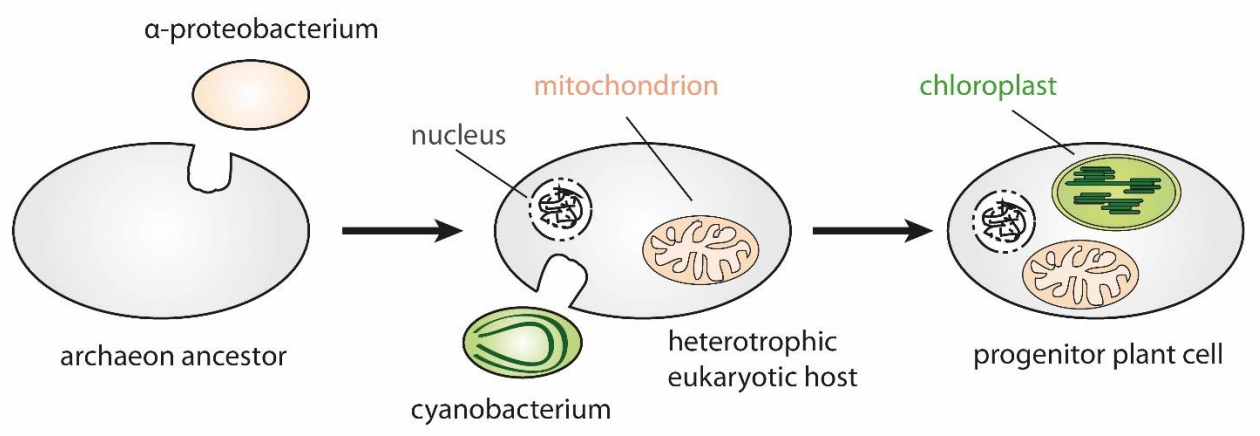


Figure 1. Evolution of organelles via primary endosymbiosis. Chloroplasts are descendants of a cyanobacterium engulfed by a heterotrophic eukaryotic host that previously incorporated an α-proteobacterium which developed into mitochondria.

From this initial "primary endosymbiosis" event three main lineages derived: the Glaucophyta, the Rhodophyta (red algae) and – what is of utter importance for the further evolution of plants – the Chlorophyta (green algae) (Chan et al., 2011).

The acquisition of these organelles had tremendous effects on the overall cellular metabolism of the host (Bhattacharya et al., 2007). Harbouring an intracellular "light-energy-converting" and thus energy providing chloroplast which performed oxygenic photosynthesis was a

prerequisite to grow and spread with only few starting materials: light, water, carbon and some basic nutrients (Gould et al., 2008). This laid the foundation to occupy new habitats, radiate and evolve into new species.

6.2. Adaptations towards an integrated organelle

Evolving from a free-living unicellular organism into an organelle that is tightly integrated and regulated within the host cell led to several adaptations. Even though organelles such as chloroplasts and mitochondria kept part of their original genome, massive gene transfer, loss and re-arrangements occurred. The chloroplast genome of the model plant *Arabidopsis thaliana* for example harbours 87 protein-coding genes (Sato et al., 1999) in comparison to present-day cyanobacteria such as *Synechocystis sp.* PCC6803 which possesses 3168 protein-coding genes (Kaneko and Tabata, 1997). In total roughly 4500 genes are assumed to have been transferred from the cyanobacterial endosymbiont to the host, but only one third of the corresponding proteins is predicted to be targeted into chloroplasts, the rest underwent redistribution into other compartments (See also Figure 2) (Leister, 2003).

With this outsourcing of genetic information, the host cell became able to gain control over chloroplast function. In consequence, this required intracellular communication and signalling to fine-tune and coordinate the expression of proteins encoded in the chloroplast genome and the nucleus. These communication processes are also referred to as anterograde signalling (from nucleus to chloroplasts) and retrograde signalling (from chloroplasts to nucleus) (Börner, 2017). Regarding, for instance, the composition of the photosynthetic complexes within the chloroplast, the different subunits represent a mosaic of either plastid- or nucleus-encoded proteins (Leister, 2003). Hence, obtaining functional photosynthesis complexes in the correct stoichiometry requires accurate coordination of gene expression, availability of the needed proteins as well as a controlled assembly process.

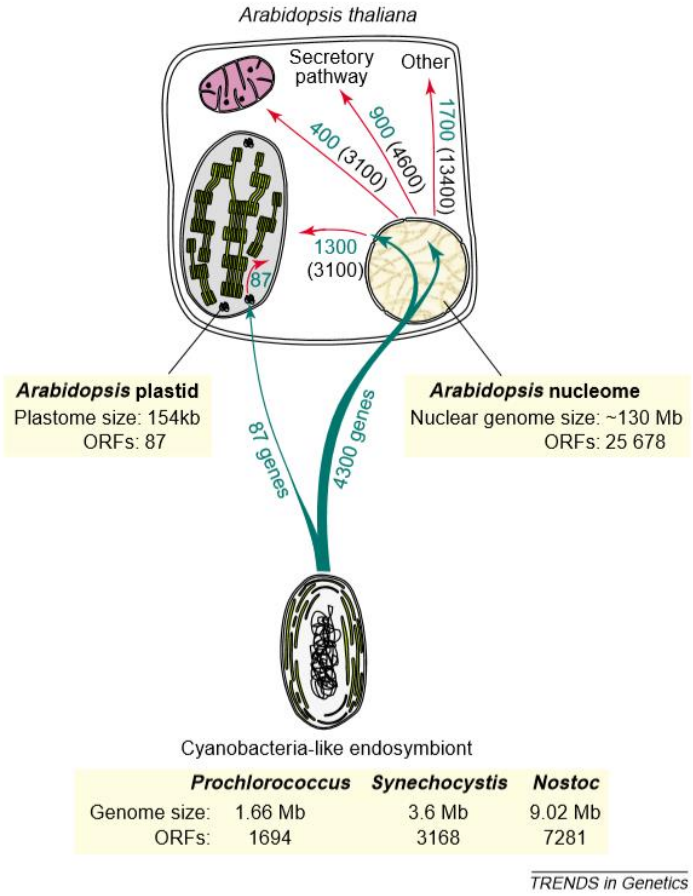


Figure 2. Distribution and estimated numbers of genes derived from the cyanobacterial ancestor which evolved into the semiautonomous chloroplast. Permission for reprint is attached (taken from Leister, 2003).

As nucleus-encoded genes are translated from RNA into proteins in the cytosol, new mechanisms for protein targeting and transport into the organelles needed to evolve. To overcome the two membranes – inherited from the prokaryotic ancestor – the so-called TOC-TIC (translocons at the outer/ inner envelope membrane of chloroplasts) system was developed allowing the recognition of targeting peptides and the import of cytosolic translated proteins to their final destination within the chloroplast (Nakai, 2018). Also in this case the protein import machinery is a mixture of ancient prokaryotic proteins such as Toc75, which partially gained new functions as well as newly evolved factors such as the GTPase receptors Toc159 and Toc34 (Day and Theg, 2018).

6.3. Ancestral and newly evolved features

Concerning the combination of inherited eubacterial and newly evolved features in plastid gene expression, it sometimes remains challenging to strictly classify a process in the gene expression machinery into these categories. More than one billion years of evolution left their marks even on ancient and highly conserved mechanisms. The next section summarizes the different steps of plastid gene expression starting from nucleotide biosynthesis over transcription, post-transcriptional regulations and finally translation and focuses on the origin of the involved factors and machineries.

6.3.1. Nucleotide biosynthesis

The prerequisite for storing genetic information and subsequent gene expression are nucleotides as the basic constituents of DNA and RNA. Their synthesis is complex and tightly regulated but the principle chemical reactions are quite conserved and only the enzymes involved differ in their origin, evolution and localization. Any nucleotide *de novo* synthesis pathway, regardless if synthesizing purines or pyrimidines, is based on the fusion of the sugar part of PRPP (5-phosphoribosyl-1-pyrophosphate) with the formed purine or pyrimidine ring structure (Zrenner et al., 2006). In case of purine biosynthesis this involves 12 plastid-localized enzymatic reactions (Moffatt and Ashihara, 2002). In contrast, UMP and thus pyrimidine *de novo* synthesis only needs six enzymatic steps which are distributed over different plant cell compartments (Witte and Herde, 2020).

As depicted in Figure 3, the initial reaction of pyrimidine *de novo* biosynthesis, also called "orotate pathway", involves HCO_3^-, glutamine and ATP which are converted into carbamoylphosphate by the plastid-localized CPS (carbamoyl-phosphate synthetase) (Moffatt and Ashihara, 2002). Transforming carbamoylphosphate into carbamoylaspartate is accomplished by the plastid ATC (aspartate transcarbamoylase) (Chen and Slocum, 2008). Carbamoylaspartate is then transported into the cytosol – by a yet unknown mechanism – where it gets converted into dihydroorotic acid by DHO (dihydroorotase) (Zrenner et al., 2006). The next enzyme, DHOD (dihydroorotate dehydrogenase), which is localized at the outer surface of the inner membrane of mitochondria subsequently turns dihydroorotic acid into orotic acid, which is transported back to the cytosol where finally UMPS (UMP synthase) connects the orotate ring structure with PRPP (5-phosphoribosyl-1-pyrophosphate) and decarboxylates the OMP to UMP (Witte and Herde, 2020).

Further phosphorylation reactions by UMP/CMP kinase (U/CMPK) and nucleoside diphosphate kinase (NDPK) finally lead to UTP which can be converted into CTP by the addition of an amino group from glutamine, which is catalysed by CTPS (CTP synthase) (Moffatt and Ashihara, 2002).

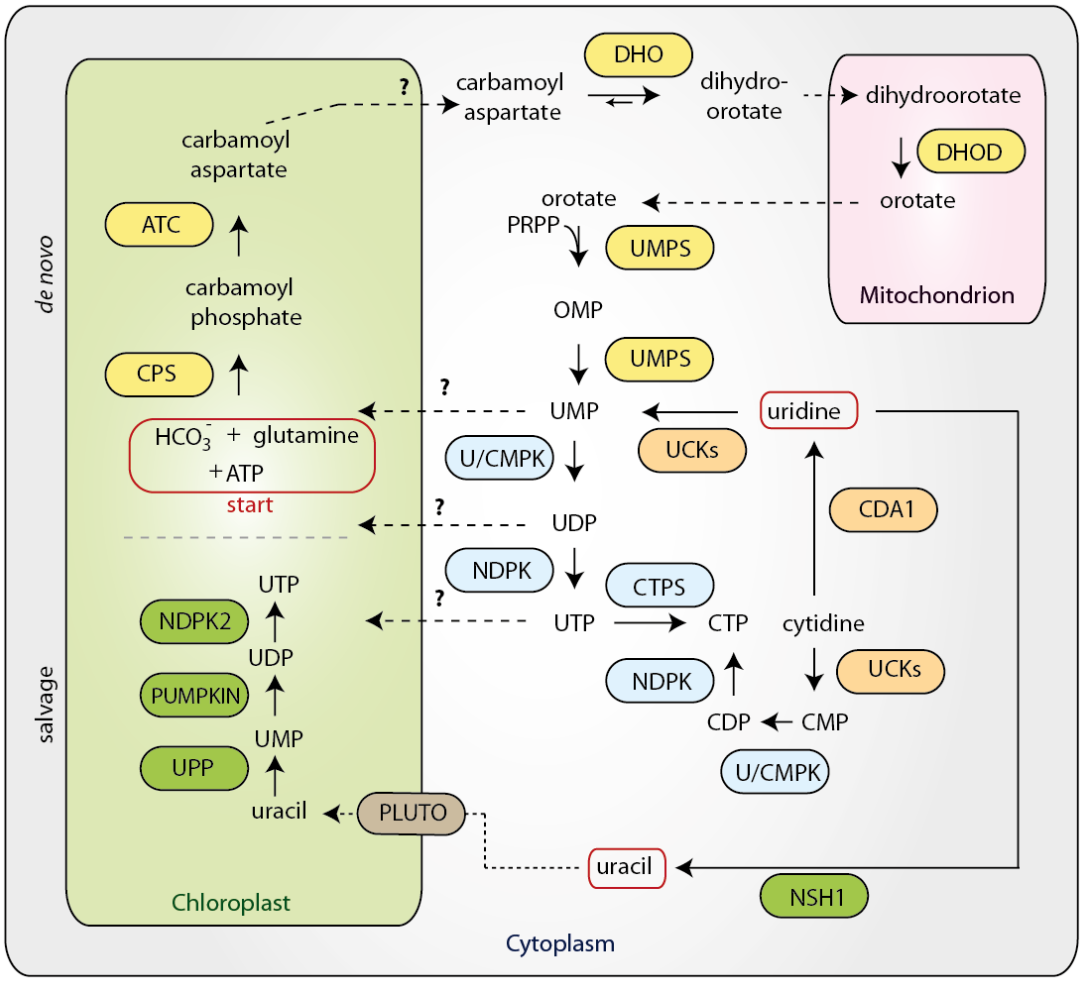


Figure 3. Pyrimidine biosynthesis pathways in plants. The enzymes involved in the six steps of *de novo* biosynthesis are coloured in yellow: CPS (carbamoyl-phosphate synthetase), ATC (aspartate transcarbamoylase), DHO (dihydroorotase), DHOD (dihydroorotate dehydrogenase), UMPS (UMP synthase). Further phosphorylation or interconversion steps are depicted in blue: U/CMPK (UMP/CMP kinase), NDPK (nucleoside diphosphate kinase), CTPS (CTP synthase). Enzymes which are part of the cytosolic uridine salvage pathway are marked in orange: UCK (uridine-cytidine kinase), CDA1 (cytidine deaminase), while the ones involved in the uracil salvage pathway are shown in green: NSH1 (nucleoside hydrolase 1), UPP (uracil phosphoribosyltransferase), PUMPKIN (plastid UMP kinase), NDPK2 (nucleoside diphosphate kinase 2). Dashed arrows with question marks represent yet unknown mechanisms of possible nucleotide (UMP/ UDP/ UTP) transfer between the compartments. Important starting points for either *de novo* synthesis or salvaging pathways are encircled in red for better orientation.

In addition to *de novo* synthesis, there are also salvaging pathways which are less energy consuming than the *de novo* synthesis as they make use of degradation products (Mainguet et al., 2009). One of them is localized in the cytosol and starts with uridine, which can also be interconverted by CDA1 from cytidine (Faivre-Nitschke et al., 1999). Uridine or cytidine can then be converted into UMP by UCKs (uridine-cytidine kinases) (Ohler et al., 2019).

A second salvage pathway can be found in chloroplasts which makes use of uracil. Uracil originates from uridine catalysed by NSH1 (uridine nucleosidase) and gets transported into the chloroplast via PLUTO (Jung et al., 2011; Girke et al., 2014). Three enzymatic reactions subsequently catalyse the formation of UTP, using uracil as the starting molecule: UPP (uracil phosphoribosyltransferase) accomplishes the conversion of uracil into UMP, PUMPKIN (plastid UMP kinase) phosphorylates UMP to UDP, and NDPK2 (nucleoside diphosphate kinase 2) finally generates UTP (Ohler et al., 2019). To which extent nucleotides can be exchanged between the different compartments is unclear (therefore represented by the dashed lines with question marks in Figure 3). Nevertheless, the viability of the *pumpkin* mutant presented below, argues for a certain degree of nucleotide transport from the cytosol to plastids to at least partially compensate the disturbed salvage pathway (Schmid et al., 2019).

Retracing the evolution of the different enzymes involved in the *de novo* pyrimidine synthesis is rather challenging, as their origin seems to be chimeric and might have involved horizontal gene transfer, especially in plants. While CPS is closely related to the respective enzymes in cyanobacteria, arguing for an endosymbiotic origin, plant ATC is not found in the same clade with cyanobacteria but is found in a monophyletic group containing also virus sequences, speaking in favour of horizontal gene transfer, while DHO seems to originate from proteobacteria (Nara et al., 2000).

6.3.2. Chloroplast gene expression

Even though chloroplasts kept some of their ancestral features, plastid gene expression has developed into a much more complex and tightly regulated process. Figure 4 summarizes ancient (beige) as well as newly evolved (red) attributes in chloroplast gene expression.

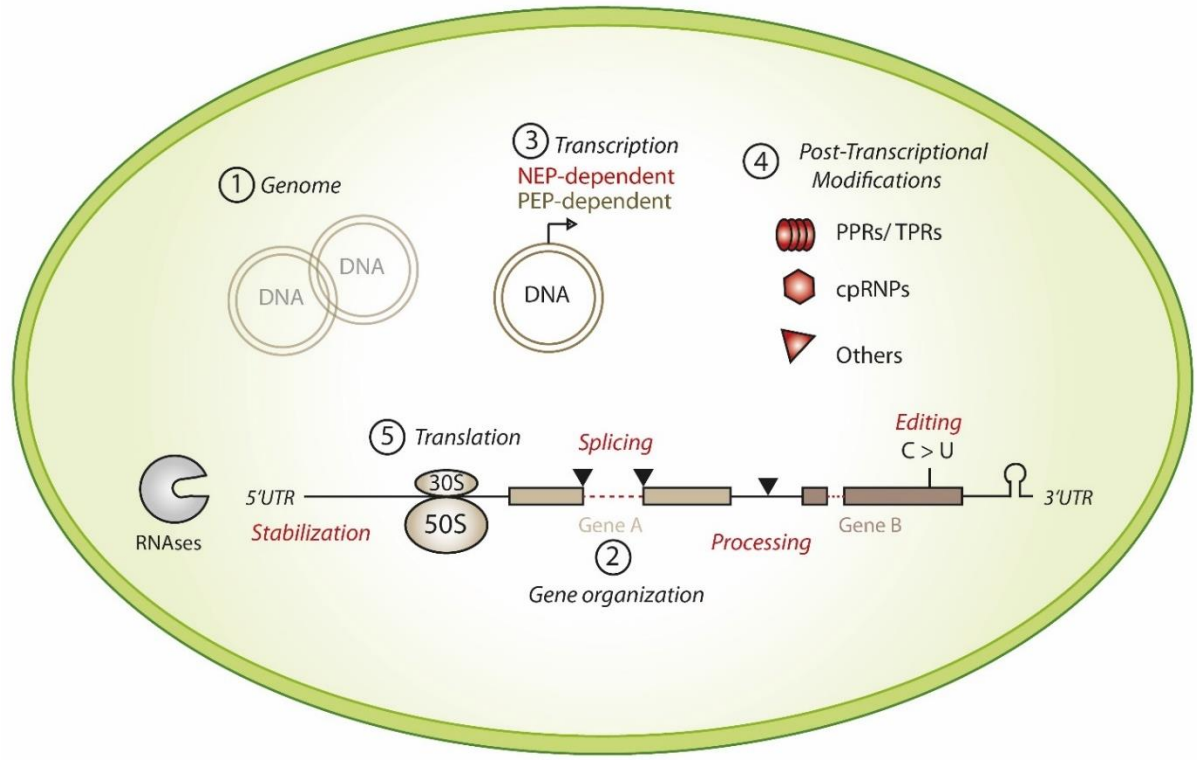


Figure 4. Overview of chloroplast gene expression with a focus on the origin of the different processes. Ancestral features and processes of eubacterial origin are depicted in beige colours while newly evolved machineries and mechanisms are marked in red. Numbers refer to the sections in the following paragraph explaining function and evolution of the depicted steps and mechanisms in detail.

According to the numbers indicated in Figure 4, the following five paragraphs describe each of the processes and their evolution in detail.

(1) In strong contrast to the eukaryotic host, plastidial DNA is stored in a circular structure, resembling eubacterial plasmids, moreover each chloroplast contains more than one copy of its genetic information (Morley and Nielsen, 2016).

(2) Similarly, the plastid-encoded genes are ordered in an operon-like structure often accompanied by eubacterial-type gene promoters and terminators leading to polycistronic precursors – a classical characteristic of eubacteria (Shahar et al., 2019).

(3) In terms of transcription one must differentiate between the two different types of RNA-polymerases which are found in the chloroplasts of angiosperms. One, namely the PEP (plastid-encoded polymerase) is from eubacterial origin and is composed of the plastid-encoded subunits RpoA/ B/ C1 and C2 (Yagi and Shiina, 2012). For promoter recognition the nucleoid-localized PEP further associates with σ^{70} sigma factors and PAPs (polymerase associated proteins), which

comprise TAC (plastid transcriptionally active chromosome) proteins (Schweer et al., 2010). While the core subunits and the sigma factors originated from the eubacterial ancestor, the associated PAP proteins do not possess eubacterial homologs (Börner et al., 2015). This adds an example of a bacterial-type and hence ancient process which was further adopted by including accessory factors. In addition, chloroplasts harbour another type of polymerase – the NEP (nuclear-encoded polymerase), also referred to as RPOTp which is a phage-type RNA polymerase (Ortelt and Link, 2014). It is composed of a single subunit and exhibits similarities to RNA polymerases from bacteriophages (Hedtke et al., 1997). Dicots even house an additional nuclear-encoded T7 phage-type polymerase (RPOTmp), which is dual-localized to plastids and mitochondria (Bohne et al., 2016). Many of the plastid-encoded genes contain promoter regions for both types of polymerases, but partly exhibit preferences for PEP- or NEP-dependent transcription (Hajdukiewicz et al., 1997). Interestingly, expression of the PEP-subunit RpoB is strictly dependent on the action of NEP polymerase (Börner et al., 2015). It is further assumed that NEP is more active in proplastids, while PEP takes over during chloroplast differentiation, even though NEP activity is still required in mature plastids (Ortelt and Link, 2014).

(4) However, in contrast to the bacterial genetic machinery, chloroplast gene expression is not predominantly regulated on the transcriptional level but gained a new layer of extensive regulation by evolving an amazingly complex system of post-transcriptional RNA processing and modification steps (Maier et al., 2008). As depicted in Figure 4, this includes protein-mediated intron splicing, processing, editing and stabilization of transcripts.

Several plastid-encoded genes acquired intron structures interrupting protein-coding, tRNA or rRNA genes (Sato et al., 1999). Depending on their structure they are classified in either group I or group II introns, whilst group II introns prevail in plant chloroplasts (de Longevialle et al., 2010). Originally, introns – as mobile genetic elements – contained open reading frames for so called IEPs (intron-encoded proteins) that possessed a maturase domain for splicing and a reverse transcriptase function essential for intron mobility (Schmitz-Linneweber et al., 2015). In present chloroplasts almost all introns lost the internally encoded maturase and also their self-splicing activity and are now dependent on several nuclear-encoded RNA-binding splicing factors (Schmitz-Linneweber et al., 2015).

Even though bacteria are capable of RNA editing – the alteration of nucleotide identities – so far only adenosine (A) to inosine (I) editing in a comparably small subset of targets was observed (Bar-Yaacov et al., 2018). In chloroplast of flowering plants, however, only C-to-U (cytidine to uridine) editing events occur and editing of the about 40 sites in Arabidopsis

chloroplast genomes is of utter importance for proper gene expression (Hammani et al., 2009). In all seed plant lineages C-to-U editing events were reported, but no editing processes were observed in green algae, thus emphasizing their independent evolution (Takenaka et al., 2013).

As most, if not all, chloroplast genes are encoded in an operon-like manner, polycistronic precursors need to be cut into mature and thus translationally competent gene units. These processing events rely on the action of endonucleases, cutting within RNA sequences as well as the action of different 5'→3' or 3'→5' exonucleases chopping off 5' or 3' overhangs from transcript ends until they are either stopped by secondary RNA-structures or due to the hindrance by RNA-binding proteins (Manavski et al., 2018). The action of these RNA-binding proteins as RNA protectors does therefore not only result in the stabilization of the transcripts by preventing their decay, but is also essential to define transcript lengths and termini (Zhelyazkova et al., 2012). The different tasks and types of RNA-binding proteins are described in more detail in section 6.4.

(5) Finally, another conserved property is the translational machinery as chloroplasts possess bacterial type 70S ribosomes, composed of the small 30S and the large 50S ribosomal subunits – in contrast to the eukaryotic 80S ribosomes consisting of a 40S and 60S subunit (Harris et al., 1994). It still needs to be mentioned that 30S and 50S ribosomal subunits differ in their composition between eubacteria and chloroplasts as plastid-specific ribosomal proteins were identified, while orthologues to the *E. coli* subunits L25 and L30 are missing (Yamaguchi and Subramanian, 2000; Yamaguchi et al., 2000).

6.3.3. Ribosome biogenesis and translation

In general – and similar to eubacteria – translation can be divided into three main steps: initiation, elongation and termination. To finally build a functional ribosome consisting of a small (SSU) and a large (LSU) ribosomal subunit several complex assembly steps are necessary, including the correct processing and maturation of ribosomal RNAs and the coordinated localization and folding of the numerous ribosomal proteins (Zoschke and Bock, 2018). The two subunits SSU and LSU need to be independently assembled before they are recruited to form the mature ribosome and start their translational activity (Tiller and Bock, 2014). Similar to eubacteria, plastid ribosomes contain 23S, 16S and 5S rRNA, but clearly show differences in their processing, which results in an additional 4.5S rRNA in land plants deriving from the 3' end of the 23S rRNA precursor (Edwards and Kossel, 1981; Stoppel and Meurer, 2012). The plethora of ribosomal proteins has a eubacterial origin, still chloroplast ribosomes

possess additional and plant-specific ribosomal subunits (PSRP2-6) (Xu et al., 2013). The correct assembly of the subunits furthermore requires the help of extraribosomal proteins, some of them have bacterial homologs but also many newly evolved ribogenesis factors (Chi et al., 2012; Jeon et al., 2014; Reiter et al., 2020).

While cyanobacteria use the Shine-Dalgarno (SD) sequence for efficient translation initiation, the plastid translation machinery requires additional *trans*-acting factors and RNA stretches with reduced secondary structures as Shine-Dalgarno sequences are not always present and well conserved (Hirose and Sugiura, 2004). For initiation, the 30S subunit is recruited with the help of initiation factors and the initiator tRNA which binds to the – mostly used – AUG start codon (Sugiura et al., 1998). Subsequently the 50S ribosomal subunit is recruited and the complete ribosome gets assembled to start translation elongation in concert with several elongation factors possessing bacterial origin (Albrecht et al., 2006). Termination of translation is triggered by stop codons, a process which requires help of auxiliary factors (Meurer et al., 2002) and plastid ribosomes are disassembled by a ribosome recycling factor (RRF) homologous to the RRF of *E. coli* (Rolland et al., 1999).

6.4. The co-evolution of chloroplast RNA metabolism and regulating proteins

As illustrated above, the extensive post-transcriptional regulation of gene expression triggered the need for numerous factors to fulfil these tasks. Especially the occurrence of new transcript termini due to processing events as well as the introduction of various editing sites and introns unable to self-splice *in vivo* led to the evolution and expansion of many nuclear-encoded RNA-binding protein families.

One outstanding family of these RNA-binding proteins are the so-called PPR (pentatricopeptide repeat) proteins, which are closely related to the family of TPR (tetratricopeptide repeat) and OPR (octotricopeptide repeat) proteins (Barkan and Small, 2014). They all share a certain architecture of – often degenerated – amino acid helical repeats, which are capable to bind DNA, RNA or other proteins. In PPR proteins, these repeats typically consist of 35 amino acids – as indicated by their name “pentatricopeptide” – while the number of repeats varies from 2 up to 30 (Rovira and Smith, 2019). Intense studies on their RNA-binding capability, since their discovery twenty years ago, could meanwhile even elucidate their recognition “code” as certain amino acids at defined positions within the repeat determine which nucleotide is bound (Barkan et al., 2012; Shen et al., 2016). In comparison to green alga such as *Chlamydomonas reinhardtii* which only possesses 14 PPR proteins, a massive expansion of this protein family happened in

land plants such as Arabidopsis (~450 PPRs) enabling the individual proteins to develop into very specific factors with defined RNA targets and mode of actions (Wang et al., 2015). In contrast, the OPR protein family underwent a diversification and expansion in *C. reinhardtii*, thus increasing complexity of RNA metabolism seems to trigger the need and evolution of appropriate regulatory factors (Bohne et al., 2009). PPR proteins display a surprisingly diverse portfolio: Besides functioning in RNA editing, which is one of the most frequent function of hitherto characterized PPR proteins in Arabidopsis, they are also involved in RNA stabilization, processing, intron splicing and often influence the translation of their RNA targets (Barkan and Small, 2014).

Another class of typical RNA-binding proteins are the chloroplast ribonucleoproteins (cpRNPs) which comprise only few members but with a broad spectra of RNA targets (Kupsch et al., 2012). They contain two RRM (RNA recognition motif) domains which confer binding to ribosome-free mRNAs, thereby contributing to mRNA stabilization and protection against degradation, thus cpRNPs rather represent general RNA-protectors (Li and Sugiura, 1991; Nakamura et al., 2001).

Concerning the splicing of plastid introns, a family of CRM-containing (Chloroplast RNA splicing and ribosome maturation) proteins was identified which derived from ribosome-associated proteins (Barkan et al., 2007). While some factors exhibit a clear preference to only one or few targets, such as maize CRS1 which assists in *atpF* intron splicing, other factors such as CRS2 or CFM3 bind to a larger set of targets often belonging to the same intron-type subclass (de Longevialle et al., 2010).

The slight differences in ribosome composition and translation provoked the evolution of several ribosome biogenesis factors involved in rRNA-processing or maturation and ribosome assembly. Some of them belong to well-characterized protein classes such as DEAD-box RNA helicases, also present in eubacteria, while many others lack eubacterial counterparts and are only found in the green lineage or are even specific for land plants (Paieri et al., 2018; Reiter et al., 2020).

In general, most of these RNA-binding proteins possess classical RNA-binding motifs, including the helical repeats of P/T/OPRs, the RRM (RNA recognition motif) or zinc finger domains (Yan et al., 2018) which facilitate their identification based on protein sequences. Especially forward genetic approaches making use of the characterization of mutant collections could elucidate other factors with either previously unknown RNA-binding motifs or simply

lacking RNA-binding domains (Manavski et al., 2012; Manavski et al., 2018). One of these examples is HCF145, which was initially identified in a screen for "high chlorophyll fluorescence" mutants, impaired in photosynthesis (Feldmann, 1991; Meurer et al., 1996). In-depth analysis of the appropriate mutant led to the identification of a defect in *psaA/B* transcript accumulation and the conclusion that the corresponding protein HCF145 is essential in stabilizing the 5'UTR of the tricistronic *psaA/psaB/rps14* precursor (Lezhneva and Meurer, 2004). *In silico* analysis further identified two tandem repeat motifs termed TMR (transcript motif binding repeat) adding a previously unknown RNA-binding domain to the existing repertoire (Manavski et al., 2015). Latest studies of the RNA-binding proteome could further contribute to the identification to previously unknown RNA-binding proteins (Reichel et al., 2016). Amongst the identified classes especially enzymes seem to play an important role (Marondedze et al., 2016), as described by the REM (RNA-Enzyme-Metabolite) hypothesis (Hentze and Preiss, 2010). These moonlighting enzymes catalyse diverse reactions within metabolic pathways such as glycolysis and additionally possess RNA-binding function which allows fine-tuning and feedback regulation of gene expression (Castello et al., 2015). Despite some well-characterized examples in the mammalian system, such as GAPDH which converts glyceraldehyde-3-phosphate to D-glycerate-1,3-bisphosphate and was proven to act as an RNA-binding protein probably via its Rossmann fold (Nagy and Rigby, 1995), characterization of moonlighting enzymes remains scarce *in planta*. Thus, the hereafter presented functional characterization of two non-classical RNA-associated factors will add another layer to this network, especially providing more details about the interaction of different metabolic processes with respect to the dual-functionality of PUMPKIN in plastid RNA and nucleotide metabolism.

7. RESULTS

7.1. Publication I: PUMPKIN, the Sole Plastid UMP Kinase, Associates with Group II Introns and Alters Their Metabolism

Schmid LM, Ohler L, Möhlmann T, Brachmann A, Muiño JM, Leister D, Meurer J, Manavski N (2019), Plant Physiol 179: 248-264

Supplementary data available at: http://www.plantphysiol.org/content/179/1/248

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

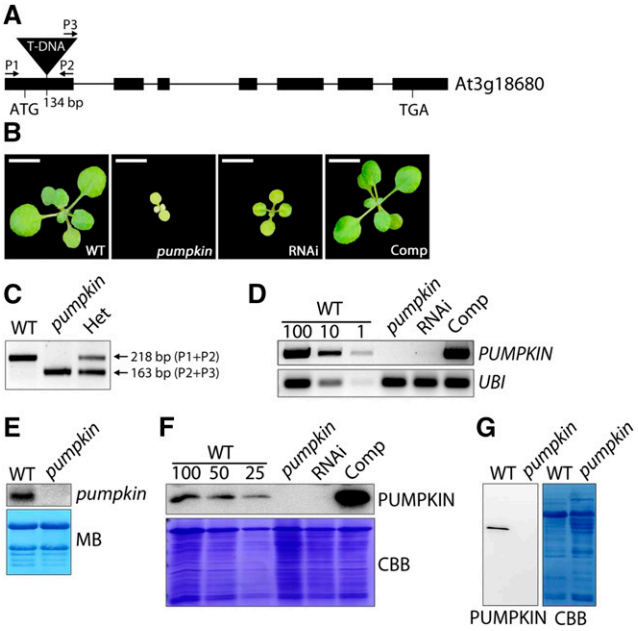


Figure 1. Phenotype and Analysis of the *pumpkin* T-DNA Insertion Line. A, Schematic representation of the Arabidopsis *PUMPKIN* gene (At3g18680). Exons are depicted as black bars, introns as black lines. The T-DNA insertion is represented by the black triangle, and its position is indicated relative 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

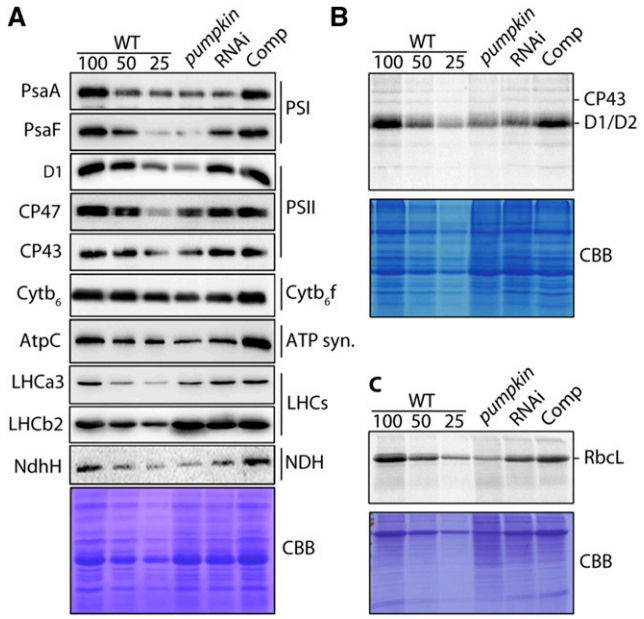


Figure 2. Protein Accumulation and Translation Rates in *pumpkin*. A, Immunotitration of thylakoid proteins in *pumpkin*. Thylakoids were extracted from the wild-type, *pumpkin,* RNAi, and Comp plants, separated on SDS gels and immunodecorated with indicated antibodies. In (A), 100, 50, and 25 of the 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

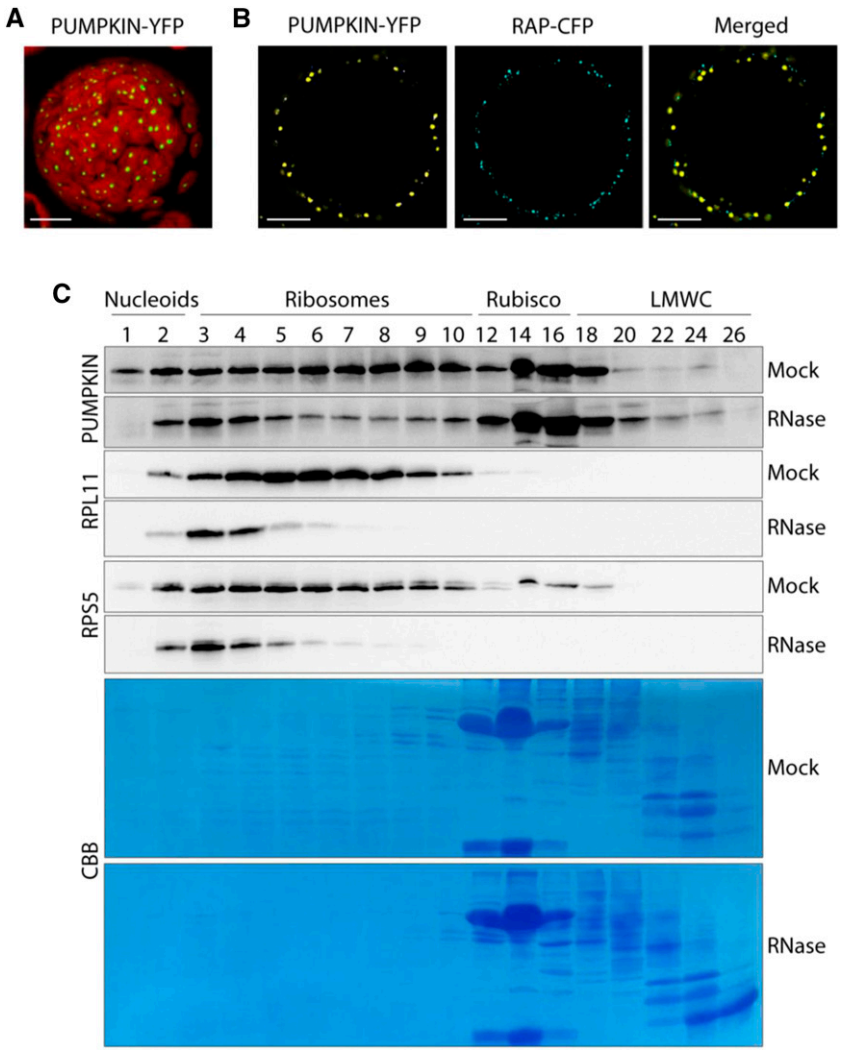


Figure 3. Subcellular Localization and in Vivo Complex Formation of PUMPKIN. A, Confocal image of a PUMPKIN-YFP-transfected *N. benthamiana* protoplast. Merged signals of chloroplast autofluorescence (red) and YFP fluorescence (yellow) are shown. The white scale bar corresponds to 10 μm. B, Macrographs 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,

