



MAX-PLANCK-INSTITUT FÜR MOLEKULARE PFLANZENPHYSIOLOGIE

ORGANELLENBIOLOGIE, BIOTECHNOLOGIE UND MOLEKULARE ÖKOPHYSIOLOGIE

Identification of chloroplast translational feedback regulation and establishment of aptamer-based mRNA purification to unravel involved regulatory factors

Dissertation

zur Erlangung des akademischen Grades

"doctor rerum naturalium"

(Dr. rer. nat.)

in der Wissenschaftsdisziplin "Molekularbiologie"

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät

Institut für Biologie und Biochemie

der Universität Potsdam

von

Rabea Ghandour

Disputation am 05. Oktober 2020

Hauptbetreuer: Prof. Dr. Ralph Bock

Betreuer: Dr. Reimo Zoschke

Gutachterinnen: Prof. Dr. Ralph Bock
Prof. Dr. Christian Schmitz-Linneweber
Prof. Dr. Francis-André Wollman

Published online on the
Publication Server of the University of Potsdam:
<https://doi.org/10.25932/publishup-48289>
<https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-482896>

What seems today inconceivable will appear one day, from a higher stand point, quite simple and harmonious.

Max Planck

Summary

After endosymbiosis, chloroplasts lost most of their genome. Many former endosymbiotic genes are now nucleus-encoded and the products are re-imported post-translationally. Consequently, photosynthetic complexes are built of nucleus- and plastid-encoded subunits in a well-defined stoichiometry. In *Chlamydomonas*, the translation of chloroplast-encoded photosynthetic core subunits is feedback-regulated by the assembly state of the complexes they reside in. This process is called Control by Epistasy of Synthesis (CES) and enables the efficient production of photosynthetic core subunits in stoichiometric amounts. In chloroplasts of embryophytes, only Rubisco subunits have been shown to be feedback-regulated. That opens the question if there is additional CES regulation in embryophytes. I analyzed chloroplast gene expression in tobacco and *Arabidopsis* mutants with assembly defects for each photosynthetic complex to broadly answer this question. My results (i) confirmed CES within Rubisco and hint to potential translational feedback regulation in the synthesis of (ii) cytochrome *b₆f* (Cyt *b₆f*) and (iii) photosystem II (PSII) subunits. This work suggests a CES network in PSII that links *psbD*, *psbA*, *psbB*, *psbE*, and potentially *psbH* expression by a feedback mechanism that at least partially differs from that described in *Chlamydomonas*. Intriguingly, in the Cyt *b₆f* complex, a positive feedback regulation that coordinates the synthesis of PetA and PetB was observed, which was not previously reported in *Chlamydomonas*. No evidence for CES interactions was found in the expression of NDH and ATP synthase subunits of embryophytes. Altogether, this work provides solid evidence for novel assembly-dependent feedback regulation mechanisms controlling the expression of photosynthetic genes in chloroplasts of embryophytes.

In order to obtain a comprehensive inventory of the *rbcL* and *psbA* RNA-binding proteomes (including factors that regulate their expression, especially factors involved in CES), an aptamer-based affinity purification method was adapted and refined for the specific purification these transcripts from tobacco chloroplasts. To this end, three different aptamers (MS2, Sephadex-, and streptavidin-binding) were stably introduced into the 3' UTRs of *psbA* and *rbcL* by chloroplast transformation. RNA aptamer-based purification and subsequent chip analysis (RAP-Chip) demonstrated a strong enrichment of *psbA* and *rbcL* transcripts and currently, ongoing mass spectrometry analyses shall reveal potential regulatory factors. Furthermore, the suborganellar localization of MS2-tagged *psbA* and *rbcL* transcripts was analyzed by a combined affinity, immunology, and electron microscopy approach and demonstrated the potential of aptamer tags for the examination of the spatial distribution of chloroplast transcripts.

Zusammenfassung

Nach der Endosymbiose wurde der größte Teil des Chloroplastengenoms in das Kerngenom transferiert. Die entsprechenden Genprodukte werden posttranslational wieder in die Chloroplasten importiert. Dementsprechend sind photosynthetische Proteinkomplexe aus plastidär- und kernkodierten Untereinheiten in definierter Stöchiometrie zusammengesetzt. In der einzelligen Grünalge *Chlamydomonas* ist die Translation von chloroplastenkodierten photosynthetischen Untereinheiten durch einen Rückkopplungsmechanismus in Abhängigkeit vom Assemblierungsstatus der entsprechenden Komplexe reguliert. Dieser „Control by Epistasy of Synthesis“ (CES) genannte Mechanismus erlaubt die effiziente Synthese von photosynthetischen Untereinheiten in den stöchiometrischen Mengen, die für die Assemblierung der Komplexe benötigt werden. In den Chloroplasten der Embryophyten wurde bisher nur die Translation von Rubisco als CES-reguliert beschrieben. Daher stellt sich die Frage, ob derartige CES-Regulationen in Embryophyten auch in anderen Photosynthesekomplexen stattfinden. Um diese Frage zu beantworten, habe ich die chloroplastidäre Genexpression in Tabak- und Arabidopsismutanten mit Defekten in der Assemblierung photosynthetischer Komplexe untersucht. Meine Ergebnisse bestätigen (i) die bekannte CES-Regulation von Rubisco und zeigen mögliche weitere assemblierungsabhängige Rückkopplungsregulationen in der Synthese des (ii) Cytochrom *b₆f* (Cyt *b₆f*) Komplexes sowie des (iii) Photosystems II (PSII). Insbesondere weisen meine Ergebnisse auf ein CES-Netzwerk hin, welches die Expressionen von *psbD*, *psbA*, *psbB*, *psbE* und wahrscheinlich auch *psbH* steuert und teilweise von der beschriebenen linearen CES-Kaskade in *Chlamydomonas* abweicht. Für die Synthese des Cyt *b₆f* Komplexes wurde zudem eine positive Feedback-Regulation der Untereinheiten PetA und PetB beobachtet, die in *Chlamydomonas* nicht gezeigt wurde. Dagegen wurden für die NDH- und ATP Synthase-Komplexe keine Hinweise auf CES-Regulation in Embryophyten gefunden. Zusammenfassend zeigen meine Ergebnisse klare Belege für bisher unbekannte CES-Regulationen, welche die Expression von photosynthetischen Genen in Embryophyten steuern.

Um das mRNA-Protein-Interaktom von *rbcL* und *psbA* zu bestimmen (einschließlich Faktoren, welche CES regulieren), wurde eine aptamer-basierte Affinitätsreinigungsmethode für die Anreicherung dieser Transkripte aus Tabakchloroplasten adaptiert und optimiert. Dazu wurden mittels Chloroplastentransformation drei verschiedene Aptamere (MS2, Sephadex- und Streptavidin-bindende Aptamere) stabil in den 3'UTR der Transkripte integriert. Die aptamer-basierte RNA-Aufreinigung und anschließende Chip-Analyse (RAP-Chip) zeigte die spezifische Anreicherung der *psbA*- bzw. *rbcL*-Transkripte. Die aktuell ausgeführte Massenspektrometrie zur Analyse der transkriptgebundenen Proteine soll potenziell regulatorische Faktoren identifizieren. Des Weiteren wurde die Lokalisation der MS2-markierten *psbA*- und *rbcL*-Transkripte innerhalb des Chloroplasten mittels Affinitätsimmunologie und Elektronenmikroskopie untersucht und dabei gezeigt, dass die Aptamer-Markierung geeignet ist, um die Transkriptverteilung innerhalb von Organellen zu untersuchen.

Table of contents

Summary	I
Zusammenfassung	II
Table of contents	III
List of figures	VII
List of supplemental figures	IX
List of tables	X
List of supplemental tables	X
Abbreviations	XI
1 Introduction	1
1.1 Origin and differentiation of the plastids	1
1.2 Chloroplast genome	1
1.2.1 Co-evolution with the nuclear genome	1
1.2.2 Engineering of the chloroplast genome.....	3
1.2.3 Oxygenic photosynthesis within chloroplasts	4
1.3 Chloroplast gene expression	4
1.3.1 Co-evolutionary innovations: transcription and post-transcriptional processing	4
1.3.2 The bacterial-like chloroplast translation machinery	6
1.3.3 Translation: the major step regulating chloroplast gene expression?.....	7
1.4 Coordinated accumulation of photosynthetic complexes	8
1.4.1 Cotranslational targeting of plastid-encoded photosynthetic subunits.....	8
1.4.2 Assembly of multimeric photosynthetic protein complexes	9
1.4.3 Translational feedback regulation: a common theme in all domains of life	10
1.4.4 CES, a major translational feedback regulation mechanism in chloroplast and mitochondria	10
1.4.5 Protease-dependent regulation	13
1.5 Ribosome profiling: genome-wide analysis of translation.....	13
1.6 Nucleus-encoded translation factors: a brief summary	15
1.7 Purification of chloroplast ribonucleoproteins (RNP) complexes	16
1.7.1 Classical methods for RNP purification.....	16
1.7.2 Aptamer-based affinity purification.....	17
1.8 Aptamer-based RNA tracking.....	19
1.9 Aim and strategies.....	19
1.9.1 Search for feedback regulation in the expression of the subunits of photosynthetic complexes in land plants	19