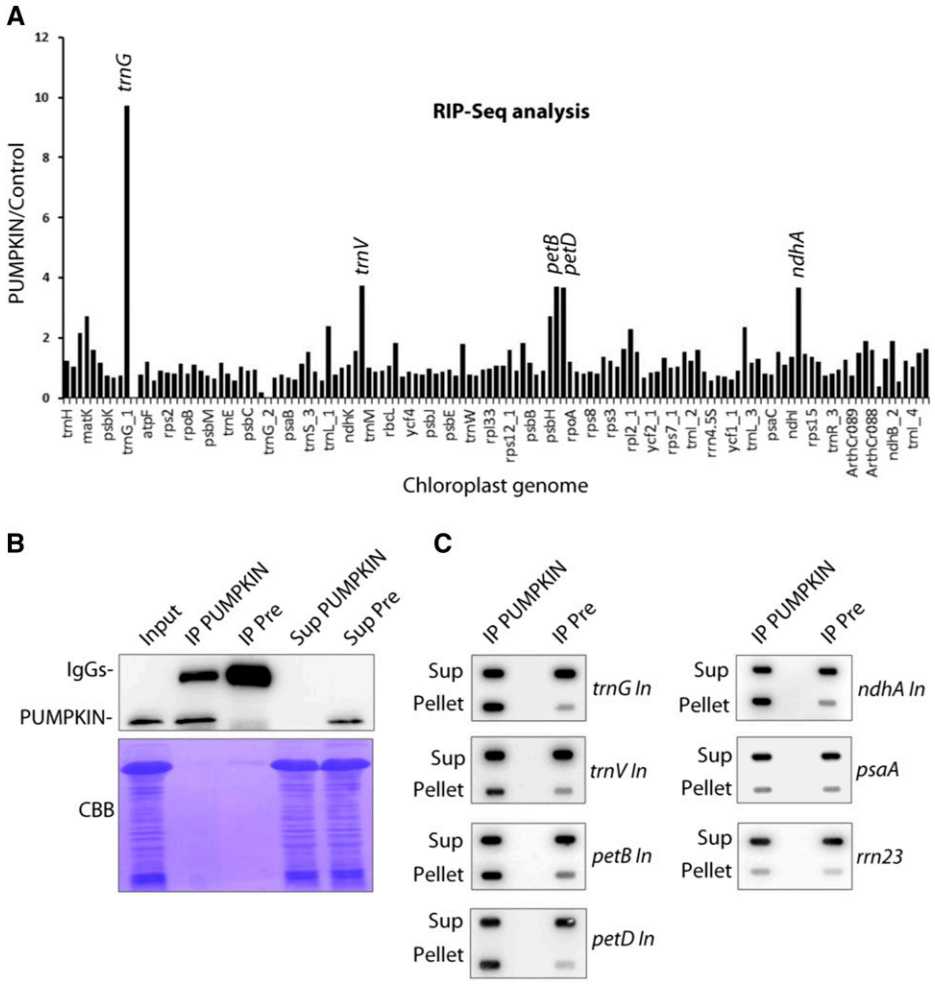


Figure 4. In Vivo Association of PUMPKIN with Chloroplast RNAs. A, RIP-Seq analysis of PUMPKIN. The RPKM (Reads Per Kilobase per Million) matrix normalizes reads for sequencing depth and gene length. The ratio of RPKM-normalized reads from PUMPKIN- and Control-IP (performed with the corresponding preimmune 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

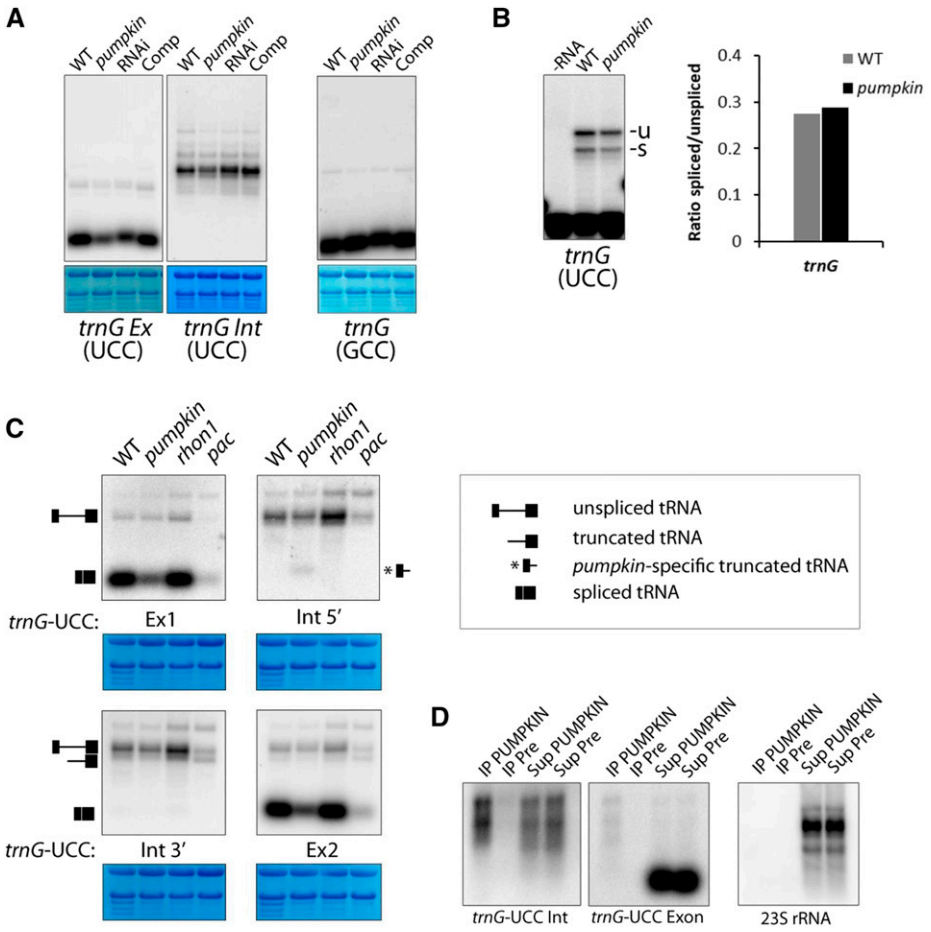


Figure 5. PUMPKIN Associates with the Intron of *trnG*-UCC and Influences Its Stability. A, RNA-gel–blot analysis of *trnG*-UCC and *trnG*-GCC in the wild-type, *pumpkin,* RNAi, and complemented (Comp) lines using 5 μg of leaf RNA. Ex, Exon; Int, Intron. Methylene blue staining is shown as loading controls (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

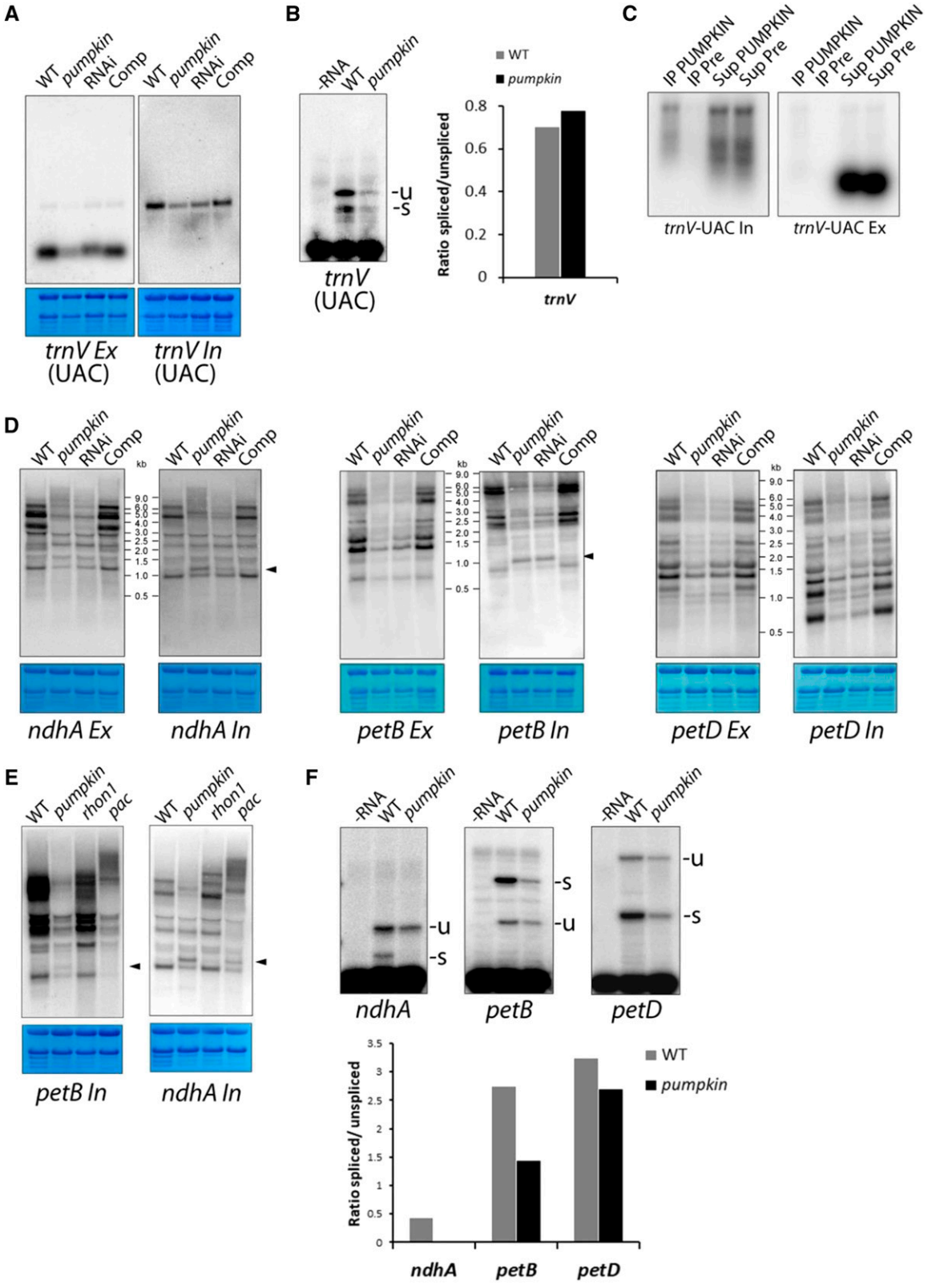


Figure 6. Analysis of the RNA Targets *trnV, ndhA, petB,* and *petD* in *pumpkin* Mutants. A, RNA-gel–blot analysis of *trnV*-UAC in the wild type, *pumpkin,* RNAi, and complemented (Comp) lines using 5 μg of leaf RNA. Ex, exon; In, intron. Methylene blue staining is shown as loading controls (lower 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.

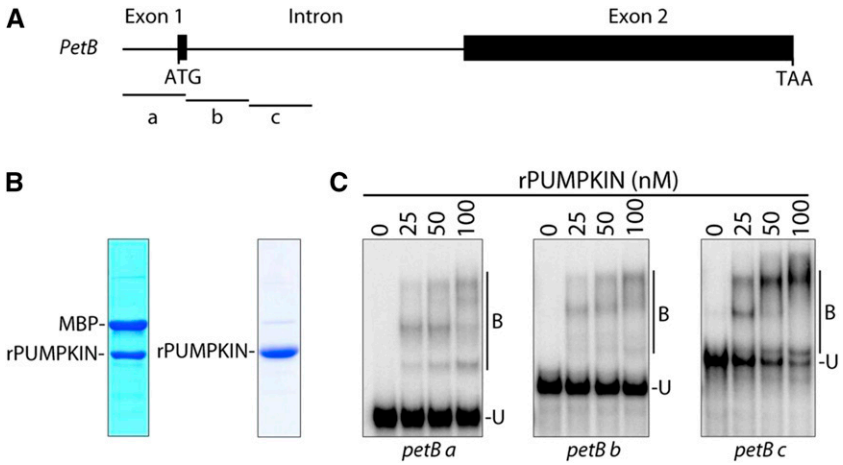


Figure 7. Recombinant PUMPKIN Binds to the *petB* Intron in Vitro. A, Schematic representation of the *petB* gene and probes (a, b, c) used in (B). ATG, start codon; TAA, stop codon. B, Recombinant MBP-PUMPKIN was purified and MBP was cleaved off (left panel). The right panel shows rPUMPKIN, obtained 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,

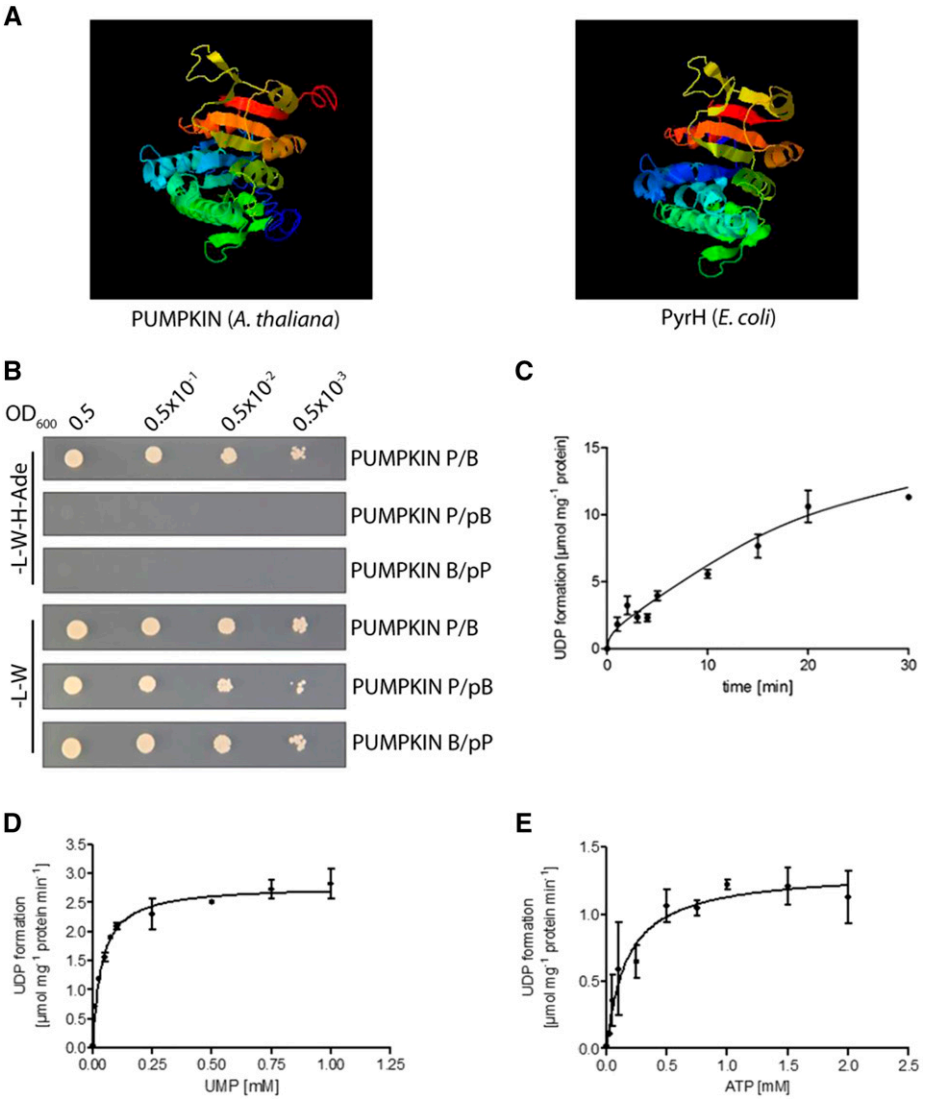


Figure 8. PUMPKIN Interacts with Itself and Exhibits Intrinsic UMP Kinase Activity in Vitro. A, Structure prediction of the UMP (Uridine monophosphate) kinases PUMPKIN (Arabidopsis) and PyrH (*E. coli*) by I-TASSER algorithm. B, Yeast-two-hybrid assay demonstrating that PUMPKIN interacts with itself. 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

7.2. Manuscript I: PUMPKIN and its Importance in the Plastid Pyrimidine Salvage Pathway

Schmid LM, Möhlmann T, John A, Vicente A, Grüner F, Marino G, Lehmann M, Manavski N, Leister D, Meurer J

Short title: PUMPKIN is an active plastid UMP kinase *in vivo*

Title: PUMPKIN and its Importance in the Plastid Pyrimidine Salvage Pathway

Lisa-Marie Schmid[a], Torsten Möhlmann[b], Annalisa John[b], Alexandre Vicente[a], Florian Grüner[a], Giada Marino[a], Martin Lehmann[a], Nikolay Manavski[a,c], Dario Leister[a], and Jörg Meurer[a,*]

[a]Plant Molecular Biology, 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]present address: Centre National de la Recherche Scientifique (CNRS), Institut de Biologie Moléculaire des Plantes, 12 rue du Général Zimmer, 67084 Strasbourg, France

ORCID IDs: 0000-0002-4834-2603 (L.S.); 0000-0002-5676-2042 (T.M.); 0000-0001-9142-675X (A.V.); 0000-0002-1525-4244 (G.M.); 0000-0003-2740-5991 (N.M.); 0000-0003-1897-8421 (D.L.); 0000-0003-2973-9514 (J.M.)

One-sentence summary

PUMPKIN functions as the sole plastid UMP kinase and clusters with up- and downstream enzymes of the pyrimidine salvage pathway

Footnotes:

List of author contributions

L.S., T.M., J.M., N.M. designed the research. L.S., T.M., A.J., A.V., F.G., G.M., and M.L. performed the research. L.S., T.M., D.L., and J.M. prepared the article with contributions from all co-authors. L.S. and J.M. supervised the whole study.

Responsibilities of the Author for Contact

It is the responsibility of the author for contact to ensure that all scientists who have contributed substantially to the conception, design or execution of the work described in the manuscript are included as authors, in accordance with the guidelines from the Committee on Publication

Ethics (COPE) (http://publicationethics.org/resources/guidelines). It is the responsibility of the author for contact also to ensure that all authors agree to the list of authors and the identified contributions of those authors.

***Corresponding author**

Jörg Meurer
Plant Sciences, Faculty of Biology, Ludwig-Maximilians-University Munich
Großhaderner Street 2-4, 82152, Planegg-Martinsried, Germany
Email address: meurer@bio.lmu.de; phone: 0049 +89 218074556

Funding information

This research was supported by Grants from the German Science Foundation (Deutsche Forschungsgemeinschaft; ME 1794/10 to J.M. and TRR 175 A05 to D.L. and A03 to J.M.)

Keywords: UMP kinase, nucleotide metabolism, *Arabidopsis thaliana*, chloroplast, RNA metabolism, PUMPKIN, PYRH, pyrimidine

Abstract

Nucleotides are fundamental compounds for each living cell, as they do not only function as building blocks for nucleic acids, but also serve to transfer energy for biochemical processes and represent substrates for many derivatives, which are active and important in numerous metabolic processes. Pyrimidine nucleotides can be either produced by *de novo* synthesis or by energy-saving salvage pathways. This work focuses on the plant uracil salvage pathway located in chloroplasts, involving several enzymes converting uracil into the final nucleotide product UTP. One major enzyme in this pathway is PUMPKIN, the plastid UMP kinase, which converts UMP to UDP. Beside its enzymatic activity, PUMPKIN was assigned an additional function in chloroplast RNA metabolism, as it associated with several intron-containing chloroplast transcripts. To further analyze which of the two functions is of major importance and if and how they influence each other, we dissected the original enzymatic function from the RNA-association capacity and vice versa, and studied the metabolic consequences of the defective plastid UMP kinase function or impaired RNA association. It appeared that loss of the enzymatic function mostly contributes to the *pumpkin* mutant phenotype. We further focused

on the interplay between the different chloroplast enzymes involved in the pyrimidine salvage pathway and provide first evidence for the existence of a chloroplast "pyrimidinosome".

INTRODUCTION

Nucleotides represent one of the most essential cellular components. They are the basic units for the synthesis of DNA and RNA, and thus are indispensable for gene expression in each living cell. Beside their function in cellular information storage, especially pyrimidine derivatives such as UDP-glucose, are the starting point for many other cellular compounds involved in multiple processes ranging from secondary metabolites for plant defense as well as cell-wall synthesis, starch and lipid metabolism (Kafer et al., 2004).

To obtain pyrimidine nucleotides, the plant cell can use the *de novo* synthesis pathway starting from the basic molecules glutamine, ATP and CO_2, which are finally converted into energy-rich nucleotides (Zrenner et al., 2006). As this represents an energy-consuming process, the plant evolved another route – the so-called salvage pathway – in order to re-use nucleotides by degrading and converting them back to uracil or uridine that can be transformed into UMP again (Mainguet et al., 2009). The two uridine-cytidine kinases UCK1 and UCK2 are of major importance for the uridine pyrimidine salvage pathway in the cytosol, but plants possess an additional route for pyrimidine salvage, which takes place in the chloroplast. This pathway starts with the import of uracil which is further converted to UTP (Witz et al., 2012; Ohler et al., 2019). This involves three key enzymes: the uracil phosphoribosyl transferase UPP converting uracil to UMP, the plastid UMP kinase PUMPKIN which phosphorylates UMP to yield UDP and the nucleoside-diphosphate kinase NDPK2 finally generating UTP (Ohler et al., 2019).

This research mainly focused on the importance of PUMPKIN in the pyrimidine salvage cascade. Besides its enzymatic activity, PUMPKIN was assigned to possess an additional function in plastid RNA metabolism. It specifically associated with intron-containing plastid transcripts and influenced their stability (Schmid et al., 2019). In this study, we extended our previous studies and focused on deciphering the roles of these two functions – the catalytic activity as an enzyme and the RNA-association capability – and their contribution to the mutant phenotype in more detail. To answer this question, we decided to use enzymatically impaired PUMPKIN versions or the *E. coli* homolog PYRH lacking the RNA-association capability and re-introduced the constructs into the *pumpkin* mutant background to analyze their rescue

competency. We additionally focused on the interplay of the three key enzymes in the plastid pyrimidine salvage pathway and provide first hints for the existence of a chloroplast "pyrimidinosome".

RESULTS

The Enzymatically Impaired PUMPKIN Is Unable to Complement the Mutant Phenotype

Previously, we have shown that PUMPKIN has UMP kinase activity and is associated with several plastid RNA introns (Schmid et al., 2019). It remained unclear to which extend the two functions contribute to the *pumpkin* mutant phenotype. In a first attempt we tried to complement the *pumpkin* mutant with enzymatically impaired PUMPKIN versions still retaining the RNA association. Thereto, two different point mutations were introduced to replace the conserved amino acids arginine at position 149 (R149) or aspartate at position 165 (D165) by alanine (A), respectively. In the *E. coli* UMP kinase, both amino acids were reported previously to be residues interacting with UMP, and thus directly influencing the enzymatic activity (Briozzo et al., 2005). The positions of the amino acids within the predicted three-dimensional structure of PUMPKIN are depicted in Figure 1A. As tested by *in vitro* studies, replacing aspartate 165 by alanine (D165A) reduced the enzymatic activity to about half of the wild-type (WT) PUMPKIN protein (Figure 1B). When exchanging the arginine at position 149 (R149) to alanine, the enzymatic activity was basically abolished, thus leading to a virtually inactive PUMPKIN enzyme. Constructs containing either the D165A or the R149A mutation were subsequently used to complement the Arabidopsis *pumpkin* mutant.

Pumpkin mutants which were complemented with PUMPKIN D165A, displayed varying phenotypes (Figure 1C). While D165A line #1 showed a partially complemented phenotype, but retained retarded development and pale leaves, the D165A#2 line behaved basically like the WT. These observations were also true concerning their photosynthetic performance (Figure 1C). Especially the newly developing leaves of line D165A#1 exhibited defects in photosynthesis as revealed by reduced F_v/F_m values, in contrast to D165A#2 plants which almost reached wild-type performance.

Mutants complemented with the enzymatically inactive PUMPKIN version R149A did not show any improvement of their phenotype nor their photosynthetic performance (Figure 1C). The plants remained small and pale-green and displayed massive defects in photosynthesis, similar to the *pumpkin* mutant. Thus, re-introducing the RNA-association capability of PUMPKIN alone did not lead to any visible rescue. Genotyping confirmed the presence of the

pumpkin mutant background in all the complemented lines (Figure 1D). To verify the expression of the two constructs, soluble protein extracts from WT, *pumpkin*, D165A#1, D165A#2 and R149A plants were subjected to immunoblot analyses with a PUMPKIN-specific antiserum (Figure 1E). As expected, the PUMPKIN protein was found in the WT but was absent in the *pumpkin* knock-out mutant. The PUMPKIN antibody was still able to detect the two mutated PUMPKIN versions D165A and R149A since they represent point mutations. A comparison revealed that R149A almost reached the expression level of the native protein in the WT. The expression in the D165A lines varied, which was in line with the observed phenotype. While D165A#1 showed a lower expression, when compared with the wild-type PUMPKIN abundance, D165A#2 plants overexpressed the D165A protein, which seemed to compensate for the reduced enzymatic activity. Taken together, we conclude that the observed *pumpkin* phenotype is rather caused by the lack of the enzymatic function of PUMPKIN as an UMP kinase.

PYRH – the *E. coli* Homolog of PUMPKIN – Largely Rescues the *pumpkin* Phenotype

In a separate approach we aimed to complement the *pumpkin* mutant with an enzymatically functional protein, which does not possess the RNA-association function. For this reason, we chose the *E. coli* homolog PYRH, which was shown to be a functional UMP kinase converting UMP to UDP (Serina et al., 1995) but which is lacking any known RNA-association capability. When comparing the enzymatic performance of PYRH and PUMPKIN by *in vitro* enzyme assays (Figure 2A), PYRH was found to be enzymatically even more active than the PUMPKIN protein (~25-fold) under the chosen conditions, thus making the *E. coli* protein a suitable UMP kinase for our complementation approach. To ensure correct localization, the corresponding gene was N- terminally fused with the chloroplast targeting peptide of RBCS and the construct was transferred into the *pumpkin* mutant background.

Figure 2B shows the phenotypes of three independently complemented PYRH-lines, displaying a gradually improving phenotype between *pumpkin* and wild-type plants. While PYRH lines #1 and #2 are still partially impaired in their growth and show pale-green leaves in the middle of the rosette, PYRH line #3 phenotypically resembled the WT.

The corresponding photosynthetic performance of all lines is represented by the measured F_v/F_m values captured by imaging chlorophyll fluorescence (Figure 2B). As expected, wild-type plants showed an F_v/F_m value of roughly 0.8. In contrast, *pumpkin* mutants were massively impaired in their photosynthetic performance. Similarly, the complemented PYRH lines #1 as well as #2 displayed defects in photosynthetic performance, especially in developing leaves.

Nevertheless, depending on the grade of complementation - and in accordance with the observed phenotype - photosynthetic performance was partially restored. The PYRH line #3 was even able to essentially reach wild-type levels in its photosynthetic as well as its developmental performance.

Genotyping confirmed that the *pumpkin* mutation was still present in all the lines (Figure 2C). We next analyzed the protein expression of PYRH in the different complemented PYRH plants using a PYRH-specific antibody. The PYRH protein was only found in the PYRH complemented lines (Figure 2D), also demonstrating the specificity of the antiserum. Interestingly, the protein expression levels of PYRH directly correlated with the observed plant phenotypes. While PYRH line #1 only displayed a weak expression of the PYRH protein and had a phenotype more resembling the *pumpkin* mutant, PYRH line #3 reached the highest expression levels of the PYRH protein, which led to an almost wild-type-like phenotype.

Next, we analyzed the capability of PYRH to also complement the molecular defects observed in the *pumpkin* mutant. Figure 3A shows RNA-gel blots probed with the known targets that are associated with PUMPKIN, using RNA derived from WT, *pumpkin*, PYRH#1, PYRH#2, and PYRH#3. As shown previously (Schmid et al., 2019), the *pumpkin* mutant displayed aberrant transcript patterns, when probed with the *petB* and *ndhA* intron as well as strongly decreased levels of *trnG*, *trnV* and *petD* precursors. In line with the gradual improvement of the phenotype in the different complemented PYRH lines, also the aberrant transcript patterns (*petB*, *ndhA*) as well as the reduction of RNA precursors (*petD, trnG, trnV*) could be gradually restored and were comparable to wild-type levels and patterns in the PYRH#3 line. As mentioned previously, *pumpkin* mutants showed a decrease in photosynthetic subunits to roughly 25% of wild-type levels (Schmid et al., 2019). When comparing the protein levels of several photosynthetic subunits as representatives for the different photosynthetic complexes in the PYRH complemented lines, a gradual increase in protein levels was observed, ranging from roughly 25-50% in PYRH line #1 to almost wild-type levels in PYRH line #3 (Figure 3B). Again, this molecular complementation is in line with the observed PYRH plant phenotypes.

PYRH Does not Associate with PUMPKIN RNA Targets *in vivo*

As prove of concept, we needed to re-assure that PYRH is indeed only enzymatically active but unable to associate with chloroplast intron RNAs – in contrast to PUMPKIN. For this purpose, we performed RNA-immunoprecipitation (RNA-IP) with PYRH-specific antibodies, using stroma extracts of PYRH complemented plants and WT as respective control (Figure 4A). The

PYRH protein was successfully pulled down in the PYRH-IP fraction but was absent in the WT, as expected, also proving the specificity of the PYRH antiserum. Subsequently, RNA was isolated from the supernatant and IP pellet fractions and subjected to slot blot hybridization with the indicated probes, representing the RNA-association targets of PUMPKIN (Figure 4B). None of the tested targets was found to be enriched in the PYRH-IP pellet compared to the WT-IP pellet. The 23S rRNA served as a control and was also not enriched in the PYRH-IP from PYRH containing plants. This clearly indicated that PYRH was unable to associate with RNA targets of PUMPKIN, and thus fulfills the criteria for a PUMPKIN version which only possesses the enzymatic activity. In our previous study we found PUMPKIN to be part of RNA-containing protein complexes ranging up to megadalton sizes (Schmid et al., 2019). Using size-exclusion chromatography of stroma extracts from PYRH complemented plants, we analyzed if PYRH can be found as part of these complexes *in vivo* as well (Figure 4C). To further investigate if these complexes may contain RNA, a second run using RNase treatment was performed. The obtained fractions were then subjected to immunoblot analyses with the indicated antibodies. Unlike the PUMPKIN protein, PYRH was predominantly found in fractions 10 to 22, which roughly corresponds to complex sizes between 75-158 kDa (fractions 18-20) and up to roughly RubisCO-sized 550 kDa (fractions 10-14) (Meurer et al., 2017; Schmid et al., 2019). These fractions were not susceptible to RNase treatment, again arguing that PYRH does not associate with chloroplast RNA *in vivo*. There is only one very weak and therefore negligible signal of PYRH in fraction 2, especially considering that only 50% of fractions 14 to 28 were loaded to avoid overloading of the gel. In strong contrast, PUMPKIN was found to be distributed over all high molecular weight fractions, which showed a clear shift when treated with RNase (Schmid et al., 2019). Expectedly, RNase treatment was successful, as proven by probing with antibodies for components of the large and small ribosomal subunit (RPL11 and RPS5), which showed a clear shift upon RNase application. PYRH was furthermore not found in its monomeric form (approx. 26 kDa), thus arguing that not only in bacteria but also *in planta* PYRH forms a hexameric structure (approx. 156 kDa), which is essential for its proper functionality (Briozzo et al., 2005). Together, these results prove that PYRH is neither associated with RNA-containing proteinaceous complexes nor with the known PUMPKIN RNA targets.

PUMPKIN Associates with Proteins Involved in Plastid RNA-Metabolism

Even though our results suggest that the lack of the enzymatic function is almost exclusively the cause for the *pumpkin* phenotype, we wanted to clarify why PUMPKIN was found to

specifically associate with a subset of intron-containing plastid transcripts. When checking the capability of the point mutated PUMPKIN versions R149A and D165A as well as the PYRH protein in binding to the previously determined binding site of PUMPKIN within the *petB* intron, we obtained binding to this region with all these proteins (Supplemental Figure S1A). As we excluded any *in vivo* association of PYRH with PUMPKIN RNA targets before (Figure 4), we further tested the specificity of this interaction by including another unrelated RNA-binding protein – PAC – in the electrophoretic mobility shift assay (EMSA). As shown in Supplemental Figure S1B, the *petB c* probe – previously identified to be shifted by the PUMPKIN protein – was as well shifted by the PAC protein, which was shown to not target *petB* by RIP-Seq analyses *in vivo* (Meurer et al., 2017). We therefore concluded that the previously observed binding to *petB c* is indeed rather unspecific. Thus, we suggest that PUMPKIN may interact with its RNA targets indirectly via other proteins and cannot be further considered to be a true RNA-binding protein. In order to identify these interaction partners, co-immunoprecipitation (Co-IP) coupled with mass spectrometry was performed. Table 1 lists all significant plastid-localized hits ($\log_2$ FC > 1 and p-value adj < 0.05) which were further checked for co-expression with PUMPKIN using the ATTEDII database (Obayashi et al., 2018). Intriguingly, we identified 33 proteins out of 70 significantly enriched plastid-localized hits ($\log_2$ FC > 1), which are involved in different steps of chloroplast RNA metabolism (highlighted in grey in Table 1). Beside classical RNA-binding proteins, such as PPR and TPR proteins (e.g. RAP, MRL1, SOT5), which fulfill various but mostly very specific tasks in posttranscriptional gene regulation (Barkan and Small, 2014), we also found CP29A, a family member of the cpRNPs, which act as general factors for RNA stabilization of large transcript pools (Kupsch et al., 2012). Moreover, we found several proteins involved in RNA processing and degradation (RNaseE and its interaction partner RHON1 as well as RNaseJ) (Stoppel et al., 2012; Halpert et al., 2019) or splicing of specific introns (RH3, ISE2 and SOT5) (Asakura et al., 2012; Carlotto et al., 2016; Huang et al., 2018). Another interesting co-immunoprecipitated candidate was PNPase. It was shown to be important for efficient 3' end maturation of chloroplast mRNAs and tRNA degradation (Walter et al., 2002). We further found two NusB domain-containing proteins (AT4G26370, AT3G13180), whose eubacterial relatives are known to regulate rRNA biosynthesis and antitermination (Greive et al., 2005). Additionally, we identified a number of proteins, assigned in transcript stability and processing or splicing of the *psbB* operon, which comprises the PUMPKIN RNA targets *petB* and *petD*: RNC1, HCF152, HCF107, WTF1, CAF2, CFM3A, CAF1 (Stoppel and Meurer, 2013). This finding could explain the observed association of PUMPKIN with *petB* and *petD* intron-containing

transcripts. Amongst the identified potential interaction partners, we discovered several components which are direct members of the pTAC family or proteins associated with "plastid transcriptionally active chromosomes", namely pTAC7 and pTAC16, as well as GYRA (DNA gyrase subunit A) (Pfalz et al., 2006). Additionally, we found CPSAR1, which is closely related to OBG (Spo0B-associated GTP-binding) proteins which were shown to be involved in organellar ribosome biogenesis (Bang et al., 2012). When comparing our hits with the lately analyzed mRNA-binding proteome in etiolated seedlings (Reichel et al., 2016) several more – yet functionally uncharacterized – proteins can be categorized into being "candidate RNA-binding proteins", such as PPR3, two RRM-motif containing proteins (AT1G01080, AT4G09040) and several ribosomal proteins (RPS7, RPS5, RPS1 and RPS11). As expected, we also found some matches of already characterized proteins from our IP-hits in the proteome data from Reichel et. al (e.g. RHON1, CP29A), thus re-enforcing their connection to the chloroplast RNA metabolism. Due to the numerous identified RNA-binding proteins which we captured in our IP approach, we readjust our previous assumption of PUMPKIN representing a true RNA-binding protein but propose that it associates with plastid introns in concert with other RNA-binding factors.

PUMPKIN Co-Localizes with the Pyrimidine Pathway Enzymes UPP and NDPK2

Our previous results revealed a spotty localization pattern of PUMPKIN in chloroplasts, which could only be partially assigned to a potential nucleoid localization (Schmid et al., 2019). Interestingly, also UPP, the upstream enzyme in the plastid pyrimidine salvage pathway displayed this spotty pattern (Ohler et al., 2019), which prompted us to investigate if enzymes of the pyrimidine salvage pathway can cluster together and thus co-localize. In a transient triple co-transformation approach into tobacco leaf tissue, using different fluorescent tags for the three enzymes (PUMPKIN-CFP, UPP-YFP and NDPK2-RFP), we analyzed their cellular localization and clustering (Figure 5A). All three tagged proteins specifically localized to chloroplasts, where we observed again the punctuated distribution, which perfectly overlapped. Additionally, co-transformation experiments using PUMPKIN-CFP and UPP-YFP (Figure 5B) or PUMPKIN-CFP and NDPK2-RFP were performed, again displaying this spotty signal within chloroplasts, which accurately overlapped. To conclude, UPP, PUMPKIN and NDPK2 – the three plastid enzymes in the pyrimidine salvage pathway – co-localized to form clusters. To further analyze the ability of these enzymes to interact with each other, immunoprecipitation (IP) experiments were performed using PUMPKIN-specific antibodies to pull-down the native PUMPKIN protein from wild-type stroma extracts and to analyze if UPP co-purifies. With

standard IP conditions as well as with an additional crosslinking-step, PUMPKIN was successfully pulled-down and detected in the IP fraction (Supplemental Figure S2) but UPP remained in the supernatant. In our previous IP-approach coupled with LC/MS we were also unable to detect UPP or NDPK2 pointing to a rather weak and/or reversible interaction between the three proteins.

DISCUSSION

An Impaired Nucleotide Metabolism Is Mainly Responsible for the *pumpkin* Phenotype

Using the enzymatically active *E. coli* homolog PYRH which is not associated with plastid RNAs when expressed in Arabidopsis, led to a clear complementation phenotype, depending on the PYRH protein level and thus the resulting recovery of nucleotide salvage in the chloroplast. PYRH was not only able to complement on the developmental level but could also restore the observed molecular defects, including the aberrant and reduced transcripts in the *pumpkin* mutant background. This clearly indicates that these observed defects are a direct consequence of impaired and thus imbalanced pyrimidine biosynthesis. Accordingly, expression of the enzymatically inactive PUMPKIN R149A version, could not restore the mutant phenotype. Interestingly, also mutants defective in the purine synthesis pathway displayed pale green to white-striped phenotypes accompanied by severe developmental defects. Mutations in ATase2 – the initial enzyme in purine synthesis – showed decreased levels of ATP and GTP which led to small, albinotic plants (Hung et al., 2004). Similar to PUMPKIN, a rice mutant impaired in GARS – the second important plastid-localized enzyme in purine biosynthesis, showed a decrease of 23S and 16S rRNA as well as effects on the transcript levels of plastid-encoded genes resulting in white-striped plants (Cao et al., 2019). Thus, an imbalance or shortage in nucleotide metabolism has far reaching consequences for the whole metabolism and development, which is a widespread phenomenon. In some bacteria, such as the gram-negative *E. coli* or the gram-positive *Streptococcus pneumoniae*, knock-out of the PUMPKIN homologs even resulted in lethality (Fassy et al., 2004). Also in humans, disturbance in the pyrimidine metabolism is linked to several diseases. Mutations in the DHODH (dihydroorotate dehydrogenase) enzyme, which is part of *de novo* pyrimidine synthesis, induced the Miller syndrome, causing severe craniofacial, limb, and eye malformations (Duley et al., 2016). Mutations in another enzyme of pyrimidine *de novo* synthesis – UMP synthase – led to orotic aciduria which was accompanied by retardation in growth and intellectual development (Suchi et al., 1997). In plants, knockdown of pyrimidine *de novo* synthesis enzymes led to reduced

growth rates (Zrenner et al., 2006; Chen and Slocum, 2008). Nevertheless, plants seem to be able to partly compensate a downregulation in *de novo* synthesis by using the uridine-salvaging pathway (Geigenberger et al., 2005). In summary, this underlines the importance of nucleotide biosynthesis, balance and the far-reaching consequences on numerous metabolic processes in case of dysfunction.

Still it remains challenging to pinpoint how exactly a shortage or imbalance of nucleotides leads to the observed molecular defects in the *pumpkin* mutant. On the one hand, nucleotides are essential as building units for nucleic acids, thus, it is reasonable to expect direct consequences on gene expression. In addition, nucleotides and their precursors are also major components for the synthesis of many derivatives such as UDP-glucose. Accordingly, also many other metabolic processes might be largely impaired in *pumpkin*. For instance, double knockouts of the two genes coding for UDP-glucose pyrophosphorylase (UGPase), which catalyzes the reaction of glucose 1-phosphate (Glc 1-P) and UTP to UDP-glucose, exhibited growth defects and male sterility (Park et al., 2010). To further gain insight if indeed shortage of a certain intermediate or derivative of the pyrimidine salvage pathway is responsible for the observed defects, rescue experiments of the *pumpkin* mutant on MS media supplemented with the downstream products of the pyrimidine salvage pathway - UDP, UTP and UDP-glucose - were performed. As shown in Supplemental Figure S3, feeding the plants with UDP-glucose could indeed improve the plant growth and photosynthetic performance. As a control plants were also fed with glucose only, which did not lead to any effect. None of the supplements showed any effects on wild-type plants. Supplementing the media with UDP or UTP could barely improve the *pumpkin* phenotype. Nevertheless, it remains challenging to assess whether the plants are capable to utilize UDP or UTP from the media and transport appropriately. Furthermore, they might be quickly metabolized (Wormit et al., 2004). Recent reports argue for transport routes for dNTPs, as the medium supplemented with 2 mM of dCTP was capable to suppress the *ven4* mutant phenotype (Xu et al., 2020). In contrast, UDP-glucose is a commonly used nucleotide-sugar transport form, hence the plant might be able to uptake and indeed distribute UDP-glucose wherever needed before it gets metabolized (Norambuena et al., 2002; Seino et al., 2010). Interestingly, an overall slight improvement in the photosynthetic performance of *pumpkin* mutants at later developmental stages was shown (Supplemental Figure S3B), consistent with the previously observed lower expression and need of the PUMPKIN protein in later developmental stages (Schmid et al., 2019). This slight improvement may indicate that as soon as a certain nucleotide pool has been established and several energy and nucleotide-consuming

processes might be slowed down compared to earlier plant developmental stages, the plant can cope better with nucleotide shortage.

RNA Association of PUMPKIN Is Mediated via RNA-Binding Protein Interaction Partners

In vitro RNA-binding studies previously suggested that PUMPKIN might directly interact with its RNA targets (Schmid et al., 2019). When performing similar RNA-binding studies with additional proteins, including the PYRH protein as well as PAC (PALE CRESS) – another unrelated RNA-binding protein – we observed comparable binding of the previously identified binding site within the *petB* intron to all these proteins (Supplemental Figure S1). As we could clearly prove that PYRH does not associate with *petB* or other PUMPKIN RNA targets *in vivo* (Figure 4) and RIP-Seq analyses of PAC excluded *petB* as one of its targets (Meurer et al., 2017), we concluded that the observed binding is rather unspecific. How this unspecific interaction is mediated and why this certain probe does somehow stick to proteins remains unclear. One might speculate that certain stretches of this probe can interact with protein surfaces via ionic interactions. It further needs to be mentioned that we were unable to prove *in vitro* binding of PUMPKIN to other identified targets (*trnG* and *ndhA*) which led to the hypothesis that the RNA association of PUMPKIN is rather accomplished in concert with additional RNA-binding factors.

When searching for protein interaction partners via Co-IP coupled with mass spectrometry, numerous plastid-localized RNA-binding proteins were identified among the top candidates. In addition, co-expression analysis using ATTEDII revealed a close relation between PUMPKIN and other genes involved in the plastid RNA metabolism, such as pTACs, PPRs and helicases, amongst them were also several candidates identified by IP/MS. Thus, PUMPKIN seems to fulfill an additional role in chloroplast RNA metabolism. In contrast to its homolog PYRH, PUMPKIN was found in RNA-containing megadalton protein complexes *in vivo* as shown by previous size-exclusion chromatography experiments (Schmid et al., 2019). Proteinaceous assemblies of this size distribution, ranging from different ribosomal assembly stages up to nucleoid-containing fractions are typical for proteins involved in chloroplast RNA metabolism (Beick et al., 2008; Stoppel et al., 2011; Meurer et al., 2017; Paieri et al., 2018). One could also speculate that PUMPKIN – as part of the salvage pathway - makes use of RNA-degradation products as we found several proteins involved in RNA decay such as RNaseE, RNaseJ and PNPase in our IP approach. Bypassing the complete route of the salvage pathway by directly

using the mononucleotides which arise through RNA decay could represent an effective way of saving and converting energy (Andrade et al., 2009).

In addition, the observed molecular defects in the *pumpkin* mutant appear to be specific. Unlike *upp* mutants, which are impaired in the upstream enzymatic reaction in the pyrimidine salvage pathway, *pumpkin* mutants displayed a reduction in plastid rRNA levels and were impaired in plastid translation (Schmid et al., 2019). Furthermore RNA-Seq analyses of *upp* did not identify these dramatic downregulation of PEP- and upregulation of NEP-dependent plastid genes in contrast to *pumpkin* (Ohler et al., 2019; Schmid et al., 2019). Altogether, the close relation to chloroplast RNA-binding proteins as well as the specificity of the defects in comparison with the *upp* mutant, argues for additional functions of PUMPKIN in chloroplast RNA metabolism. Under which conditions this potential moonlighting function in RNA metabolism becomes crucial for the plant remains to be investigated.

Plastid Enzymes of the Pyrimidine Salvage Pathway Cluster Together

Several plastid enzymes involved in pyrimidine and purine biosynthesis display a spotty signal pattern within the chloroplast when examining their localization by fluorescence microscopy (Bölter et al., 2007; Yang et al., 2015; Ohler et al., 2019). In some cases, this punctuated signal could be attributed to a nucleoid-specific localization (Bölter et al., 2007; Yang et al., 2015). Nevertheless, in the case of PUMPKIN, the observed spots did not completely overlap with the signals of co-transformed nucleoid controls (Schmid et al., 2019). A partial distribution into nucleoids is still reasonable as we found several nucleoid-localized pTAC and pTAC-associated proteins in our IP-approach and the downstream enzyme NDPK2 was previously co-localized with nucleoids as well (Bölter et al., 2007). However, when comparing the pattern with localization studies of purine biosynthetic enzymes in mammalian cells, another hypothesis arose. In recent years, numerous papers provided evidence that enzymes involved in *de novo* purine synthesis can cluster together to form the so called “purinosome” (An et al., 2008; Baresova et al., 2018). The formation of this complex allows substrate channeling leading to increased reaction rates as well as the protection of susceptible intermediate products such as ammonia (Chitrakar et al., 2017). Furthermore, the establishment of this protein assembly provides another layer of regulation of purine metabolism, in addition to allosteric feedback inhibition and thus allows fine-tuning of this biosynthetic pathway (Pedley and Benkovic, 2017).

In case of pyrimidine biosynthesis, final evidence for a "pyrimidinosome" is still lacking. Nevertheless, co-precipitation of the first three enzymes in *de novo* pyrimidine synthesis was shown in *Trypanosoma cruzi*, speaking in favor that also enzymes required for pyrimidine biosynthesis can cluster together (Nara et al., 2012). Thus, it seems reasonable to assume that also enzymes of the salvage pathways can assemble to form such a metabolon. Regarding pyrimidine *de novo* synthesis, substrate channeling seems to be favorable, as in some species fusion proteins such as the animal CAD evolved, which contains the enzymatic domains of the first three enzymes in pyrimidine *de novo* synthesis (Nara et al., 2000). We could now provide first evidence that UPP, PUMPKIN, and NDPK2 – the plastid localized enzymes of the pyrimidine salvage pathway – cluster together and co-localize as shown by co-transformation of tagged versions and fluorescence microscopy. Still, this interaction seems to be rather transient, as neither UPP nor NDPK2 could be found via PUMPKIN Co-IP approaches either followed by immunoblot analyses or by mass spectrometry. As shown in the mammalian system, the clustering of purine synthesis enzymes is a reversible interaction, which is only visible under certain conditions (Baresova et al., 2018). Thus, it remains questionable if "pyrimidinosomes" were present in the stroma used for our pull-down experiments.

MATERIALS AND METHODS

Plant Material and Growth Conditions

Arabidopsis plants were grown on soil under the following conditions: 16-h-light (20°C)/8-h-dark (18°C), 100 µmol photons m^{-2} s^{-1}. The *pumpkin* knock-out mutant was obtained from the GabiKat collection (D02154) (Kleinboelting et al., 2012). *N. benthamiana* plants were grown on soil under short day conditions (120 µmol photons m^{-2} s^{-1} in a 10 h light/14 h dark regime at 20°C).

Generation of Plant Lines

The complete coding sequence of PUMPKIN was amplified using primers listed in Supplemental Table S1 (PUMPKIN_fwd/rev) and inserted into pBluescript (Stratagene, Heidelberg). Mutations R149A and D165A were introduced by site directed mutagenesis with corresponding primers given in Table S1 using the Quickchange II Kit (Agilent Technologies, Santa Clara, USA). Primers containing Gateway attB sites were used to amplify the point mutated PUMPKIN versions from pBluescript vectors. The amplified products were inserted

into pDONR and finally into pUB-Dest (Grefen et al., 2010). Constructs were transformed into *Agrobacterium tumefaciens* strain GV3101 (Koncz et al., 1994). Floral inoculation was used to transform *pumpkin* floral buds (Narusaka et al., 2010). PYRH was amplified from *E. coli* genomic DNA with primers listed in Table S1 and the resulting product was translationally fused at the N-terminus with the chloroplast target sequence of RBCS (first 80 AA, primers Table S1). Subsequent cloning and transformation steps were carried out as described above.

PAM Measurements

The photosynthetic performance was measured using an Imaging PAM chlorophyll fluorimeter (Imaging PAM, M-Series; Walz, Effeltrich, Germany). The presented maximum PSII quantum yield was determined as the ratio of F_v/F_m (F_v= variable fluorescence, F_m= maximal fluorescence) and is displayed in false colors. Prior to the measurement plants were incubated in the dark for 30 min. The parameters used for the measurement were specified previously (Pulido et al., 2018).

Protein Extraction and Immunoblot Analyses

The PYRH antibody was raised against the synthetic peptides of amino acids 167-181 and 227-241 and used in 1:1,000 dilutions (Eurogentec, Cologne, Germany). The PUMPKIN antiserum was described recently (Schmid et al., 2019). RPL11 and RPS5 sera were used in previous studies (Meurer et al., 2017). The NdhH antibody was kindly supplied by Tsuyoshi Endo (Kyoto University). All other antibodies (CP43, D1, PsaA, PsaF, Cytb_6, AtpC) were purchased from Agrisera (Umeå, Sweden).

RNA-immunoprecipitation and Slot Blot Analysis

RNA-immunoprecipitation and slot blot hybridization were performed as stated recently (Meurer et al., 2017) using the PYRH-complemented line and Col-0 wild-type plants as control and PYRH-specific antibodies. SiMAG-Protein G beads were replaced by SureBeads™ Protein A Magnetic Beads (BioRad, Feldkirchen, Germany). Primers used to generate the indicated probes are listed in Table S1.

RNA Extraction and RNA Gel Blot Analyses

Leaf RNA extraction and blotting, probe labeling and hybridization was performed essentially as described previously (Manavski et al. 2015). Probes used are listed in Table S1.

In Vitro Enzyme Kinetics Measurements

Purified, recombinant PYRH, PUMPKIN and corresponding mutant variants were tested for UMP kinase activity in a coupled photometric assay as described in Blondin et al., (1994) with modifications in (Schmid et al., 2019).

Size-Exclusion Chromatography, Co-Immunoprecipitation, and Mass Spectrometry

Size-exclusion chromatography was performed as described recently (Schmid et al., 2019). Briefly, 3 mg of isolated stroma from PYRH-complemented plants was fractionated on a Superose 6 10/300 GL column using buffer A as described previously (Olinares et al., 2010). For RNase treatment 300 µg of RNase A (QIAGEN) were added and the stroma was incubated on ice for 30 min before the run. After a void volume of 2.5 mL, fractions were collected and precipitated using 100% w/v trichloroacetic acid. Fractions were loaded on a 15% denaturing polyacrylamide gel (from fraction 14 to 28 only half was loaded to avoid overloading), blotted, and decorated with the indicated antibodies.

Co-immunoprecipitation of proteins was performed similarly as described for RNA-co-immunoprecipitation. For the additional cross-linking step for the PUMPKIN-IP shown in Figure S2, DPS (Dithiobis(succinimidyl propionate)) was added to the chloroplast lysis buffer to a final concentration of 2.5 mM. IP fractions were either subjected to SDS-PAGE or were eluted in 0.1 M glycine, pH 2.5 and subsequently prepared for mass spectrometry (Smaczniak et al., 2012). Briefly, after reduction (2.25 mM DTT, 30 min at 37°C) and alkylation (5 mM iodoacetamide, 30 min at room temperature in the dark), the samples were digested overnight at 37°C with 1.5 µg trypsin (Pierce Trypsin Protease, MS Grade). The resulting peptides were subsequently desalted with home-made C18 stage tips (Rappsilber et al., 2003) and vacuum dried to near dryness before storing at -80°C. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) was performed on a nano-LC system (Ultimate 3000 RSLC, ThermoFisher Scientific, Waltham, MA, USA) coupled to an Impact II high-resolution Q-TOF (Bruker Daltonics, Bremen, Germany) via a CaptiveSpray nano ESI source (Bruker Daltonics). Peptides were loaded on a 2-cm nano-trap column (Acclaim Pepmap, C18, 100 Å, 100-µm inner diameter, Thermo Fisher Scientific) and separated on a 50-cm analytical column (Acclaim Pepmap RSLC, C18, 100 Å, 75-µm inner diameter, Thermo Fisher Scientific). Chromatographic separation was carried out in a 50°C column oven using a 30-min linear gradient of 5-45% acetonitrile at a flow rate of 250 nl per min. MS1 spectra were acquired at 3

Hz with a mass range m/z 200–2000, and with the 18 most intense peaks being selected for fragmentation. The dynamic exclusion duration was set to 0.5 min.

MS raw files were processed using MaxQuant software (version 1.6.3.4 (Cox and Mann, 2008)) and peak lists were searched against the *Arabidopsis thaliana* UniProt database (version April 2019) using the built-in Andromeda search engine (Cox et al., 2011) with the default settings. False discovery rate (FDR) was set to 1% for peptides and proteins. The label-free quantification (LFQ) algorithm (Cox et al., 2014) with default settings was used to quantify proteins across samples. Downstream bioinformatic and statistical analysis were performed in Perseus (version 1.6.2.3 (Tyanova et al., 2016)). After excluding potential contaminants, reverse hits, and proteins identified only by site modification, protein groups were retained if they had been quantified in all 3 replicates in one genotype. Protein LFQ intensities were $\log_2$-transformed and missing values imputed from a normal distribution. Significantly enriched proteins were determined by t-test statistics and p-values adjusted for multiple comparisons according to the Benjamini-Hochberg approach (Benjamini and Hochberg, 1995).

Co-Expression and Localization Prediction Analysis

To identify genes which display co-expression with PUMPKIN, the ATTEDII database "CoExSearch" tool was used (http://atted.jp/). The localizations of the potential PUMPKIN protein interaction partners were predicted according to the SUBA4 database (Hooper et al., 2017).

Protein Overexpression, Purification, and Electrophoretic Mobility Shift Assay

Coding sequences (lacking chloroplast targeting sequences) of the indicated genes (PUMPKIN/ R149A/ D165A and PYRH) were cloned into pMAL-TEV using BamHI/SalI restriction sites. For the point mutated PUMPKIN versions (D165A and R149A) the previously described complementation clones were used as template. Constructs were overexpressed and affinity purified as stated previously (Schmid et al., 2019). To obtain MBP as control, purified MBP-PUMPKIN was digested with TEV protease (Invitrogen, Waltham, USA) according to the manufacturer's protocol followed by size-exclusion chromatography on a Superdex 200 10/300 Column in running buffer containing (25 mM HEPES-KOH, 10 mM NaCl, 10 mM EDTA, pH 8.0). Appropriate fractions were collected and concentrated on Amicon filter units (Amicon Ultracel 3K, Schondorf, Germany). The MBP-PAC clone, as well as the 23S rRNA probe were described recently (Meurer et al., 2017). EMSA reactions were composed as specified before using the previously described *petB b* and *c* RNA probes (Schmid et al., 2019).

Localization Experiments

To obtain a C-terminal CFP fusion construct, the coding sequence of PUMPKIN was transferred into the Gateway vector pB7CWG2 via LR-reaction using the previously described pENTR^{TM}/SD/D_TOPO®-PUMPKIN clone (Schmid et al., 2019). The C-terminal UPP-YFP construct was generated recently (Ohler et al., 2019). For localization analysis with NDPK2, the coding sequence (At5g63310) was cloned into pENTR™/SD/D-TOPO® using primers Fw_cacc_NDPK2 and Rev_NDPK2_o.stop and transferred into pK7RWG2 via LR reaction to obtain a C-terminal RFP-fusion construct.

Transient transformation into *N. benthamiana* leaves was performed as described (Schmid et al., 2019). Fluorescence signals were detected with a Leica TCS SP5II microscope through an HCX PL APO 63x1.2W water immersion objective. For detection of fluorophores, the following settings were used: CFP: 465-507 nm; YFP: 520-555 nm; RFP: 592-620 nm; Chlorophyll: 721-774 nm at 418 and 514 nm excitation wavelength.

Accession Numbers

PUMPKIN (At3g18680/ NCBI Gene ID: 821399), UPP (At3g53900/ NCBI Gene ID: 824557), NDPK2 (At5g63310/ NCBI Gene ID: 836451), PYRH in *E. coli* (NCBI Gene ID: 944989)

Supplemental Data

Supplemental Figure S1. Electrophoretic Mobility Shift Assays (EMSAs) with Recombinant PUMPKIN, R149A, D165A, PYRH and PAC Protein.

Supplemental Figure S2: Co-immunoprecipitation of PUMPKIN.

Supplemental Figure S3. Rescue Experiment of Wild-type and *pumpkin* Plants on Media Supplemented with UDP, UTP, UDP-glucose and Glucose.

Supplemental Table S1. List of Oligonucleotides.

Acknowledgements

We wish to thank Alice Barkan for providing the pMAL-TEV vector. We thank Tsuyoshi Endo for the NdhH antiserum.

Figures and Figure Legends

Figure 1

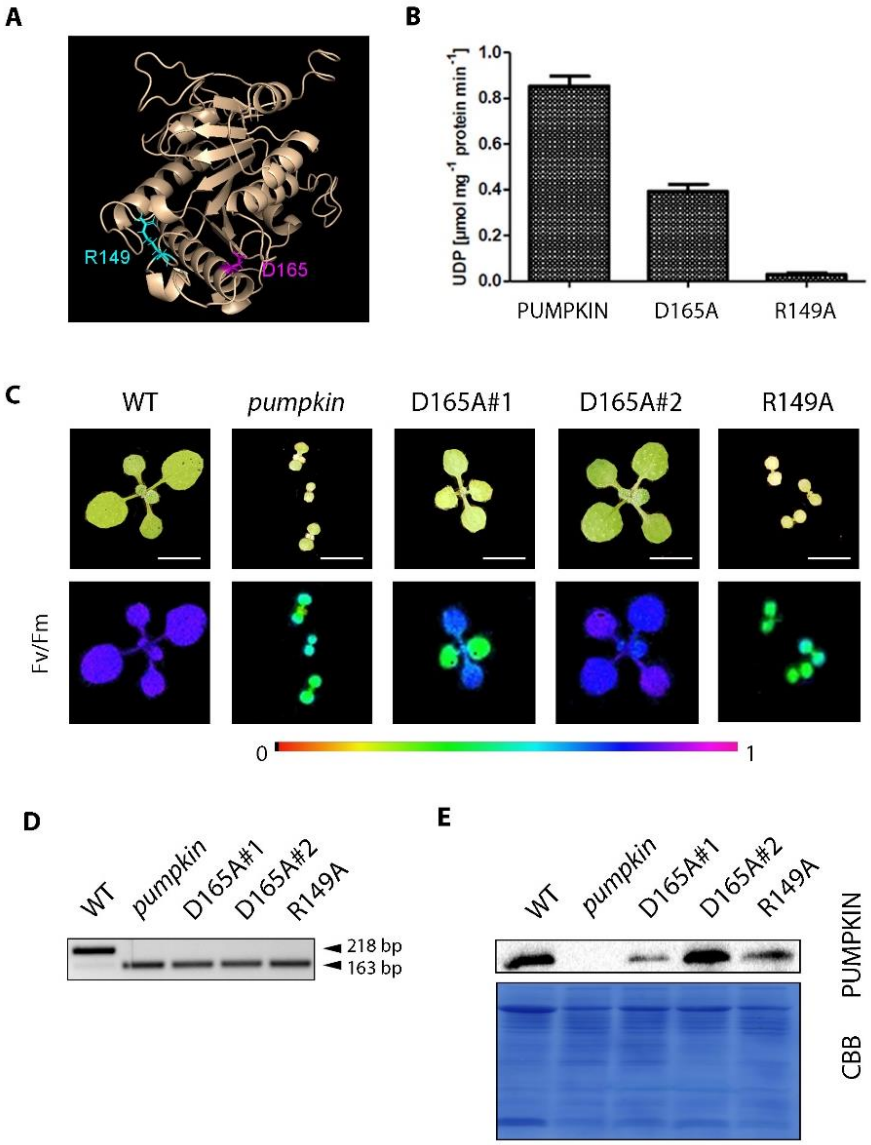


Figure 1. Complementation of *pumpkin* Mutants with Enzymatically Impaired PUMPKIN Versions. A) Three-dimensional structural prediction of the PUMPKIN protein lacking the chloroplast targeting sequence, generated by iTASSER (Yang et al., 2015). The two mutated residues were colored using PyMol (R149 in cyan, D165 in magenta). B) *In vitro* enzymatic assay measuring UDP-formation of recombinant PUMPKIN, D165A and R149A. The activity was determined in a coupled photometric test with UMP (0.03 mM), ATP (0.2 mM) and GTP (0.5 mM). Data represent the mean of three replicates. C) Phenotype of wild type (WT), *pumpkin* mutant and *pumpkin* lines complemented with point mutated versions of PUMPKIN, (D165A and R149A respectively) changing either D (aspartate) at position 165 to A (alanine) or R (arginine) at position 149 to A (alanine) to obtain PUMPKIN versions with impaired enzymatic activity (upper lane). Photosynthetic performance (F_v/F_m) of WT, *pumpkin* and D165A or R149A complemented mutants measured by the Imaging PAM. Color range indicates the measured values (lower lane). D) Genotyping of respective lines with three-primer PCR yielding a 218 bp product in the WT and a 163 bp product representing the mutant allele. E) Immunodetection of PUMPKIN in WT, *pumpkin* and complemented lines D165A#1 and #2 and R149A to prove successful expression of the complementation products. CBB = Coomassie Brilliant Blue.

Figure 2

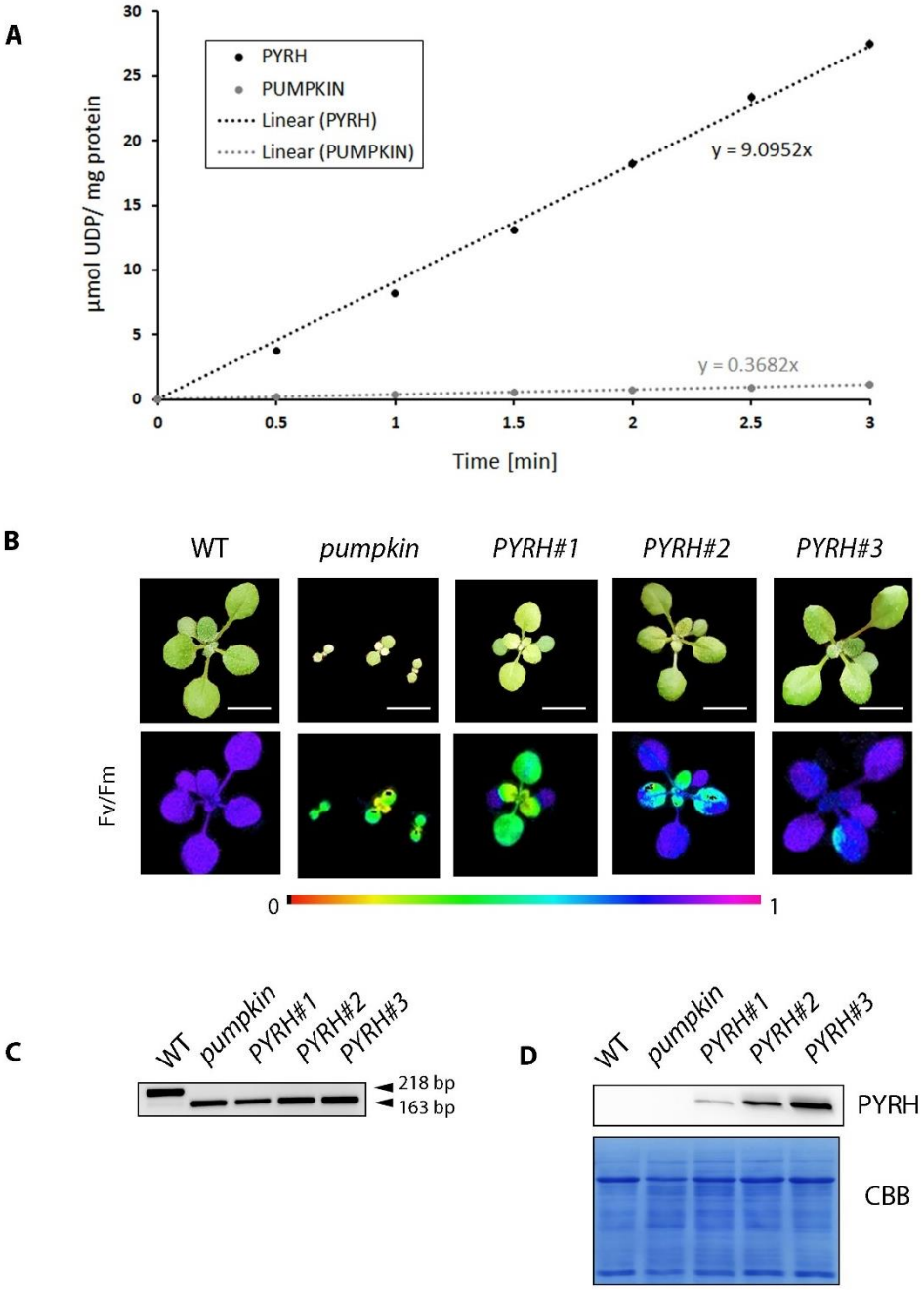


Figure 2. Complementation of *pumpkin* Mutants with the *E. coli* Homolog PYRH. A) *In vitro* enzymatic assay measuring time-dependent UDP-formation of recombinant PUMPKIN and PYRH. The substrates UMP and ATP were kept at 1 mM, GTP was provided at 0.5 mM. Data represent the mean of 7 independent replicates. Standard deviations are included but are too small to recognize. B) Phenotype of wild type (WT), *pumpkin* mutant and independent *pumpkin* lines complemented with PYRH (upper lane). Photosynthetic performance (F_v/F_m) of WT, *pumpkin* and PYRH complemented mutants measured by Imaging PAM (lower lane). C) Genotyping of respective lines with three-primer PCR yielding a 218 bp product in the WT and a 163 bp product representing the mutant allele. D) Immunodetection of PYRH in WT, *pumpkin*, and PYRH complemented lines to prove successful expression of PYRH. CBB = Coomassie Brilliant Blue.

Figure 3

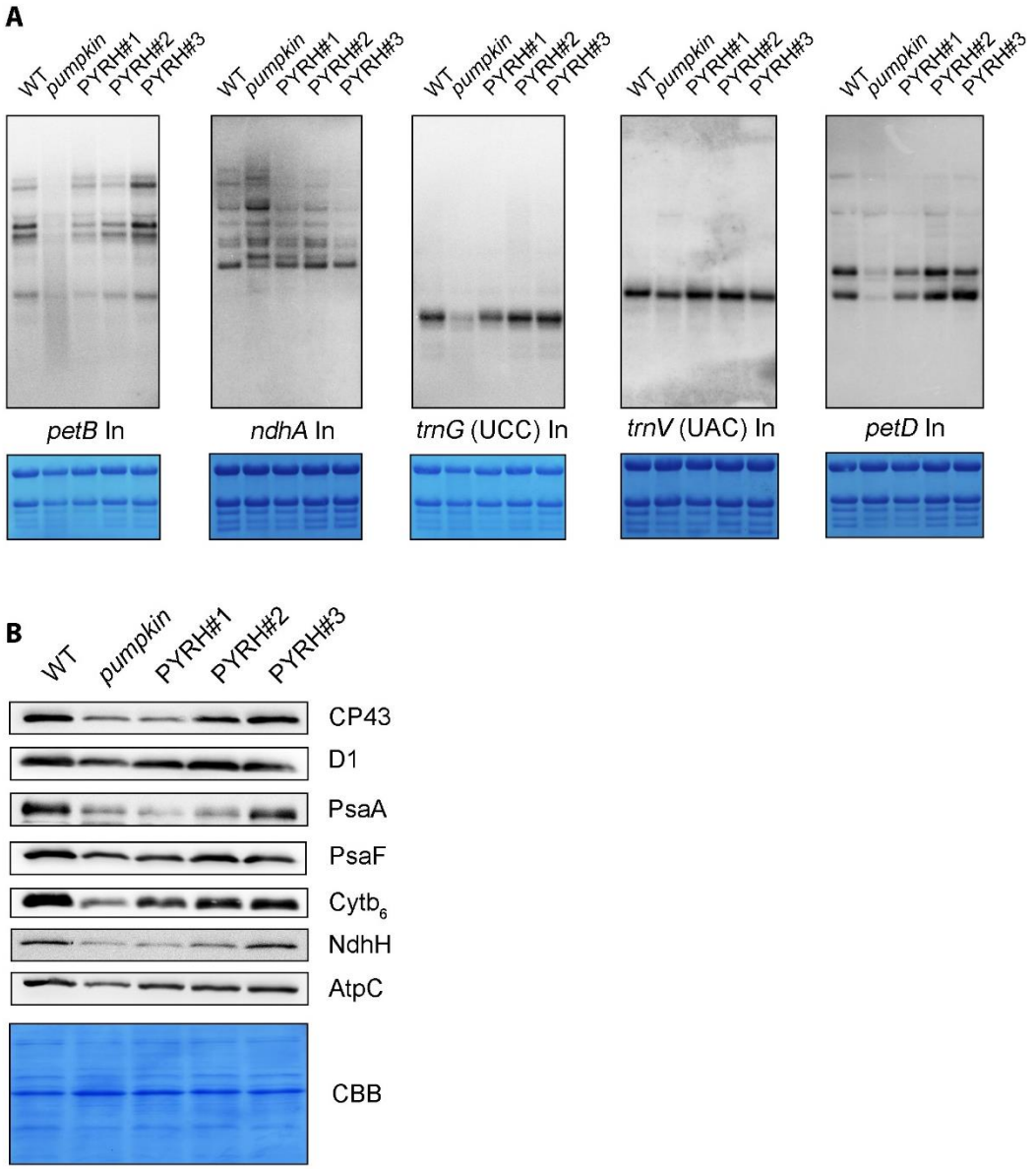


Figure 3. Molecular Analysis of *pumpkin* Mutants Complemented with the *E. coli* Homolog PYRH. A) RNA-gel blot analyses of known PUMPKIN RNA targets using WT, *pumpkin*, and three independent PYRH complemented *pumpkin* lines. A total of 5µg RNA was used, the panel below shows Methylene Blue staining as loading control. B) Exemplary immunoblot analyses of extracted thylakoids (corresponding to 5 µg chlorophyll) from WT, *pumpkin*, and PYRH complemented lines using the indicated antibodies.

Figure 4

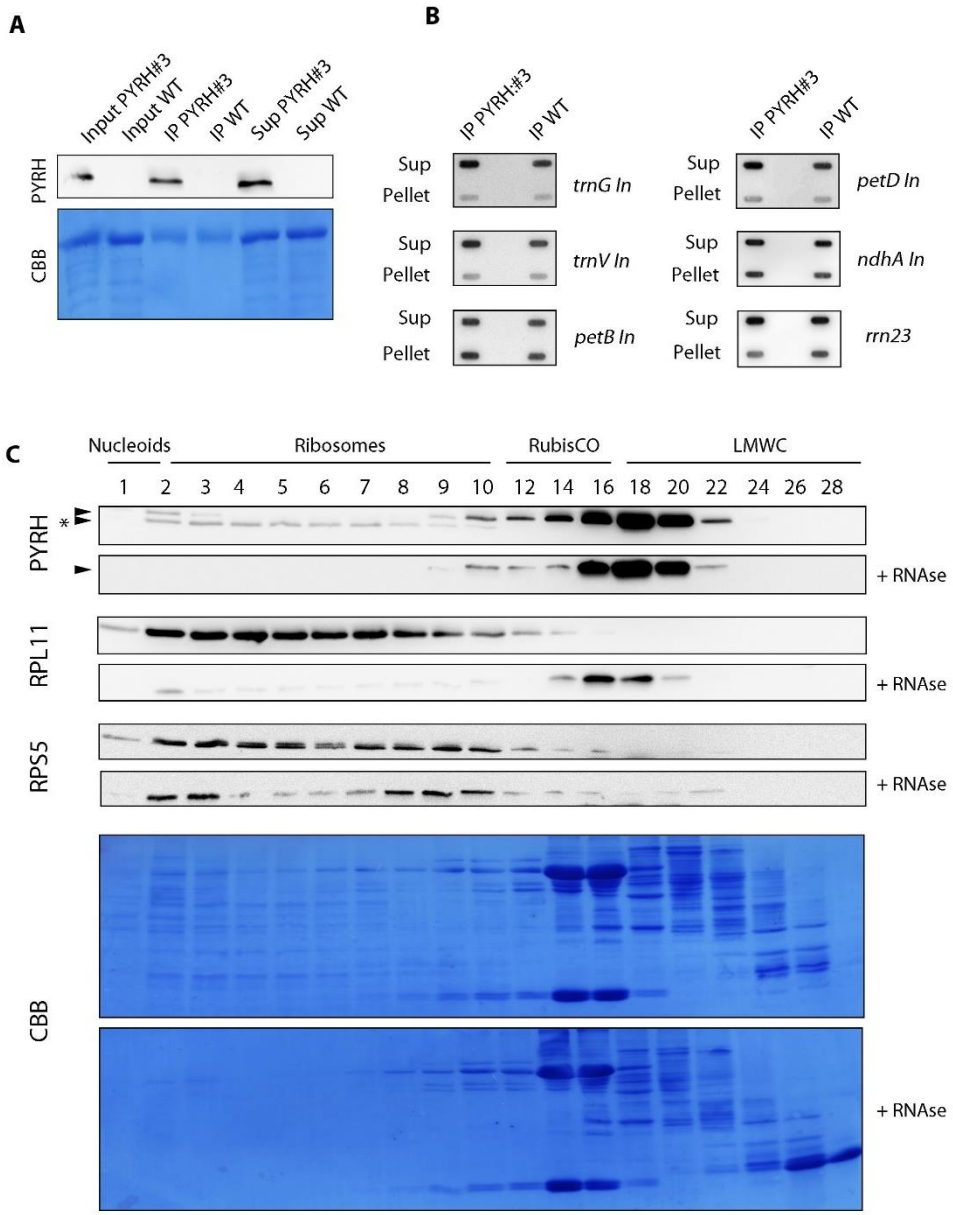


Figure 4. Analysis of *in vivo* RNA Association of PYRH in PYRH-complemented *pumpkin* Plants. A) Immunoblot of immunoprecipitated (IP) PYRH protein, pulled down from stroma extract of PYRH complemented *pumpkin* plants using PYRH-specific antibodies. Wild-type stroma extract served as control. IP = Immunoprecipitate, Sup = Supernatant, CBB = Coomassie Brilliant Blue. B) Slot Blot hybridization of immunoprecipitated RNAs (from PYRH IP, shown in panel A). Blots were hybridized with indicated probes representing PUMPKIN RNA targets as well as 23S rRNA as control. No enrichment of RNA targets was observed in the PYRH-IP. IP = Immunoprecipitate, Sup = Supernatant, In = Intron, WT = wild type. C) Size-exclusion chromatography (SEC) of native stroma extracts from PYRH-complemented *pumpkin* plants. Untreated stroma and RNAse-treated stroma were separated on a Superose 6 10/300 GL column. Fractions (indicated by numbers) were precipitated and subjected to SDS PAGE (only 50% of fractions 14-28 were loaded). Blots were probed with the indicated antibodies. The indicated complexes were derived from Olinares et al. (2010). LMWC = low molecular weight complexes. The asterisk marks an unspecific band. CBB (Coomassie Brilliant Blue) staining is shown as loading control.

Table 1. List of Putative PUMPKIN Interactors Identified by Co-immunoprecipitation Followed by Mass Spectrometry Analysis Using Three Independent Biological Replicates. The table contains the significant hits (log_2 FC > 1 and p-value adj < 0.05) and lists if the identified proteins are co-expressed with PUMPKIN according to the ATTEDII database. The subcellular localizations were predicted according to SUBA4 (http://suba.live/). Hits for cytosolic and other cellular compartments were removed from this table. Proteins involved in plastid RNA metabolism, either proven by functional characterization or occurrence of known RNA-binding motifs are highlighted in grey.