1.9.2	Identification of chloroplast translation factors using an aptamer-based affinity purification approach	20
2	Material and methods	21
2.1	Material	21
2.1.1	Suppliers of chemicals and consumables	21
2.1.2	Specialized chemicals	21
2.1.3	Consumables	22
2.1.4	Reaction kits.....	23
2.1.5	Common enzymes.....	24
2.1.6	Molecular weight markers.....	24
2.1.7	Equipment	24
2.1.8	Antibodies	26
2.1.9	Oligonucleotides	28
2.2	Methods	31
2.2.1	Growth, cultivation, and transformation of plants and bacteria.....	31
2.2.2	Nucleic acid analysis.....	34
2.2.3	Microarray-based ribosome profiling	40
2.2.4	Ribo-seq	42
2.2.5	Protein analysis	44
2.2.6	Expression and purification of the affinity-tagged adapter protein MS2-MBP	48
2.2.7	Aptamer-based affinity purification.....	49
2.2.8	Mass spectrometry	51
2.2.9	Transmission electron microscopy (TEM).....	52
3	Results	54
3.1	Identification of assembly-dependent translational feedback regulation in photosynthetic complexes of embryophytes.....	54
3.1.1	Selection of mutants with assembly defects in photosynthetic complexes	54
3.1.2	Analysis of feedback regulation in chloroplast gene expression	56
3.1.3	Confirmation of the only known CES in embryophytes: translation of RbcL is regulated by the availability of its assembly partner, RbcS	57
3.1.4	The translational feedback regulation of <i>rbcL</i> occurs at the initiation level	59
3.1.5	MRL1 is a potential translation factor involved in the CES regulation of <i>rbcL</i>	60
3.1.6	Identification of potential CES regulation in PSII	62
3.1.7	Evidence for a PSII CES cascade in embryophytes	64
3.1.8	Reduced PsbD levels induce significant alterations of chloroplast ribosome occupancy on PSII and PSI transcripts	69

3.1.9	Impaired <i>psbA</i> expression does not cause alterations in the expression of other PSII subunits	72
3.1.10	Impaired cytochrome <i>b₅₅₉</i> expression induces reduced <i>psbA</i> translation	73
3.1.11	The knockout of a PSII assembly factor causes a decrease in the translation output of <i>psbB</i>	74
3.1.12	The knockout of a PSII assembly factor evokes significant alteration of ribosome distribution on <i>psbA</i> , <i>psbB</i> , and <i>rbcL</i> transcripts at the cotyledon stage.....	75
3.1.13	Knockdown of <i>psaA</i> and <i>psaD</i> does not induce substantial translation feedback regulation in PSI.....	76
3.1.14	Knockout of a PSI assembly factor drives a potential CES regulation of <i>psaC</i> and <i>psaI</i>	78
3.1.15	Potential positive feedback regulation between <i>petA</i> and <i>petB</i> in cytochrome <i>b_{6f}</i> complex.....	80
3.1.16	Confirmation of enhanced synthesis of PetB upon overexpression of <i>petA</i> by pulse labeling	82
3.1.17	Enhancement of <i>petB</i> translation output in Δ <i>petL</i> mutant.....	83
3.1.18	Lack of CES interaction between AtpB and AtpA and potential side effects upon reduction of AtpC.....	84
3.1.19	Study of gene expression in NDH dehydrogenase mutants does not reveal any substantial feedback regulation	86
3.1.20	Ribo-seq enables analysis of translation in knockout mutants of photosynthetic subunits	88
3.2	Identification and characterization of <i>psbA</i> and <i>rbcL</i> translation factors using an aptamer-based affinity purification	93
3.2.1	Establishment of an MS2 aptamer-based affinity purification in chloroplast.....	94
3.2.2	Streptavidin-binding aptamer enables the specific affinity purification of <i>psbA</i> but not <i>rbcL</i>	107
3.2.3	<i>rbcL</i> and <i>psbA</i> can be purified using the Sephadex-binding aptamer	109
3.2.4	Comparison of the MS2, streptavidin- and Sephadex-binding aptamers in regard to the enrichment efficiency and integrity of the purified <i>psbA</i> and <i>rbcL</i> mRNAs	111
3.2.5	Mass spectrometry analysis of the RNA-binding proteome of <i>psbA</i>	114
3.2.6	Mass spectrometry analysis of the RNA-binding proteome of <i>rbcL</i>	115
4	Discussion and outlook	116
4.1	Identification of assembly-dependent translational feedback regulation in photosynthetic complexes of embryophytes.....	116
4.1.1	Ribosome profiling readily detects CES in the assembly of Rubisco in embryophytes	116
4.1.2	Evidence for a PSII CES network in embryophytes	117

4.1.3	Identification of potential positive feedback regulation in Cyt <i>b₆f</i> complex.....	121
4.1.4	Search for potential CES regulation in the ATP synthase complex.....	122
4.1.5	Preliminary analyses do not provide evidence for CES regulation in the NDH complex 123	
4.1.6	Towards the confident identification of CES regulation and the examination of its interplay with protein degradation	123
4.1.7	General lessons on chloroplast translation gained from the analysis of numerous chloroplast mutants	124
4.1.8	Are there no true “neutral” insertion sites in the chloroplast genome?.....	125
4.2	Towards unraveling of the RNA-binding proteomes and localizations of <i>psbA</i> and <i>rbcL</i> mRNAs using an aptamer-based affinity purification.....	127
4.2.1	Chloroplast transformation and homoplasmy of transplastomic plants	127
4.2.2	Establishment of an aptamer-based affinity purification for chloroplast transcripts	127
4.2.3	Pros and cons of the aptamer-based affinity purification of chloroplast RNPs	129
4.2.4	Using an aptamer to study the suborganellar localization of chloroplast transcripts	130
5	Supplemental information.....	132
6	References	150
7	Publications and conference attendance	171
7.1	Publications.....	171
7.2	Conference and workshops attendance	171
8	Eidesstattliche Erklärung.....	172
9	Acknowledgements.....	173

List of figures

Figure 1.1: Physical map of the tobacco chloroplast genome.	3
Figure 1.2: Overview of essential steps in chloroplast gene expression.	6
Figure 1.3: CES contributes to the biogenesis of the photosynthetic apparatus in <i>Chlamydomonas</i>	13
Figure 3.1: Reduced levels of RBCS provoke a decrease in the transcript and ribosome footprint abundances of <i>rbcL</i> and a defect in the protein accumulation of RbcL.	58
Figure 3.2: Pulse labeling of as- <i>RBCS</i> mutant validates the CES regulation of <i>rbcL</i>	60
Figure 3.3: MRL1 is a potential mediator of <i>rbcL</i> CES regulation.	61
Figure 3.4: Transcriptome and ribosome footprint profiling of <i>psbD</i> translation mutants reveals potential CES regulation of <i>psbA</i>	64
Figure 3.5: Transcriptome and ribosome footprint analyses reveals potential CES regulation of <i>psbB</i> in KD- <i>psbD</i> mutant.	66
Figure 3.6: Knockdown of <i>psbD</i> results in similar translation regulatory effects in embryophytes in comparison to <i>Chlamydomonas</i>	68
Figure 3.7: PsbA synthesis is downregulated when PsbD protein level is reduced.	69
Figure 3.8: Transcriptome-wide analysis of ribosome occupancy in KD- <i>psbD</i> mutant reveals significant alterations in PSII and PSI transcripts.	71
Figure 3.9: Defect in the expression of <i>psbA</i> has no effect on the translation of <i>psbB</i> or other transcripts encoding PSII subunits.	72
Figure 3.10: Reduced expression of <i>cyt b559</i> results in the decrease of the translation output of <i>psbA</i>	73
Figure 3.11: Decrease of the translation output of <i>psbB</i> in Δ <i>psbN</i> mutant.	75
Figure 3.12: Altered ribosome footprint distribution in Δ <i>psbN</i> at the cotyledon stage.	76
Figure 3.13: Reduced expression of <i>psaA</i> and <i>psaD</i> does not trigger CES regulation within PSI.	78
Figure 3.14: <i>ycf4</i> -dependent assembly defect of PSI provokes numerous alterations on translation level.	79
Figure 3.15: Enhanced translation output of <i>petA</i> triggers the translation of <i>petB</i> in Δ <i>psaI</i> and Δ <i>ycf10</i> mutants.	81
Figure 3.16: PetB synthesis is upregulated in Δ <i>psaI</i>	82
Figure 3.17: Overexpression of <i>petG</i> enhances <i>petB</i> translation.	84