AGI Code	Protein Name	log_2 Fold Change	Co-Expression	Uniprot ID
AT4G01037	WTF1 (WHATS THIS FACTOR 1 homolog)	7.99	no	A0MFS5
AT1G10700	PRS3 (Ribose-phosphate pyrophosphokinase 3)	7.86	no	Q93Z66; A0A1P8ANQ1
AT3G46780	PLASTID TRANSCRIPTIONALLY ACTIVE 16	7.48	no	Q9STF2
AT2G04270	RNE (Ribonuclease E/G-like protein)	6.58	no	F4IV66; Q8GVF1
AT3G18680	PUMPKIN, Amino acid kinase family protein	6.53	–	Q9LSA9; A0A1I9LRT3
AT1G79080	Pentatricopeptide repeat-containing protein	6.42	no	A3KPF8
AT4G37510	RNC1 (Ribonuclease III domain-containing protein)	6.34	yes	Q9SZV0
AT5G18570	OBGC/ CPSAR1 (GTP-binding protein OBGC)	5.95	yes	Q8L7L0
AT3G17170	RFC3 (Translation elongation factor/ putative S6-type protein of small ribosomal subunit)	5.31	yes	Q948R9
AT3G18390	CRS1 / YhbY/ EMB1865 (CRM domain-containing protein)	5.31	yes	Q9LS51
AT3G03630	CS26 (S-sulfo-L-cysteine synthase)	5.10	yes	O22682
AT1G06190	RHON1 (Rho termination factor)	4.90	yes	Q94K75; A0A1P8AM12; A0A1P8AM09
AT2G43750	OASB (Cysteine synthase)	4.89	no	P47999
AT1G01080	RNA-binding (RRM/RBD/RNP motifs) family protein	4.79	yes	Q8W592; F4HQH8; A0A1P8APK3
AT3G23070	CFM3A (CRM-domain containing factor)	4.74	yes	F4J2U9
AT5G63420	RNaseJ/EMB2746 (Ribonuclease J)	4.53	yes	Q84W56
AT3G04760	Pentatricopeptide repeat (PPR-like) superfamily protein	4.17	no	Q9SR00; A0A1I9LPC7
AT4G09040	RNA-binding (RRM/RBD/RNP motifs) family protein	3.80	no	Q6NPL0; A8MS54
AT5G22510	INVE (Alkaline/neutral invertase E)	3.77	no	Q9FK88
AT4G38460	GGR/ GGPPS11 (Heterodimeric geranylgeranyl pyrophosphate synthase small subunit)	3.50	no	Q39108
AT5G23310	FSD3 (Superoxide dismutase [Fe] 3)	3.37	yes	Q9FMX0

AT1G70070	ISE2/ EMB25 DExH-box ATP-dependent RNA helicase DExH15	3.34	yes	B9DFG3
AT4G02990	MTERF4/ RUG2 (Transcription termination factor)	3.32	yes	Q9ZT96
AT1G02150	PPR3 Pentatricopeptide repeat-containing protein	3.18	yes	Q8LPS6
AT5G66470	GTPase ERA-like	3.17	yes	Q8VZ74
AT4G14510	CFM3B (CRM-domain containing factor)	3.12	no	F4JVH1
AT3G10690	GYRA (DNA gyrase subunit A)	3.03	no	Q9CAF6
AT3G03710	PNP1/ PNPase1/ RIF10 (Polyribonucleotide nucleotidyltransferase 1)	3.01	yes	Q8GZQ3
ATCG01120	RPS15 (30S ribosomal protein S15)	2.97	no	P56805
AT2G37220	CP29B (RNA-binding protein CP29B)	2.93	yes	Q9ZUU4
ATCG00650	RPS18 (30S ribosomal protein S18)	2.83	no	P56807
AT2G20020	CAF1 (CRS2-associated factor 1)	2.81	yes	Q9SL79
AT3G25690	CHUP1 (Hydroxyproline-rich glycoprotein family protein)	2.76	no	F4JA42; Q9LI74; A0A1I9LT48
AT5G46580	SOT1 (Pentatricopeptide repeat-containing protein)	2.69	yes	Q9LS25
AT3G17040	HCF107 (High chlorophyll fluorescent 107)	2.52	no	F4J3Z3; Q8RWG2; A0A1I9LTR4; A0A1I9LTR5
ATCG00900; ATCG01240	RPS7A (30S ribosomal protein S7)	2.45	no	P61841
AT5G16715	EMB2247 (Valine--tRNA ligase)	2.43	yes	F4KE63
AT5G59750	RIBA3 (Monofunctional riboflavin biosynthesis protein)	2.42	no	Q9FN89; F4KJA1; A0A1P8BHH7
AT5G64300	RIBA1 (Bifunctional riboflavin biosynthesis protein)	2.42	no	P47924
AT2G31890	RAP (RAP domain-containing protein)	2.39	yes	Q8VZE7
AT3G53460	CP29A (Chloroplast RNA-binding protein 29A)	2.38	yes	F4JAF3
AT3G53460	PP287 (Tetratricopeptide repeat (TPR)-like superfamily protein)	2.38	yes	Q43349
AT5G14320	RPS13 (30S ribosomal protein S13)	2.36	no	B3H631; P42732
AT3G20230	putative L18/L5-type protein of ribosomal subunit	2.32	yes	Q9LJX6
AT1G30610	EMB2279/ SOT5 (Pentatricopeptide (PPR) repeat-containing protein)	2.30	yes	F4I6D9; Q9SA76
AT3G15190	RPS20 (30S ribosomal protein S20)	2.25	yes	Q9ASV6
AT5G66055	AKRP/ EMB2036 (Ankyrin repeat domain-containing protein)	2.18	no	Q05753
AT3G12080	EMB2738 (GTP-binding family protein)	2.12	yes	F4J8M2; B5X565
AT5G24314	PTAC7 (PLASTID TRANSCRIPTIONALLY ACTIVE 7)	2.04	no	Q8VZV9
AT2G33800	RPS5 (30S ribosomal protein S5)	2.04	yes	P93014

AT3G09650	HCF152 (Pentatricopeptide repeat-containing protein)	2.03	no	Q9SF38
AT1G12900	GAPA-2 (Glyceraldehyde-3-phosphate dehydrogenase)	1.86	no	F4HNZ6; Q9LPW0; A0A1P8APR6
AT4G26370	Antitermination NusB domain-containing protein	1.85	yes	F4JUX9; Q93XY7
AT1G23400	CAF2 (CRS2-associated factor 2)	1.71	no	Q9LDA9
AT1G79850	RPS17 (30S ribosomal protein S17)	1.65	yes	P16180
AT3G13120	RPS10 (30S ribosomal protein S10)	1.63	yes	Q9LK61
AT1G16720	HCF173 (High chlorophyll fluorescence phenotype 173)	1.62	no	Q8W4D6; A0A1P8AWY1; A0A1P8AWU9
ATCG00905; ATCG00065; ATCG01230	RPS12A/B (30S ribosomal protein S12)	1.59	no	P62126
ATCG00160	RPS2 (30S ribosomal protein S2)	1.47	no	P56797
AT3G26650	GAPA1 (Glyceraldehyde-3-phosphate dehydrogenase)	1.46	no	P25856
AtCg00330	RPS14 (30S ribosomal protein S14)	1.41	no	P56804
AT5G30510	RPS1 (30S ribosomal protein S1)	1.40	yes	Q93VC7
AT4G34830	MRL1 (Pentatricopeptide repeat-containing protein, Maturation of RBCL 1)	1.30	no	Q0WLC6
ATCG00770	RPS8 (30S ribosomal protein S8)	1.29	no	P56801
AT3G13180	NOL1/NOP2 (antitermination NusB domain-containing protein)	1.24	yes	Q8VYC4
AT2G31810	Acetolactate synthase small subunit 2	1.21	no	Q93YZ7
ATCG00760	RPL36 (50S ribosomal protein L36)	1.14	no	P62117
AT5G26742	RH3/ EMB1138 (DEAD-box ATP-dependent RNA helicase 3)	1.12	no	F4K180; Q8L7S8
ATCG00750	RPS11 (30S ribosomal protein S11)	1.11	no	P56802
AT5G54600	RPL24 (50S ribosomal protein L24)	1.02	no	P92959; F4K1S8

Figure 5

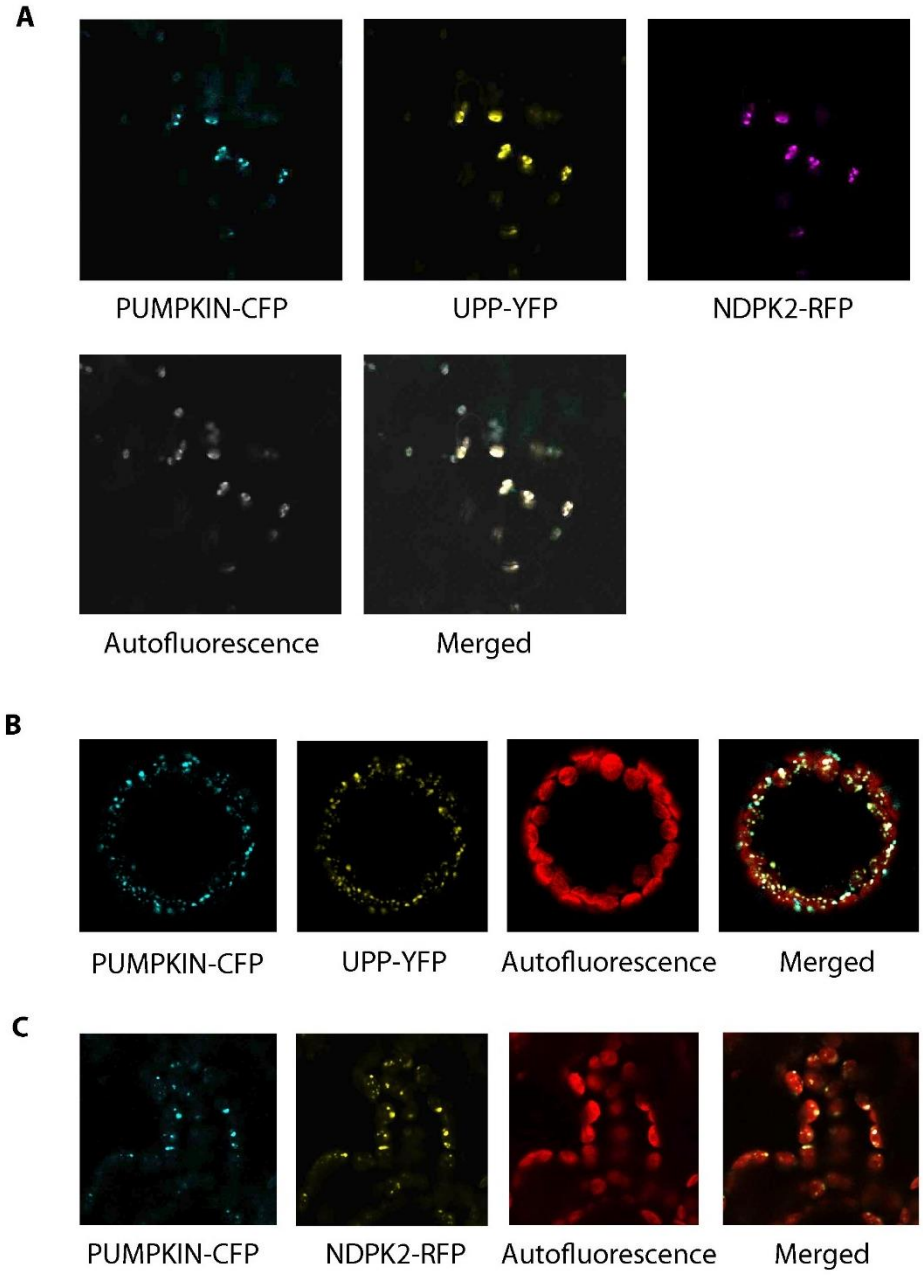
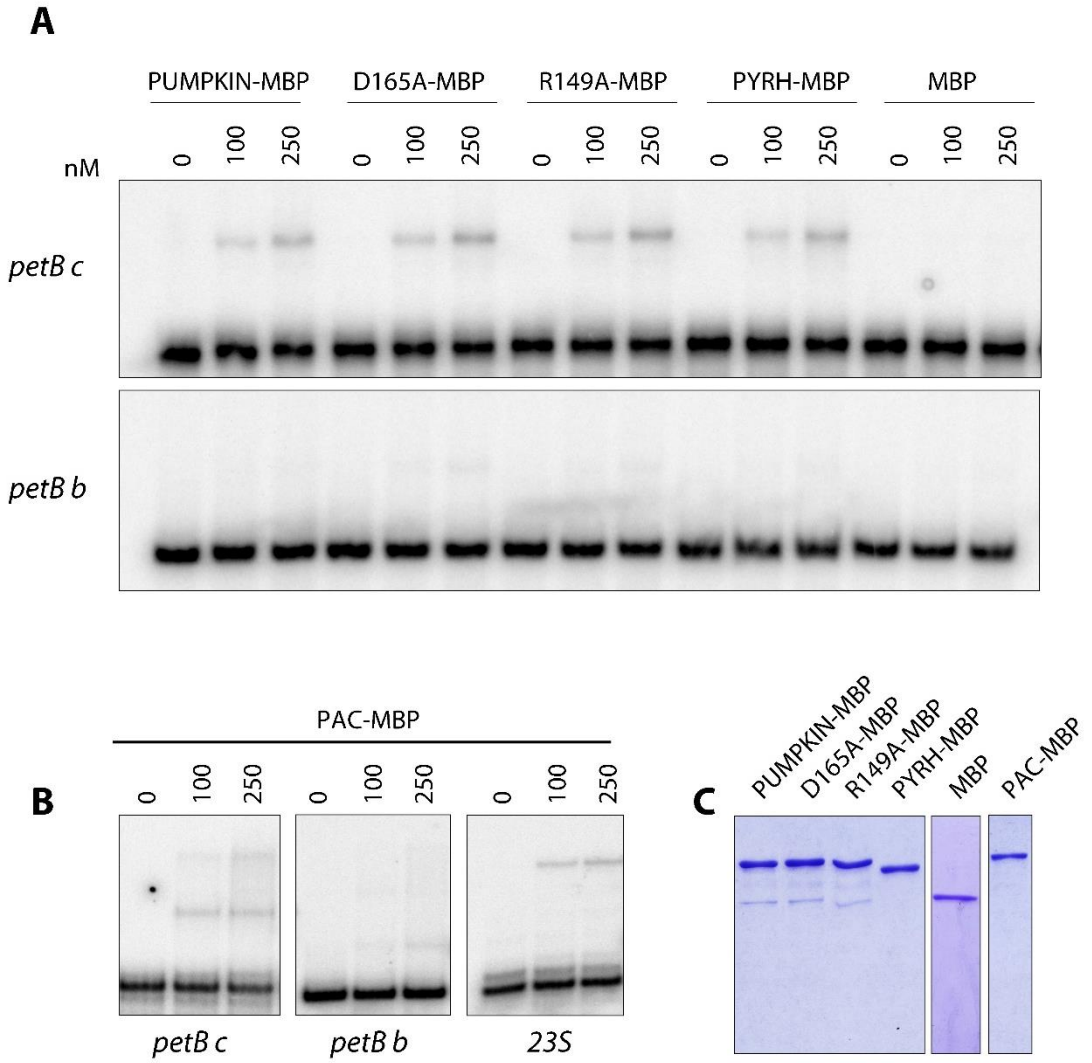


Figure 5. Co-localization Analyses of Enzymes in the Pyrimidine Salvage Pathway Using Fluorescence Microscopy. A) Microscopy images of the spotty triple co-localization of PUMPKIN-CFP, UPP-YFP, and NDPK2-RFP in tobacco leaf tissue. Single fluorescent pictures, chlorophyll autofluorescence and a merged picture are shown. B) Co-localization of PUMPKIN-CFP and UPP-YFP in transiently transformed tobacco protoplasts. C) Co-localization of PUMPKIN-CFP and NDPK2-RFP in transiently transformed tobacco leaf tissue.

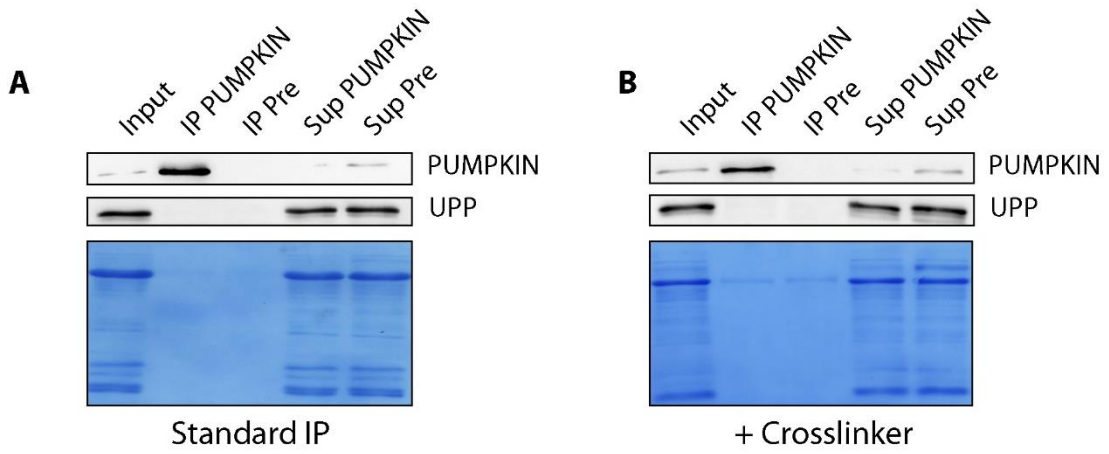
SUPPLEMENTAL DATA

Supplemental Figure S1



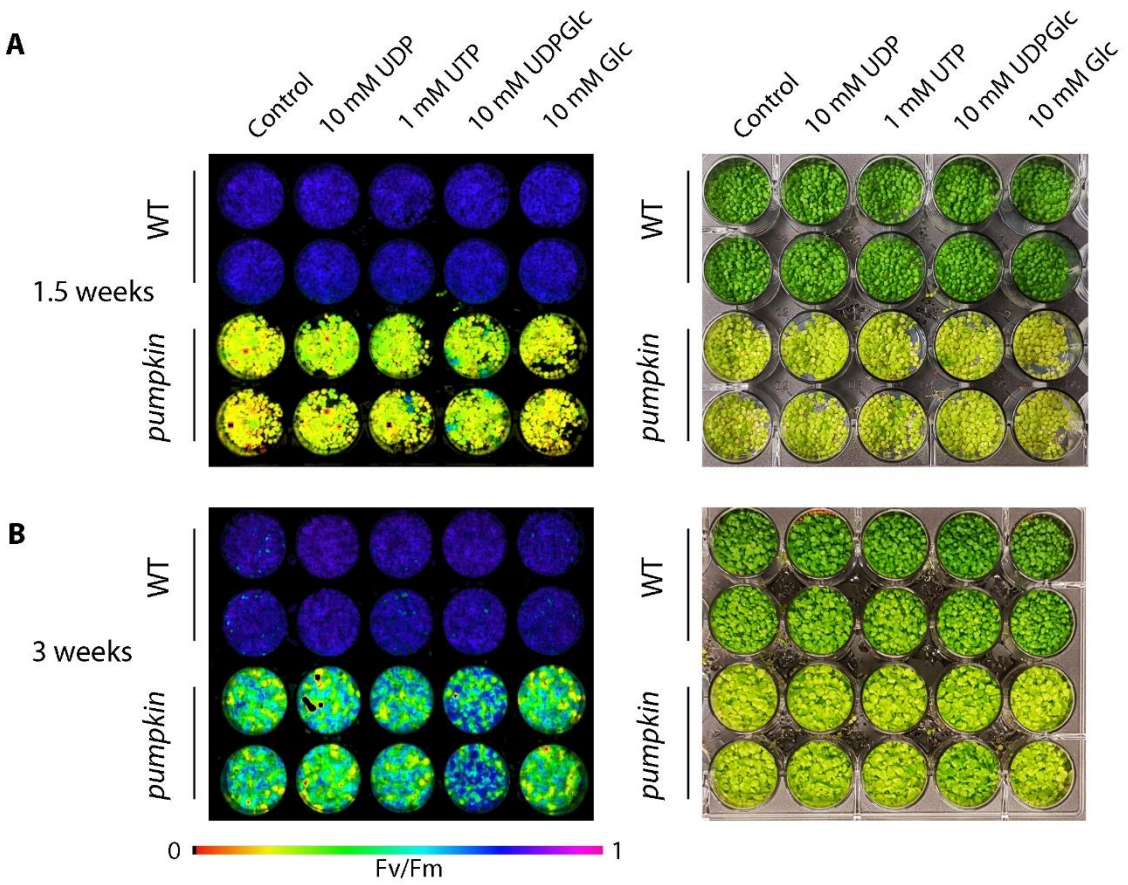
Supplemental Figure S1. Electrophoretic Mobility Shift Assays (EMSAs) with Recombinant PUMPKIN, R149A, D165A, PYRH and PAC Protein. A) EMSA with recombinant MBP-tagged PUMPKIN, R149A, D165A, and PYRH protein using two previously described *petB* intron RNA probes (Schmid et al., 2019). To exclude any binding of the MBP-tag, MBP protein was used as negative control. Increasing protein concentrations (nM) are indicated. B) EMSA with recombinant MBP-tagged PAC protein using two previously described *petB* intron RNA probes (Schmid et al., 2019). To ensure the purification of functional PAC protein, an EMSA with a 23S rRNA probe served as positive control (Meurer et al., 2017). C) SDS-PAGE of affinity purified recombinant proteins used for the EMSA studies. Staining: Coomassie Brilliant Blue.

Supplemental Figure S2



Supplemental Figure S2. Co-immunoprecipitation of PUMPKIN. A) Co-immunoprecipitation of PUMPKIN from wild-type stroma extracts using PUMPKIN-specific antibody and pre-immune serum as control. Fractions were subjected to immunoblot analyses using PUMPKIN- and UPP-specific antibodies, respectively. IP = Immunoprecipitate, Sup = Supernatant, CBB = Coomassie Brilliant Blue. B) See A), the stroma extract was cross-linked with DSP before performing the IP experiment.

Supplemental Figure S3



Supplemental Figure S3. Rescue Experiment of Wild-type and *pumpkin* Plants on Media Supplemented with UDP, UTP, UDP-glucose and Glucose. A) Pictures and PAM measurements (F_v/F_m) of wild-type (WT) and *pumpkin* plants grown on different supplemented MS media for 1.5 weeks. As control, plants were grown on normal ½ MS medium. Concentrations of supplements: UDP 10 mM, UTP 1 mM, UDP-glucose (UDPGlc) 10 mM, glucose 10 mM. B) Pictures and PAM measurements of plants grown for 3 weeks with and without supplements. Experimental set-up and descriptions are similar to panel A).

Supplemental Table S1. List of Oligonucleotides

Name	**Sequence 5'-3'**	**Experiment**
Pumpkin_fwd	TCATGGCAATTCCGTTGCCTCT	Full length construct for introducing point mutations
Pumpkin_rev	TCATGAGGTTGTTGTAACAATGG	Full length construct for introducing point mutations
Pumpkin_R149A_fwd	CCAGGTAGATCCAGCAAAGATATTTCCTCCT CCAACAACT	Mutagenesis Primer
Pumpkin_R149A_rev	AGTTGTTGGAGGAGGAAATATCTTTGCTGGA TCTACCTGG	Mutagenesis Primer
Pumpkin_D165A_fwd	CAGCAGAGGAGCGAGCAAGGCCACTGCAC	Mutagenesis Primer
Pumpkin_D165A_rev	GTGCAGTGGCCTTGCTCGCTCCTCTGCTG	Mutagenesis Primer
PyrH_BamHI_fwd	GGCGGCCGCGGATCCATGGCTACCAATGCAA AAC	Full length construct for complementation
PyrH_XhoI_rev	CAGCCCGGGCTCGAGTTATTCCGTGATTAAA GTCCC	Full length construct for complementation
RBCS_XbaI/ BamHI_fwd	GGCGGCCGCTCTAGAAGAACAATGGCTTCCT CTATGCTC	N-terminal fusion of RBCS targeting peptide to PYRH
RBCS_XbaI/ BamHI_rev	CAGCCCGGGGGATCCTTCGGAATCGGTAAGG	N-terminal fusion of RBCS targeting peptide to PYRH
Pumpkin_attB_fwd	GGGGACAAGTTTGTACAAAAAAGCAGGCTTC ATGGCAATTCCGTTGCCTCT	Gateway cloning R149A/D165A
Pumpkin_attB_rev	GGGGACCACTTTGTACAAGAAAGCTGGGTTC ATGAGGTTGTTGTAACAATGG	Gateway cloning R149A/D165A
pyrH_attB_fwd	GGGGACAAGTTTGTACAAAAAAGCAGGCTTC ATGCAAAAC	Gateway cloning RBCScTP-PYRH
pyrH_attB_rev	GGGGACCACTTTGTACAAGAAAGCTGGGTTT ATTCCGTGATTAAAGTCCC	Gateway cloning RBCScTP-PYRH
Fw_Pumpkin_5'UTR	CGGCGACACTTCTCTACTGA	Genotyping
Rev_Pumpkin_Signal	GGACAAAGAAGAAGAGCAGCTG	Genotyping
T-DNA_PCR_Gabi-kat	ATAATAACGCTGCGGACATCTACATTTT	Genotyping
ndhA_int-f_34Q	CGATTCAGAGTATGCTCCTATCCACCGAC	Slot blot, RNA-gel blot (Intron)
ndhA_int-rev_35Q	CAGAACCGTACATGAGGTCTTGGC	Slot blot, RNA-gel blot (Intron)
petBintFw	ACCTCGTAGCCGGACTTCTT	Slot blot, RNA-gel blot (Intron)

petBintRev	ACTCCCCCTCCAAGAACTGT	Slot blot, RNA-gel blot (Intron)
Fw_trnG_Intron	ATTGTTCGAGTTGGCCCGAT	Slot blot, RNA-gel blot (Intron)
Rev_trnG_Intron	ACTTTGGTTTCATTCGGCTCC	Slot blot, RNA-gel blot (Intron)
Fw_trnV_Intron	CAGAGGAGTCCATCACGCAA	Slot blot, RNA-gel blot (Intron)
Rev_trnV_Intron	CCTAGCATTGAGTAGACTGGGT	Slot blot, RNA-gel blot (Intron)
Fw_petD_intron	TGGGCTCCGTAAAATCCAGT	Slot blot, RNA-gel blot (Intron)
Rev_petD_intron	GTGCTTTCTGGGTCGTCTCA	Slot blot, RNA-gel blot (Intron)
T7 23S RNA full	TAATACGACTCACTATAGGGAGATTC AAACGAGGAAAGGCTTACG	Slot blot
Rev 23S RNA full	AGGAGAGCACTCATCTTGG	Slot blot
Fw_T7_petBin_1	TAATACGACTCACTATAGGGTGTGTGACTTG TTAGAATTGACCC	EMSA probe *petB b*
Rev_petBin_1	ATAAGATAGAACGGAAAGAAGTCCG	EMSA probe *petB b*
Fw_T7_petBin_2	TAATACGACTCACTATAGGGAAACTTTAATA AATCAAAAAAATTTTCTGC	EMSA probe *petB c*
Rev_petBin_2	AACCTCAGATTCATACGAG	EMSA probe *petB c*
Fw_cacc_NDPK2	CACCATGGTGGGAGCGACTGTA	Localization study, NDPK2-RFP
Rev_NDPK2_o.stop	CTCCCTTAGCCATGTAGCTAGAG	Localization study, NDPK2-RFP
Fw_Pumpkin_CACC	CACCATGGCAATTCCGTTGCCTCTTACC	Localization study, PUMPKIN-CFP
Rev_Pumpkin_o.Stop	TGAGGTTGTTGTAACAATGGAGTTC	Localization study, PUMPKIN-CFP
Fw_Pumpkin_BamHI	ATATGGATCCTCAGACAATGGTTCATCTCCA GAC	Cloning of PUMPKIN/ D165A/ R149A in pMAL for Overexpression
Rev_Pumpkin_stop_ SalI	AATAGTCGACTTATGAGGTTGTTGTAACAAT GGAGTTC	Cloning of PUMPKIN/ D165A/ R149A in pMAL for Overexpression
Fw_BamHI_pyrH	ATATGGATCCGCTACCAATGCAAAACCCGTC	Cloning of PYRH in pMAL for Overexpression
Rev_pyrH_stop_SalI	AATAGTCGACTTATTCCGTGATTAAAGTCCCT TC	Cloning of PYRH in pMAL for Overexpression

REFERENCES

An S, Kumar R, Sheets ED, Benkovic SJ (2008) Reversible compartmentalization of de novo purine biosynthetic complexes in living cells. Science **320:** 103-106

Andrade JM, Pobre V, Silva IJ, Domingues S, Arraiano CM (2009) The role of 3'-5' exoribonucleases in RNA degradation. Prog Mol Biol Transl Sci **85:** 187-229

Asakura Y, Galarneau E, Watkins KP, Barkan A, van Wijk KJ (2012) Chloroplast RH3 DEAD box RNA helicases in maize and Arabidopsis function in splicing of specific group II introns and affect chloroplast ribosome biogenesis. Plant Physiol **159:** 961-974

Bang WY, Chen J, Jeong IS, Kim SW, Kim CW, Jung HS, Lee KH, Kweon HS, Yoko I, Shiina T, Bahk JD (2012) Functional characterization of ObgC in ribosome biogenesis during chloroplast development. Plant J **71:** 122-134

Baresova V, Skopova V, Souckova O, Krijt M, Kmoch S, Zikanova M (2018) Study of purinosome assembly in cell-based model systems with de novo purine synthesis and salvage pathway deficiencies. Plos One **13**(7): e0201432.

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

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

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) **57:** 289-300

Bölter B, Sharma R, Soll J (2007) Localisation of Arabidopsis NDPK2--revisited. Planta **226:** 1059-1065

Briozzo P, Evrin C, Meyer P, Assairi L, Joly N, Bārzu 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

Cao P, Ren Y, Liu X, Zhang T, Zhang P, Xiao L, Zhang F, Liu S, Jiang L, Wan J (2019) Purine nucleotide biosynthetic gene GARS controls early chloroplast development in rice (Oryza sativa L.). Plant Cell Rep **38:** 183-194

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:** 165-173

Chen CT, Slocum RD (2008) Expression and functional analysis of aspartate transcarbamoylase and role of de novo pyrimidine synthesis in regulation of growth and development in Arabidopsis. Plant Physiol Biochem **46:** 150-159

Chitrakar I, Kim-Holzapfel DM, Zhou W, French JB (2017) Higher order structures in purine and pyrimidine metabolism. J Struct Biol **197:** 354-364

Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics **13:** 2513-2526

Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol **26:** 1367-1372

Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. J Proteome Res **10:** 1794-1805

Duley JA, Henman MG, Carpenter KH, Bamshad MJ, Marshall GA, Ooi CY, Wilcken B, Pinner JR (2016) Elevated plasma dihydroorotate in Miller syndrome: Biochemical, diagnostic and clinical implications, and treatment with uridine. Mol Genet Metab **119:** 83-90

Fassy F, Krebs O, Lowinski M, Ferrari P, Winter J, Collard-Dutilleul V, Salahbey Hocini K (2004) UMP kinase from Streptococcus pneumoniae: evidence for co-operative ATP binding and allosteric regulation. Biochem J **384:** 619-627

Geigenberger P, Regierer B, Nunes-Nesi A, Leisse A, Urbanczyk-Wochniak E, Springer F, van Dongen JT, Kossmann J, Fernie AR (2005) Inhibition of de novo pyrimidine synthesis in growing potato tubers leads to a compensatory stimulation of the pyrimidine salvage pathway and a subsequent increase in biosynthetic performance. Plant Cell **17:** 2077-2088

Grefen C, Donald N, Hashimoto K, Kudla J, Schumacher K, Blatt MR (2010) A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. Plant J **64:** 355-365

Greive SJ, Lins AF, von Hippel PH (2005) Assembly of an RNA-protein complex. Binding of NusB and NusE (S10) proteins to boxA RNA nucleates the formation of the antitermination complex involved in controlling rRNA transcription in Escherichia coli. J Biol Chem **280:** 36397-36408

Halpert M, Liveanu V, Glaser F, Schuster G (2019) The Arabidopsis chloroplast RNase J displays both exo- and robust endonucleolytic activities. Plant Mol Biol **99:** 17-29

Hooper CM, Castleden IR, Tanz SK, Aryamanesh N, Millar AH (2017) SUBA4: the interactive data analysis centre for Arabidopsis subcellular protein locations. Nucleic Acids Res **45:** D1064-D1074

Huang W, Zhu Y, Wu W, Li X, Zhang D, Yin P, Huang J (2018) The Pentatricopeptide Repeat Protein SOT5/EMB2279 Is Required for Plastid rpl2 and trnK Intron Splicing. Plant Physiol **177:** 684-697

Hung WF, Chen LJ, Boldt R, Sun CW, Li HM (2004) Characterization of Arabidopsis glutamine phosphoribosyl pyrophosphate amidotransferase-deficient mutants. Plant Physiol **135:** 1314-1323

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

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:** D1211-1215

Koncz C, Martini N, Szabados L, Hrouda M, Bachmair A, Schell J (1994) Specialized vectors for gene tagging and expression studies. *In* SB Gelvin, RA Schilperoort, eds, Plant Molecular Biology Manual. Springer Netherlands, Dordrecht, pp 53-74

Kupsch C, Ruwe H, Gusewski S, Tillich M, Small I, Schmitz-Linneweber C (2012) Arabidopsis chloroplast RNA binding proteins CP31A and CP29A associate with large transcript pools and confer cold stress tolerance by influencing multiple chloroplast RNA processing steps. Plant Cell **24:** 4266-4280

Mainguet SE, Gakière B, Majira A, Pelletier S, Bringel F, Guérard F, Caboche M, Berthomé R, Renou JP (2009) Uracil salvage is necessary for early Arabidopsis development. Plant J **60:** 280-291

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

Nara T, Hashimoto M, Hirawake H, Liao CW, Fukai Y, Suzuki S, Tsubouchi A, Morales J, Takamiya S, Fujimura T, Taka H, Mineki R, Fan CK, Inaoka DK, Inoue M, Tanaka A, Harada S, Kita K, Aoki T (2012) Molecular interaction of the first 3 enzymes of the de novo pyrimidine biosynthetic pathway of Trypanosoma cruzi. Biochem Biophys Res Commun **418:** 140-143

Nara T, Hshimoto T, Aoki T (2000) Evolutionary implications of the mosaic pyrimidine-biosynthetic pathway in eukaryotes. Gene **257:** 209-222

Narusaka M, Shiraishi T, Iwabuchi M, Narusaka Y (2010) The floral inoculating protocol: a simplified *Arabidopsis thaliana* transformation method modified from floral dipping. Plant Biotechnology **27:** 349-351

Norambuena L, Marchant L, Berninsone P, Hirschberg CB, Silva H, Orellana A (2002) Transport of UDP-galactose in plants. Identification and functional characterization of AtUTr1, an Arabidopsis thaliana UDP-galactos/UDP-glucose transporter. J Biol Chem **277:** 32923-32929

Obayashi T, Aoki Y, Tadaka S, Kagaya Y, Kinoshita K (2018) ATTED-II in 2018: A Plant Coexpression Database Based on Investigation of the Statistical Property of the Mutual Rank Index. Plant Cell Physiol **59:** 440

Ohler L, Niopek-Witz S, Mainguet SE, Möhlmann T (2019) Pyrimidine Salvage: Physiological Functions and Interaction with Chloroplast Biogenesis. Plant Physiol **180:** 1816-1828

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

Paieri F, Tadini L, Manavski N, Kleine T, Ferrari R, Morandini P, Pesaresi P, Meurer J, Leister D (2018) The DEAD-box RNA Helicase RH50 Is a 23S-4.5S rRNA Maturation Factor that Functionally Overlaps with the Plastid Signaling Factor GUN1. Plant Physiol **176:** 634-648

Park JI, Ishimizu T, Suwabe K, Sudo K, Masuko H, Hakozaki H, Nou IS, Suzuki G, Watanabe M (2010) UDP-glucose pyrophosphorylase is rate limiting in vegetative and reproductive phases in Arabidopsis thaliana. Plant Cell Physiol **51:** 981-996

Pedley AM, Benkovic SJ (2017) A New View into the Regulation of Purine Metabolism: The Purinosome. Trends in Biochemical Sciences **42:** 141-154

Pfalz J, Liere K, Kandlbinder A, Dietz KJ, Oelmuller 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

Pulido P, Zagari N, Manavski N, Gawronski P, Matthes A, Scharff LB, Meurer J, Leister D (2018) CHLOROPLAST RIBOSOME ASSOCIATED Supports Translation under Stress and Interacts with the Ribosomal 30S Subunit. Plant Physiol **177:** 1539-1554

Rappsilber J, Ishihama Y, Mann M (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal Chem **75:** 663-670

Reichel M, Liao Y, Rettel M, Ragan C, Evers M, Alleaume AM, Horos R, Hentze MW, Preiss T, Millar AA (2016) In Planta Determination of the mRNA-Binding Proteome of Arabidopsis Etiolated Seedlings. Plant Cell **28:** 2435-2452

Schmid LM, Ohler L, Möhlmann T, Brachmann A, Muiño JM, Leister D, Meurer J, Manavski N (2019) PUMPKIN, the Sole Plastid UMP Kinase, Associates with Group II Introns and Alters Their Metabolism. Plant Physiol **179:** 248-264

Seino J, Ishii K, Nakano T, Ishida N, Tsujimoto M, Hashimoto Y, Takashima S (2010) Characterization of rice nucleotide sugar transporters capable of transporting UDP-galactose and UDP-glucose. J Biochem **148:** 35-46

Serina L, Blondin C, Krin E, Sismeiro O, Danchin A, Sakamoto H, Gilles AM, Barzu 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

Smaczniak C, Li N, Boeren S, America T, van Dongen W, Goerdayal SS, de Vries S, Angenent GC, Kaufmann K (2012) Proteomics-based identification of low-abundance signaling and regulatory protein complexes in native plant tissues. Nat Protoc **7:** 2144-2158

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

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

Suchi M, Mizuno H, Kawai Y, Tsuboi T, Sumi S, Okajima K, Hodgson ME, Ogawa H, Wada Y (1997) Molecular cloning of the human UMP synthase gene and characterization of point mutations in two hereditary orotic aciduria families. Am J Hum Genet **60:** 525-539

Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods **13:** 731-740

Walter M, Kilian J, Kudla J (2002) PNPase activity determines the efficiency of mRNA 3'-end processing, the degradation of tRNA and the extent of polyadenylation in chloroplasts. EMBO J **21:** 6905-6914

Witz S, Jung B, Fürst S, Möhlmann T (2012) De novo pyrimidine nucleotide synthesis mainly occurs outside of plastids, but a previously undiscovered nucleobase importer provides substrates for the essential salvage pathway in Arabidopsis. Plant Cell **24:** 1549-1559

Wormit A, Traub M, Flörchinger M, Neuhaus HE, Möhlmann T (2004) Characterization of three novel members of the Arabidopsis thaliana equilibrative nucleoside transporter (ENT) family. Biochem J **383:** 19-26

Xu D, Leister D, Kleine T (2020) VENOSA4, a Human dNTPase SAMHD1 Homolog, Contributes to Chloroplast Development and Abiotic Stress Tolerance. Plant Physiol **182:** 721-729

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER Suite: protein structure and function prediction. Nat Methods **12:** 7-8

Yang Z, Shang Z, Wang L, Lu Q, Wen X, Chi W, Zhang L, Lu C (2015) Purine biosynthetic enzyme ATase2 is involved in the regulation of early chloroplast development and chloroplast gene expression in Arabidopsis. Photosynth Res **126:** 285-300

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

7.3. Publication II: PALE CRESS binds to plastid RNAs and facilitates the biogenesis of the 50S ribosomal subunit

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

Supplementary data available at: https://onlinelibrary.wiley.com/doi/10.1111/tpj.13662

the plant journal

SEB Society for Experimental Biology

The Plant Journal (2017) **92,** 400–413 doi: 10.1111/tpj.13662

PALE CRESS binds to plastid RNAs and facilitates the biogenesis of the 50S ribosomal subunit

Jörg Meurer[1], Lisa-Marie Schmid[1], Rhea Stoppel[1,†], Dario Leister[1], Andreas Brachmann[2] and Nikolay Manavski[1,*]

[1]Plant Sciences, Faculty of Biology, Ludwig-Maximilians-University Munich, Großhaderner Straße 2–4, 82152, Planegg-Martinsried, Germany, and

[2]Genetics, Faculty of Biology, Ludwig-Maximilians-University Munich, Großhaderner Straße 2–4, 82152, Planegg-Martinsried, Germany

Received 28 June 2017; revised 4 August 2017; accepted 8 August 2017; published online 14 August 2017.

*For correspondence (e-mail n.manavski@biologie.uni-muenchen.de).

[†]Present address: Bayer CropScience AG, Industriepark Höchst, 65926 Frankfurt/Main, Germany.