Figure 3.18: Reduced AtpC levels causes alterations in the expression of several chloroplast-encoded genes.....	86
Figure 3.19: Lack of substantial CES interactions in <i>ndhC/K/J</i> and <i>ndhA/H/I</i> mutants.....	88
Figure 3.20: Phenotype of the knockout mutants.....	89
Figure 3.21: Truncation of the C-terminus of PsbD induces a decrease in the translation output of <i>psbA</i>	90
Figure 3.22: Translation of <i>psbA</i> is compromised by the PSII assembly defect in $\Delta psbB$ operon mutant whereas <i>petA</i> translation is independent of PetB or PetD.	92
Figure 3.23: The knockout of the ATP synthase complex results in global alteration of translation in the chloroplast.	93
Figure 3.24: Insertion of the selected aptamers does not disrupt the structures of the 3' UTR of <i>psbA</i> and <i>rbcL</i> <i>in silico</i>	95
Figure 3.25: Physical map of the chloroplast genome in wild type and in <i>psbA</i> and <i>rbcL</i> transplastomic lines.	97
Figure 3.26: RFLP analysis of the <i>psbA</i> -MS2 and <i>rbcL</i> -MS2 primary transformants.....	98
Figure 3.27: Tagging of <i>psbA</i> with three different aptamers by stable transformation of the plastome.	99
Figure 3.28: Tagging of <i>rbcL</i> with three different aptamers by stable transformation of the plastome.	100
Figure 3.29: Double maltose and heparin column purification yields a pure MS2-MBP recombinant protein.....	102
Figure 3.30: <i>psbA</i> and <i>rbcL</i> mRNAs are efficiently and specifically purified with the MS2 aptamer.	105
Figure 3.31: Aptamer-based measurements of suborganellar localization of <i>psbA</i> and <i>rbcL</i> transcripts.	107
Figure 3.32: Purification of <i>psbA</i> mRNA with the streptavidin-binding aptamer.....	109
Figure 3.33: Sephadex-binding aptamer enables the purification of <i>rbcL</i> mRNA.....	111
Figure 3.34: Comparison of the purification of <i>psbA</i> and <i>rbcL</i> with different aptamers.	113
Figure 4.1: Structure of PSII complex.....	119
Figure 4.2: PSII CES network in embryophytes	120
Figure 4.3: Proposed model for translational feedback regulation of <i>psbA</i>	121

List of supplemental figures

Supplemental Figure 1: Transcriptome-wide analysis of the total RNA coverage in KD- <i>psbD</i> mutant reveals no significant alterations.	132
Supplemental Figure 2: Assembly defect of PSII in Δ <i>psbN</i> causes <i>psbB</i> downregulation at the cotyledon stage.....	132
Supplemental Figure 3: Transcriptome-wide analysis of the total RNA coverage in Δ <i>psbN</i> mutant at the cotyledon stage.....	133
Supplemental Figure 4: Ribosome profiling of <i>atpB</i> -GTG at the cotyledon stage.	133
Supplemental Figure 5: Ratio of the translation output in as- <i>AtpC</i> in comparison to WT.....	134
Supplemental Figure 6: Truncated PsbD is produced by the <i>aadA</i> insertion.	134
Supplemental Figure 7: Ratio of the translation output in Δ <i>atpB</i> in comparison to pRB8.	134

List of tables

Table 1: Overview of antibodies used in this work	26
Table 2: List of oligonucleotides used in this work	28
Table 3: Overview of plasmids used or generated in this work	34
Table 4: PCR basic reactions mix and program	37
Table 5: Analyzed tobacco and <i>Arabidopsis</i> mutants	55

List of supplemental tables

Supplemental Table 1: Reproducibility of ORF average of transcripts between the biological replicates	135
Supplemental Table 2: Reproducibility of ORF average of ribosome footprint between the biological replicates.....	137
Supplemental Table 3: Reproducibility of the probe signals of the transcript between the biological replicates.....	139
Supplemental Table 4: Reproducibility of the probe signals of ribosome footprints between the biological replicates.....	141
Supplemental Table 5: Mapping statistics of Ribo-seq data	143
Supplemental Table 6: Proteins enriched in the <i>psbA</i> -MS2 affinity purification.....	145
Supplemental Table 7: Proteins enriched in the <i>psbA</i> affinity purification experiments (with MS2 aptamer, with streptavidin-binding aptamer without elution and with 30-min elution).....	147
Supplemental Table 8: Summary of the proteins enriched in the <i>rbcL</i> -MS2 purification	148

Abbreviations

%	percent
°C	degree Celsius
μg	microgram
μL	microliter
μm	micrometer
μmol m ⁻² s ⁻¹	micromol per square meter per second (light intensity unit)
A	adenosine
<i>aadA</i>	gene encoding streptomycin/spectinomycin adenylyl transferase
Amp	ampicillin
aSD	anti-Shine-Dalgarno
ATP	adenosine triphosphate
BAP	<i>N</i> ⁶ -benzylamino purine
bp	base pair
BSA	bovine serum albumin
CLIP	cross-linking immunoprecipitation
cm	centimeter
cp	chloroplast
cpm	counts per minute
cpRNP	chloroplast ribonucleoproteins
<i>Cr</i>	<i>Chlamydomonas reinhardtii</i>
CTAB	cetyltrimethyl ammonium bromide
Cyt <i>b_{6f}</i>	cytochrome <i>b_{6f}</i>
dCTP	deoxycytidine triphosphate
dNTP	deoxyribonucleotide triphosphate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
<i>E.coli</i>	<i>Escherichia coli</i>
e.g.	<i>exempli gratia</i> (for example)
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
et al.	<i>et alia</i> (and others)
EtBr	ethidium bromide
etc.	et cetera
FDR	false discovery rate
FPLC	fast protein liquid chromatography
g	acceleration of gravity or gram
gDNA	genomic DNA
h	hour
HAT	half a tetratricopeptide

HCl	hydrogen chloride
i.e.	<i>id est</i> (in other words)
IPTG	isopropyl β -D-1-thiogalactopyranoside
UGV	integrated genome viewer
kb	kilobase
kDa	kilodalton
KOH	potassium hydroxide
L	liter
LB	Luria-Bertani
LC-MS	liquid chromatography mass spectrometry
M	molar
MBP	maltose binding protein
MCP	MS2 coat protein
mg	milligram
min	minute
mL	milliliter
mM	millimolar
mTERF	mitochondrial transcription termination factor
MNase	micrococcal nuclease
MQ-H ₂ O	Milli-Q water
mRNA	messenger ribonucleic acid
MS	Murashige and Skoog media
NAA	1-Naphthylacetic acid
NaOH	sodium hydroxide
NDH	NADH dehydrogenase
NEP	nucleus-encoded plastid RNA polymerase
ng	nanogram
nt	nucleotide
<i>Nt</i>	<i>Nicotiana tabacum</i>
OEC	oxygen-evolving complex
ORF	open reading frame
P	promoter or phosphate
PCR	polymerase chain reaction
PEP	plastid-encoded plastid RNA polymerase
PH	Petit Havana
PPR	pentatricopeptide repeat
Prrn	rRNA promoter
PSI	Photosystem I
PSII	Photosystem II
ptDNA	plastid DNA
RBD	RNA-binding domain

Ribo-seq	deep-sequencing-based ribosome profiling
RIP	RNA immunoprecipitation
RMOP	revised medium for organogenesis of <i>Nicotiana plumbaginifolia</i>
RNA	ribonucleic acid
RNase	ribonuclease
RNA-Seq	RNA sequencing
RNP	ribonucleoprotein
RPKM	reads per kilobase per million
rpm	revolutions per minute
RRM	RNA recognition motif
rRNA	ribosomal RNA
RT	room temperature
RuBisCO	Ribulose-1,5-bisphosphate-carboxylase/oxygenases
S	sulfur
SD	Shine-Dalgarno
SDS	sodium dodecyl sulfate
sec	second(s)
Spec	spectinomycin
SSC	saline-sodium citrate buffer
ssDNA	single-stranded DNA
ssRNA	single-stranded RNA
Strep	streptomycin
T	terminator
TAE	tris-acetate-EDTA
TBST	tris-buffered saline with Tween20
TEM	transmission electron microscopy
TEMED	tetramethylethylenediamine
tris	tris (hydroxymethyl) aminomethane
U	unit
UTR	untranslated region
UV	ultra violet
V	volt
v/v	volume/volume
w/v	weight/volume
WT	wild-type
<i>ycf</i>	gene encoding hypothetical chloroplast open reading frame

1 Introduction

1.1 Origin and differentiation of the plastids

Chloroplasts are subcellular organelles characteristic for photosynthetic organisms. They are highly complex and essentially involved in photosynthesis as well as many other metabolic processes (Neuhaus and Emes, 2000). Chloroplasts are part of a large group of organelles, named plastids, which includes proplastids, etioplasts, amyloplasts, elaioplasts, gerontoplasts, and chromoplasts (Jarvis and Lopez-Juez, 2013; Lopez-Juez and Pyke, 2005). Members of this group have a wide range of important specialized roles, for example, starch storage in amyloplasts (Miyazawa et al., 1999), accumulation of carotenoids in chromoplasts (Egea et al., 2010) and oil accumulation in elaioplasts. As first suggested by Mereschkowski (1905), chloroplasts were acquired through endosymbiosis of an oxygenic photosynthetic cyanobacterium into a non-photosynthetic eukaryotic host that already possessed a mitochondrion (Palmer, 2003; Sagan, 1967).

1.2 Chloroplast genome

1.2.1 Co-evolution with the nuclear genome

Throughout more than one billion years following the endosymbiotic event, the chloroplast lost most of its genetic material (Timmis et al., 2004). Most of the lost genes were relocated to the nucleus by lateral gene transfer (Bock and Timmis, 2008; Martin, 2003). Today's embryophytes chloroplasts encode a small set of 100 to 120 genes (Figure 1.1), which represents only ~ 5 % of the cyanobacterial genetic information (Martin et al., 2002). About 3000 proteins are located in chloroplasts, most of which are nucleus-encoded and imported to the chloroplast post-translationally. Not only that this complex system of host-endosymbiont interaction accentuates the nuclear primacy over the chloroplast but it poses also a dilemma about the reasons for retaining a small conserved cluster of genes in a separate compartment in the plant cell. Allen (2015) proposed the CoRR (colocation for redox regulation) hypothesis according to which the genes retained in the chloroplast are those whose expression needs to be under the direct regulation of the redox state of their gene products. However, the decisive reason that accounts for the retention of few genes in the chloroplast remains ambiguous.

Chloroplast genes can be separated into three groups: 'photosynthesis-related', 'genetic system', and 'miscellaneous' (Figure 1.1). Approximately 50 of the genes retained in the chloroplast are photosynthesis-related and encode for the subunits and assembly factors of the photosynthetic complexes, namely photosystem II (PSII), photosystem I (PSI), cytochrome *b₆f* (Cyt *b₆f*), ATP synthase and the NAD(P)H dehydrogenase-like (NDH) complexes as well as the large subunit of the ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). Almost 60 genes are genetic-system genes encoding for products involved in the different steps of the chloroplast gene expression including transcription, RNA processing, translation, and protein degradation. Among these genes are those

encoding for a complete set of tRNAs (Alkatib et al., 2012; Rogalski et al., 2008) and rRNAs, the core subunits of the plastid-encoded RNA polymerase (PEP) (RpoA, RpoB, RpoC1, and RpoC2) and approximately one-third of the ribosomal proteins (Zoschke and Bock, 2018). In addition, *matK*, encoding a putative splicing factor of group II introns (Zoschke et al., 2010), *clpP*, encoding the P subunit of the caseinolytic protease (Clp) (Shanklin et al., 1995; Shikanai et al., 2001) and CcsA, a *c*-type cytochrome biogenesis protein involved in heme attachment (Orsat et al., 1992; Xie and Merchant, 1996) are chloroplast-encoded. The remaining genes represent the small group of miscellaneous genes. In dicots this group includes *accD*, encoding a subunit of the acetyl-CoA carboxylase (Kode et al., 2005; Sasaki et al., 1993a; Sasaki et al., 1995) and three conserved open reading frames (hypothetical chloroplast open reading frame, *ycf*). The *ycf10* gene encodes a chloroplast inner envelope protein involved in the Calvin-Benson-Bassham (CBB) cycle reactions (Rolland et al., 1997; Sasaki et al., 1993b), the *ycf1* gene product is proposed to function in protein import (Kikuchi et al., 2013) while the function of the *ycf2* gene product remains unclear (Drescher et al., 2000).

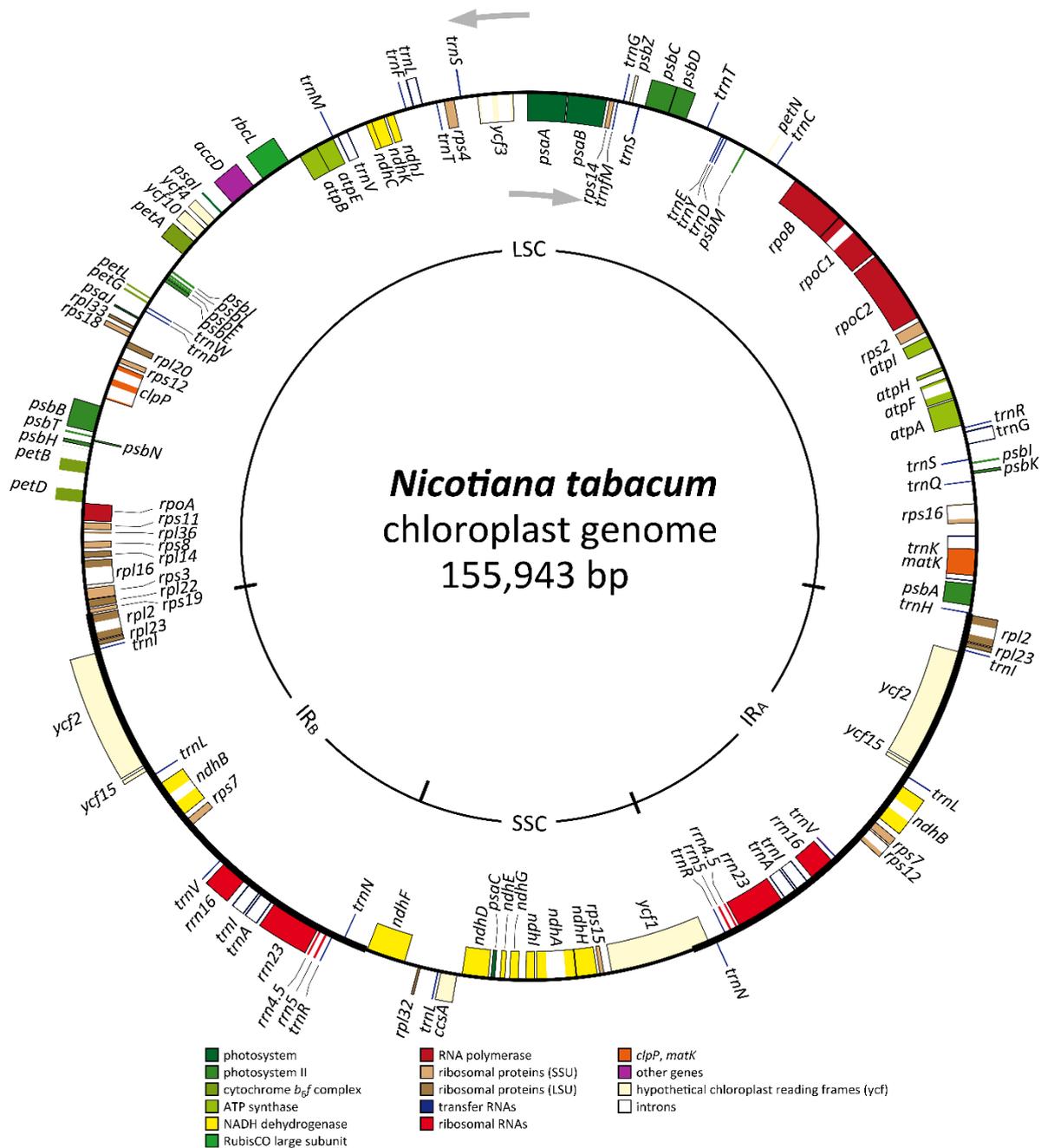


Figure 1.1: Physical map of the tobacco chloroplast genome.

The map was drawn with OGDRAW (Greiner et al., 2019) using the reference sequence Z00044.2 from NCBI. Genes inside and outside the circle are transcribed following the direction indicated by the gray arrows. Colored boxes indicate the coding regions. The color code is indicated at the bottom. LSC: large single-copy region; SSC: small single-copy region; IRA: inverted repeat A; IRB: inverted repeat B.

1.2.2 Engineering of the chloroplast genome

Transformation of the chloroplast genome was first established in *Chlamydomonas reinhardtii* (referred to as *Chlamydomonas* hereafter) (Boynton et al., 1988) followed by *Nicotiana tabacum* (referred to as tobacco hereafter) (Svab et al., 1990). In the following years, chloroplast genome engineering was enabled for several species including rice (Lee et al., 2006), tomato (Ruf et al., 2001), potato (Sidorov et al., 1999), sugar beet (De Marchis et al., 2009) and many others. Most recently, chloroplast

transformation protocols for the model plant *Arabidopsis thaliana* (referred to as *Arabidopsis* hereafter) were established (Ruf et al., 2019; Yu et al., 2017). The most prominent advantage of chloroplast transformation over nuclear transformation is that, in most species, chloroplasts are maternally inherited which facilitates the containment of transgenes (Ruf et al., 2007). Furthermore, transgenes can be precisely integrated into the chloroplast genome by homologous recombination, are not silenced (Verma and Daniell, 2007), reach exceptionally high expression levels (De Cosa et al., 2001; Oey et al., 2009), and can be arranged and co-expressed in operon-like structures (De Cosa et al., 2001; Krichevsky et al., 2010). Chloroplast transformation has been very successfully used for basic and applied research (Bock, 2015). In the context of this work, tobacco chloroplast transformation was used to create knockdown and knockout mutants of chloroplast genes, and to tag chloroplast-encoded transcripts.

1.2.3 Oxygenic photosynthesis within chloroplasts

The chloroplast is the site where photosynthesis takes place in the plant cell, i.e., where the solar energy is converted into energy-rich molecules. Photosynthesis is divided into two processes, the light reactions, and the carbon fixation reactions or the CBB cycle. The light reactions take place in the thylakoids while the CBB cycle occurs in the stroma. The thylakoid membrane is an intricate system that houses the major photosynthetic multiprotein complexes: PSII, Cyt *b₆f*, PSI, ATP synthase, and NDH complex. The light reactions start with one chlorophyll pigment in PSII absorbing one photon and releasing one electron, which is then passed to a pheophytin molecule. This pheophytin passes the electron to a quinone molecule, which is thus reduced to plastoquinol and starts the electron transport chain in the thylakoid membrane. The electrons are then transferred to the Cyt *b₆f* complex causing transport of protons to the lumen. Plastocyanin transfers the electrons from Cyt *b₆f* complex to PSI. Ferredoxin is the last electron acceptor that assists PSI to ultimately reduce NADP⁺ to NADPH. Finally, the ATP synthase complex harnesses the proton electrochemical gradient to produce ATP. The ATP and NADPH produced by the light reactions and released in the stroma are used for carbon (CO₂) fixation by Rubisco in the CBB cycle. The CBB cycle produces not only the precursors for sucrose and starch synthesis but also intermediates for other biosynthetic processes in the chloroplast such as the shikimic acid pathway (Lichtenthaler, 1999).