SUMMARY

The plant-specific PALE CRESS (PAC) protein has previously been shown to be essential for photoautotrophic growth. Here we further investigated the molecular function of the PAC protein. PAC localizes to plastid nucleoids and forms large proteinaceous and RNA-containing megadalton complexes. It co-immunoprecipitates with a specific subset of chloroplast RNAs including *psbK–psbI*, *ndhF*, *ndhD*, and 23S ribosomal RNA (rRNA), as demonstrated by RNA immunoprecipitation in combination with high throughput RNA sequencing (RIP-seq) analyses. Furthermore, it co-migrates with premature 50S ribosomal particles and specifically binds to 23S rRNA *in vitro*. This coincides with severely reduced levels of 23S rRNA in *pac* leading to translational deficiencies and related alterations of plastid transcript patterns and abundance similar to plants treated with the translation inhibitor lincomycin. Thus, we conclude that deficiency in plastid ribosomes accounts for the *pac* phenotype. Moreover, the absence or reduction of PAC levels in the corresponding mutants induces structural changes of the 23S rRNA, as demonstrated by *in vivo* RNA structure probing. Our results indicate that PAC binds to the 23S rRNA to promote the biogenesis of the 50S subunit.

Keywords: *Arabidopsis thaliana*, chloroplast, ribogenesis, RNA metabolism, RNA-binding proteins.

INTRODUCTION

In spite of the low coding capacity of the genomes in chloroplasts and mitochondria, the organellar RNA metabolism requires a surprisingly high number of nuclear encoded factors (Germain *et al.*, 2013; Stoppel and Meurer, 2013; Belcher *et al.*, 2015). Some plastid RNA metabolism factors are highly specific for individual RNA processing or stability events; others affect a plethora of transcripts. Due to the pleiotropic nature of knockout plants it is often difficult to pinpoint the primary function of factors involved in plastid gene expression (Stoppel *et al.*, 2012). This situation is particularly true when general plastid translation is affected due to the lack of ribosome processing and/or assembly, which indirectly causes transcriptional deficiencies due to a loss of expression of the plastid-encoded polymerase (PEP) leading to strong effects in RNA processing patterns and levels (Hajdukiewicz *et al.*, 1997; Legen *et al.*, 2002).

Remarkably, in yeast approximately 250 ribosome biogenesis factors required for rRNA processing and assembly of ribosomes were identified, from which multiple orthologues can be found in the genomes of plants (Weis *et al.*, 2015). Biogenesis of chloroplast eubacterial-like 70S ribosomes consisting of 30S and 50S subunits is comparably complex in particular because it involves processing, maturation, modification, and assembly of the four co-transcribed rRNAs, 16S, 23S, 4.5S, and 5S, together with at least 57 ribosomal proteins, five of which, PSPR2–PSPR6, do not have eubacterial homologs (Yamaguchi and Subramanian, 2000; Yamaguchi *et al.*, 2000; Shajani *et al.*, 2011).

Ribosome biogenesis necessitates accurate processing of the rRNAs and occurs in an assembly-assisted manner (Stoppel and Meurer, 2012; Tiller and Bock, 2014). Various factors required for chloroplast ribosome biogenesis were described but the precise molecular function often remained unclear (e.g. Bellaoui *et al.*, 2003; Bisanz *et al.*, 2003; Bollenbach *et al.*, 2005; Komatsu *et al.*, 2010; Nishimura *et al.*, 2010; Lu *et al.*, 2011; Bang *et al.*, 2012; Stoppel *et al.*, 2012; Lee *et al.*, 2013; Fristedt *et al.*, 2014). The observed defects in these and other previously characterized mutants with altered chloroplast rRNA expression

were accompanied by a relative increase of unprocessed or partially processed precursors of one or more of the four rRNAs (Bohne, 2014), suggesting that the affected proteins promote rRNA processing rather than stabilization or assembly of the processed rRNAs in later stages of ribosome biogenesis.

Recent proteomics and localization studies suggested that initial ribosome biogenesis and rRNA processing take place similar as in eubacteria in association with the chloroplast nucleoids (Majeran *et al.*, 2012; Bohne, 2014). This factor may be important to spatially separate early rRNA processing and ribosome assembly processes from highly regulated translationally active sites in the stroma or in association with thylakoid membranes (Zerges, 2000).

PAC is a plant-specific protein important for leaf and chloroplast development (Reiter *et al.*, 1994; Grevelding *et al.*, 1996; Meurer *et al.*, 1998; Tirlapur *et al.*, 1999). Null alleles of *PAC* are seedling lethal and exhibit an albino phenotype (Reiter *et al.*, 1994; Grevelding *et al.*, 1996). The *pac* phenotype could be partially rescued by treatment with cytokinin (Grevelding *et al.*, 1996). This might be due to the fact that cytokinin treatment leads to accumulation of plastid transcripts (Zubo *et al.*, 2008) and has a protective role on the photosynthetic apparatus (Cortleven *et al.*, 2014), which might contribute to the greening of the pale *pac* leaves. Investigations done on the molecular level showed that the absence of PAC led to alterations of the maturation patterns and/or abundance of several plastid mRNAs (Meurer *et al.*, 1998), suggesting that PAC might be involved in plastid gene expression. However, the sublocalisation within the chloroplast, a potential RNA-binding capability, the endogenous RNA targets, protein complex formation as well as the underlying function of PAC remained entirely unknown.

The results presented here provide evidence that PAC is a nucleoid-localized, RNA-binding protein associated *in vivo* with several plastid RNAs including 23S rRNA. PAC co-migrates with immature 50S ribosomal particles and induces structural changes of the 23S rRNA when it is absent or its levels are reduced. *Pac* mutants show only uniformly reduced levels but no processing alterations of rRNAs pointing to a role of PAC in the assembly of the large ribosomal subunit.

RESULTS

PAC is required for accumulation of plastid ribosomal RNAs and translation

The albinotic *pac* mutants are only viable on sucrose-supplemented medium. In order to achieve photoautotrophic growth and to obtain sufficient amounts of plant material for thorough molecular studies, we generated RNAi plants, in which *PAC* gene expression was substantially downregulated (Figure S1a). For the same reason, partially complemented plants (pComp) were selected, which, similar to the *PAC* RNAi plants, were able to grow on soil but exhibited a pale-green phenotype and retarded growth as compared to the wild type (WT; accession Col-0) and to fully complemented mutant plants (Figure S1b).

RNA gel blot analysis revealed that one of the most obvious changes in *pac* mutants was related to the plastid ribosomal RNA. Chloroplast *rrn23* and *rrn16* transcripts were barely detectable in *pac* mutants and their levels were notably reduced in the partially complemented *pac* plants as well as in *PAC* RNAi lines (Figure 1a). Quantification of the plastid rRNA in total leaf RNA using Agilent Bioanalyzer showed that the decrease of *rrn23–1.1*, *-1.3*, and *rrn16* correlated with the phenotype severity (Figures 1a and S1). When compared with the WT, levels of the 1.1 and 1.3 kb fragments of the 23S rRNA were reduced to 15 and 23%, respectively, while the reduction of the 16S rRNA amounted to 32% suggesting that *rrn23* is affected stronger than *rrn16* in *pac* mutants (Figure 1a).

Reduction of rRNA levels is likely to result in less ribosomes, which in turn reduces translation efficiency. Therefore, polysome loading of plastid transcripts was examined. In accordance with the observed rRNA decrease, polysome loading of the randomly chosen *psaA* and *psbK* transcripts was clearly affected in *pac* mutants pointing to a general translation deficiency (Figure 1b). Also *de novo* protein synthesis was severely impaired, as *in vivo* labeling of plastid proteins resulted in very poor ^{35}S-methionine incorporation in *pac*, *PAC* RNAi and pComp plants as compared to the WT and fully complemented lines (Figure 1c).

As impaired protein synthesis would lead to insufficient transcription due to reduced levels of the PEP, patterns and abundance of plastid transcripts in *pac* were compared to plants treated with the plastid translation inhibitor lincomycin and to two other very pale mutants, *rhon1* (Stoppel *et al.*, 2012; Chi *et al.*, 2014) and *ins11* (Cho *et al.*, 2009), with obvious and slight reductions of plastid rRNAs, respectively (Figure 2). Abolished plastid translation in the lincomycin-treated plants led to severe alterations of plastid transcripts levels and/or maturation patterns, as has been demonstrated in other translation deficient mutants (Cho *et al.*, 2009; Wang *et al.*, 2016). Interestingly, some of the tested transcripts in *pac* (*psaA*, *psbB*, *psbA*) were comparable with the ones in lincomycin-treated plants in terms of their processing patterns and accumulation, whereas others showed mild (*petA*, *psbK*, *ycf2*, *rbcL*) to pronounced (*rrn16*, *rrn23*, *ndhD*, *ycf1*, *ndhF*) differences, indicating that the observed transcript alterations in *pac* mutants might not entirely be secondary effects or a mere consequence of compromised translation. Notably, two of the tested transcripts (*ndhD*, *ycf1*) were found at much higher levels in *pac* when compared with the lincomycin-treated and WT control plants (Figures 2 and S1c), indicating that transcription is not primarily impaired in *pac*.

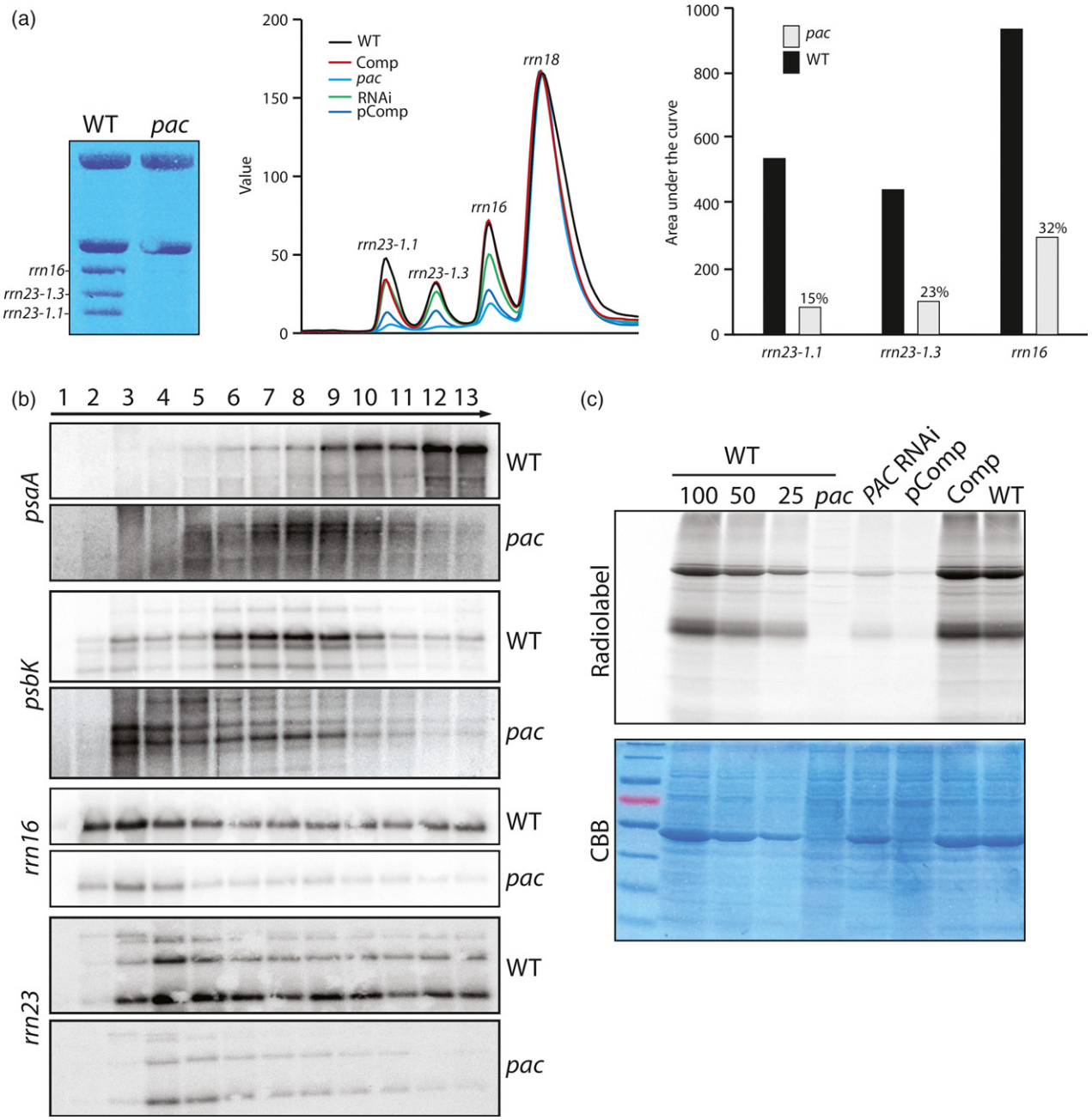


Figure 1. Quantification of rRNA and investigation of translation in *pac* mutants.
(a) Analyses of rRNA. Methylene blue (MB) staining of total RNA from 3-week-old WT and *pac* plants grown on sucrose-supplemented MS medium (left panel). Quantification of total RNA from WT, *pac*, *PAC* RNAi, pComp, and Comp plants using Bioanalyzer (middle panel). Values in all RNA samples were normalized to the cytosolic 18S rRNA. Amounts of plastid rRNA in *pac* and the WT are graphed (right panel).
(b) Polysome loading analysis of *psaA* and *psbK* mRNA. Numbers of fractions and used probes are indicated. As levels of *psaA* and *psbK* transcripts are reduced in *pac* mutants (Figure 2), panels showing polysome loading in *pac* were overexposed for better comparison.
(c) *In vivo* labeling of newly synthesized proteins in WT, *pac*, *PAC* RNAi, pComp, and Comp seedlings using [^{35}S]-Met in the presence of cycloheximide. Equal amounts of total protein (100 corresponds to 50 μg of total protein) were separated on an SDS-gel, stained with Coomassie Brilliant Blue (CBB), and analyzed by fluorography.

PAC is found in large RNA-containing complexes in chloroplast nucleoids and forms homomultimers

To examine the localization of the full-length protein, the entire coding region of PAC was fused to YFP and expressed transiently in tobacco protoplasts under the control of the cauliflower 35S promoter. YFP signals co-localized with chlorophyll fluorescence confirming the chloroplast localization (Figure 3a). Moreover, the signals

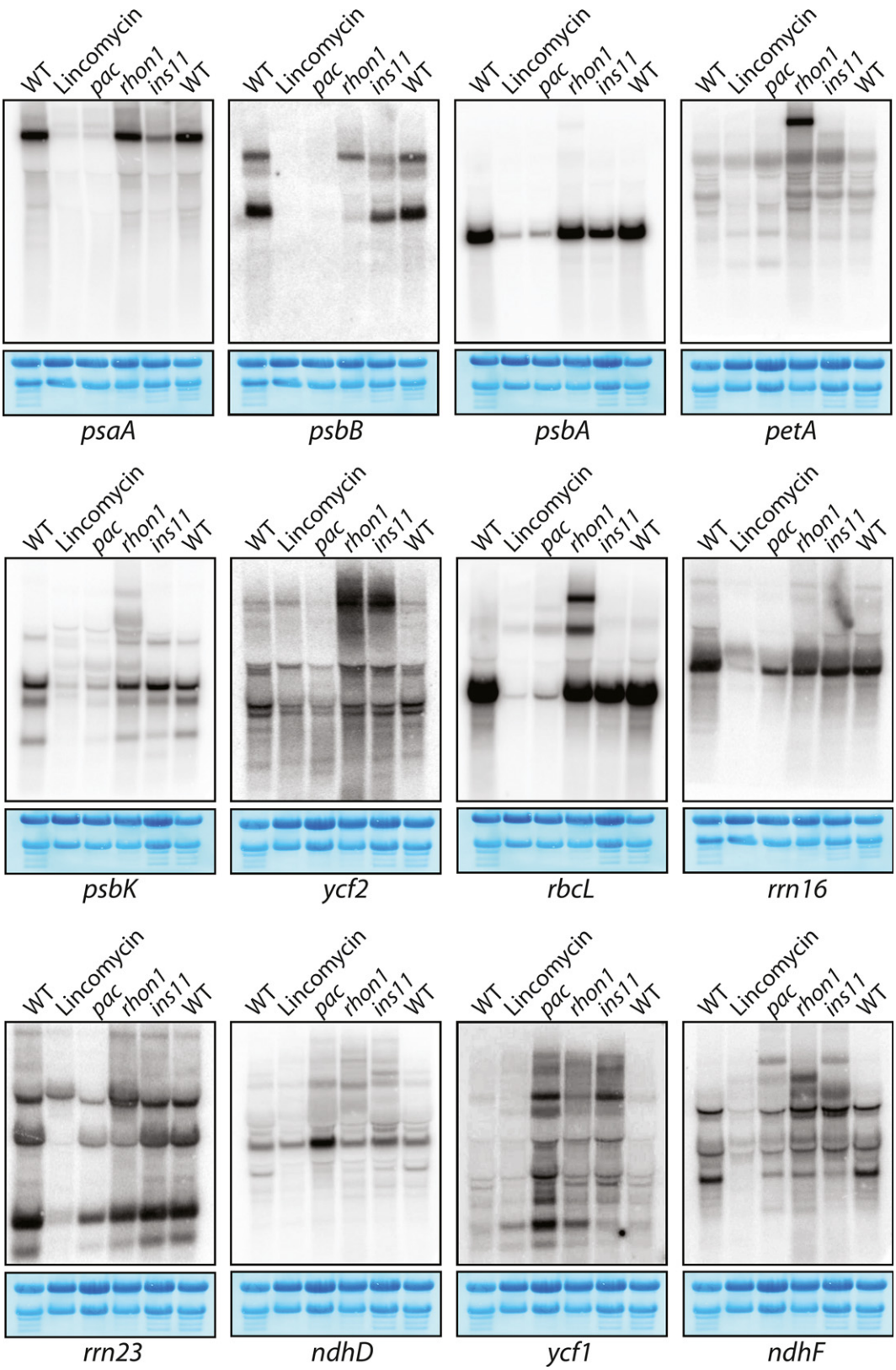


Figure 2. RNA gel blot analyses demonstrating chloroplast transcript defects in *pac* mutants.
Seedling leaf RNA (5 μg) from WT, lincomycin-treated plants, *pac*, *rhon1* and the pale mutant *ins11* was hybridized with the indicated probes. Blots were stained with methylene blue. [Colour figure can be viewed at wileyonlinelibrary.com].

within the chloroplasts appeared spotty implying that PAC might be localized in nucleoids. To check this finding, the chloroplast nucleoids-targeted protein RAP (Kleinknecht *et al.*, 2014) was fused to CFP and co-expressed with PAC. The overlap of the fluorescence signals confirmed the nucleoid localization (Figure 3a).

When mature PAC was expressed in *E. coli* as a fusion to the maltose binding protein (MBP) and purified by affinity and subsequent size exclusion chromatography (SEC) in the presence of β-mercaptoethanol, it did not elute from the gel filtration column at the expected size of 77 kDa but in a fraction corresponding to the size of ~158 kDa (aldolase) (Figure 3b), indicating that PAC could form redox-stable homodimers. When reducing agents were omitted, a second protein peak of size between 440 kDa (ferritin) and 158 kDa (aldolase) appeared, arguing for redox-sensitive multimerization of PAC.

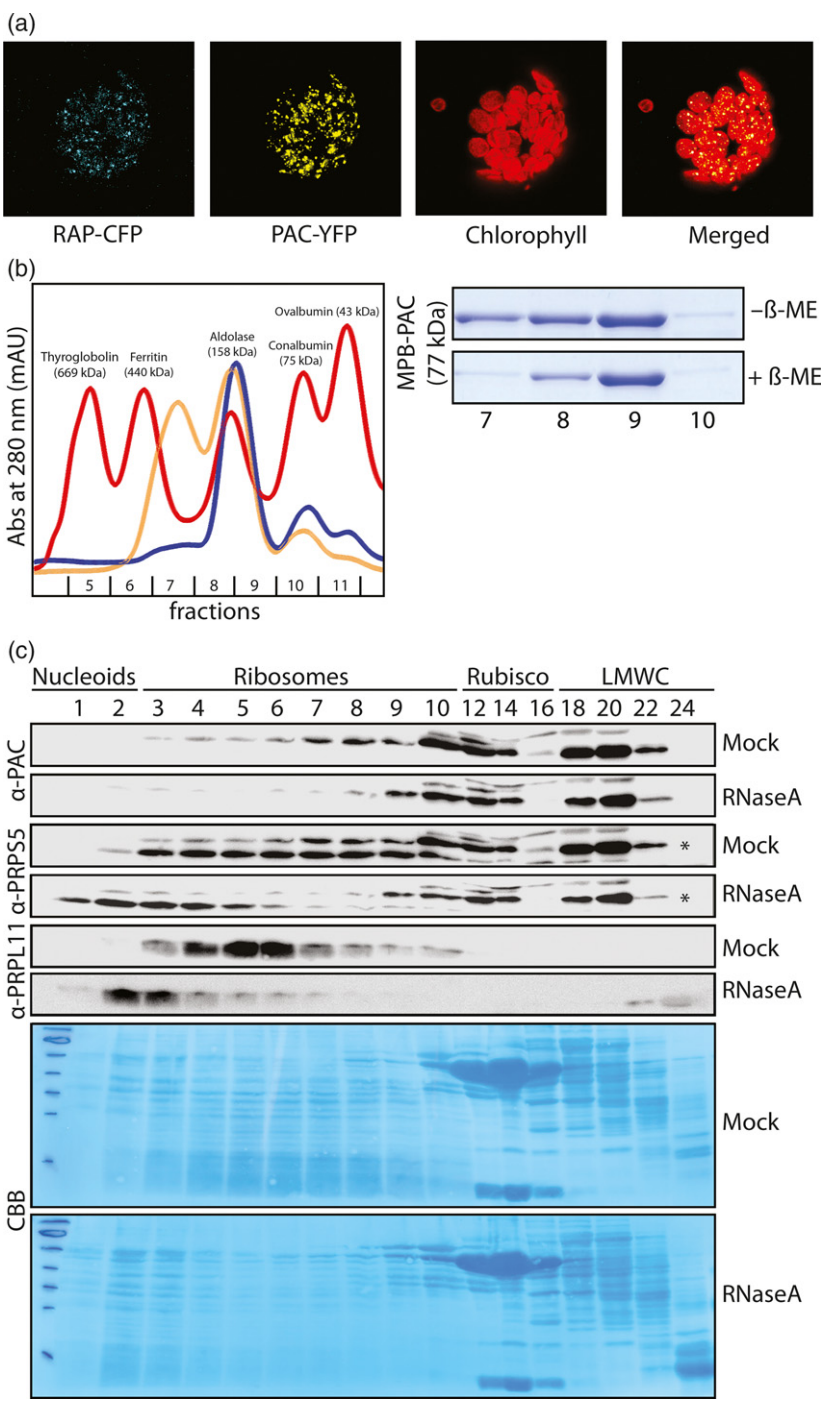


Figure 3. Subcellular localization, multimerization and size distribution of PAC proteins
(a) PAC localizes to chloroplast nucleoids. Confocal laser microscopy of tobacco protoplasts transiently co-expressing PAC-YFP and RAP-CFP (nucleoid-localized protein). Single and merged fluorescence signals are shown.
(b) Homomultimerization of PAC. Chromatogram of SEC-fractionated marker proteins (red), β-ME-treated (yellow-beige), and untreated (blue) recombinant MBP-PAC proteins.
(c) Size distribution of PAC complexes. Native stroma extracts were fractionated on Superose 6 10/300 GL column using an ÄKTA FPLC system before and after RNase treatment. Precipitated fractions were separated on SDS gels (fractions 1–10, 100%, fractions 12–24, 33%) and blots were stained with CBB. Complexes are indicated above and were deduced from Olinares *et al.* (2010). The used antibodies are shown on the left. Membranes were reprobed. PAC signals from previous immunodetection are marked by asterisks. LMWC, low-molecular-weight complexes.

To investigate PAC complexes *in vivo*, native stroma extracts were fractionated by SEC and investigated by immunoblotting. Columns and conditions were chosen to separate complexes in the range up to 5 MDa allowing chloroplast ribosomes to be resolved in different assembly and functional states (Olinares *et al.*, 2010). Under those conditions, PAC was found to form complexes ranging from ~75 kDa to ~3 MDa (Figure 3c). Such broad size range migration behaviour has been reported for other well characterized RNA-binding proteins with distinct molecular function in chloroplasts (Beick *et al.*, 2008; Pfalz *et al.*, 2009; Stoppel *et al.*, 2012; Hammani and Barkan, 2014). Three PAC protein peaks were observed and sizes were estimated according to marker protein migration: the first in the low-molecular-weight range (~75–158 kDa, fractions 18–20), the second roughly corresponding to the size of Rubisco (~550 kDa, fractions 10–14), and the third ranging from 1 to 3 MDa (fractions 3–8) (Figure 3c). The latter disappeared upon RNase treatment indicating that it contains RNA. Moreover, these PAC megadalton complexes co-migrated with ribosomal particles, as shown by immunoblot analyses with PRPL11 and PRPS5 antibodies, indicating that high-molecular-weight PAC complexes are likely to contain ribosomal RNA. Interestingly, the PAC peak in the MDa range (fractions 7 and 8) did not correspond to the peak of the 50S subunit (fractions 5 and 6) but co-migrated with RNase-sensitive ribosomal particles, indicating that PAC is not associated with mature 50S subunits but with rRNA of premature, not fully assembled and RNase-sensitive ribosomes. In general, the highly structured nature of rRNAs and the fact that mature ribosomal subunits are tightly packed with proteins makes ribosomal RNA less accessible for RNases (Williams and Barkan, 2003). On the contrary, naked rRNAs and rRNA in partially assembled ribosomal subunits are more accessible, thus more sensitive to nucleolytic attacks. Hence, ribosomal particles that co-migrated with PAC are likely to resemble pre-assembled states of the 50S ribosomal subunit, as they are RNase-sensitive and were found in fractions, in which immature ribosomes were identified using SEC and proteomics approaches (Olinares *et al.*, 2010). The other two PAC complexes, migrating at ~550 and ~75–158 kDa, were not affected by the RNase treatment, suggesting that those are either RNA-free homo- and/or heteromultimers or contain tightly bound and/or compactly structured RNA fragments, which are barely accessible by RNases.

PAC Associates with a subset of plastid RNAs *in vivo*

Nucleoids are the site of ribosome biogenesis (Bohne, 2014) and RNA metabolism in general. Given that *pac* mutants show alterations of plastid RNAs and that PAC resides in plastid nucleoids, where it forms RNA-containing complexes, points to a direct role of PAC in RNA metabolism and/or ribosome biogenesis. Therefore, RNA-coimmunoprecipitation coupled with RNA deep sequencing (RIP-seq) was conducted to find RNAs associated with PAC *in vivo*. Two replicates were included to identify bound RNAs using WT stroma extracts and PAC-specific, immunoprecipitation (IP)-capable antibodies. Antibodies against HCF101, another soluble plastid protein or the pre-immune serum were used in negative control experiments. RNAs from *psbK–psbI*, *psbM–trnD*, *rbcL–accD*, *rrn23*, *ndhF/D*, and *ycf1* were enriched in the precipitate (Figure 4a). Slot-blot analysis of PAC-immunoprecipitated RNAs confirmed strong association with *psbK–psbI*, *rrn23*, *ndhF* and *ndhD*, as well as a slight enrichment of *psbM* and *ycf1*, but failed to reproduce association with *rbcL* (Figure 4b), indicating that enrichment of *rbcL–accD* is probably an artefact. Such artefacts have been described also in RIP-chip assays (Ferrari *et al.*, 2017). The enrichment of the target RNAs *psbK–psbI*, *rrn23*, *ndhF* and *ndhD* was also unambiguously confirmed by a repeated RIP-seq experiment (Figure S2).

PAC binds RNA *in vitro*

As PAC associates with 23S rRNA of most likely pre-assembled 50S ribosomal particles and *pac* mutants exhibited pronounced defects in ribosomal RNA abundance along with translation impairments that mostly account for the mutant phenotype, it is reasonable to assume that *rrn23* is one of the primary targets of PAC. Thus, we next aimed to fine-map *in vivo* the region within *rrn23* which is associated with PAC by immunoprecipitation and slot-blot hybridization using five 80-nt probes along the *rrn23* sequence. The strongest enrichment was detected with the internal *rrn23-3* probe, indicating that the PAC interaction site is located in the central part of the *rrn23* transcript (Figure 5a). As this association might be indirect via another protein, we next tested whether PAC can bind to this rRNA directly *in vitro*. For that purpose, recombinant PAC (rPAC) was proteolytically cleaved from the MBP moiety, subjected to a size exclusion purification step and used afterwards in an electrophoretic mobility shift assay (EMSA) with RNAs of ~150 nt length produced by *in vitro* transcription. rPAC was able to bind directly and with specificity to the middle part of *rrn23*, but did not recognize a probe from the 3′ region of *rrn23* (Figure 5b), indicating that PAC is indeed an RNA-binding protein. This result is further supported by the structural model of PAC generated by I-TASSER (http://zhanglab.ccmb.med.umich.edu), which shows that the mature PAC protein consists of helical hairpins (Figure 5c), a common characteristic of RNA-binding proteins containing PPR, OPR, half-a-tetratricopeptide (HAT) and PUF domains (Paulson and Tong, 2012; Hall, 2016).

PAC contributes to the biogenesis of the 50S ribosomal subunit

The above results demonstrated that PAC is associated with the 23S rRNA *in vivo* and can bind to it *in vitro*.

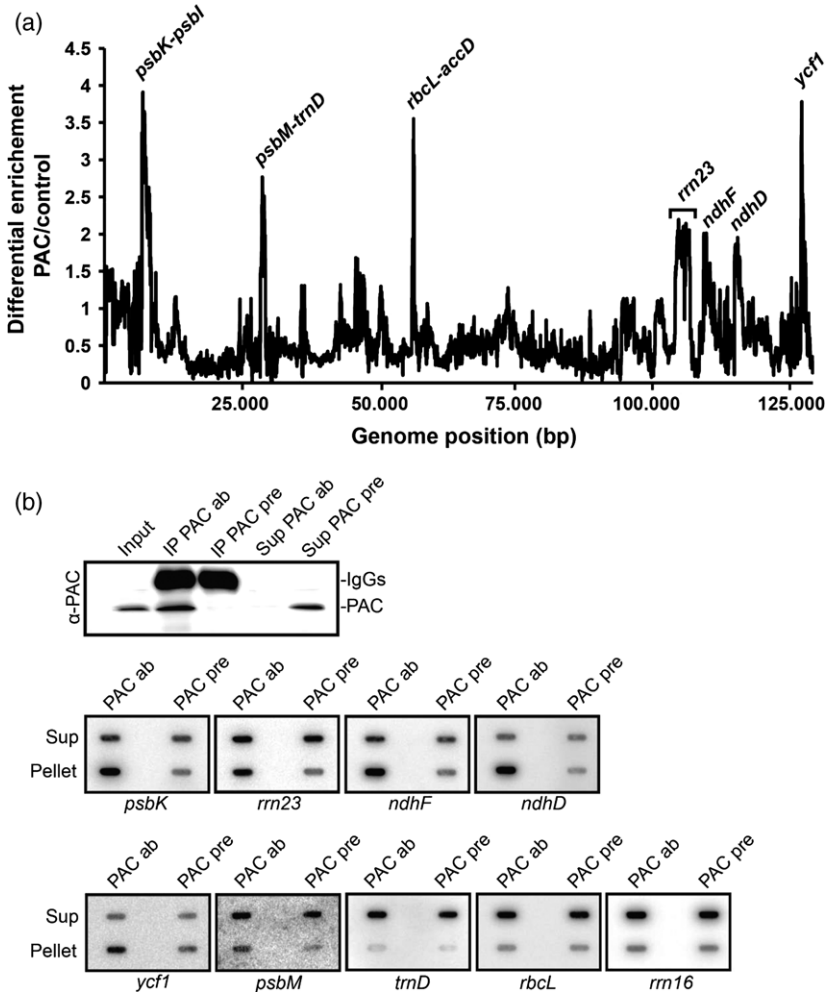


Figure 4. Co-Immunoprecipitation to identify RNAs associated with PAC *in vivo*.
(a) RIP-seq analysis. The differential enrichment, calculated by dividing the reads resulting from the IP with PAC antibodies by the reads of the control (pre-serum), is plotted according to the plastome position. Highly enriched RNAs are labeled with the locus name. The data used to generate this figure are provided in Table S2. The reads for the rRNAs (genome position 101.000–100.750) were normalized to the 16S rRNA.
(b) Slot-blot validation of identified targets. To confirm the RIP-seq data, slot-blot hybridizations were performed. Following immunoprecipitation (upper panel), pellet and supernatant RNAs were isolated, slot-blotted and hybridized with the indicated probes. Antibodies (ab); Pre-immune serum (pre); Supernatant (Sup).

Importantly, in *pac* mutants, levels of ribosomal RNAs are reduced but not altered in their processing patterns (Figures 2 and S3), indicating that PAC is not an rRNA processing factor. Besides rRNA processing, ribosome assembly is another important step of the ribosome maturation process, which requires auxiliary factors that are either associated or interact transiently with the rRNA or proteins of ribosomal subunits. To address the question whether biogenesis of the 50S subunit is affected, stroma extracts from WT and *PAC* RNAi plants, in which *PAC* transcripts were reduced to less than 10%, while PAC proteins accumulated to less than 25% of WT levels (Figure S1a), were subjected to SEC under the same condition as mentioned above. In the stroma of *PAC* RNAi plants, the ~550 kDa PAC complex (fractions 10–14) accumulated to similar levels as in the WT but the ~75–158 kDa complex (fractions 18–20) was greatly reduced (Figure 6). Interestingly, the RNA-containing high-molecular-weight PAC complexes (fractions 3–8), which co-migrated with premature ribosomal particles, were under the limit of detection. This might suggest that Rubisco-sized PAC complexes are rapidly and early formed and remain stable, whereas the other two complexes might exhibit some dynamics, e.g. the ~74–158 kDa PAC particles could represent a pool of homomers, which only associate with 23S rRNA to form MDa complexes when they are needed. As levels of this pool are highly reduced in PAC RNAi plants also the MDa RNA-containing complexes could not be formed.

Also amounts of PRPL11 as part of the 50S ribosomal subunit were significantly reduced, while the PRPS5 protein of the 30S particle accumulated to WT levels, suggesting that the biogenesis of the 50S ribosomal subunit is affected in *pac*. It is worth noting that levels of PAC and proteins of the 50S ribosomal subunit might be even stronger reduced than deduced from immunoblot analysis, as equal amounts of stromal protein were loaded and Rubisco levels in *PAC* RNAi plants are reduced to roughly 50% (see loading controls in Figures 1c and S1a). Given that Rubisco is the most abundant protein in the stroma, normalizing to total protein causes overloading of otherwise underrepresented proteins in *PAC* RNAi lines.

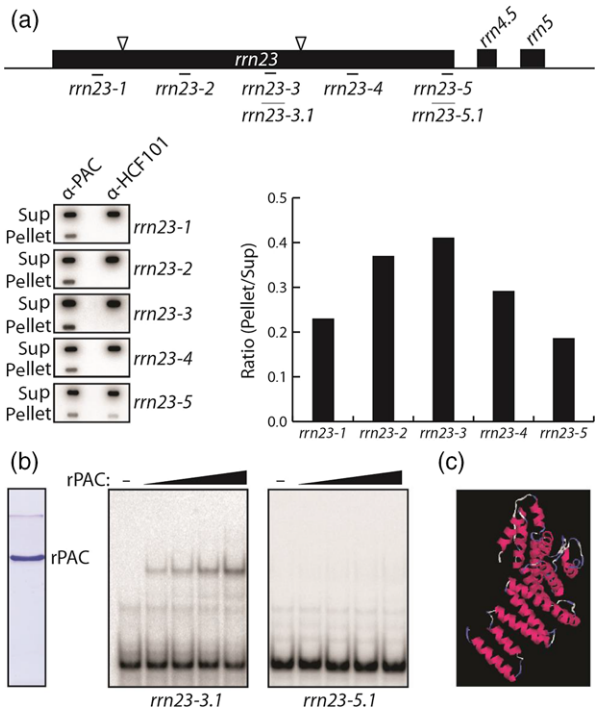


Figure 5. Fine mapping of the 23S RNA target and *in vitro* binding.
(a) Slot-blot analysis of immunoprecipitated RNAs. Probes are shown as black lines and positions along the 23S RNA are indicated; triangles indicate the hidden breaks (upper panel). Hybridization was performed with 80-mer probes specific for the 23S RNA (lower left panel). Signals were quantified (ImageQuant) and the ratio of pellet and supernatant is graphed (Lower right panel).
(b) Recombinant PAC binds with specificity to the middle part of the 23S RNA *in vitro*. Recombinant MBP-PAC was purified and MBP was cleaved off (left panel). Increasing concentrations of mature recombinant PAC were used in EMSA experiments with the RNA probes *rrn23-3.1* and *rrn23-5.1*, produced by *in vitro* transcription (for primer information see Table S1) (right panel). Positions of 23S RNA probes are shown in (a) as gray lines.
(c) 3D model of PAC generated by I-TASSER.

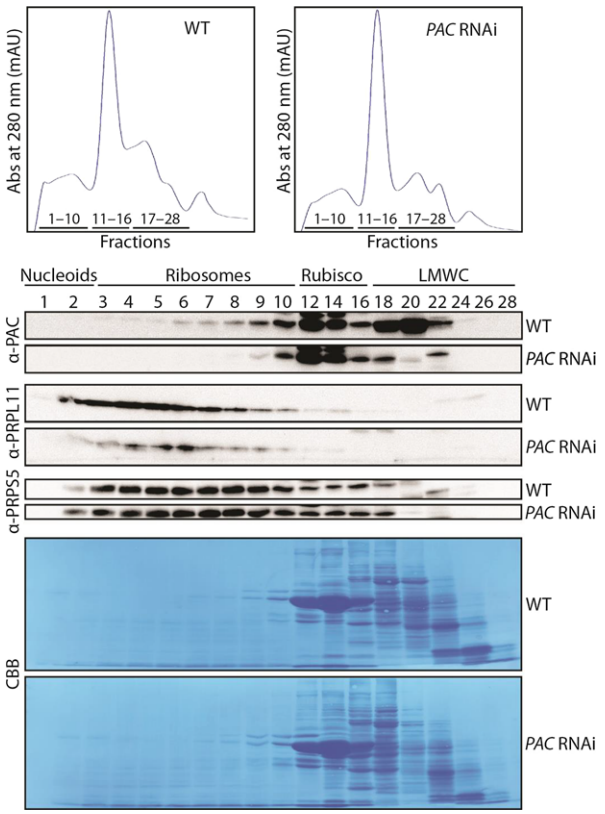


Figure 6. Size exclusion chromatography of stroma extracts from *PAC* RNAi plants.
WT and *PAC* RNAi stroma was fractionated as described in Figure 3(c). Chromatograms of the runs are shown above. Used antibodies, fractions and deduced complexes are indicated. Filters were stained with Coomassie Brilliant Blue (lower panel). [Colour figure can be viewed at wileyonlinelibrary.com].

To further test the hypothesis that PAC influences the biogenesis of the 50S ribosomal subunit, we assayed the effect of PAC on the structure of the 23S rRNA *in vivo* by probing with dimethyl sulfate (DMS) – a small reactive chemical, which can penetrate plant tissues and membranes quickly and easily (Hoffmann, 1980). DMS methylates adenines and cytosines only in single stranded RNAs. RNA regions that are less structured or not protected by proteins are more accessible, hence, show higher DMS reactivity (Lempereur *et al.*, 1985). In primer extension assays, DMS-methylated nucleotides prevent base pairing, thus reverse transcription is inhibited and strand extension pauses, resulting in accumulation of shorter primer extension products.