1.3 Chloroplast gene expression

1.3.1 Co-evolutionary innovations: transcription and post-transcriptional processing

Some prokaryotic features have been retained in the chloroplast from its cyanobacterial ancestor, for instance, the clustering of many chloroplast genes into polycistronic operon-like units and the prokaryotic-type 70S ribosomes. However, during endosymbiont-host co-evolution many processes that are rare or absent in bacterial gene expression have evolved in chloroplasts.

One of these distinguishable features is the usage of two different types of RNA polymerases, a bacterial-type plastid-encoded polymerase (PEP) and one (monocots) or two (dicots) phage-type nucleus-encoded polymerases (NEP) (Börner et al., 2015). The core subunits of the PEP polymerase are plastid-encoded and associate with one out of six nucleus-encoded sigma factors to build the holoenzyme (Allison, 2000; Lysenko, 2007; Schweer et al., 2010). It is generally thought that PEP transcribes photosynthesis-related genes (Mullet, 1993) and is dominant in mature chloroplasts. Moreover, the transcript accumulation of several photosynthesis-related genes has been linked to sigma factors (Ishizaki et al., 2005; Nagashima et al., 2004; Privat et al., 2003). In *Arabidopsis*, NEP polymerase is encoded by three nuclear *RPOT* (RNA polymerase of the phage T3/T7 type) genes whose products are localized in chloroplasts (*RPOTp*), mitochondria (*RPOTm*), and both organelles (*RPOTmp*) (Hedtke et al., 1997; Hedtke et al., 2000; Hess and Börner, 1999; Kühn et al., 2009) (Figure 1.2). NEP transcribes genetic-system genes and is more active at early developmental stages in non-green cells. However, the functional classification of PEP and NEP is oversimplified as most of the chloroplast genes possess promoters of both polymerases (Swiatecka-Hagenbruch et al., 2008; Zhelyazkova et al., 2012). Additionally, NEP was shown indirectly to be able to transcribe some photosynthesis-related genes (with low efficiency) (Allison et al., 1996).

Furthermore, chloroplasts are characterized by a complex post-transcriptional RNA metabolism, where its transcripts regularly undergo several processing steps, including RNA splicing, RNA editing, and intercistronic and end-processing (Barkan, 2011). All of these steps are uncommon in bacteria (Barkan, 2011; Lyska et al., 2013). The chloroplast transcripts are either monocistronic or polycistronic, the latter of which undergo intercistronic processing (Figure 1.2). The 5' and 3' untranslated regions (UTRs) of the chloroplast transcripts are processed by endo- and exo-ribonucleic activity (Stern et al., 2010). RNA stability factors including pentatricopeptide repeat (PPR) proteins were shown to protect the RNA from nuclease digestion and thereby direct the end-maturation of the transcript (Stern et al., 2010). The chloroplast genome encloses 20 introns classified into two groups (group I and II). Most of the chloroplast introns are group II introns except that in *trnL-UAA*. RNA splicing is reported to be enabled by nucleus-encoded factors as well as the chloroplast-encoded splicing factor, MatK (Schmitz-Linneweber et al., 2015). In plants, RNA editing consists of C-to-U substitution (Figure 1.2) and it was first identified in *rpl2* transcript in maize chloroplast where it led to the creation of a canonical ATG start codon (Hoch et al., 1991). RNA editing has been reported in the chloroplast of all embryophytes except the *Marchantiidae* clade in which RNA editing has been lost during evolution (Freyer et al., 1997). Among the examples reported in the tobacco chloroplast is the implication of RNA editing in the restoration of the C-terminus of the *psbE* transcript (Hayes and Hanson, 2008). In all cases, chloroplast RNA editing is enabled by the editosome whose components are encoded in the nucleus and post-translationally imported to the chloroplast (Small et al., 2020).

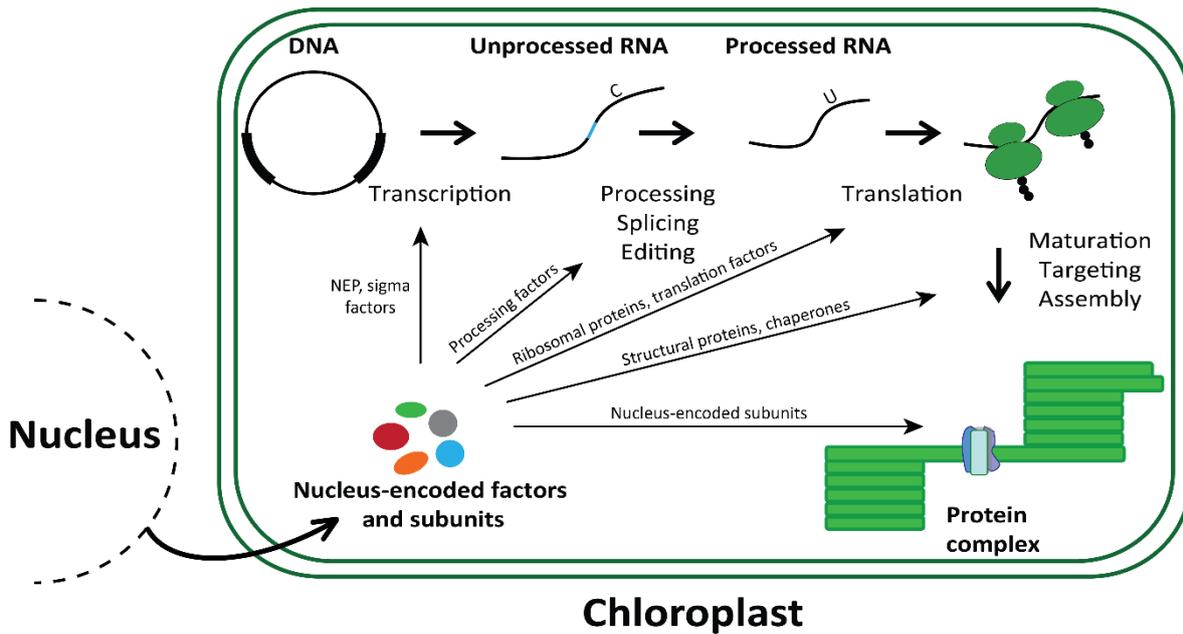


Figure 1.2: Overview of essential steps in chloroplast gene expression.

All steps of chloroplast gene expression are controlled by nucleus-encoded factors that are synthesized in the cytosol and imported to the chloroplast. These nucleus-encoded proteins assist transcription, RNA processing, translation, protein metabolism as well as targeting and assembly of proteins. Adapted from Zoschke and Bock (2018) and Lyska et al. (2013).

1.3.2 The bacterial-like chloroplast translation machinery

As expected from its prokaryotic origin, chloroplast translation is carried out by a bacterial-like translation machinery whose components are dually encoded in the chloroplast and the nucleus (Tiller and Bock, 2014; Zoschke and Bock, 2018). Chloroplast ribosomes exhibit high similarity to 70S bacterial ribosomes (Bieri et al., 2017; Graf et al., 2016) and consist of two subunits: a small 30S subunit and a large 50S subunit. Both of these subunits are ribonucleoprotein complexes that consist of rRNAs and ribosomal proteins. Most chloroplast ribosomal proteins and rRNAs (23S, 16S, and 5S) have orthologs in *Escherichia coli* (*E.coli*) further reaffirming its prokaryotic origin. Nonetheless, the chloroplast translation machinery acquired some distinct features. For instance, six nucleus-encoded ribosomal proteins were described in chloroplast ribosomes and were found to lack bacterial orthologs (Yamaguchi and Subramanian, 2000; Yamaguchi et al., 2000). These proteins were named plastid-specific ribosomal proteins (PSRPs) 1-6. Further investigations have shown later that PSRP1 is an ortholog of the cold-shock protein pY in *E.coli* and is not a part of the small ribosomal subunit (Sharma et al., 2010; Sharma et al., 2007). Moreover, Yamaguchi and Subramanian (2003) have shown that PSRP4 is a homolog to the ribosomal protein THX in *Thermus thermophilus*. Some of the chloroplast ribosomal proteins possess extensions that change the conformation of the ribosomes at the interaction sites with the mRNA and the nascent polypeptide (Bieri et al., 2017). These extensions are thought to compensate for the modifications in the rRNA structure (Ahmed et al., 2016; Bieri et al.,

2017; Graf et al., 2016). The chloroplast possesses four rRNAs species: 16S rRNA located in the 30S subunit and 23S, 5S, and 4.5S rRNAs in the 50S subunit. The chloroplast *rrn23* gene is split into two genes: the 5' fragment encodes the 23S rRNA and the 3' fragment corresponds to the 4.5S rRNA, which is missing in bacteria (Whitfeld et al., 1978). Interestingly, the chloroplast 23S rRNA is posttranscriptionally processed into three fragments in what is known as the “hidden breaks” processing (Nishimura et al., 2010). In addition, there are minor structural differences between the rRNAs of chloroplasts and bacteria, however, the anti-Shine-Dalgarno (aSD) sequence in the 16S rRNA and the catalytic domain in the 23S rRNA are notably conserved. The chloroplast genome also encodes the complete set of tRNA used in chloroplast translation. The other components of the translation machinery such as the initiation, elongation, termination, ribosome recycling factors, and the tRNA synthetases are nucleus-encoded and have bacterial orthologs.