WT, pComp, *PAC* RNAi and *pac* plants were treated with DMS. Total RNA was isolated and modified 23S rRNA was assayed by primer extension using two oligos: one binding at the very 3′ end of the processed 23S rRNA, the second binding near the central region of the 23S rRNA. Reduced levels of PAC in pComp, *PAC* RNAi, and its absence in *pac* plants resulted in differential DMS methylation of several nucleotides as compared with the DMS-treated as well as untreated WT. Although levels of rRNAs are considerably reduced in pComp, *PAC* RNAi, and *pac* (Figure S3), distinctly higher reactivity in those lines was observed near the 3′ region of processed 23S rRNA, indicating that this region is less structured and/or less associated with ribosomal proteins compared with the WT (Figure 7) pointing to 50S maturation defects in *pac* mutant lines.

DISCUSSION

In our previous study, we showed that knockouts of the nuclear PAC gene are defective in general chloroplast gene expression with pleiotropic effects in the RNA transcript patterns and abundance (Meurer *et al.*, 1998). Generally, in pale mutants defective in gene expression, NEP and PEP transcripts are up- or downregulated, respectively, and processing patterns differ from the WT as it is the case in

pac (Cho *et al.*, 2009). Therefore, it is difficult to determine the primary defects in such mutants unless the direct RNA targets are known. Here we made effort to identify the main RNA targets and deduce the function of the PAC protein.

The genetic and molecular data presented here pointed out that PAC is an RNA-binding protein localized in plastid nucleoids. It is associated *in vivo* with the 23S rRNA as well as with *psbK–psbI*, *ndhF* and *ndhD* mRNAs, and forms high-molecular-weight complexes. Importantly, the pronounced impact of PAC on the biogenesis of plastid ribosomes severely influences translation, leaf coloration, plant growth and leads to seedling lethality.

PAC binds to the plastid 23S rRNA and assists in the biogenesis of 50S ribosomes

RIP-seq and slot-blot hybridization identified 23S rRNA as well as *psbK–psbI*, *ndhF* and *ndhD* messages as RNA targets of PAC. The albino phenotype, the dramatic decrease of especially the plastid 23S rRNAs, the severely impaired *de novo* translation, the aberrant polysome loading and the altered 23S rRNA structure in chloroplasts, all point to an important role of PAC in ribosome biogenesis. Several

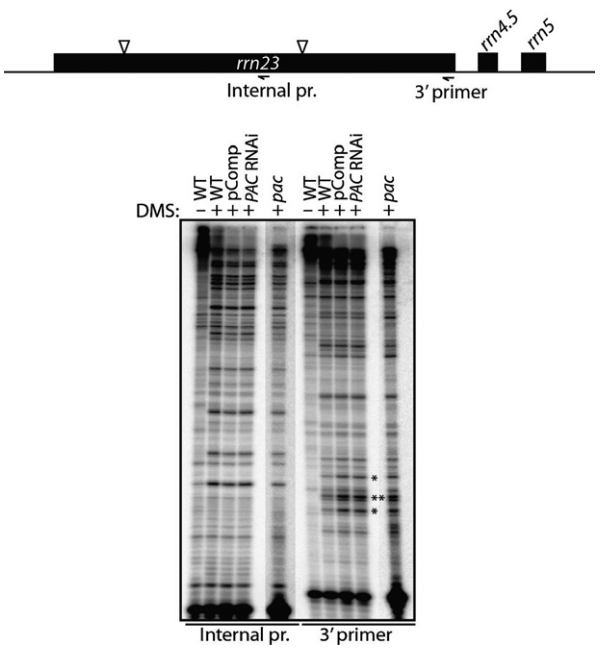


Figure 7. DMS structure probing of 23S RNA.
RNA was extracted from DMS-treated leaves and used in a primer extension assay. Untreated WT leaf RNA was used as a negative control. Fragments present in all samples represent reverse transcription abortion products resulting from strong RNA secondary structures or processing sites. Bands occurring only in DMS-treated samples correspond to methylation events in unstructured or free-of-protein RNA segments. Differences between DMS-treated WT and pComp, *PAC* RNAi as well as *pac* samples are marked by asterisks. Strong effects are indicated by double asterisks. Due to reduced levels of plastid rRNA in *pac*, the *pac* lane was overexposed for a better comparison. Primer positions are indicated in the graph above. Internal primer: 6_23S_Rev; 3′ primer: Rev_7_23S.

mutants, e.g. *ppr2*, *ppr4*, *ppr5*, *crs2-1*, and *rnc1* (Jenkins and Barkan, 2001; Williams and Barkan, 2003; Schmitz-Linneweber *et al.*, 2006; Watkins *et al.*, 2007; Beick *et al.*, 2008), exhibiting pale to albinotic phenotypes have been reported to lack plastid ribosomes, as it is the case in *pac* mutants. In many of these mutants, it is difficult to assess whether the lack of chloroplast ribosomes is a primary or a secondary effect resulting from defects in rRNA processing, splicing or stability of tRNAs or of transcripts encoding proteins of the translational machinery. In contrast, in other mutants such as *rap*, *rh22*, *rhon1* and *sot1*, a direct interaction with chloroplast rRNAs has been demonstrated (Chi *et al.*, 2012, 2014; Stoppel *et al.*, 2012; Kleinknecht *et al.*, 2014; Wu *et al.*, 2016). Similarly, in *pac* mutants the lack of plastid rRNAs appears to be a direct consequence of the absence of PAC proteins, as we found PAC to be associated with the 23S rRNA *in vivo* and to be able to bind to it *in vitro* (Figures 4 and 5). Defects in plastid rRNAs would lead to impaired translation of chloroplast mRNAs, including those of the PEP, which in turn results in general aberrant transcript abundance and patterns consequently causing albinism (Legen *et al.*, 2002). Thus, based on our results, it stands to reason that the major target of PAC is the 23S rRNA and one of its functions is the assistance in the biogenesis of the 50S ribosomal subunit, as aberrant ribosomes are likely to account for the pleiotropic *pac* mutant phenotype.

Five proteins have been reported so far to play a role in ribosome assembly in chloroplasts, three GTPases: DER (Jeon *et al.*, 2014), ObgC (Bang *et al.*, 2012), and RIF1 (Liu *et al.*, 2010), and two DEAD-box RNA helicases: RH3 (Asakura *et al.*, 2012) and RH22 (Chi *et al.*, 2012). While homologs of the GTPases exist in bacteria, no clear homologs could be found for the two helicases, which were both important for the biogenesis of the large ribosomal subunit. Similar to RH3 and RH22, and to the so called plastid-specific ribosomal proteins (PSRPs) (Tiller *et al.*, 2012), which basically represent a subset of ribosome-associated proteins identified by proteomics approaches of purified plastid ribosomes from spinach (Yamaguchi and Subramanian, 2000, 2003), PAC seems to have been acquired rather late during endosymbiosis. One reason for that may be the fact that higher plants, in contrast to bacteria, are exposed to changing environmental conditions and need to adjust the plastid RNA metabolism and translation according to external stimuli (Cho *et al.*, 2009; Chotewutmontri and Barkan, 2016). It is therefore possible that plant-specific RNA-binding proteins occurred to regulate gene expression and translation according to the need of the plant cell in response to environmental changes. This assumption would imply that the expression of RNA-binding factors themselves must be under environmental control. In preliminary experiments we investigated a presumable redox-, temperature- and light-dependent expression of *PAC*. It

appeared that *PAC* transcripts were severely downregulated upon oxidative stress, heat treatment under different light regimes, such as red, far red and high light conditions, while cold and DCMU treatment had no measurable effect on *PAC* expression (Figure S4).

Generally, chloroplasts contain protein-synthesizing systems that are considered to be very similar to those of bacteria, which are by far much better studied and understood. Beside the ribosomal proteins, three other non-ribosomal protein classes are needed for correct folding of rRNAs in *E. coli*: RNA chaperones, RNA helicases and ribosome-dependent GTPases (Kaczanowska and Ryden-Aulin, 2007). As PAC does not possess any motifs characteristic of DEAD-box helicases or GTPases, it is possible that it might belong to a novel class of RNA chaperones. In general, *E. coli* ribosomal subunits can be reconstituted *in vitro* without the need of any other factors (Williamson, 2003). However, to facilitate, speed up and optimize ribosome assembly additional non-ribosomal proteins are required in eubacteria and obviously also in chloroplasts. RNA chaperones are needed to destabilize misfolded structures of long RNA molecules such as 23S RNA allowing the RNA to refold and form the correct structure (Lorsch, 2002). As PAC impacts the 23S rRNA structure *in vivo*, it is possible that PAC represent an 23S rRNA chaperon. Although the PAC binding site is located in the central part of the 23S rRNA, the most pronounced structural changes were located at its 3′ end as demonstrated by *in vivo* DMS probing (Figure 7). Given that the assembly of the *E. coli* 23S rRNA does not follow the 5′ to 3′ transcription direction (Kaczanowska and Ryden-Aulin, 2007), it is therefore possible that during assembly in the tertiary context PAC bound to the central part of the 23S rRNA is in close proximity to 3′ end of the 23S rRNA and can influence the RNA structure through interaction with other factors or ribosomal proteins, which act on the 3′ end of the 23S rRNA. Alternatively, changes in the structure at the 3′ end of the 23S rRNA could also be a direct consequence of the binding of PAC farther upstream, as such structural changes of RNAs downstream of protein binding sites have already been demonstrated (Prikryl *et al.*, 2011; Hammani *et al.*, 2012).

We next considered the function of PAC with respect to the other RNA targets: *psbK–psbI*, *ndhF* and *ndhD*. Effects of the disruption of the *psbK* gene have only been studied in cyanobacteria and green algae so far. *PsbK* mutants in *Synechocystis sp.* PCC 6803 and in *Thermosynechococcus elongatus* BP-1 showed photoautotrophic growth comparable to the WT (Zhang *et al.*, 1993; Sugimoto and Takahashi, 2003). In *Chlamydomonas reinhardtii*, disruption of the *PsbK* gene led to destabilization of the PSII reaction center complex (Takahashi *et al.*, 1994). In cyanobacteria, *psbK* is transcribed as a monocistronic message (Zhang *et al.*, 1993), while in tobacco it is part of the *psbK–psbI* Rubisco *trnG* gene cluster (Meng *et al.*, 1991). In *pac* mutants, levels of all four transcript forms – the tricistronic *psbK–psbI–trnG* (2.6 kb), the dicistronic *psbK–psbI* (1.3 kb), and the two forms of monocistronic *psbK* transcripts (0.4 and 0.7 kb) – were notably reduced as compared with the WT (Figure S3). Also levels of unprocessed and processed *psbI* and *trnG* were strongly reduced (Figure S5a). To consider the effect of reduced *psbI* and *psbK* transcript levels, we aimed to investigate the formation of photosynthetic complexes in *PAC* RNAi lines by blue native PAGE. It appeared that in *PAC* RNAi lines the PSII supercomplexes and the PSI-NDH complex were slightly less abundant than in the WT, which would be consistent with reduced levels of *psbI* (Schwenkert *et al.*, 2006), while PSII monomer, cytochrome b_6f complex, LHCII assembly, and LHCII trimer accumulated to nearly WT levels (Figure S5b). Nevertheless, it seems quite unlikely that effects on photosynthetic complex stability could account significantly for the *pac* phenotype. An additional contribution to the severe *pac* phenotype is rather likely to result from a decrease of *trnG*, which encodes the tRNA(Gly) (UCC). In maize, disruption of PPR5 leads to decrease in both, spliced and non-spliced *trnG*-UCC RNAs accompanied by deficient chloroplast ribosomes and seedling lethality (Beick *et al.*, 2008). Therefore, we cannot exclude that reduced levels of this tRNA beside impaired 50S ribosome biogenesis, at least to a certain extent, also affects translation in *pac* and contributes to the phenotype.

Interestingly, levels of the third RNA target of PAC, *ndhD*, were significantly increased in *pac* mutants, therefore, it is possible that PAC regulates amounts of NdhD by destabilizing the *ndhD* RNA, so that a decrease or an absence of PAC would lead to increased *ndhD* transcript levels (Figures 2 and S3). Alternatively, decrease of PEP due to impaired translation might be compensated by increasing the amount and/or activity of the NEP, as has been described in transplastomic *Δrpo* and plastid gene expression mutants (Legen *et al.*, 2002; Cho *et al.*, 2009). This up-regulation could be triggered by retrograde signals when translation is strongly reduced (*pac*) but not completely abolished as is the case in lincomycin-treated plants. Thus, RNAs generated by the NEP, such as *ndhD*, would overaccumulate. In favor of this theory, levels of *rpoB* transcripts – one of the few plastid genes transcribed exclusively by the NEP (Börner *et al.*, 2015) – were greatly increased as well. Although levels of polycistronic *ndhD* transcripts (~2.6 kb) were considerably higher than in the WT, the monocistronic message (~1.4 kb) was decreased (Figure S3), rather pointing to a role of PAC in *ndhD* processing. Levels of *ndhF* were reduced in *pac* and a high-molecular-weight band appeared, indicating stability or processing defects. However, given that Arabidopsis and tobacco mutants defective in NDH complex accumulation exhibited only a very subtle phenotype, even when subjected to various stress conditions (Barth and Krause, 2002; Muraoka *et al.*, 2006) it seems unlikely that defective NDH

complexes in *pac* contributed significantly to the overall phenotypic alterations in *pac*.

The fact that retrograde signals are not inhibited in *pac* is further supported by the observation that the RNAs of the nucleus-encoded light-regulated and photosynthesis-related genes *rbcS*, *psbO* and *petC* accumulated in *pac* to WT levels (Meurer *et al.*, 1998) implying that a certain developmental stage of the chloroplast was reached in *pac* presumably due to a residual PAC-independent translation. PAC is expressed at early stages of development (Figure S5c), in which highest translational activity is needed to achieve the required expression of plastid genes. Binding of PAC to *psbK*, *ndhF* and *ndhD* mRNAs adds a further layer of complexity to the function of PAC in RNA metabolism in addition to its function in biogenesis of the 50S subunits, which could be important for fine tuning of plastid gene expression.

EXPERIMENTAL PROCEDURES

Plant cultivation

The Arabidopsis *pac* allele used in this work corresponds to the *pac-2* T-DNA insertion line described previously (Grevelding *et al.*, 1996). Surface-sterilized seeds were grown on " MS medium supplemented with 1.5% sucrose in continuous light with a photon flux density (PFD) of 20 ± 30 μmol photons m^{-2} sec^{-1} at 23°C. For inhibition of plastid translation, 0.45 g L^{-1} lincomycin was added to the MS medium. For regulation experiments, Arabidopsis WT plants were grown on soil in a 16 h-light (20°C)/8 h-dark (18°C) cycle with a PFD of 50 μmol photons m^{-2} sec^{-1} for 3 weeks. For the treatments, the plants were carefully transferred to reaction tubes, with the roots into liquid medium and subjected to growth light (GL 24 h), dark (24 h), red light (660 nm, 20 μmol photons m^{-2} sec^{-1}, 24 h), far red light (720 nm, 1 μmol photons m^{-2} sec^{-1}, 24 h), cold (4°C, GL, 24 h), heat (40°C, GL, 1 h), DCMU (50 μM, GL, 24 h), Paraquat (MV 50 μM, GL, 24 h). For co-immunoprecipitation analyses WT plants and complemented lines were grown on soil in a 12 h-light (20°)/12 h-dark (18°) cycle for 2 weeks with a PFD of 100 μmol photons m^{-2} sec^{-1}.

EMSA

EMSA experiments were performed as described previously (Manavski *et al.*, 2015) with some changes. Binding reactions (20 μl) contained 40 mM Tris–HCl pH 8.0; 30 mM KCl; 1 mM $MgCl_2$; 0.01% w/v NP40; 1 mM DTT, 50 μg ml^{-1} heparin, trace amounts of radiolabelled probes and increasing concentrations of recombinant PAC (100, 200, 400, 800 nM). Probes were produced by *in vitro* transcription using PCR products as described (Manavski *et al.*, 2015). For primer information see Table S1.

In vivo labeling

Two-week-old Arabidopsis plants grown at 30 μmol photons m^{-2} sec^{-1} were harvested and roots were removed. After incubation in 400 μl labeling buffer (1 mM phosphate buffer KH_2PO_4/K_2HPO_4 pH 6.3; 20 μg ml^{-1} cycloheximide; 0.1% w/v tween 20) for 30 min at RT, 40 μl of 35S-methionine were added. Samples were vacuum infiltrated three times for 20 sec, 500 μl of labeling buffer were added and incubated for 1 h at room temperature in ambient light. Plants were washed in labeling buffer and water, and frozen in liquid nitrogen. Total protein was extracted from ground tissue with the same volume of isolation buffer containing 100 mM Tris pH 8.0; 5 mM EGTA pH 8.0; 5 mM EDTA; 150 mM NaCl; 0.5% Triton X-100, 10 mM DTT and a protease inhibitor cocktail (Roche, Basel, Switzerland). Following centrifugation (4°/15 min/13 000 *g*), protein concentrations were determined via Bradford and equal amounts of total protein were separated on a 10% polyacrylamide gel. Proteins were visualized by Coomassie staining. Gel was dried and signals were detected using the Typhoon Variable Mode Imager (GE Healthcare, Little Chalfont, UK).

Size exclusion chromatography (SEC) and sucrose gradient fractionation

Chloroplasts were isolated from 4-week-old plants as described previously (Stoppel *et al.*, 2012). Chloroplasts were lysed in extraction buffer (10 mM HEPES-KOH, pH 8.0, 5 mM $MgCl_2$, and protease inhibitor cocktail (Roche, Basel, Switzerland) by passing the suspension 20 times through a 0.5 mm needle. Membranes were pelleted by centrifugation at 45 000 *g* for 30 min at 4°C. One milligram of RNase A-treated (100 μg RNase A (Qiagen, Valencia, CA, USA), 1 h on ice.) or untreated stroma extracts were fractionated by SEC using Superose 6 10/300 GL column (GE Healthcare, Little Chalfont, UK) and an ÄKTA FPLC system (Amersham Biosciences, Little Chalfont, Buckinghamshire, UK) as described previously (Olinares *et al.*, 2010). Fractions (0.5 ml) were precipitated with TCA and separated on 10% SDS-PAGE.

SEC of recombinant PAC proteins was performed on a Superdex 200 10/300 GL column (GE Healthcare, Little Chalfont, UK) with buffer containing 100 mM Tris–HCL pH 8.0, 150 mM NaCl, 1 mM EDTA and flow rate of 0.7 ml min^{-1}. Size of complexes were determined using the Gel Filtration HMW Calibration Kit (GE Healthcare, Little Chalfont, UK).

RIP-seq and slot-blot

For RIP-seq, 2-week-old WT plants were used. Chloroplast isolation and stroma extract preparations were performed as described previously (Schmitz-Linneweber *et al.*, 2005). For deep sequencing of RNAs co-immunoprecipitating with PAC, libraries were generated with the ScriptSeq v2 RNA-seq Library Preparation Kit (Epicentre) according to manufacturer's instructions. Sequencing (2× 150 bp, v2 chemistry) was performed on a MiSeq sequencer (Illumina) yielding 4.9 Mio and 5.4 Mio primary reads for PAC antibodies and pre-immune serum control, respectively. Primary reads were aligned to the Arabidopsis chloroplast genome (accession number NC_000932.1) using CLC Genomics Workbench 6.5.1 (Qiagen, Valencia, CA, USA) 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. Numeric data (reads/nucleotides) are provided in Table S2. The analogous second RIP-seq experiment yielded 12 Mio and 8.4 Mio primary reads for HCF101 control and PAC antibodies, respectively.

Slot-blot experiments were performed as described recently (Manavski *et al.*, 2015). Primers for PCR probes and 80-mers are listed in Table S1.

Polysome loading assay

Polysome loading experiments were performed as described (Barkan, 1993).

RNA gel blot analysis and rRNA quantification

RNA extraction, electrophoresis, transfer, and probe labeling were performed as described recently (Manavski *et al.*, 2015). Blots

were stripped and reprobed. For primer information see Table S1. rRNA quantification was performed on Bioanalyzer using 1 μg of total RNA and RNA 6000 Nano Kit (Agilent, Santa Clara, CA, USA).

Antibodies

PAC antibodies (Stoppel *et al.*, 2011) were used at 1:5000 dilutions and plastid (P)RPS5 antibodies (Uniplastomic) at dilution 1:1000. Anti-PRPL11 antibodies were raised against peptides GVEKGSKDPQQDKVG and MGIDIDPPILEPKKK. Peptide production, rabbit immunization and IgG purification were preformed by Eurogentec. For immunodetection, dilutions of 1:1000 were used. PAC pre-immune serum and HCF101 antibodies (Schwenkert *et al.*, 2010) were used as negative controls in RIP-seq experiments.

Generation of *PAC* RNAi lines and mutant complementation

A 299-bp fragment of the *PAC* sequence was PCR-amplified with primers (PAC RNAi 1c/PAC RNAi 2), cloned into pENTR/D-TOPO (Invitrogen, Carlsbad, CA, USA), and transferred into the binary Gateway vector pB7GWIWG2(I) (Plant System Biology). The resulting construct was introduced into *Agrobacterium tumefaciens* GV3101 (pMP90RK) and used for transformation of WT plants via floral dip method (Clough and Bent, 1998). Transformed seedlings were selected with BASTA. Plants able to grow on soil and showing a pale-green phenotype were selected.

The PAC cDNA starting six nucleotides upstream of the start codon and lacking the stop codon was amplified with primers (Fw PAC pAUL11/Rev PAC pAUL11) and cloned into pAUL11 (Lyska *et al.*, 2013) using the Gateway technique (Invitrogen, Carlsbad, CA, USA). This allowed HA-Strep-tagging *in vivo* (Stoppel *et al.*, 2012). Heterozygous *pac* plants were transformed by floral dip method (Clough and Bent, 1998) and transgenic T_1-plants were selected with BASTA.

Blue native PAGE

Thylakoid isolation, solubilization, and BN-PAGE were performed as described (Schwenkert *et al.*, 2006). Lanes were denatured for 30 min in denaturation buffer (2% SDS; 66.7 mM Na_2CO_3; 100 mM DTT) and run in the second dimension in Tris-Tricine PAGE (10% acrylamide and 4 M urea). Gels were stained with colloidal Coomassie Brilliant Blue as described previously (Dyballa and Metzger, 2009).

Recombinant protein expression and purification

The coding sequence of PAC lacking the transit peptide for chloroplast import (aa 23–313) was amplified from an Arabidopsis cDNA library using oligonucleotides (Fw *Bam*HI PAC Rev/Rev *Sal*I PAC Strep), which introduced *Bam*HI and *Sal*I restriction sites as well as a Strep-Tag II. PCR products were digested (*Bam*HI/*Sal*I) and cloned into the BamHI/*Sal*I sites of pMAL-TEV vector (kindly provided by Alice Barkan). Expression and purification were performed as described (Chi *et al.*, 2014). Purified proteins were used directly for SEC, EMSA or stored up to 2 weeks in the fridge.

Subcellular localization studies

The entire coding region of PAC was PCR-amplified from an Arabidopsis cDNA library using oligonucleotides (Fw PAC CACC/Rev PAC – stop). Products were cloned into pENTR/D-TOPO (Invitrogen, Carlsbad, CA, USA) and subcloned into the binary Gateway vector pB7YWG2 (Plant System Biology). pENTR/D-TOPO::RAP (kindly provided by Jörg Nickelsen (Kleinknecht *et al.*, 2014)) was subcloned into pH7CWG2 via the Gateway technique.

Protoplasts were prepared as described previously (Schweiger *et al.*, 2012). 10 μg of pB7YWG2:PAC and pH7CWG2:RAP, each, were mixed with 100 μl of protoplast suspension and 125 μl freshly-made PEG solution (2 g PEG4000, 1.75 ml H_2O, 1 ml 1 M mannitol, 0.5 ml 1 M $Ca(NO_3)_2$) was added. After incubation for 7.5 min, W5 buffer (150 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 2 mM MES pH 5.7, Osm. 550–580) was added to 2.5 ml. Protoplasts were incubated at 23°C overnight in the dark and fluorescence signals were analyzed on a confocal laser scanning microscope as described (Schweiger *et al.*, 2012).

DMS probing

Leaves from 2-week-old WT, pComp, *PAC* RNAi, and *pac* plants were harvested and incubated for 4 min in 1% of DMS freshly prepared in water. DMS solution was then replaced by ice-cold water and samples were chilled on ice for 5 min. Leaves were frozen in liquid nitrogen and total RNA was extracted using Tri-Pure Isolation Reagent (Roche, Basel, Switzerland) as recommended by the manufacturer. In total, 1 μg of total RNA was used for primer extension assay with primers Rev_7_23s and 4_23S_Rev, and transcriptor reverse transcriptase (Roche, Basel, Switzerland). Primer extension reactions were performed as described (Barkan, 2011) except that ddNTPs were not added. Extension was performed at 50°C. Samples were separated on a pre-run denaturing 8% polyacrylamide gel in 1× Tris/borate/EDTA (TBE) buffer.

ACKNOWLEDGEMENTS

We wish to thank Jörg Nickelsen for providing the RAP entry clone and Kamel Hammani for critical reading of the manuscript. We are grateful to Christian Schmitz-Linneweber for conducting preliminary RIP-chip experiments. We are thankful to Laura Kleinknecht for performing preliminary *in vitro* nucleic acid-binding experiments and Alice Barkan for providing lab space for those experiments. This research was supported by the German Science Foundation (Deutsche Forschungsgemeinschaft; ME1794/6–1 to J.M. and TRR175 A03 to J.M.).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Analysis of *PAC* RNAi, partially complemented and fully complemented *pac* plants.

Figure S2. Enrichment of PAC RNA targets identified by RIP-seq.

Figure S3. RNA gel blot analysis.

Figure S4. Expression of *PAC* under various conditions.

Figure S5. Analysis of *psbI* and *trnG* transcripts. Investigation of photosynthetic complexes in PAC RNAi plants.

Table S1. Primer sequences and probes used in this work.

Table S2. Depth of coverage (reads/nucleotides) of mapped reads on the chloroplast genome.

REFERENCES

Asakura, Y., Galarneau, E., Watkins, K.P., Barkan, A. and van Wijk, K.J. (2012) Chloroplast RH3 DEAD box RNA helicases in maize and

Arabidopsis function in splicing of specific group II introns and affect chloroplast ribosome biogenesis. *Plant Physiol.* **159**, 961–974.

Bang, W.Y., Chen, J., Jeong, I.S. *et al.* (2012) Functional characterization of ObgC in ribosome biogenesis during chloroplast development. *Plant J.* **71**, 122–134.

Barkan, A. (1993) Nuclear mutants of maize with defects in chloroplast polysome assembly have altered chloroplast RNA metabolism. *Plant Cell*, **5**, 389–402.

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

Barth, C. and Krause, G.H. (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. and 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. and Barkan, A. (2015) Large-scale genetic analysis of chloroplast biogenesis in maize. *Biochim. Biophys. Acta*, **1847**, 1004–1016.

Bellaoui, M., Keddie, J.S. and Gruissem, W. (2003) DCL is a plant-specific protein required for plastid ribosomal RNA processing and embryo development. *Plant Mol. Biol.* **53**, 531–543.

Bisanz, C., Begot, L., Carol, P., Perez, P., Bligny, M., Pesey, H., Gallois, J.L., Lerbs-Mache, S. and Mache, R. (2003) The Arabidopsis nuclear DAL gene encodes a chloroplast protein which is required for the maturation of the plastid ribosomal RNAs and is essential for chloroplast differentiation. *Plant Mol. Biol.* **51**, 651–663.

Bohne, A.V. (2014) The nucleoid as a site of rRNA processing and ribosome assembly. *Front. Plant Sci.* **5**, 257.

Bollenbach, T.J., Lange, H., Gutierrez, R., Erhardt, M., Stern, D.B. and Gagliardi, D. (2005) RNR1, a 3′-5′ exoribonuclease belonging to the RNR superfamily, catalyzes 3′ maturation of chloroplast ribosomal RNAs in *Arabidopsis thaliana*. *Nucleic Acids Res.* **33**, 2751–2763.

Börner, T., Aleynikova, A.Y., Zubo, Y.O. and Kusnetsov, V.V. (2015) Chloroplast RNA polymerases: role in chloroplast biogenesis. *Biochim. Biophys. Acta*, **1847**, 761–769.

Chi, W., He, B., Mao, J., Li, Q., Ma, J., Ji, D., Zou, M. and Zhang, L. (2012) The function of RH22, a DEAD RNA helicase, in the biogenesis of the 50S ribosomal subunits of Arabidopsis chloroplasts. *Plant Physiol.* **158**, 693–707.

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

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

Chotewutmontri, P. and Barkan, A. (2016) Dynamics of chloroplast translation during chloroplast differentiation in maize. *PLoS Genet.* **12**, e1006106.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743.

Cortleven, A., Nitschke, S., Klaumunzer, M., Abdelgawad, H., Asard, H., Grimm, B., Riefler, M. and Schmulling, T. (2014) A novel protective function for cytokinin in the light stress response is mediated by the Arabidopsis histidine kinase2 and Arabidopsis histidine kinase3 receptors. *Plant Physiol.* **164**, 1470–1483.

Dyballa, N. and Metzger, S. (2009) Fast and sensitive colloidal coomassie G-250 staining for proteins in polyacrylamide gels. *J. Vis. Exp.* **30**, e1431.

Ferrari, R., Tadini, L., Moratti, F., Lehniger, M.K., Costa, A., Rossi, F., Colombo, M., Masiero, S., Schmitz-Linneweber, C. and Pesaresi, P. (2017) CRP1 Protein: (dis)similarities between *Arabidopsis thaliana* and Zea mays. *Front. Plant Sci.* **8**, 163.

Fristedt, R., Scharff, L.B., Clarke, C.A., Wang, Q., Lin, C., Merchant, S.S. and Bock, R. (2014) RBF1, a plant homolog of the bacterial ribosome-binding factor RbfA, acts in processing of the chloroplast 16S ribosomal RNA. *Plant Physiol.* **164**, 201–215.

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

Grevelding, C., Suter-Crazzolara, C., von Menges, A., Kemper, E., Masterson, R., Schell, J. and Reiss, B. (1996) Characterisation of a new allele of pale cress and its role in greening in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **251**, 532–541.

Hajdukiewicz, P.T., Allison, L.A. and Maliga, P. (1997) The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J.* **16**, 4041–4048.

Hall, T.M. (2016) De-coding and re-coding RNA recognition by PUF and PPR repeat proteins. *Curr. Opin. Struct. Biol.* **36**, 116–121.

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

Hammani, K., Cook, W.B. and Barkan, A. (2012) RNA-binding and RNA remodeling activities of the half-a-tetratricopeptide (HAT) protein HCF107 underlie its effects on gene expression. *Proc. Natl Acad. Sci. USA*, **109**, 5651–5656.

Hoffmann, G.R. (1980) Genetic effects of dimethyl sulfate, diethyl sulfate and related compounds. *Mutat. Res.* **75**, 63–129.

Jenkins, B.D. and Barkan, A. (2001) Recruitment of a peptidyl-tRNA hydrolase as a facilitator of group II intron splicing in chloroplasts. *EMBO J.* **20**, 872–879.

Jeon, Y., Ahn, C.S., Jung, H.J., Kang, H., Park, G.T., Choi, Y., Hwang, J. and Pai, H.S. (2014) DER containing two consecutive GTP-binding domains plays an essential role in chloroplast ribosomal RNA processing and ribosome biogenesis in higher plants. *J. Exp. Bot.* **65**, 117–130.

Kaczanowska, M. and Ryden-Aulin, M. (2007) Ribosome biogenesis and the translation process in Escherichia coli. *Microbiol. Mol. Biol. Rev.* **71**, 477–494.

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

Komatsu, T., Kawaide, H., Saito, C. *et al.* (2010) The chloroplast protein BPG2 functions in brassinosteroid-mediated post-transcriptional accumulation of chloroplast rRNA. *Plant J.* **61**, 409–422.

Lee, K.H., Park, J., Williams, D.S., Xiong, Y., Hwang, I. and Kang, B.H. (2013) Defective chloroplast development inhibits maintenance of normal levels of abscisic acid in a mutant of the Arabidopsis RH3 DEAD-box protein during early post-germination growth. *Plant J.* **73**, 720–732.

Legen, J., Kemp, S., Krause, K., Profanter, B., Herrmann, R.G. and Maier, R.M. (2002) Comparative analysis of plastid transcription profiles of entire plastid chromosomes from tobacco attributed to wild-type and PEP-deficient transcription machineries. *Plant J.* **31**, 171–188.

Lempereur, L., Nicoloso, M., Riehl, N., Ehresmann, C., Ehresmann, B. and Bachellerie, J.P. (1985) Conformation of yeast 18S rRNA. Direct chemical probing of the 5′ domain in ribosomal subunits and in deproteinized RNA by reverse transcriptase mapping of dimethyl sulfate-accessible. *Nucleic Acids Res.* **13**, 8339–8357.

Liu, H., Lau, E., Lam, M.P. *et al.* (2010) OsNOA1/RIF1 is a functional homolog of AtNOA1/RIF1: implication for a highly conserved plant cGTPase essential for chloroplast function. *New Phytol.* **187**, 83–105.

Lorsch, J.R. (2002) RNA chaperones exist and DEAD box proteins get a life. *Cell*, **109**, 797–800.

Lu, Y., Li, C., Wang, H., Chen, H., Berg, H. and Xia, Y. (2011) AtPPR2, an Arabidopsis pentatricopeptide repeat protein, binds to plastid 23S rRNA and plays an important role in the first mitotic division during gametogenesis and in cell proliferation during embryogenesis. *Plant J.* **67**, 13–25.

Lyska, D., Engelmann, K., Meierhoff, K. and Westhoff, P. (2013) pAUL: a gateway-based vector system for adaptive expression and flexible tagging of proteins in Arabidopsis. *PLoS ONE*, **8**, e53787.

Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., Watkins, K.P., Barkan, A. and van Wijk, K.J. (2012) Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant Physiol.* **158**, 156–189.

Manavski, N., Torabi, S., Lezhneva, L., Arif, M.A., Frank, W. and 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.

Meng, B.Y., Wakasugi, T. and Sugiura, M. (1991) Two promoters within the *psbK–psbI–trnG* gene cluster in tobacco chloroplast DNA. *Curr. Genet.* **20**, 259–264.

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

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

Nishimura, K., Ashida, H., Ogawa, T. and Yokota, A. (2010) A DEAD box protein is required for formation of a hidden break in Arabidopsis chloroplast 23S rRNA. *Plant J.* **63**, 766–777.

Olinares, P.D., Ponnala, L. and van Wijk, K.J. (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.

Paulson, A.R. and Tong, L. (2012) Crystal structure of the Rna14-Rna15 complex. *RNA*, **18**, 1154–1162.

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

Prikryl, J., Rojas, M., Schuster, G. and Barkan, A. (2011) Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. *Proc. Natl Acad. Sci. USA*, **108**, 415–420.

Reiter, R.S., Coomber, S.A., Bourett, T.M., Bartley, G.E. and Scolnik, P.A. (1994) Control of leaf and chloroplast development by the Arabidopsis gene pale cress. *Plant Cell*, **6**, 1253–1264.

Schmitz-Linneweber, C., Williams-Carrier, R. and Barkan, A. (2005) RNA immunoprecipitation and microarray analysis show a chloroplast Pentatricopeptide repeat protein to be associated with the 5′ region of mRNAs whose translation it activates. *Plant Cell*, **17**, 2791–2804.

Schmitz-Linneweber, C., Williams-Carrier, R.E., Williams-Voelker, P.M., Kroeger, T.S., Vichas, A. and Barkan, A. (2006) A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast rps12 pre-mRNA. *Plant Cell*, **18**, 2650–2663.

Schweiger, R., Muller, N.C., Schmitt, M.J., Soll, J. and Schwenkert, S. (2012) AtTPR7 is a chaperone-docking protein of the Sec translocon in Arabidopsis. *J. Cell Sci.* **125**, 5196–5207.

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

Schwenkert, S., Netz, D.J., Frazzon, J., Pierik, A.J., Bill, E., Gross, J., Lill, R. and Meurer, J. (2010) Chloroplast HCF101 is a scaffold protein for [4Fe-4S] cluster assembly. *Biochem J.* **425**, 207–214.

Shajani, Z., Sykes, M.T. and Williamson, J.R. (2011) Assembly of bacterial ribosomes. *Annu. Rev. Biochem.* **80**, 501–526.

Stoppel, R. and 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. and 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. and 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. and 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.

Sugimoto, I. and Takahashi, Y. (2003) Evidence that the PsbK polypeptide is associated with the photosystem II core antenna complex CP43. *J. Biol. Chem.* **278**, 45004–45010.

Takahashi, Y., Matsumoto, H., Goldschmidt-Clermont, M. and Rochaix, J.D. (1994) Directed disruption of the Chlamydomonas chloroplast psbK gene destabilizes the photosystem II reaction center complex. *Plant Mol. Biol.* **24**, 779–788.

Tiller, N. and Bock, R. (2014) The translational apparatus of plastids and its role in plant development. *Mol Plant*, **7**, 1105–1120.

Tiller, N., Weingartner, M., Thiele, W., Maximova, E., Schottler, M.A. and Bock, R. (2012) The plastid-specific ribosomal proteins of *Arabidopsis thaliana* can be divided into non-essential proteins and genuine ribosomal proteins. *Plant J.* **69**, 302–316.

Tirlapur, U.K., Dahse, I., Reiss, B., Meurer, J. and Oelmuller, R. (1999) Characterization of the activity of a plastid-targeted green fluorescent protein in Arabidopsis. *Eur. J. Cell Biol.* **78**, 233–240.

Wang, Y., Wang, C., Zheng, M. *et al.* (2016) WHITE PANICLE1, a Val-tRNA synthetase regulating chloroplast ribosome biogenesis in rice, is essential for early chloroplast development. *Plant Physiol.* **170**, 2110–2123.

Watkins, K.P., Kroeger, T.S., Cooke, A.M., Williams-Carrier, R.E., Friso, G., Belcher, S.E., van Wijk, K.J. and Barkan, A. (2007) A ribonuclease III domain protein functions in group II intron splicing in maize chloroplasts. *Plant Cell*, **19**, 2606–2623.

Weis, B.L., Kovacevic, J., Missbach, S. and Schleiff, E. (2015) Plant-specific features of ribosome biogenesis. *Trends Plant Sci.* **20**, 729–740.

Williams, P.M. and Barkan, A. (2003) A chloroplast-localized PPR protein required for plastid ribosome accumulation. *Plant J.* **36**, 675–686.

Williamson, J.R. (2003) After the ribosome structures: how are the subunits assembled? *RNA*, **9**, 165–167.

Wu, W., Liu, S., Ruwe, H. *et al.* (2016) SOT1, a pentatricopeptide repeat protein with a small MutS-related domain, is required for correct processing of plastid 23S-4.5S rRNA precursors in *Arabidopsis thaliana*. *Plant J.* **85**, 607–621.

Yamaguchi, K. and Subramanian, A.R. (2000) The plastid ribosomal proteins – identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast). *J. Biol. Chem.* **275**, 28466–28482.