The translation mechanism in chloroplasts is highly similar to that in bacteria given the conservation of the components of the translation machinery. The chloroplast translation starts with the formation of the pre-initiation complex, consisting of the small ribosomal 30S subunit and the initiator tRNA. It was suggested that, in chloroplasts, a scanning mechanism could enable the recognition of the start codon (Drechsel and Bock, 2011). Similar to bacteria, several chloroplast genes possess a Shine-Dalgarno (SD) sequence (Shine and Dalgarno, 1974), which by interacting with the conserved aSD sequence in the 16S rRNA enables the correct positioning of the pre-initiation complex (Drechsel and Bock, 2011; Hirose et al., 1998). Almost 30 % of the chloroplast genes in *Arabidopsis* lack a SD sequence upstream of their start codon (Gawroński et al., 2020; Scharff et al., 2011). The transcripts of these genes were proposed to have less local secondary structure around the start codon, which facilitates their translation initiation (Scharff et al., 2011). The SD-aSD interaction and its relevance in the chloroplast translation were further substantiated in Scharff et al. (2017). In this study, the mutation of the aSD sequence resulted in a massive chloroplast translation defect in genes with SD sequence. The codon ATG is the canonical start codon used in the chloroplast, however, it is evident that GTG and TTG can serve as alternative start codons (Hirose et al., 1999; Hirose and Sugiura, 2004b), although this is thought to be determined by the sequence context (Boeck and Kolakofsky, 1994). After recognition of the start codon, the 50S ribosomal subunit associates to the pre-initiation complex to form the 70S ribosome that proceeds with translation elongation. According to the universal genetic code, the triplets TAA, TGA, and TAG are recognized as stop codons in the chloroplast.

1.3.3 Translation: the major step regulating chloroplast gene expression?

Lots of evidence have shown that the chloroplast gene expression is primarily regulated at post-transcriptional and translational levels in contrast to its cyanobacterial ancestor (Zoschke and Bock, 2018). For example, the plastid transcripts are more stable than their bacterial counterparts whose half-lives are in the range of minutes (Klaff and Gruissem, 1991; Klug, 1993), hence transcriptional responses in chloroplasts are slower. However, in a recent study where metabolic labeling was used to assess RNA

stability in *Arabidopsis*, the measured RNA half-lives were shorter in comparison to previous reports (Szabo et al., 2020). Likewise, chloroplast transcripts are processed to smaller oligo- or monocistronic units (Zhelyazkova et al., 2012), which precludes a bacterial-like transcriptional co-regulation of subunits of protein complexes. Many reports conveyed that the translation of core subunits of the photosynthetic machinery is regulated by light such as the D1 subunit (PsbA), whereas the transcript accumulation remains largely unchanged (Chotewutmontri and Barkan, 2018; Schuster et al., 2020). Moreover, in the unicellular green alga *Chlamydomonas*, the translation of many chloroplast mRNAs is rate-limiting for gene expression and feedback regulation mechanisms control the translation rate of subunits of photosynthetic complexes (Choquet and Wollman, 2009). In addition, many of the factors involved in chloroplast gene expression are proposed to act in post-transcriptional regulation (Barkan and Small, 2014; Lyska et al., 2013).

In general, translation is regulated at the initiation, elongation, and termination levels (Hershey et al., 2012). However, the regulation at the initiation level is the most common mechanism. Most of the reported translation regulatory factors act on the initiation level (Zoschke and Bock, 2018). However, the elongation might also be regulated as shown in Chotewutmontri and Barkan (2018), where the global translation elongation rate in the chloroplast increases in response to light. Other processes might affect the ribosome dynamics, for instance, protein targeting, protein folding, or complex assembly. Ribosome pausing was also proposed to be dependent on the SD sequence (Zoschke et al., 2013).

1.4 Coordinated accumulation of photosynthetic complexes

1.4.1 Cotranslational targeting of plastid-encoded photosynthetic subunits

In chloroplasts, the photosynthetic complexes reside in the thylakoid membrane, an extremely intricate membrane system that houses the light reactions of photosynthesis (Pribil et al., 2014). The ribosomes in chloroplasts are partitioned between the stroma and the thylakoid membrane. The majority of the proteins that reside in the thylakoid membrane are subunits of the photosynthetic machinery and many of these subunits are integral membrane proteins (Figure 1.3). Roughly half of the photosynthetic subunits in the thylakoid membrane are nucleus-encoded, produced in the cytosol, and then post-translationally imported into the chloroplast before being further targeted to the thylakoid membrane. In-depth studies of the targeting mechanisms of the nucleus-encoded proteins to the thylakoid membrane revealed four main pathways: cpSec (chloroplast Secretory), cpTAT (Twin arginine translocation), cpSRP (Signal Recognition Particle) and the “spontaneous” pathway (Celedon and Cline, 2013). On the other hand, less is known about the cotranslational targeting of the plastid-encoded proteins, however, it is thought that these subunits utilize one of the abovementioned pathways (Zoschke and Bock, 2018). Some of the chloroplast-encoded subunits were demonstrated to be translated by thylakoid-bound ribosomes (Jagendorf and Michaels, 1990) and cotranslationally integrated into the thylakoid membrane such as subunits of PSII, PSI, and Cyt *b₆f* complex (Röhl and

van Wijk, 2001; Uniacke and Zerges, 2009; van Wijk et al., 1995; Zoschke and Barkan, 2015). The most thoroughly studied mechanisms of cotranslational targeting are those of PsbA and cytochrome *f* (PetA). Cotranslational targeting of PetA is mediated by cpSecA that binds to its N-terminus (Röhl and van Wijk, 2001). For PsbA, *in vitro* studies suggested the involvement of cpFtsY, cpSecY, ALB3 and Vipp1 in the cotranslational targeting based on their interaction with the nascent peptide (Walter et al., 2015; Zhang et al., 2001). A pioneering study revealed a novel chloroplast mechanism of protein targeting mediated by the interaction between cpSRP54 (chloroplast signal recognition particle) subunit and the ribosomal protein uL4 that initiates cotranslational membrane targeting (Hristou et al., 2019).

In the thylakoid-localized complexes, the plastid-encoded subunits are likely to be assembled cotranslationally given the rapid degradation of unassembled subunits (Zoschke and Bock, 2018). The cotranslational targeting and assembly raise another level of complication of the chloroplast gene expression.

1.4.2 Assembly of multimeric photosynthetic protein complexes

Multimeric protein complexes consist of multiple proteins that assemble according to stoichiometric ratios (Marsh and Teichmann, 2015). Accumulation of a subunit outside of the complex might have negative effects since the subunit is then usually not functional or even deleterious. The stoichiometric accumulation of the constituent subunits of protein complexes is controlled by two major regulatory processes: 1) proteolysis of unassembled subunits, 2) assembly-dependent feedback regulation of the synthesis of the subunit. Evidence from ribosome profiling data has shown that the production of subunits of multiprotein complexes is proportional to their stoichiometry within the complex in bacteria (Li et al., 2014). Recently, a prominent study reported that the synthesis of the subunits in multimeric complexes is proportional to their stoichiometry in yeast and higher eukaryotes (Taggart and Li, 2018). This finding indicates that the protein abundance of components of multimeric complexes in prokaryotes and eukaryotes is adjusted already at the protein synthesis level.

In plants, energy transduction is performed in the chloroplast by the photosynthetic protein complexes embedded in the thylakoid membrane. The assembly of such oligomeric complexes requires a temporal and spatial organization within the cell to produce and deliver the various subunits of a given complex in the stoichiometry required for its functional assembly. The dual genetic origin of photosynthetic complexes in chloroplasts adds a level of complexity compared to prokaryotes. Given the complexity of these processes, it is likely that the regulation required for the biogenesis and repair of photosynthetic complexes is intricate.

It was shown that the stoichiometry of photosynthetic complexes is fine-tuned at the synthesis level in *Chlamydomonas*, tobacco, *Arabidopsis*, and maize (Chotewutmontri and Barkan, 2016; Trösch et al., 2018). The question remains whether in case of perturbations or during complex biogenesis, the

stoichiometry of subunits is achieved by readjusting the synthesis levels of different subunits in a given complex or by the degradation of unassembled subunits.

1.4.3 Translational feedback regulation: a common theme in all domains of life

Translational autoregulation or feedback regulation permits the fine-tuning of protein synthesis levels in response to changing conditions. Many studies documented that negative feedback regulation regulates the synthesis of subunits that assemble into protein complexes in bacteria (Freedman et al., 1987; Mattheakis and Nomura, 1988; Shen-Orr et al., 2002). Some examples of proteins whose translation is feedback regulated include translation initiation factor IF3 (Butler et al., 1986), the β -subunit of RNA polymerase (Peacock et al., 1983), SecA protein in *E.coli* (Schmidt and Oliver, 1989) and a handful of ribosomal proteins such as ribosomal protein S15 (Portier et al., 1990), L4 (Li et al., 1996; Yates and Nomura, 1980) and S3 (Hendrick et al., 2001). In most instances, the autoregulated proteins, if produced in excess and not assembled into their complexes, bind to the 5' UTR of their mRNAs and inhibit translation initiation causing a negative feedback regulation of translation.

Such intricate regulation is rare in yeast, even for the ribosomal proteins. One of the very few cases of identified translation feedback regulation in yeast is that of Dbp2p RNA helicase (Barta and Iggo, 1995). Springer et al. (2010) have shown that, in general, the protein abundance quantitatively reflects the gene copy number in yeast. Another study reported that almost 10 % of the yeast genome consists of dosage-compensated genes, i.e., genes whose protein level doesn't correlate with the gene copy number (Ishikawa et al., 2017). Interestingly, the dosage-compensated genes mostly encoded for subunits of multiprotein complexes, which led to the speculation that the production of protein complexes could be regulated by translational feedback regulation. However, ribosome profiling analysis showed no change of translation efficiency indicating that translation is not the mechanism underlying dosage compensation (Ishikawa et al., 2017). A prominent study further supported that eukaryotes lack a feedback regulation (Taggart and Li, 2018). In this ribosome profiling study it was shown that, after modifying the gene copy number, most of the proteins were still produced in proportion to their copy number. Strikingly, the synthesis rate of ribosomal proteins did not compensate for gene dosage as was reported previously (Dephoure et al., 2014). Although this study excludes a general feedback regulation in eukaryotes, a possible indirect effect caused by the perturbation that was used in this study cannot be excluded.

1.4.4 CES, a major translational feedback regulation mechanism in chloroplast and mitochondria

In photosynthetic organisms, the synthesis of some subunits of the photosynthetic complexes was reported to be feedback-regulated by the assembly of the complex. This negative assembly-dependent feedback regulation of translation was termed control by epistasy of synthesis (CES) (Choquet and Wollman, 2009). Insights into CES regulation of the assembly of photosynthetic complexes were first

described in *Chlamydomonas* (Figure 1.3), where it was shown that the absence of a certain core subunit in specific photosynthetic mutants results in a decreased synthesis rate of another chloroplast-encoded subunit from the same complex. The latter subunit whose synthesis rate is controlled by the availability of its assembly partner is designated as a CES subunit. CES regulation was identified for all photosynthetic complexes in *Chlamydomonas*, and sometimes more than one feedback loop was found per complex. In PSI, the presence of PsaB was shown to be required for efficient translation of the *psaA* transcript (Stampacchia et al., 1997). Likewise, *psaA* and *psaB* mutants showed reduced synthesis of PsaC (Takahashi et al., 1991). These feedback loops define a ‘CES cascade’, where PsaB is required for the translation of *psaA*, which in turn is required for PsaC synthesis (Wostrikoff et al., 2004) (Figure 1.3). Another CES cascade was found in PSII, where *psbD* mutants displayed a reduction in the translation of *psbA* and *psbB* (Minai et al., 2006; Trösch et al., 2018) and *psbA* mutants showed a decrease in the translation of *psbB* but not of *psbD* (Minai et al., 2006). These findings in *Chlamydomonas* led to the conclusion that PsaB initiates the CES cascade within PSI by affecting the translation of *psaA*, which subsequently alters the translation of *psbB* (Figure 1.3). These CES cascades define the hierarchical synthesis of chloroplast-encoded subunits and depict a major role of CES in the sequential assembly of photosynthetic complexes in *Chlamydomonas*. The sole CES case where the molecular regulation mechanism is identified is that of PetA. PetA is a CES subunit in the Cyt *b₆f* complex whose translation is reduced in the absence of PetB (cytochrome *b₆*) or PetD (subunit IV) (Kuras and Wollman, 1994) (Figure 1.3). MCA1 and TCA1 are protein factors known to bind the 5’ UTR of the *petA* transcript, thereby stabilizing it (MCA1), and promoting its translation initiation (TCA1) (Loiselay et al., 2008; Wostrikoff et al., 2001). Boulouis et al. (2011) showed that the C-terminus of the unassembled PetA binds to MCA1 and triggers its proteolytic degradation. The degradation of MCA1 then causes a decrease in both the transcript accumulation and the translation of *petA* mRNA. As long as the PetA assembly partners (PetB and PetD) are available, PetA’s C-terminus is occluded by the assembly, MCA1 is stable, and the *petA* mRNA is stable and expressed.

Theoretically, the reduced synthesis of a CES subunit, when its assembly is compromised, could be explained by two different mechanisms: 1) the unassembled CES subunit could exert negative feedback on its translation, 2) the assembly partner mediates, directly or indirectly, the translation of the CES subunit. In most of the cases studied in *Chlamydomonas*, the synthesis of CES subunits is auto-regulated by negative feedback loops (Choquet and Wollman, 2009). The only described exception is the α subunit (AtpA) of the ATP synthase complex, which is transactivated by its assembly partner, the β subunit (AtpB) (Drapier et al., 2007) (Figure 1.3).

Altogether, CES is a common feature in the expression of photosynthetic genes in *Chlamydomonas* chloroplast. However, its occurrence in embryophytes is unclear. The lack of plant mutants defective for the expression of a single chloroplast-encoded subunit as well as technical limitations for the analysis of protein synthesis by pulse labeling in multicellular organisms have hampered the assessment of CES

in embryophytes. In tobacco, only the CES regulation of the plastid-encoded large subunit of Rubisco by the abundance of the nucleus-encoded small subunit of Rubisco has been described (Quick et al., 1992; Rodermel et al., 1996; Rodermel et al., 1988). In agreement with the RNA-binding capacity of RbcL (Yosef et al., 2004), Wostrikoff and Stern (2007) provided evidence that the decrease of *rbcL* translation is due to direct binding of the unassembled RbcL to its mRNA. While the CES regulation of Rubisco in embryophytes resembles that in *Chlamydomonas* (Khrebtukova and Spreitzer, 1996), differences were found in the CES regulation of the Cyt *b₆f* complex. Knockout mutants of *petD* and *petB* exhibited only a mild reduction in the translation of the polycistronic *petA* transcript based on polysome analysis in tobacco (Monde et al., 2000). Given the polycistronic nature of *petA* transcription unit, it is hard to have a firm conclusion regarding CES regulation of *petA*. However, this finding hints at differences between embryophytes and *Chlamydomonas*. Moreover, in contrast to *Chlamydomonas*, ribosome profiling analysis of AtpB mutants in plants did not show any defect in the translation of AtpA (Trösch et al., 2018; Zoschke et al., 2013). Interestingly, a potential CES regulation within PSII between PsbB and PsbH was observed where the translation of *psbB* was affected by the availability of PsbH (Felder et al., 2001; Levey et al., 2014). Hints for this potential CES regulation in PSII was recently reported in *Chlamydomonas* (Trösch et al., 2018). In summary, a systematic investigation to examine how common is the CES regulation in the chloroplasts of higher plants remained to be performed.

CES regulation is also involved in the biogenesis of the mitochondrial respiratory complexes in yeast. Several studies analyzing the expression of the core subunit of the cytochrome oxidase complex (COX), Cox1p, encoded in the yeast mitochondria, have proven that this protein is a CES subunit (Barrientos et al., 2004; Cabral and Schatz, 1978; Calder and McEwen, 1991; Poutre and Fox, 1987). The molecular mechanism underlying this CES regulation was unraveled (Barrientos et al., 2004; Perez-Martinez et al., 2003): upon compromised assembly of COX, the unassembled Cox1p binds to Mss51p (Perez-Martinez et al., 2003) and blocks its function as translation activator of the *coxI* mRNA (Perez-Martinez et al., 2003; Zambrano et al., 2007). Another instance of CES regulation in yeast mitochondria is that of Atp6p and Atp8p whose synthesis is dependent on the availability of Atp9p (Jean-Francois et al., 1986; Ooi et al., 1987).

Strikingly, despite the wide contribution of the CES process in the biogenesis of protein complexes in chloroplasts and mitochondria, no CES regulation was reported in cyanobacteria (Choquet and Wollman, 2009). This finding poses the question of whether the assembly-dependent regulation of translation was established after endosymbiosis.