Yamaguchi, K. and Subramanian, A.R. (2003) Proteomic identification of all plastid-specific ribosomal proteins in higher plant chloroplast 30S ribosomal subunit – PSRP-2 (U1A-type domains), PSRP-3 alpha/beta (ycf65 homologue) and PSRP-4 (Thx homologue). *Eur. J. Biochem.* **270**, 190–205.

Yamaguchi, K., von Knoblauch, K. and Subramanian, A.R. (2000) The plastid ribosomal proteins. Identification of all the proteins in the 30 S subunit of an organelle ribosome (chloroplast). *J. Biol. Chem.* **275**, 28455–28465.

Zerges, W. (2000) Translation in chloroplasts. *Biochimie*, **82**, 583–601.

Zhang, Z.H., Mayes, S.R., Vass, I., Nagy, L. and Barber, J. (1993) Characterization of the psbK locus of Synechocystis sp. PCC 6803 in terms of Photosystem II function. *Photosynth. Res.* **38**, 369–377.

Zubo, Y.O., Yamburenko, M.V., Selivankina, S.Y. *et al.* (2008) Cytokinin stimulates chloroplast transcription in detached barley leaves. *Plant Physiol.* **148**, 1082–1093.

8. DISCUSSION

8.1. Pleiotropic mutants: Primary versus secondary defects

In general, working with Arabidopsis mutant lines which exhibit massive growth and developmental defects leading to albinotic or even lethal phenotypes often complicates the identification of the underlying primary defects caused by the loss of the respective protein, especially due to the restricted amount of plant material and occurring secondary defects. Many mutants somehow impaired in proper chloroplast functions display defects in plastid morphology, photosynthetic performance and thus energy metabolism and often growth defects (Taylor et al., 1987; Romani et al., 2012; Manavski et al., 2015). In several cases such mutants are even embryo-lethal, making an in-depth functional characterization and biochemical analysis almost impossible. However, the dramatic effects of impaired chloroplast function underline the necessity of proper chloroplast performance for the entire plant (Devic, 2008; Bryant et al., 2011). In particular disturbance of plastid translation causes far reaching consequences as the expression of the plastid-encoded polymerase (PEP) is affected which then causes dysregulated transcript levels of plastid-encoded genes, also those for photosynthesis and translation (e.g. genes for ribosomal proteins or tRNAs) (Williams-Carrier et al., 2014). This down-regulation of preferentially PEP-dependent transcripts is observed in PEP-reduced mutants as well as *pep* knockout mutants, in which NEP-dependent genes, such as *RpoB* are upregulated (Börner et al., 2015). Finally, this leads to impaired photosynthesis and the observed pale phenotypes.

Characterizations of *pac* mutants, detected numerous defects including massively impaired photosynthesis and development even leading to seedling lethality (Reiter et al., 1994). Still, for more than 20 years it remained unclear which function the corresponding PAC protein fulfils. Based on the discovery of abnormal plastid transcript patterns and abundance, PAC was proposed to function in chloroplast mRNA maturation and accumulation (Meurer et al., 1998). Nevertheless, the finally identified role as an RNA chaperone acting on the stability of the 23S rRNA and thereby ribosome biogenesis could only be addressed by proving the association of PAC with RNA in megadalton size proteinaceous complexes and its co-migration with ribosomal subunits in different assembly states as well as verifying its association with RNA using RIP-Seq, *in vitro* RNA-binding and structuring studies (Meurer et al., 2017).

Similarly, *pumpkin* mutants exhibited pronounced developmental and pleiotropic defects e.g. improper plastid translation causing a decrease of preferentially PEP-generated transcripts as

well as aberrant transcript patterns. This was initially attributed to the impaired stabilization of transcripts due to the missing PUMPKIN protein acting in RNA metabolism (Schmid et al., 2019). Still it was impossible to estimate if and to which extent the lacking enzymatic function of PUMPKIN in the pyrimidine biosynthesis pathway might contribute to these defects. Only by separately re-introducing the two functions – RNA association or enzymatic activity – we were able to discriminate between real primary defects directly caused by the mutation and secondary defects which accumulated due to interplay with other metabolic processes.

To summarize, even though both mutants show a similar pleiotropic phenotype and disturbance in chloroplast development including impaired translation and abnormal plastid transcript abundance and patterns, in *pac* mutants the defects in translation are indeed a direct consequence of the absence of PAC protein acting in ribosome biogenesis. In contrast, in *pumpkin* mutants, the disturbed nucleotide biosynthesis – due to the lack of the enzymatic function of PUMPKIN – leads to these dramatic secondary defects. Even though challenging, understanding the underlying molecular function of such proteins is of utter importance as they possess such a great impact on the whole plant development.

8.2. Chloroplast RNA metabolism and its interplay with other metabolic pathways – new possibilities for regulation

PUMPKIN, as the sole plastid UMP kinase acting as an enzyme in the plastid pyrimidine pathway also possesses RNA-association capability. Even though this association might be mediated via other interacting RNA-binding proteins, PUMPKINs` enzymatic activity and RNA-association capability also point to a dual and thus integrating function of PUMPKIN by connecting nucleotide biosynthesis and plastid gene regulation. The enzymatic properties of PUMPKIN open even further layers for regulation and sensing of nucleotide pools. As reported for the *E. coli* homologue PYRH, the kinase function is regulated by feed-back mechanisms as UTP inhibits the conversion of UMP to UDP while GTP acts as an allosteric activator to balance pyrimidine and purine synthesis (Meyer et al., 2008). The negative regulation by UTP and the positive stimulation by GTP were also observed for the enzymatic activity of PUMPKIN in *in vitro* studies (data not shown). These effectors might not only influence the enzymatic function but also the RNA-association ability of PUMPKIN in response to the existing nucleotide and energy state.

Interestingly, also other proteins involved in nucleotide biosynthesis seem to have moonlighting functions. Mutants defective in *UPP* – the upstream enzyme in the plastidial pyrimidine salvage pathway converting uracil to UMP – also displayed massive developmental defects which could not be reversed by introducing the homologues enzyme from *E. coli*, even though it was proven to be active and correctly localized in these complemented plants *in vivo* (Ohler et al., 2019). Vice-versa *upp* mutants could be rescued by the complementation with enzymatically impaired UPP versions, clearly demonstrating the importance of another – still unclear – function mainly responsible for the observed phenotype (Ohler et al., 2019).

In general, the influence or interplay of plastid gene expression with other metabolic pathways becomes more and more obvious (Bohne and Nickelsen, 2017). According to the REM (RNA-enzyme-metabolite) hypothesis especially enzymes seem to be an optimal platform to gain new functions but still retaining their regulatory mechanism (e.g. allosteric activation/inhibition, or non/competitive inhibitors) thus allowing tight control of their activity (Castello et al., 2015).

One very well studied example of a dual-functioning RNA-binding protein is the *Chlamydomonas* DLA2 subunit of the chloroplast pyruvate dehydrogenase which was shown to disassemble from the complex under mixotrophic growth conditions to bind the *psbA* mRNA, encoding the central D1 subunit of photosystem II (Bohne et al., 2013). Recent experiments could further verify this functional switch between fatty acid synthesis and gene expression control and identified an E3-binding domain within the protein as well as parts of the catalytic enzyme domain to confer specific RNA binding (Neusius et al., 2019). This way the protein integrates the energetic status of the cell into a regulatory output on lipid and protein biosynthesis and photosynthesis, respectively. Due to the nature of the RNA targets the action of RNA-binding proteins also affects different metabolic pathways. The PPR protein PDM2 for instance is involved in the editing and thus the post-transcriptional control of the chloroplast-encoded *accD* gene, whose protein product, the acetyl-CoA carboxylase beta subunit is essential in fatty acid synthesis (Kode et al., 2005; Du et al., 2017).

Sometimes proteins are also capable to regulate their own expression in a feedback-loop mechanism under certain conditions. In the case of the RubisCO large subunit (LSU) this leads to a dual functionality. While RubisCO large subunit is predominantly active as part of the enzyme for carbon fixation in the Calvin-Benson cycle, its self-regulation capacity also makes it a *bona fide* RNA-binding protein, binding to its own transcript via an RRM-motif and stopping translation of the *rbcL* mRNA (Cohen et al., 2006). Thereby RubisCO expression can be fine-tuned, especially under oxidative stress conditions (Cohen et al., 2006).

In addition, several RNA-binding factors possess overlapping targets but only some studies analysed if they indeed cooperate and form complexes, which would represent another layer of regulation. One recent example is the interplay of the RNA chaperone-like BSF, the processing factor CRP1 and the ribosomal release factor-like protein PrfB3 which act in concert on the *petB/petD* mRNA and regulate its stability and translation (Jiang et al., 2019). The identified protein interaction partners of PUMPKIN also deserve further research to elucidate which of the candidates are direct or indirect interaction partners to gain further insight into the possibly regulatory role of PUMPKIN within these proteinaceous complexes. Advances in big data analyses and computational network modelling are additional promising tools to further follow up these questions and gain insight into the complex regulatory networks and responses reacting to internal and external stimuli. As outlined in the following paragraph, identification of yet unknown RNA-binding proteins will further contribute to our understanding of the interplay of different metabolic pathways and the regulation of chloroplast gene expression.

8.3. The unexplored versatility of RNA-binding or associated proteins

Recent advances in whole-transcriptome analyses allowed a new approach to search for RNA-binding proteins using cross-linking and the pull-down of RNA-binding proteins via fishing the RNA molecules. Subsequent mass spectrometry approaches led to the identification of large proteome datasets representing RNA-binding proteins. Besides the identification of already known RNA-binding factors – which also validate the method – an unexpected amount of non-classical RNA-binding proteins was identified. They lack so-far known RNA-binding motifs and are often involved in other metabolic processes, thus supporting the paradigm of so-called enigmRBPs (RNA-binding proteins) – referring to their mysterious double life (Hentze et al., 2018).

Unfortunately, the studies were performed using an oligo(dT)-pull-down strategy based on poly(A) tails, which is not suitable for the plethora of plastid-encoded transcripts – as they lack stable polyadenylated termini – thus explaining the very low coverage of plastid RNA-binding proteins identified. If anything, poly(A) tails at plastid-encoded RNAs serve to mark transcripts which should undergo degradation (Schuster et al., 1999). Additionally, Reichel et. al (2016) used etiolated seedlings – more suitable for UV-crosslinking – as input material. As they are lacking mature chloroplasts, also the large pool of RNA-binding proteins involved in chloroplast gene expression is probable not fully established in this type of starting material.

Hence, it appears reasonable that plastid-localized RNA-associated proteins such as PUMPKIN were not identified with these approaches. To circumvent this underrepresentation of chloroplast RNA-binding factors, the idea arose to introduce suitable RNA-tags for pulldown, so called aptamers, which could replace the missing poly(A) stretches (Panchapakesan et al., 2017). Still, this technique is not optimal in plastid research as the stable introduction of such stretches into the chloroplast genome is time-consuming and labour-intensive and in addition such modifications can cause unforeseen effects on RNA stability or processing. Watkins et al. (2019) made use of another slightly different technique by using antisense oligonucleotides for the specific pulldown of *psbA* mRNA and additionally exploited the already known *psbA*-binding protein HCF173 to pulldown *psbA* and its associated proteome. Nevertheless, latest achievements in tailoring proteins for certain purposes opened another new possibility to search for yet unknown RNA-binding or RNA-associated factors acting on plastid transcripts. McDermott et al. (2019) designed an artificial PPR-protein targeted towards Arabidopsis *psbA* mRNA, allowing the analysis of the *psbA* associated proteome.

With this programmable proteins in our hand, systematic study of the plastid RNA-binding proteome will not only identify new factors important for chloroplast gene expression but will also contribute to understand the concerted action of different RNA-binding factors on the same target, which was recently reported (Jiang et al., 2019). Regarding the promising results from cytosolic RNA-binding proteome studies it is more than likely that especially chloroplasts, which are dependent on a tight regulation of their gene expression and metabolism, also harbour many – yet uncharacterized – dual-functioning RNA-binding proteins, representing interesting candidates for reverse genetic screens.

8.4. Engineering of RNA-binding proteins – future perspectives

Besides the engineering of RNA-binding proteins for pull-down strategies, in-depth understanding of RNA-binding proteins provides an excellent platform for additional applications. The elucidated recognition modes of e.g. PPR, TALE and PUF proteins offer a toolbox of different RNA-binding motifs with specificity towards a certain target (Filipovska and Rackham, 2012). This first RNA-binding modular unit can then be combined with other domains to introduce a desired function into the designer protein (see Figure 5).

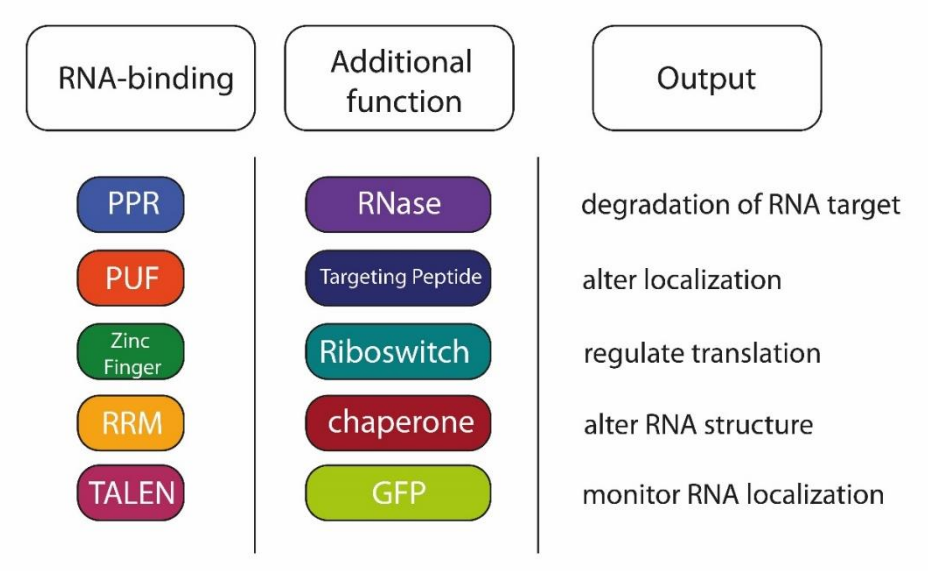


Figure 5. The modular toolbox of engineered RNA-binding proteins. RNA-binding specificity is achieved using the known RNA-binding motifs. Additional domains can confer additional functions leading to different actions on the desired target RNA. PPR = pentatricopeptide repeat, PUF = Pumilio and FBF protein family, RRM = RNA recognition motif, TALEN = Transcription Activator-like Effector Nuclease, GFP = green fluorescent protein.

This includes signal peptides for correct localization as well as functional domains which directly modify the chosen target. Such functional modules could be e.g. nucleases to cleave and thus inactivate a transcript, domains to edit certain nucleotides or fluorescent tags to track RNA localization (Chen and Varani, 2013). Especially inactivation of disease-related RNAs represents a promising approach to treat genetic disorders or cancer, as the use of RNA interference for silencing of disease-related RNAs already raised encouraging results (Mansoori et al., 2014). Another publication described the successful modulation of miRNA function by an engineered RNA-binding protein in cancer cells (Chen et al., 2016). Not only in human health, synthetic RBPs can contribute to fight diseases but as also *in planta* RNA-binding proteins are important for immune response. Hence, designing RNA-binding proteins which specifically bind virus RNA and degrade it might for instance help in developing virus-resistant plants – an emerging issue especially for crop plants (Huh and Paek, 2013). Lately, a synthetic PPR protein was used to regulate the *in vivo* expression of a transgene stably integrated into chloroplasts that could increase the induced expression of the transgene up to 40-fold, thus being a promising approach for optimizing transgenic protein production *in planta* (Rojas et al., 2019). Similarly, controlling the translation of transgenes in plastids by inducible expression is a useful approach to circumvent negative effects on plant growth when transgenes are constitutively expressed (Hennig et al., 2007). Making use of redesigned bacterial riboswitches, which can switch on or off translation in a ligand-dependent manner, Verhounig et al. (2010)

were already successful to control plastid translation of a GFP control gene by targeting these riboswitches to Shine Dalgarno sequences. To figure out the optimal design in terms of specificity, localization, interplay with other processes and regulation, fundamental research also on unusual RBPs, such as PAC or PUMPKIN will broaden the repertoire and improve the success in using these engineered proteins.

8.5. The origin and conservation of PUMPKIN and PAC

Besides understanding the molecular action of proteins, tracing their origin and evolution further provides valuable information. Especially comparing the sequence, structure or other features of proteins with their close or remote relatives, helps to elucidate if these proteins emerged early in evolution and fulfil essential tasks in ancient processes or if they evolved at a certain time point to adopt to changing conditions or mechanisms.

UMP kinases resembling PUMPKIN can be found from archaea, eubacteria over cyanobacteria, green algae, mosses and ferns up to flowering plants, speaking in favour for a common prokaryotic ancestor (Figure 6). Besides UMP kinases of other land plant species, the closest relatives to PUMPKIN are UMP kinases from cyanobacteria, arguing that PUMPKIN may have originated from the cyanobacterial ancestor which developed into the chloroplast. Nevertheless, this type of UMP kinase seems to be very ancient, as it is not only found in eubacteria but also in archaea. Neither the UMP kinase representatives from archaea (see blue dots in the phylogenetic tree) nor the examples from proteobacteria (red dots) strictly clustered into distinct groups but appeared as a mixed paraphyletic group, impeding the reconstruction of their evolution. Still, it can be assumed that this type of UMP kinase was already present in the last universal common ancestor (LUCA). Similarly, CTPase (PyrG), another gene involved in nucleotide biosynthesis, was reported to reach back to LUCA (Leipe et al., 2002), underlining the importance of nucleotide biosynthesis, which is a prerequisite for genetic information storage.

The closest relatives to PUMPKIN found in rhodophyta, phaeophyceae or diatomea are fusion proteins of aspartate kinase and homoserine dehydrogenase, two enzymes in pyrimidine *de novo* synthesis (not included in the phylogeny). Hence, these organisms might only possess *de novo* synthesis pathways but not the uracil salvage pathway. For comparison also the relation to the other three UMP kinases of Arabidopsis which are located in the cytosol, were analysed. As they belong to another type of UMP kinases which show specificity for both UMP and CMP, they expectedly belong to a separate cluster (Zhou et al., 1998).

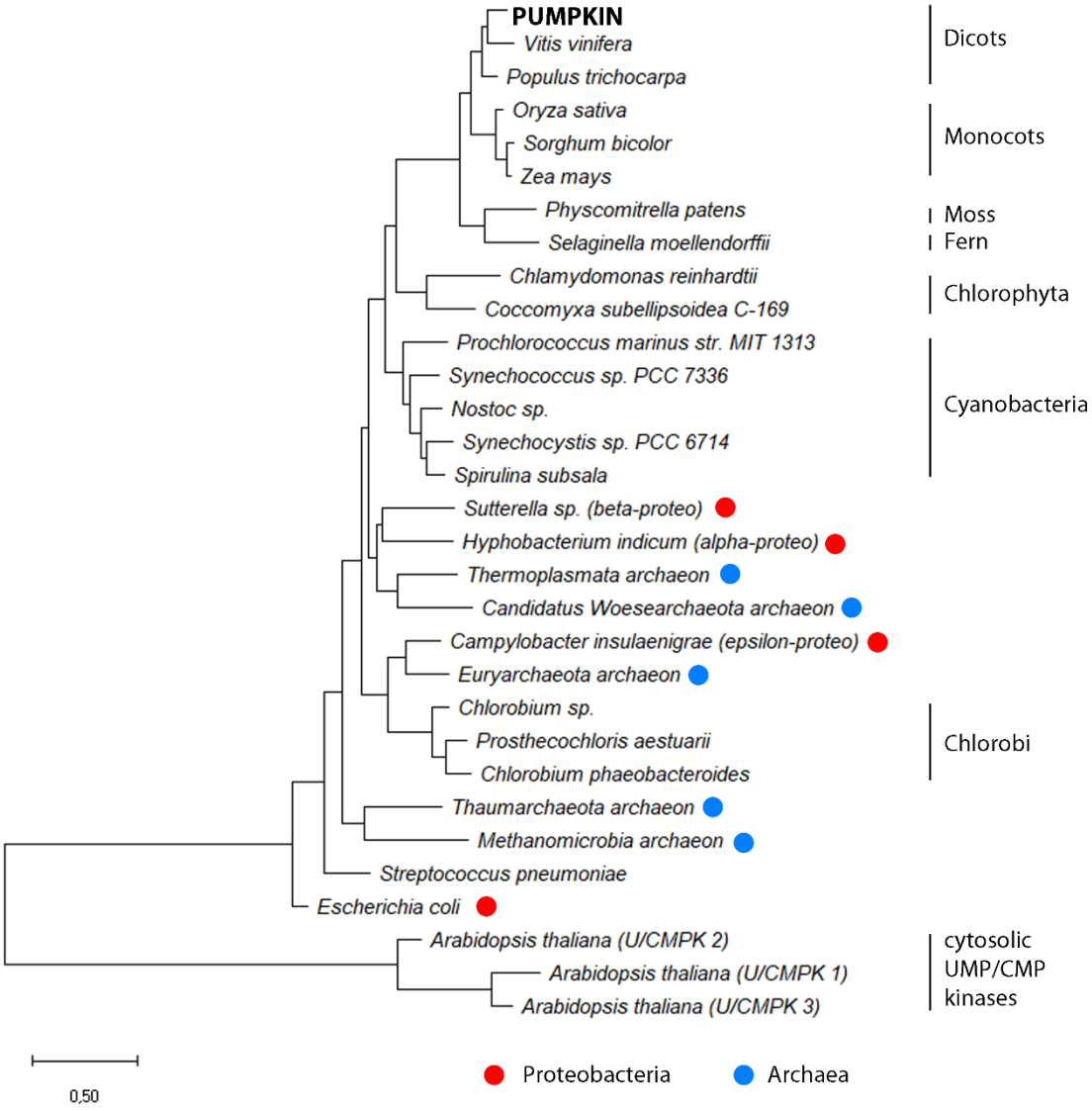


Figure 6. Phylogenetic tree of the UMP kinases. The alignment of representative proteins of different species was analysed by the maximum likelihood method using MEGA (Jones et al., 1992; Kumar et al., 2018). The scale bar indicates the evolutionary distance. NCBI accessions of all sequences are provided in the Appendix.

Comparing the amino acid sequences of different orthologues, certain residues essential for the enzymatic function, e.g. for substrate binding, are strongly conserved. The residues chosen for point mutations in order to obtain PUMPKIN versions with altered enzymatic activity (aspartic acid 165 and arginine 149, see manuscript) are indeed conserved over all analysed species, with the exception of the aspartic acid in the eukaryotic-type UMP/CMP kinases of Arabidopsis, which might be also explained by a suboptimal alignment due to their remote relation. This strong conservation further confirms the importance of the kinase function in nucleotide

biosynthesis as this enzyme is essential in several eubacteria (Fassy et al., 2004). At which time point during evolution PUMPKIN acquired its additional role in plastid RNA metabolism remains challenging to define. On protein sequence level no additional stretches or domains were acquired compared to eubacteria such as *E. coli* or also to cyanobacterial members except targeting peptides to correctly localize the protein into the chloroplast. Only examining the phenotypes and defects of knockouts and detailed molecular analyses of the respective proteins in different clades such as chlorophyta or cyanobacteria might shed light on the evolution of its dual functionality. So far, detailed experiments are only available for Arabidopsis and rice besides many eubacteria (Ingraham and Neuhard, 1972; Bucurenci et al., 1998; Zhu et al., 2016; Schmid et al., 2019).

In strong contrast, relatives of the PAC protein only exist in land plants (Figure 7). Similarly, also CRASS (CHLOROPLAST RIBOSOME ASSOCIATED) orthologs are only present in embryophytes but are neither found in green algae nor in cyanobacteria. Hence, PAC and CRASS represent newly evolved factors which might have been necessary to adopt to the new conditions when conquering land and leaving water. One might further hypothesize that these factors were also indispensable to cope with changing temperatures that occur in terrestrial habitats, in comparison to water environments (de Vries and Archibald, 2018). Especially processes connected to translation are known to react sensitive to cold-treatment when such factors are missing (Paieri et al., 2018; Pulido et al., 2018) as presumably the biogenesis of ribosomes might need more assistance for correct RNA-folding and assembly under these conditions (Bevilacqua et al., 2016). CGL20 (CONSERVED ONLY IN THE GREEN LINEAGE20), another factor assisting in 50S ribosome biogenesis is conserved in chloroplastidae, as the name already indicates, but also lacks a prokaryotic origin (Reiter et al., 2020). Thus, plastid ribosome biogenesis has acquired several new auxiliary factors, most probably also fulfilling regulatory functions and supporting the assembly of the slightly different translational machinery when compared to prokaryotes. Furthermore, this allows a tight control of chloroplast function by nuclear-encoded factors. Similarly, also other protein families involved in adjusting plastid gene expression massively increased in land plants, including PPR proteins (Barkan and Small, 2014). This recruitment of factors was either achieved by evolving new factors, as seems the case for PAC or also CGL20 or by making use of existing scaffolds which simply kept their function but underwent duplication or also developed new features (Reiter et al., 2020).

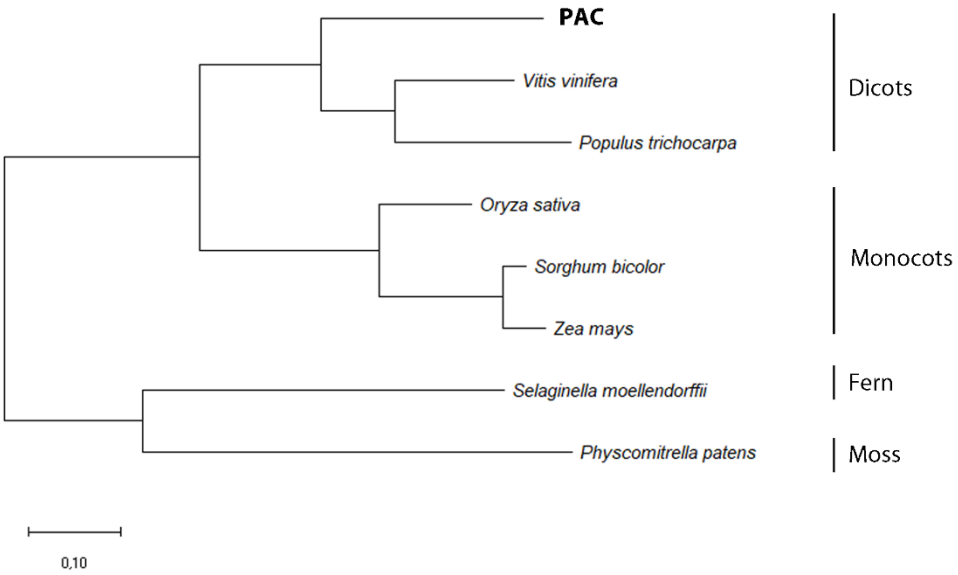


Figure 7. Phylogenetic tree of PAC. The alignment of representative proteins of different species was analysed by the maximum likelihood method using MEGA (Jones et al., 1992; Kumar et al., 2018). The scale bar indicates the evolutionary distance. NCBI accessions of all sequences are provided in the Appendix.

Even though the PAC protein does not contain known RNA-binding domains, it possesses a stretch of extraordinary many negatively charged residues. When comparing the amino acid composition of several ribosome biogenesis factors, a low isoelectric point (ranging from 4.5 to 5.5) and the occurrence of long stretches of glutamic (E) and aspartic (D) acids seems to be a common feature (Liu et al., 2015; Pulido et al., 2018; Reiter et al., 2020). In the case of RH50 and RH22 (DEAD-box RNA helicases 50 and 22, respectively), the isoelectric point is higher (7.7 and 5.9, respectively) probably due to their additional DEAD-box RNase helicase function, but both proteins contain several stretches of repetitive E and D residues (Chi et al., 2012; Paieri et al., 2018). DEAD-box RNase helicase are also found in prokaryotes, some examples in eubacteria furthermore showed that they fulfil similar functions in ribosome biogenesis (Khemici and Linder, 2016). RH22 for instance is closely related to the eubacterial SrmB protein and interacts with the same ribosomal protein (L24/ RPL24) (Chi et al., 2012). Similarly, also other ribogenesis factors e.g. GTPases such as DER (Double Era-like GTPase) or rRNA methylases such as CMAL, possess eubacterial relatives (Jeon et al., 2014; Zou et al., 2020). Thus, some auxiliary factors kept their original function while others, such as PAC, were newly recruited to aid in the process of ribosome assembly and maturation.

8.6. Features of accessory proteins aiding in ribosome biogenesis

Another shared feature between eubacteria and chloroplasts is the sublocalisation of ribosome biogenesis. In eubacteria it was long believed that translation happens co-transcriptionally and takes place in the nucleoids, but new techniques allowed the discrimination of directly nucleoid-localized ribosomes and a much higher number of ribosome-rich regions in the cytoplasm. This argues for ribosome biogenesis taking place in the nucleoids while mature ribosomes then translate mRNAs in the cytosol (Bakshi et al., 2015). Similarly, nucleoids are assumed to be the sub-compartment for ribosome biogenesis in plants as revealed by localization studies of ribosome biogenesis factors and proteomics data uncovering the composition of nucleoid fractions (Bohne, 2014). Expectedly, also PAC was identified as a nucleoid protein, as it co-localized with the nucleoid protein RAP (Meurer et al., 2017).

Regarding the function of different ribogenesis factors, many of them are necessary for correct processing and maturation of ribosomal RNAs and additionally were reported to also directly interact with ribosomal proteins. Still for several factors, such as PAC, CGL20 or CRASS, it remains challenging to predict how exactly this interaction is mediated as they lack classical RNA-binding motifs. Furthermore, crystal structures are missing so far to explain which parts of the protein surfaces might interact with the ribosomal RNAs and/or proteins (Meurer et al., 2017; Pulido et al., 2018; Reiter et al., 2020). One could further speculate about the stretches of acidic amino acids (glutamic and aspartic acid) that are found in repeats of 3 up to 10 in RH22, for instance (Chi et al., 2012). These negatively charged motifs might interact with positively charged amino acid stretches (including lysine, arginine and histidine residues) found in some ribosomal proteins such as RPL22, RPL2, RPS17 or RPS20 (https://www.uniprot.org/). A similar mechanism was shown for the small heat shock protein α-crystallin containing acidic regions which interact with the basic regions of fibroblast growth factor 1 (Edwards et al., 2001). Another hypothesis concerns the nature of glutamic acid as it was found to be the second most disorder-promoting residue (Uversky, 2013). Such intrinsic disorders influencing the folding and structure of proteins were reported to allow transient but specific interactions with multiple partners – a feature which would perfectly match to the requirement of an auxiliary factor that can be released easily after assembly (Pazos et al., 2013). Furthermore, RNA chaperone activity was another function attributed to intrinsically disordered regions, thus, these regions might be as well involved in rRNA recognition and binding (Oldfield et al., 2020).

Another relevant point to study is in which step of ribosome biogenesis a factor is active. Unfortunately, a detailed chronology of ribosome assembly is still missing in plants. At least in

eubacteria some steps and assembly intermediates have been analysed, still the precise action and involvement of extraribosomal factors remains to be determined (Fournier et al., 2010; Ridgeway et al., 2012). Some first hints towards an assembly model are provided by co-migration and co-immunoprecipitation analyses to figure out if a ribosome biogenesis factor is associated with the small (SSU) or large (LSU) ribosomal subunit (Jeon et al., 2014; Meurer et al., 2017; Paieri et al., 2018). Such co-migration experiments also allow to differentiate if a factor is predominantly linked to premature or mature ribosomes thus indicating an involvement in early assembly steps or in later steps of translation or polysome loading. A promising technique called "complexome profiling", which uses shotgun proteomics of fractions from native electrophoresis gels, was already successfully applied to identify ribosomal assembly states from human mitochondria (Wessels et al., 2013). Hence, such an approach might be very helpful to capture intermediates and develop an assembly profile of plastid ribosomes as well.

8.7. Adapting plastid gene expression to different conditions

So far characterization of proteins involved in plastid RNA metabolism almost entirely focused on characterizing their targets and functions. Another interesting question is to elucidate how the expression and activity of the RNA-binding factors themselves is influenced by transcription factors, external stimuli and/or metabolic ligands. Only some examples of differentially expressed RNA-binding factors were reported so far, which show an increase or decrease of their abundance in response to various conditions, ranging from heat or light over osmotic stress to an induction after viral infection (Naqvi et al., 1998; Stoppel et al., 2011; Ambrosone et al., 2015; Nyiko et al., 2019). The expression of PUMPKIN, for instance, was shown to change during development and unpublished data further revealed that PUMPKIN expression is light dependent. The chloroplast RNA binding proteins CP31A and CP29A were found to play an important role in conferring cold tolerance (Kupsch et al., 2012). Similarly, other factors involved in rRNA processing, ribogenesis and translation, such as RBD1 (Chloroplast RNA-Binding Protein 1), CRASS, RH50 or CGL20 were shown to be of particular importance under cold treatment (Wang et al., 2016; Paieri et al., 2018; Pulido et al., 2018; Reiter et al., 2020). RBD1 was shown to be more tightly bound to its target, the 23S rRNA, under chilling conditions to guarantee proper processing (Wang et al., 2016). It is assumed that especially RNA folding and thus also upstream processes are temperature sensitive (Bevilacqua et al., 2016). It was further speculated that ribosome run off might be enhanced under cold conditions, leaving the remaining "naked" transcripts unprotected and therefore accessible to degradation (Zoschke and Bock, 2018). Thus, protection and RNA-restructuring by RNA-

binding proteins becomes even more important. Identification of similar mechanisms in the nuclear system of Arabidopsis and even in eubacteria further support this frequently observed sensitivity of mutants impaired in ribosome biogenesis towards cold temperatures (Resch et al., 2010; Juntawong et al., 2013; Huang et al., 2016). Also PAC is involved in plastid ribosome biogenesis acting as an RNA chaperone which binds to 23S rRNA, hence, further studies on the behaviour of *pac* mutants under different conditions should be performed (Meurer et al., 2017). Initial studies already found that PAC levels are regulated by light (Reiter et al., 1994). Additional experiments revealed a strong responsiveness of *PAC* mRNA levels upon treatment of wild-type plants under different conditions (e.g. temperature, light regime), thus making PAC another excellent candidate for studies in terms of adapting the plastid translational machinery towards changing environmental conditions (Meurer et al., 2017).

9. REFERENCES

Albrecht V, Ingenfeld A, Apel K (2006) Characterization of the snowy cotyledon 1 mutant of Arabidopsis thaliana: the impact of chloroplast elongation factor G on chloroplast development and plant vitality. Plant Mol Biol **60:** 507-518

Ambrosone A, Batelli G, Nurcato R, Aurilia V, Punzo P, Bangarusamy DK, Ruberti I, Sassi M, Leone A, Costa A, Grillo S (2015) The Arabidopsis RNA-binding protein AtRGGA regulates tolerance to salt and drought stress. Plant Physiol **168:** 292-306

Bakshi S, Choi H, Weisshaar JC (2015) The spatial biology of transcription and translation in rapidly growing Escherichia coli. Front Microbiol **6:** 636

Bar-Yaacov D, Pilpel Y, Dahan O (2018) RNA editing in bacteria: occurrence, regulation and significance. RNA Biol **15:** 863-867

Barkan A, Klipcan L, Ostersetzer O, Kawamura T, Asakura Y, Watkins KP (2007) The CRM domain: an RNA binding module derived from an ancient ribosome-associated protein. RNA **13:** 55-64

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:** e1002910

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

Bevilacqua PC, Ritchey LE, Su Z, Assmann SM (2016) Genome-Wide Analysis of RNA Secondary Structure. Annu Rev Genet **50:** 235-266

Bhattacharya D, Archibald JM, Weber AP, Reyes-Prieto A (2007) How do endosymbionts become organelles? Understanding early events in plastid evolution. Bioessays **29:** 1239-1246

Bohne A-V, Schwarz C, Jalal A, Ossenbühla F, Nickelsen J (2009) Control of organellar gene expression in Chlamydomonas reinhardtii -- future perspectives. Endocytobiosis & Cell Research**:** 70-80

Bohne AV (2014) The nucleoid as a site of rRNA processing and ribosome assembly. Front Plant Sci **5:** 257

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

Bohne AV, Teubner M, Liere K, Weihe A, Börner T (2016) In vitro promoter recognition by the catalytic subunit of plant phage-type RNA polymerases. Plant Mol Biol **92:** 357-369

Börner T (2017) The discovery of plastid-to-nucleus retrograde signaling-a personal perspective. Protoplasma **254:** 1845-1855

Börner T, Aleynikova AY, Zubo YO, Kusnetsov VV (2015) Chloroplast RNA polymerases: Role in chloroplast biogenesis. Biochim Biophys Acta **1847:** 761-769

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:** 1678-1689

Bucurenci N, Serina L, Zaharia C, Landais S, Danchin A, Barzu 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

Chan CX, Gross J, Yoon HS, Bhattacharya D (2011) Plastid origin and evolution: new models provide insights into old problems. Plant Physiol **155:** 1552-1560

Chen CT, Slocum RD (2008) Expression and functional analysis of aspartate transcarbamoylase and role of de novo pyrimidine synthesis in regulation of growth and development in Arabidopsis. Plant Physiol Biochem **46:** 150-159

Chen Y, Varani G (2013) Engineering RNA-binding proteins for biology. FEBS J **280:** 3734-3754

Chen Y, Yang F, Zubovic L, Pavelitz T, Yang W, Godin K, Walker M, Zheng S, Macchi P, Varani G (2016) Targeted inhibition of oncogenic miR-21 maturation with designed RNA-binding proteins. Nat Chem Biol **12:** 717-723

Chi W, He B, Mao J, Li Q, Ma J, Ji D, Zou M, Zhang L (2012) The function of RH22, a DEAD RNA helicase, in the biogenesis of the 50S ribosomal subunits of Arabidopsis chloroplasts. Plant Physiol **158:** 693-707

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

Day PM, Theg SM (2018) Evolution of protein transport to the chloroplast envelope membranes. Photosynth Res **138:** 315-326

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

de Vries J, Archibald JM (2018) Plant evolution: landmarks on the path to terrestrial life. New Phytol **217:** 1428-1434

Devic M (2008) The importance of being essential: EMBRYO-DEFECTIVE genes in Arabidopsis. C R Biol **331:** 726-736

Du L, Zhang J, Qu S, Zhao Y, Su B, Lv X, Li R, Wan Y, Xiao J (2017) The Pentratricopeptide Repeat Protein Pigment-Defective Mutant2 is Involved in the Regulation of Chloroplast Development and Chloroplast Gene Expression in Arabidopsis. Plant Cell Physiol **58:** 747-759

Edwards K, Kössel H (1981) The rRNA operon from Zea mays chloroplasts: nucleotide sequence of 23S rDNA and its homology with E.coli 23S rDNA. Nucleic Acids Res **9:** 2853-2869

Edwards KL, Kueltzo LA, Fisher MT, Middaugh CR (2001) Complex effects of molecular chaperones on the aggregation and refolding of fibroblast growth factor-1. Arch Biochem Biophys **393:** 14-21

Eme L, Spang A, Lombard J, Stairs CW, Ettema TJG (2017) Archaea and the origin of eukaryotes. Nat Rev Microbiol **15:** 711-723

Faivre-Nitschke SE, Grienenberger JM, Gualberto JM (1999) A prokaryotic-type cytidine deaminase from Arabidopsis thaliana gene expression and functional characterization. Eur J Biochem **263:** 896-903

Fassy F, Krebs O, Lowinski M, Ferrari P, Winter J, Collard-Dutilleul V, Salahbey Hocini K (2004) UMP kinase from Streptococcus pneumoniae: evidence for co-operative ATP binding and allosteric regulation. Biochem J **384:** 619-627

Feldmann KA (1991) T-DNA insertion mutagenesis in Arabidopsis: mutational spectrum. The Plant Journal **1:** 71-82

Filipovska A, Rackham O (2012) Modular recognition of nucleic acids by PUF, TALE and PPR proteins. Mol Biosyst **8:** 699-708

Fournier GP, Neumann JE, Gogarten JP (2010) Inferring the ancient history of the translation machinery and genetic code via recapitulation of ribosomal subunit assembly orders. PLoS One **5:** e9437

Girke C, Daumann M, Niopek-Witz S, Möhlmann T (2014) Nucleobase and nucleoside transport and integration into plant metabolism. Front Plant Sci **5:** 443

Gould SB, Waller RF, McFadden GI (2008) Plastid evolution. Annu Rev Plant Biol **59:** 491-517

Hajdukiewicz PT, Allison LA, Maliga P (1997) The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. EMBO J **16:** 4041-4048

Hammani K, Okuda K, Tanz SK, Chateigner-Boutin AL, 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:** 3686-3699

Harris EH, Boynton JE, Gillham NW (1994) Chloroplast ribosomes and protein synthesis. Microbiol Rev **58:** 700-754

Hedtke B, Börner T, Weihe A (1997) Mitochondrial and chloroplast phage-type RNA polymerases in Arabidopsis. Science **277:** 809-811

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

Hennig A, Bonfig K, Roitsch T, Warzecha H (2007) Expression of the recombinant bacterial outer surface protein A in tobacco chloroplasts leads to thylakoid localization and loss of photosynthesis. FEBS J **274:** 5749-5758

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

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

Hirose T, Sugiura M (2004) Multiple elements required for translation of plastid atpB mRNA lacking the Shine-Dalgarno sequence. Nucleic Acids Res **32:** 3503-3510

Huang CK, Shen YL, Huang LF, Wu SJ, Yeh CH, Lu CA (2016) The DEAD-Box RNA Helicase AtRH7/PRH75 Participates in Pre-rRNA Processing, Plant Development and Cold Tolerance in Arabidopsis. Plant Cell Physiol **57:** 174-191

Huang J, Gogarten JP (2007) Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? Genome Biol **8:** R99

Huh SU, Paek KH (2013) Plant RNA binding proteins for control of RNA virus infection. Front Physiol **4:** 397

Ingraham JL, Neuhard J (1972) Cold-Sensitive Mutants of Salmonella-Typhimurium Defective in Uridine Monophosphate Kinase (Pyrh). Journal of Biological Chemistry **247:** 6259-6265

Jeon Y, Ahn CS, Jung HJ, Kang H, Park GT, Choi Y, Hwang J, Pai HS (2014) DER containing two consecutive GTP-binding domains plays an essential role in chloroplast ribosomal RNA processing and ribosome biogenesis in higher plants. J Exp Bot **65:** 117-130

Jiang J, Chai X, Manavski N, Williams-Carrier R, He B, Brachmann A, Ji D, Ouyang M, Liu Y, Barkan A, Meurer J, Zhang L, Chi W (2019) An RNA Chaperone-Like Protein Plays Critical Roles in Chloroplast mRNA Stability and Translation in Arabidopsis and Maize. Plant Cell **31:** 1308-1327

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci **8:** 275-282

Jung B, Hoffmann C, Möhlmann T (2011) Arabidopsis nucleoside hydrolases involved in intracellular and extracellular degradation of purines. Plant J **65:** 703-711

Juntawong P, Sorenson R, Bailey-Serres J (2013) Cold shock protein 1 chaperones mRNAs during translation in Arabidopsis thaliana. Plant J **74:** 1016-1028

Kaneko T, Tabata S (1997) Complete genome structure of the unicellular cyanobacterium Synechocystis sp. PCC6803. Plant Cell Physiol **38:** 1171-1176

Khemici V, Linder P (2016) RNA helicases in bacteria. Curr Opin Microbiol **30:** 58-66

Kode V, Mudd EA, Iamtham S, Day A (2005) The tobacco plastid accD gene is essential and is required for leaf development. Plant J **44:** 237-244

Köster T, Marondedze C, Meyer K, Staiger D (2017) RNA-Binding Proteins Revisited - The Emerging Arabidopsis mRNA Interactome. Trends Plant Sci **22:** 512-526

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol **35:** 1547-1549

Kupsch C, Ruwe H, Gusewski S, Tillich M, Small I, Schmitz-Linneweber C (2012) Arabidopsis chloroplast RNA binding proteins CP31A and CP29A associate with large transcript pools and confer cold stress tolerance by influencing multiple chloroplast RNA processing steps. Plant Cell **24:** 4266-4280

Leipe DD, Wolf YI, Koonin EV, Aravind L (2002) Classification and evolution of P-loop GTPases and related ATPases. J Mol Biol **317:** 41-72

Leister D (2003) Chloroplast research in the genomic age. Trends Genet **19:** 47-56

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

Li YQ, Sugiura M (1991) Nucleic acid-binding specificities of tobacco chloroplast ribonucleoproteins. Nucleic Acids Res **19:** 2893-2896

Liu J, Zhou W, Liu G, Yang C, Sun Y, Wu W, Cao S, Wang C, Hai G, Wang Z, Bock R, Huang J, Cheng Y (2015) The conserved endoribonuclease YbeY is required for chloroplast ribosomal RNA processing in Arabidopsis. Plant Physiol **168:** 205-221

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:** 36

Mainguet SE, Gakiere B, Majira A, Pelletier S, Bringel F, Guerard F, Caboche M, Berthome R, Renou JP (2009) Uracil salvage is necessary for early Arabidopsis development. Plant J **60:** 280-291

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

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, 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

Mansoori B, Sandoghchian Shotorbani S, Baradaran B (2014) RNA interference and its role in cancer therapy. Adv Pharm Bull **4:** 313-321

Maréchal E (2018) Primary Endosymbiosis: Emergence of the Primary Chloroplast and the Chromatophore, Two Independent Events. *In* E Maréchal, ed, Plastids: Methods and Protocols. Springer US, New York, NY, pp 3-16

Marondedze C, Thomas L, Serrano NL, Lilley KS, Gehring C (2016) The RNA-binding protein repertoire of Arabidopsis thaliana. Sci Rep **6:** 29766

Martin WF, Garg S, Zimorski V (2015) Endosymbiotic theories for eukaryote origin. Philos Trans R Soc Lond B Biol Sci **370:** 20140330

McDermott JJ, Watkins KP, Williams-Carrier R, Barkan A (2019) Ribonucleoprotein capture by in vivo expression of a designer pentatricopeptide repeat protein in Arabidopsis. Plant Cell **31:** 1723-1733

Meurer J, Grevelding C, Westhoff P, Reiss B (1998) The PAC protein affects the maturation of specific chloroplast mRNAs in Arabidopsis thaliana. Molecular and General Genetics **258:** 342-351

Meurer J, Lezhneva L, Amann K, Godel M, Bezhani S, Sherameti I, Oelmuller R (2002) A peptide chain release factor 2 affects the stability of UGA-containing transcripts in Arabidopsis chloroplasts. Plant Cell **14:** 3255-3269

Meurer J, Meierhoff K, Westhoff P (1996) Isolation of high-chlorophyll-fluorescence mutants of Arabidopsis thaliana and their characterisation by spectroscopy, immunoblotting and northern hybridisation. Planta **198:** 385-396

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, Barzu 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

Moffatt BA, Ashihara H (2002) Purine and pyrimidine nucleotide synthesis and metabolism. Arabidopsis Book **1:** e0018

Morley SA, Nielsen BL (2016) Chloroplast DNA Copy Number Changes during Plant Development in Organelle DNA Polymerase Mutants. Front Plant Sci **7:** 57

Nagy E, Rigby WF (1995) Glyceraldehyde-3-phosphate dehydrogenase selectively binds AU-rich RNA in the NAD(+)-binding region (Rossmann fold). J Biol Chem **270:** 2755-2763

Nakai M (2018) New Perspectives on Chloroplast Protein Import. Plant Cell Physiol **59:** 1111-1119

Nakamura T, Ohta M, Sugiura M, Sugita M (2001) Chloroplast ribonucleoproteins function as a stabilizing factor of ribosome-free mRNAs in the stroma. J Biol Chem **276:** 147-152

Naqvi SM, Park KS, Yi SY, Lee HW, Bok SH, Choi D (1998) A glycine-rich RNA-binding protein gene is differentially expressed during acute hypersensitive response following Tobacco Mosaic Virus infection in tobacco. Plant Mol Biol **37:** 571-576

Nara T, Hshimoto T, Aoki T (2000) Evolutionary implications of the mosaic pyrimidine-biosynthetic pathway in eukaryotes. Gene **257:** 209-222

Neusius D, Kleinknecht L, Bohne A-V, Nickelsen J (2019) RNA recognition by the E2 subunit of the chloroplast pyruvate dehydrogenase complex from Chlamydomonas. bioRxiv**:** 831339

Nyikó T, Auber A, Bucher E (2019) Functional and molecular characterization of the conserved Arabidopsis PUMILIO protein, APUM9. Plant Mol Biol **100:** 199-214

Ohler L, Niopek-Witz S, Mainguet SE, Möhlmann T (2019) Pyrimidine Salvage: Physiological Functions and Interaction with Chloroplast Biogenesis. Plant Physiol **180:** 1816-1828

Oldfield CJ, Peng Z, Kurgan L (2020) Disordered RNA-Binding Region Prediction with DisoRDPbind. Methods Mol Biol **2106:** 225-239

Ortelt J, Link G (2014) Plastid gene transcription: promoters and RNA polymerases. Methods Mol Biol **1132:** 47-72

Paieri F, Tadini L, Manavski N, Kleine T, Ferrari R, Morandini P, Pesaresi P, Meurer J, Leister D (2018) The DEAD-box RNA Helicase RH50 Is a 23S-4.5S rRNA Maturation Factor that Functionally Overlaps with the Plastid Signaling Factor GUN1. Plant Physiol **176:** 634-648

Panchapakesan SSS, Ferguson ML, Hayden EJ, Chen X, Hoskins AA, Unrau PJ (2017) Ribonucleoprotein purification and characterization using RNA Mango. RNA **23:** 1592-1599

Pazos F, Pietrosemoli N, Garcia-Martin JA, Solano R (2013) Protein intrinsic disorder in plants. Front Plant Sci **4:** 363

Pulido P, Zagari N, Manavski N, Gawronski P, Matthes A, Scharff LB, Meurer J, Leister D (2018) CHLOROPLAST RIBOSOME ASSOCIATED Supports Translation under Stress and Interacts with the Ribosomal 30S Subunit. Plant Physiol **177:** 1539-1554

Reichel M, Liao Y, Rettel M, Ragan C, Evers M, Alleaume AM, Horos R, Hentze MW, Preiss T, Millar AA (2016) In Planta Determination of the mRNA-Binding Proteome of Arabidopsis Etiolated Seedlings. Plant Cell **28:** 2435-2452

Reiter B, Vamvaka E, Marino G, Kleine T, Jahns P, Bolle C, Leister D, Rühle T (2020) The Arabidopsis Protein CGL20 Is Required for Plastid 50S Ribosome Biogenesis. Plant Physiology **182:** 1222-1238

Reiter RS, Coomber SA, Bourett TM, Bartley GE, Scolnik PA (1994) Control of leaf and chloroplast development by the Arabidopsis gene pale cress. Plant Cell **6:** 1253-1264

Resch A, Vecerek B, Palavra K, Blasi U (2010) Requirement of the CsdA DEAD-box helicase for low temperature riboregulation of rpoS mRNA. RNA Biol **7:** 796-802

Ridgeway WK, Millar DP, Williamson JR (2012) Quantitation of ten 30S ribosomal assembly intermediates using fluorescence triple correlation spectroscopy. Proc Natl Acad Sci U S A **109:** 13614-13619

Rojas M, Yu Q, Williams-Carrier R, Maliga P, Barkan A (2019) Engineered PPR proteins as inducible switches to activate the expression of chloroplast transgenes. Nature Plants **5:** 505-511

Rolland N, Janosi L, Block MA, Shuda M, Teyssier E, Miege C, Cheniclet C, Carde JP, Kaji A, Joyard J (1999) Plant ribosome recycling factor homologue is a chloroplastic protein and is bactericidal in escherichia coli carrying temperature-sensitive ribosome recycling factor. Proc Natl Acad Sci U S A **96:** 5464-5469

Romani I, Tadini L, Rossi F, Masiero S, Pribil M, Jahns P, Kater M, Leister D, Pesaresi P (2012) Versatile roles of Arabidopsis plastid ribosomal proteins in plant growth and development. Plant J **72:** 922-934

Rovira AG, Smith AG (2019) PPR proteins - orchestrators of organelle RNA metabolism. Physiol Plant **166:** 451-459

Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S (1999) Complete structure of the chloroplast genome of Arabidopsis thaliana. DNA Res **6:** 283-290

Schmid LM, Ohler L, Möhlmann T, Brachmann A, Muiño JM, Leister D, Meurer J, Manavski N (2019) PUMPKIN, the Sole Plastid UMP Kinase, Associates with Group II Introns and Alters Their Metabolism. Plant Physiol **179:** 248-264

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

Schuster G, Lisitsky I, Klaff P (1999) Polyadenylation and degradation of mRNA in the chloroplast. Plant Physiol **120:** 937-944

Schweer J, Turkeri H, Kolpack A, Link G (2010) Role and regulation of plastid sigma factors and their functional interactors during chloroplast transcription - recent lessons from Arabidopsis thaliana. Eur J Cell Biol **89:** 940-946

Shahar N, Weiner I, Stotsky L, Tuller T, Yacoby I (2019) Prediction and large-scale analysis of primary operons in plastids reveals unique genetic features in the evolution of chloroplasts. Nucleic Acids Res **47:** 3344-3352

Shen C, Zhang D, Guan Z, Liu Y, Yang Z, Yang Y, Wang X, Wang Q, Zhang Q, Fan S, Zou T, Yin P (2016) Structural basis for specific single-stranded RNA recognition by designer pentatricopeptide repeat proteins. Nat Commun **7:** 11285

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

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, Meurer J (2012) The cutting crew - ribonucleases are key players in the control of plastid gene expression. J Exp Bot **63:** 1663-1673

Sugiura M, Hirose T, Sugita M (1998) Evolution and mechanism of translation in chloroplasts. Annu Rev Genet **32:** 437-459

Takenaka M, Zehrmann A, Verbitskiy D, Hartel B, Brennicke A (2013) RNA editing in plants and its evolution. Annu Rev Genet **47:** 335-352

Taylor WC, Barkan A, Martienssen RA (1987) Use of nuclear mutants in the analysis of chloroplast development. Dev Genet **8:** 305-320

Tiller N, Bock R (2014) The translational apparatus of plastids and its role in plant development. Mol Plant **7:** 1105-1120

Uversky VN (2013) The alphabet of intrinsic disorder: II. Various roles of glutamic acid in ordered and intrinsically disordered proteins. Intrinsically Disord Proteins **1:** e24684

Verhounig A, Karcher D, Bock R (2010) Inducible gene expression from the plastid genome by a synthetic riboswitch. Proc Natl Acad Sci U S A **107:** 6204-6209

Wang F, Johnson X, Cavaiuolo M, Bohne AV, Nickelsen J, Vallon O (2015) Two Chlamydomonas OPR proteins stabilize chloroplast mRNAs encoding small subunits of photosystem II and cytochrome b6 f. Plant J **82:** 861-873

Wang S, Bai G, Wang S, Yang L, Yang F, Wang Y, Zhu JK, Hua J (2016) Chloroplast RNA-Binding Protein RBD1 Promotes Chilling Tolerance through 23S rRNA Processing in Arabidopsis. PLoS Genet **12:** e1006027

Watkins KP, Williams-Carrier R, Chotewutmontri P, Friso G, Teubner M, Belcher S, Ruwe H, Schmitz-Linneweber C, van Wijk KJ, Barkan A (2020) Exploring the proteome associated with the mRNA encoding the D1 reaction center protein of Photosystem II in plant chloroplasts. Plant J

Wessels HJ, Vogel RO, Lightowlers RN, Spelbrink JN, Rodenburg RJ, van den Heuvel LP, van Gool AJ, Gloerich J, Smeitink JA, Nijtmans LG (2013) Analysis of 953 human proteins from a mitochondrial HEK293 fraction by complexome profiling. PLoS One **8:** e68340

Williams-Carrier R, Zoschke R, Belcher S, Pfalz J, Barkan A (2014) A major role for the plastid-encoded RNA polymerase complex in the expression of plastid transfer RNAs. Plant Physiol **164:** 239-248

Witte CP, Herde M (2020) Nucleotide Metabolism in Plants. Plant Physiol **182:** 63-78

Xu T, Lee K, Gu L, Kim JI, Kang H (2013) Functional characterization of a plastid-specific ribosomal protein PSRP2 in Arabidopsis thaliana under abiotic stress conditions. Plant Physiol Biochem **73:** 405-411

Yagi Y, Shiina T (2012) Evolutionary aspects of plastid proteins involved in transcription: the transcription of a tiny genome is mediated by a complicated machinery. Transcription **3:** 290-294

Yamaguchi K, Subramanian AR (2000) The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast). J Biol Chem **275:** 28466-28482

Yamaguchi K, von Knoblauch K, Subramanian AR (2000) The plastid ribosomal proteins. Identification of all the proteins in the 30 S subunit of an organelle ribosome (chloroplast). J Biol Chem **275:** 28455-28465

Yan J, Zhang Q, Yin P (2018) RNA editing machinery in plant organelles. Sci China Life Sci **61:** 162-169

Zhelyazkova P, Hammani K, Rojas M, Voelker R, Vargas-Suárez M, Börner T, Barkan A (2012) Protein-mediated protection as the predominant mechanism for defining processed mRNA termini in land plant chloroplasts. Nucleic Acids Res **40:** 3092-3105

Zhou L, Lacroute F, Thornburg R (1998) Cloning, expression in Escherichia coli, and characterization of Arabidopsis thaliana UMP/CMP kinase. Plant Physiol **117:** 245-254

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, Bock R (2018) Chloroplast Translation: Structural and Functional Organization, Operational Control, and Regulation. Plant Cell **30:** 745-770

Zou M, Mu Y, Chai X, Ouyang M, Yu LJ, Zhang L, Meurer J, Chi W (2020) The critical function of the plastid rRNA methyltransferase, CMAL, in ribosome biogenesis and plant development. Nucleic Acids Res

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

10. ACKNOWLEDGEMENTS

First, I wish to thank my supervisor PD Dr. Jörg Meurer for the opportunity to work on these challenging and interesting projects, for letting me turn into an autonomous and critical scientist and for his trust in my skills. My very special thanks to Dr. Nikolay Manavski from whom I learned everything! Thank you for all your patience and support and of course our ABBA and Depeche Mode accompanied experimental sessions. Thank you for donating your time and effort on these projects and my training!

I would like to express my gratitude to my TAC committee for their great input and ideas and the SFB TRR175 for the nice community and of course for funding my position. Special thanks to my fantastic collaboration partners from Kaiserslautern – Dr. Torsten Möhlmann and his PhD students Lisa Ohler and Annalisa John – for their substantial contributions to the PUMPKIN project. Many thanks to the LSM – Life Science Munich – Graduate School and our great coordinators Francisca Rosa Mende and Nadine Hamze for all your support and help. I further wish to thank Prof. Dr. Leister for providing laboratory space and equipment. All the best to my two successional colleagues Alexandre and Florian, I wish you good luck with your projects and hope that you will have such a great time as me.

Now some special ones. Many thanks to my office mates Natalie and Anna! Without you my PhD would have been half as enjoyable and also thanks for some extra rounds of chocolate if experiments didn't want to work. I especially want to thank our good souls Angie, Ute and Gudrun who always did their best to help me with any issue. Also, many thanks to Cordelia Bolle for critically reading and correcting my thesis. Thanks to my former and current lunch people for entertaining lunchbreaks and the great staff from our cafeteria, helping me to survive exhausting afternoons with some extra caffeine. Without my great colleagues from AG Leister, AG Frank, AG Geigenberger and AG Soll this work and also my time here would have been much more challenging – thanks a lot for your straightforward help whenever I needed a good advice! And of course, many thanks to the gardeners for taking care of my plants! Last but not least, thanks to my research practical students, I really enjoyed to work with you and teach you.

I further wish to thank my friend and family for their unconditional support and for listening to my never-ending reports about my latest breakthroughs – or failures ;-)

11. APPENDIX

List of sequences used for phylogenetic analyses of PUMPKIN and PAC

The following paragraph lists all proteins sequences which were used to generate the phylogenetic trees shown in the discussion (Section 8.5 “The origin and conservation of PUMPKIN and PAC”). When applicable, targeting peptides were predicted via ChloroP (http://www.cbs.dtu.dk/services/ChloroP/) and removed from the sequences before the alignment.

PUMPKIN (*Arabidopsis thaliana* NP_001327269.1) tree:

Populus trichocarpa (XP_006382835.2), *Vitis vinifera* (RVW32729.1), *Oryza sativa* (XP_015638950.1), *Sorghum bicolor* (XP_002456998.1), *Zea mays* (ACG34670.1), *Physcomitrella patens* (XP_024356956.1), *Selaginella moellendorffii* (XP_002989696.2), *Chlamydomonas reinhardtii* (XP_001701996.1), *Coccomyxa subellipsoidea C-169* (XP_005649703.1), *Synechococcus sp. PCC 7336* (WP_017323937.1), *Synechocystis sp. PCC 6714* (WP_028946647.1), *Nostoc sp.* (WP_012412364.1), *Prochlorococcus marinus str. MIT 1313* (KZR69540.1), *Chlorobium sp.* (TLU86129.1), *Campylobacter insulaenigrae* (VDH06034.1), *Sutterella sp*. (WP_118275560.1), *Streptococcus pneumoniae* (COI84283.1), *Prosthecochloris aestuarii* (WP_012506353.1), *Chlorobium phaeobacteroides* (WP_012475388.1), *Escherichia coli* (MXF67036.1), *Euryarchaeota archaeon* (MBM56337.1), *Thermoplasmata archaeon* (RLF63849.1), *Arabidopsis thaliana* UMP/CMP kinase 1 (NP_567093.1), *Arabidopsis thaliana* UMP/CMP kinase 2 (NP_001320061.1), *Arabidopsis thaliana* UMP/CMP kinase 3 (NP_001031942.1), *Spirulina subsalsa* (WP_017306786.1), *Thaumarchaeota archaeon* (MBH59406.1), *Candidatus Woesearchaeota archaeon* (MAG62215.1), *Methanomicrobia archaeon* (NCA95618.1), *Hyphobacterium indicum* (WP_109260276.1)

PAC (*Arabidopsis thaliana* OAP08721.1) tree:

Populus trichocarpa (XP_024441296.1), *Vitis vinifera* (RVW54883.1), *Zea mays* (ACG37291.1), *Oryza sativa* (XP_015637477.1), *Sorghum bicolor* (XP_002439753.1), *Selaginella moellendorffii* (XP_002990128.2), *Physcomitrella patens* (XP_024376361.1)

Permissions for republishing

Permission for the Review article: Manavski N, Schmid LM, Meurer J (2018) RNA-stabilization factors in chloroplasts of vascular plants. Essays Biochem 62: 51-64

Permission to reuse content from an article published by Portland Press:

- If the content that you are seeking to re-use is in a Portland Press article that is published open access under a CC BY licence NO permissions are required, although you must cite the published article and credit the authors when you re-use it (or part of it).
- If the article you are seeking to re-use is published open access under any other type of licence (e.g. CC BY NC-ND) or a Portland Press license to publish then please complete a re-use permission-request form via copyright.com.
- To find out what licence the article is published under look for the copyright line on the published article, which can be found underneath the abstract or full text, depending on what view you are seeing for the article.
- **FOR AUTHORS:** if you are a named author on the article you wish to re-use then you will not need to seek any permissions except for re-use of non-open access papers that involves commercial re-selling or bulk distribution. For the latter, please visit copyright.com.

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Mar 24, 2020

This Agreement between LMU -- Lisa-Marie Schmid ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4795451219895
License date	Mar 24, 2020
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Plant Journal
Licensed Content Title	PALE CRESS binds to plastid RNAs and facilitates the biogenesis of the 50S ribosomal subunit
Licensed Content Author	Nikolay Manavski, Andreas Brachmann, Dario Leister, et al
Licensed Content Date	Sep 21, 2017
Licensed Content Volume	92
Licensed Content Issue	3
Licensed Content Pages	14
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Order reference number	PAC Permission
Title of your thesis / dissertation	Functional Characterization of PAC and PUMPKIN, Two Proteins Involved in Chloroplast RNA Metabolism

Expected completion date	Mar 2020
Expected size (number of pages)	120
Requestor Location	LMU Grosshaderner Str. 2-4 Planegg-Martinsried, other 82152 Germany Attn: LMU
Publisher Tax ID	EU826007151
Total	0.00 EUR

Terms and Conditions

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a"Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at http://myaccount.copyright.com).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.
- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, **and any CONTENT (PDF or image file) purchased as part of your order,** is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.
- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner.**For STM Signatory Publishers clearing permission under the terms of the STM Permissions Guidelines only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and**

does not involve the separate exploitation of the permitted figures or extracts, You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.
- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.
- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.
- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.
- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The Creative Commons Attribution License (CC-BY) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The Creative Commons Attribution Non-Commercial (CC-BY-NC)License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

The Creative Commons Attribution Non-Commercial-NoDerivs License (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online Library http://olabout.wiley.com/WileyCDA/Section/id-410895.html

Other Terms and Conditions:

v1.10 Last updated September 2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

ELSEVIER LICENSE
TERMS AND CONDITIONS

Mar 24, 2020

This Agreement between LMU -- Lisa-Marie Schmid ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4795460228403
License date	Mar 24, 2020
Licensed Content Publisher	Elsevier
Licensed Content Publication	Trends in Genetics
Licensed Content Title	Chloroplast research in the genomic age
Licensed Content Author	Dario Leister
Licensed Content Date	Jan 1, 2003
Licensed Content Volume	19
Licensed Content Issue	1
Licensed Content Pages	10
Start Page	47
End Page	56
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables /illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No

Will you be translating?	No
Title	Functional Characterization of PAC and PUMPKIN, Two Proteins Involved in Chloroplast RNA Metabolism
Institution name	n/a
Expected presentation date	Mar 2020
Order reference number	Leister 2003 Permission
Portions	Figure 1: Genes of cyanobacterial origin and the intracellular targeting of their products in Arabidopsis
Requestor Location	LMU Grosshaderner Str. 2-4 Planegg-Martinsried, other 82152 Germany Attn: LMU
Publisher Tax ID	GB 494 6272 12
Total	0.00 EUR

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be

altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing,

formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are

available at http://creativecommons.org/licenses/by/4.0.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.9

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

Copyright Clearance Center Marketplace™

American Society of Plant Biologists - License Terms and Conditions

Order Date	24-Mar-2020
Order license ID	1024681-1
ISSN	0032-0889
Type of Use	Republish in a thesis/dissertation
Publisher	AMERICAN SOCIETY OF PLANT BIOLOGISTS
Portion	Chapter/article

LICENSED CONTENT

Publication Title	Plant physiology
Article Title	PUMPKIN, the sole Plastid UMP Kinase, Associates with Group II Introns and Alters Their Metabolism.
Author/Editor	AMERICAN SOCIETY OF PLANT PHYSIOLOGISTS.
Date	01/01/1926
Language	English
Country	United States of America
Rightsholder	American Society of Plant Biologists
Publication Type	Journal
Start Page	248
End Page	264
Issue	1
Volume	179

REQUEST DETAILS

Portion Type	Chapter/article
Page range(s)	248-264
Total number of pages	17
Format (select all that apply)	Print, Electronic
Who will republish the content?	Academic institution
Duration of Use	Life of current and all future editions
Lifetime Unit Quantity	Up to 4,999
Rights Requested	Main product and any product related to main product
Distribution	Worldwide
Translation	Original language of publication
Copies for the disabled?	No
Minor editing privileges?	No
Incidental promotional use?	No
Currency	EUR

NEW WORK DETAILS

Title	Functional Characterization of PAC and PUMPKIN, Two Proteins Involved in Chloroplast RNA Metabolism
Instructor name	PD Dr. Jörg Meurer
Institution name	LMU München Department Biologie I
Expected presentation date	2020-03-31

ADDITIONAL DETAILS

Order reference number	Permission PUMPKIN
The requesting person / organization to appear on the license	Lisa-Marie Schmid

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	PUMPKIN, the Sole Plastid UMP Kinase, Associates with Group II Introns and Alters Their Metabolism
Editor of portion(s)	Schmid, Lisa-Marie; Ohler, Lisa; Möhlmann, Torsten; Muiño, Jose M; Meurer, Jörg; Manavski, Nikolay; Leister, Dario; Brachmann, Andreas
Volume of serial or monograph	179
Title of the article/chapter the portion is from	PUMPKIN, the sole Plastid UMP Kinase, Associates with Group II Introns and Alters Their Metabolism.
Author of portion(s)	Schmid, Lisa-Marie; Ohler, Lisa; Möhlmann, Torsten; Muiño, Jose M; Meurer, Jörg; Manavski, Nikolay; Leister, Dario; Brachmann, Andreas
Issue, if republishing an article from a serial	1

Page or page range of portion	248-264	**Publication date of portion**	2018-11-08

PUBLISHER TERMS AND CONDITIONS

Permission is granted for the life of the current edition and all future editions, in all languages and in all media.

CCC Republication Terms and Conditions

1. Description of Service; Defined Terms. This Republication License enables the User to obtain licenses for republication of one or more copyrighted works as described in detail on the relevant Order Confirmation (the "Work(s)"). Copyright Clearance Center, Inc. ("CCC") grants licenses through the Service on behalf of the rightsholder identified on the Order Confirmation (the "Rightsholder"). "Republication", as used herein, generally means the inclusion of a Work, in whole or in part, in a new work or works, also as described on the Order Confirmation. "User", as used herein, means the person or entity making such republication.

2. The terms set forth in the relevant Order Confirmation, and any terms set by the Rightsholder with respect to a particular Work, govern the terms of use of Works in connection with the Service. By using the Service, the person transacting for a republication license on behalf of the User represents and warrants that he/she/it (a) has been duly authorized by the User to accept, and hereby does accept, all such terms and conditions on behalf of User, and (b) shall inform User of all such terms and conditions. In the event such person is a "freelancer" or other third party independent of User and CCC, such party shall be deemed jointly a "User" for purposes of these terms and conditions. In any event, User shall be deemed to have accepted and agreed to all such terms and conditions if User republishes the Work in any fashion.

3. Scope of License; Limitations and Obligations.

    3.1. All Works and all rights therein, including copyright rights, remain the sole and exclusive property of the Rightsholder. The license created by the exchange of an Order Confirmation (and/or any invoice) and payment by User of the full amount set forth on that document includes only those rights expressly set forth in the Order Confirmation and in these terms and conditions, and conveys no other rights in the Work(s) to User. All rights not expressly granted are hereby reserved.

    3.2. General Payment Terms: You may pay by credit card or through an account with us payable at the end of the month. If you and we agree that you may establish a standing account with CCC, then the following terms apply: Remit Payment to: Copyright Clearance Center, 29118 Network Place, Chicago, IL 60673-1291. Payments Due: Invoices are payable upon their delivery to you (or upon our notice to you that they are available to you for downloading). After 30 days, outstanding amounts will be subject to a service charge of 1-1/2% per month or, if less, the maximum rate allowed by applicable law. Unless otherwise specifically set forth in the Order Confirmation or in a separate written agreement signed by CCC, invoices are due and payable on "net 30" terms. While User may exercise the rights licensed immediately upon issuance of the Order Confirmation, the license is automatically revoked and is null and void, as if it had never been issued, if complete payment for the license is not received on a timely basis either from User directly or through a payment agent, such as a credit card company.

    3.3. Unless otherwise provided in the Order Confirmation, any grant of rights to User (i) is "one-time" (including the editions and product family specified in the license), (ii) is non-exclusive and non-transferable and (iii) is subject to any and all limitations and restrictions (such as, but not limited to, limitations on duration of use or circulation) included in the Order Confirmation or invoice and/or in these terms and conditions. Upon completion of the licensed use, User shall either secure a new permission for further use of the Work(s) or immediately cease any new use of the Work(s) and shall render inaccessible (such as by deleting or by removing or severing links or other locators) any further copies of the Work (except for copies printed on paper in accordance with this license and still in User's stock at the end of such period).

    3.4. In the event that the material for which a republication license is sought includes third party materials (such as photographs, illustrations, graphs, inserts and similar materials) which are identified in such material as having been used by permission, User is responsible for identifying, and seeking separate licenses (under this Service or otherwise) for, any of such third party materials; without a separate license, such third party materials may not be used.

    3.5. Use of proper copyright notice for a Work is required as a condition of any license granted under the Service. Unless otherwise provided in the Order Confirmation, a proper copyright notice will read substantially as follows: "Republished with permission of [Rightsholder's name], from [Work's title, author, volume, edition number and year of copyright]; permission conveyed through Copyright Clearance Center, Inc. " Such notice must be provided in a reasonably legible font size and must be placed either immediately adjacent to the Work as used (for example, as part of a by-line or footnote but not as a separate electronic link) or in the place where substantially all other credits or notices for the new work containing the republished Work are located. Failure to include the required notice results in loss to the Rightsholder and CCC, and the User shall be liable to pay liquidated damages for each such failure equal to twice the use fee specified in the Order Confirmation, in addition to the use fee itself and any other fees and charges specified.

    3.6. User may only make alterations to the Work if and as expressly set forth in the Order Confirmation. No Work may be used in any way that is defamatory, violates the rights of third parties (including such third parties' rights of copyright, privacy, publicity, or other tangible or intangible property), or is otherwise illegal, sexually explicit or obscene. In addition, User may not conjoin a Work with any other material that may result in damage to the reputation of the Rightsholder. User agrees to inform CCC if it becomes aware of any infringement of any rights in a Work and to cooperate with any reasonable request of CCC or the Rightsholder in connection therewith.

4. Indemnity. User hereby indemnifies and agrees to defend the Rightsholder and CCC, and their respective employees and directors, against all claims, liability, damages, costs and expenses, including legal fees and expenses, arising out of any use of a Work beyond the scope of the rights granted herein, or any use of a Work which has been altered in any unauthorized way by User, including claims of defamation or infringement of rights of copyright, publicity, privacy or other tangible or intangible property.

5. Limitation of Liability. UNDER NO CIRCUMSTANCES WILL CCC OR THE RIGHTSHOLDER BE LIABLE FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL OR INCIDENTAL DAMAGES (INCLUDING WITHOUT LIMITATION DAMAGES FOR LOSS OF BUSINESS PROFITS OR INFORMATION, OR FOR BUSINESS INTERRUPTION) ARISING OUT OF THE USE OR INABILITY TO USE A WORK, EVEN IF ONE OF THEM HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. In any event, the total liability of the Rightsholder and CCC (including their respective employees and directors) shall not exceed the total amount actually paid by User for this license. User assumes full liability for the actions and omissions of its principals, employees, agents, affiliates, successors and assigns.

6. Limited Warranties. THE WORK(S) AND RIGHT(S) ARE PROVIDED "AS IS". CCC HAS THE RIGHT TO GRANT TO USER THE RIGHTS GRANTED IN THE ORDER CONFIRMATION DOCUMENT. CCC AND THE RIGHTSHOLDER DISCLAIM ALL OTHER WARRANTIES RELATING TO THE WORK(S) AND RIGHT(S), EITHER EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. ADDITIONAL RIGHTS MAY BE REQUIRED TO USE ILLUSTRATIONS, GRAPHS, PHOTOGRAPHS, ABSTRACTS, INSERTS OR OTHER PORTIONS OF THE WORK (AS OPPOSED TO THE ENTIRE WORK) IN A MANNER CONTEMPLATED BY USER; USER UNDERSTANDS AND AGREES THAT NEITHER CCC NOR THE RIGHTSHOLDER MAY HAVE SUCH ADDITIONAL RIGHTS TO GRANT.

7. Effect of Breach. Any failure by User to pay any amount when due, or any use by User of a Work beyond the scope of the license set forth in the Order Confirmation and/or these terms and conditions, shall be a material breach of the license created by the Order Confirmation and these terms and conditions. Any breach not cured within 30 days of written notice thereof shall result in immediate termination of such license without further notice. Any unauthorized (but licensable) use of a Work that is terminated immediately upon notice thereof may be liquidated by payment of the Rightsholder's ordinary license price therefor; any unauthorized (and unlicensable) use that is not terminated immediately for any reason (including, for example, because materials containing the Work cannot reasonably be recalled) will be subject to all remedies available at law or in equity, but in no event to a payment of less than three times the Rightsholder's ordinary license price for the most closely analogous licensable use plus Rightsholder's and/or CCC's costs and expenses incurred in collecting such payment.

8. Miscellaneous.

    8.1. User acknowledges that CCC may, from time to time, make changes or additions to the Service or to these terms and conditions, and CCC reserves the right to send notice to the User by electronic mail or otherwise for the purposes of notifying User of such changes or additions; provided that any such changes or additions shall not apply to permissions already secured and paid for.

    8.2. Use of User-related information collected through the Service is governed by CCC's privacy policy, available online here: https://marketplace.copyright.com/rs-ui-web/mp/privacy-policy

    8.3. The licensing transaction described in the Order Confirmation is personal to User. Therefore, User may not assign or transfer to any other person (whether a natural person or an organization of any kind) the license created by the Order Confirmation and these terms and conditions or any rights granted hereunder; provided, however, that User may assign such license in its entirety on written notice to CCC in the event of a transfer of all or substantially all of User's rights in the new material which includes the Work(s) licensed under this Service.

    8.4. No amendment or waiver of any terms is binding unless set forth in writing and signed by the parties. The Rightsholder and CCC hereby object to any terms contained in any writing prepared by the User or its principals, employees, agents or affiliates and purporting to govern or otherwise relate to the licensing transaction described in the Order Confirmation, which terms are in any way inconsistent with any terms set forth in the Order Confirmation and/or in these terms and conditions or CCC's standard operating procedures, whether such writing is prepared prior to, simultaneously with or subsequent to the Order Confirmation, and whether such writing appears on a copy of the Order Confirmation or in a separate instrument.

    8.5. The licensing transaction described in the Order Confirmation document shall be governed by and construed under the law of the State of New York, USA, without regard to the principles thereof of conflicts of law. Any case, controversy, suit, action, or proceeding arising out of, in connection with, or related to such licensing transaction shall be brought, at CCC's sole discretion, in any federal or state court located in the County of New York, State of New York, USA, or in any federal or state court whose geographical jurisdiction covers the location of the Rightsholder set forth in the Order Confirmation. The parties expressly submit to the personal jurisdiction and venue of each such federal or state court.If you have any comments or questions about the Service or Copyright Clearance Center, please contact us at 978-750-8400 or send an e-mail to support@copyright.com.

v 1.1