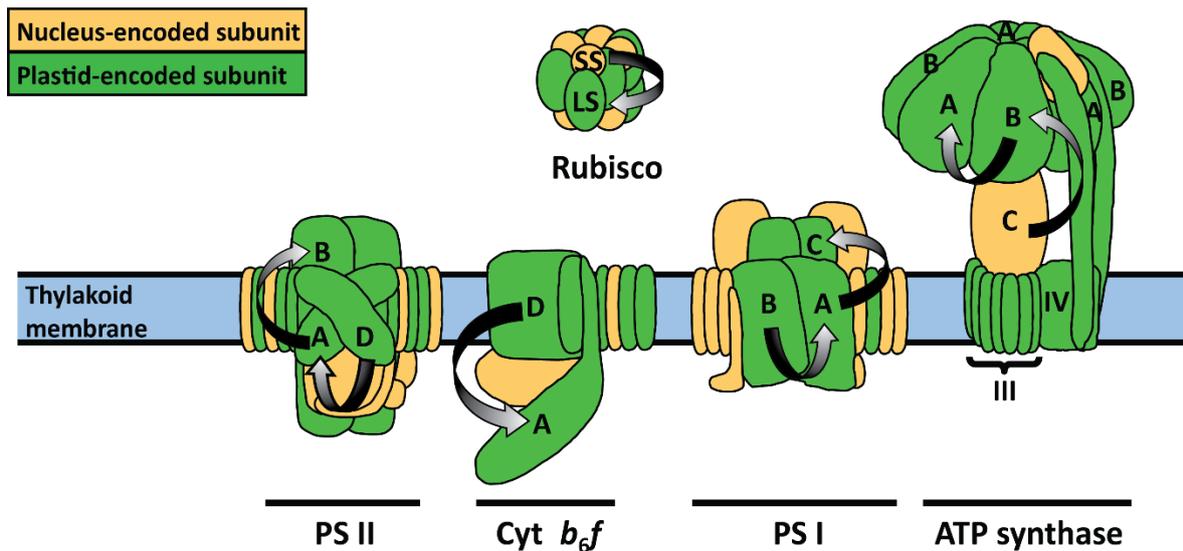


Figure 1.3: CES contributes to the biogenesis of the photosynthetic apparatus in *Chlamydomonas*. Schematic representation of photosynthetic protein complexes in the thylakoid membrane. Arrows represent the epistatic relationships between the identified CES subunits and their assembly partners. The arrowheads emphasize the CES subunits. For details see section 1.4.4. The figure is adapted from Choquet and Wollman (2009).

1.4.5 Protease-dependent regulation

Protein accumulation depends not only on the translation rate but also on post-translational proteolysis (Adam, 2000; Wollman et al., 1999). Proteolysis is triggered by protein misfolding, misassembly (Adam, 1996), excess production regarding the complex stoichiometry, and mistargeting (Halperin and Adam, 1996). In chloroplasts, more than 20 protease machinery have been described, most of which are of bacterial origin (Nishimura et al., 2016; van Wijk, 2015). Protein degradation in the chloroplast is mainly carried out by the Clp protease complex (ATP-dependent caseinolytic protease) (Nishimura and van Wijk, 2015; Shanklin et al., 1995), FtsH complex (ATP-dependent zinc metalloprotease) (Kato and Sakamoto, 2018; Lindahl et al., 1996), Lon (ATP-dependent protease) (Ostersetzer et al., 2007), and Deg (ATP-independent protease) (Itzhaki et al., 1998). Clp proteases as well as Lon proteases are located in the stroma and are involved in the degradation of stromal proteins. Thylakoid proteins on the other hand are degraded by FtsH, a membrane-associated metalloprotease, and the endopeptidase Deg (Adam, 2000; Adam et al., 2006; Nishimura et al., 2016; Sakamoto, 2006).

1.5 Ribosome profiling: genome-wide analysis of translation

Technical challenges have long hampered the study of translation. Pulse-labeling and polysome profiling were the methods of choice to study translation, but they are both labor-intensive, provide a limited resolution, and are not suited for genome-wide analyses. The development of ribosome profiling has revolutionized the study of different aspects of translation. The concept of the technique can be traced back over 50 years when polysomes were first described and it was found that endonuclease

treatment converts these polysomes into monomeric ribosomes (monosomes) that protect small fragments of mRNAs (Steitz, 1969). These ribosome-protected fragments, referred to as ribosome footprints, reflect the positions of translating ribosomes on mRNAs. Many later studies have extended the finding of ribosome footprints to uncover ribosome pausing as well as other aspects of protein synthesis (Wolin and Walter, 1988). The development of next-generation sequencing technologies allowed the sequencing-based ribosome profiling to globally map the ribosome positions at a transcriptome-wide scale (Ingolia et al., 2009). This approach does not only captures translation *in vivo* in real-time but also enables a quantitative and codon-resolved analysis of many aspects of translation. Applications of ribosome profiling include the definition of whole translomes (start codons, splice junctions, upstream ORFs, etc.), assessment of regulatory translational dynamics, and the study of ribosome behavior (e.g., ribosome pausing) (Ingolia, 2014). On average, the length of ribosome footprints is ~ 30 nucleotides, however, it differs between species. For example, in *Saccharomyces cerevisiae* (yeast), the median of the length distribution of nuclear footprints is 28 to 29 nt whereas that in mammalian cells is 30 to 31 nt. Ribosome footprints of approximately 16 nt were also obtained arising from stalled ribosomes at the 3' end of truncated mRNAs. More intriguingly, 21 nt footprints were obtained in a study by Lareau et al. (2014) and were assigned to a rotated conformation of the elongating ribosome. A recent study revealed that the 21 nt ribosome footprints correspond to ribosomes with open ribosomal A sites and are more enriched under stress conditions (Wu et al., 2019).

Theoretically, one footprint corresponds to one translating ribosome, which, in most of cases, will result in the production of one protein. Taking this into account, the abundance of ribosome footprints reflects the amount of protein synthesized. It is important to note that changes in the abundance of ribosome footprints reflect alteration either in transcript abundance or translational activity (or both). To disentangle these effects, the determination of transcript levels is usually performed in parallel to the profiling of ribosome footprints. Eventually, comparison of the transcript and ribosome footprint abundance enables the calculation of translation efficiencies for each gene (Ingolia, 2014). Accordingly, the approach measures the two major determinants of gene expression: the transcript level and its translational activity. One of the most prominent characteristics of ribosome footprints is the trinucleotide periodicity that results from the mechanism of translation elongation following the triplet genetic code. The trinucleotide periodicity has been used as quality control of ribosome profiling data in order to distinguish the mRNA fragments, which stem from actively translated transcripts rather than from transcripts that are protected for other reasons (e.g., by RNA-binding proteins). Furthermore, Hsu et al. (2016) used this criteria to determine novel translated ORFs in *Arabidopsis*.

The technique has been applied for *Chlamydomonas* (Cavaiuolo et al., 2017; Chung et al., 2015) and several plant species including *Arabidopsis* (Chotewutmontri et al., 2020; Hsu et al., 2016; Liu et al., 2013), maize (Chotewutmontri and Barkan, 2018; Rojas et al., 2018; Zoschke et al., 2013), tobacco (Kwon et al., 2016), lettuce (Kwon et al., 2016), and soybean (Shamimuzzaman and Vodkin, 2018).

Moreover, as a rapid alternative to the sequencing-based ribosome profiling, a microarray-based method was used to study chloroplast translation. In the latter approach, ribosome footprints are differentially labeled and hybridized to a high-resolution tiling microarray (Zoschke et al., 2013). This method is suitable for small genomes and has a resolution of ~ 30 nt (Scharff et al., 2017; Schuster et al., 2020; Trösch et al., 2018; Zoschke and Barkan, 2015; Zoschke et al., 2017; Zoschke et al., 2013; Zoschke et al., 2016). In these studies, different aspects of translation regulation were described in different cell types, at different developmental stages, and under different stress conditions. Furthermore, ribosome profiling not only effectively enabled the validation of known targets of PPR proteins but also facilitated the identification of novel target mRNAs, which were not found by classical methods (Chotewutmontri et al., 2020; Rojas et al., 2018; Williams-Carrier et al., 2019; Zoschke et al., 2013; Zoschke et al., 2016).

One of the most critical steps in ribosome profiling is the nuclease treatment to generate the ribosome footprints. RNase I is the most common nuclease in eukaryotic studies. It can produce precise 5' and 3' ends of ribosome footprints and thus both ends can be used for mapping the sequencing reads, which increases the visibility of the characteristic trinucleotide periodicities. The micrococcal nuclease (MNase) from *Staphylococcus aureus* is another widely used nuclease, most commonly in bacterial ribosome profiling (Mohammad et al., 2016; Oh et al., 2012). The activity of MNase is selective in that it preferentially cleaves RNA at A or U nucleotides (Dingwall et al., 1981), which results in a wider range of footprint size distribution. In the context of this work, RNase I was used in the sequencing-based ribosome profiling whereas MNase was used in the microarray-based approach.

1.6 Nucleus-encoded translation factors: a brief summary

In plant cells, gene expression in the nucleo-cytosolic and the chloroplast compartments is coordinated by a large number of nucleus-encoded proteins (Figure 1.2) and intercommunicate by retrograde and anterograde signals. All the RNA metabolism processes and translational regulation in chloroplasts rely on nucleus-encoded RNA-binding proteins (RBPs). Most of the RBPs interact with their targets in a sequence-specific manner. Several RBP families exist in the chloroplast, naming the chloroplast ribonucleoproteins (cpRNPs) family (Tillich et al., 2010), the half a tetratricopeptide protein family (HAT) family, and the mitochondrial transcription termination factors (mTERF) family (reviewed in (Hammani et al., 2014)). The largest RBP family in plants is the PPR protein family. All PPR proteins are localized in mitochondria or chloroplast (Lurin et al., 2004) and are involved in organellar gene expression. In contrast to most eukaryotes with ~ 5 to 30 PPRs, embryophytes possess an expanded PPR family with more than 400 members (O'Toole et al., 2008). PPR proteins belong to the alpha-solenoid superfamily and possess a PPR motif that consists of organized repeats of ~ 35 amino acids (Small and Peeters, 2000). In plants, the PPR family is classified into two subfamilies, P-type and PLS-type PPR proteins (Barkan and Small, 2014; Lurin et al., 2004). P-type PPR proteins harbor the canonical PPR motif (35 amino acids) and are involved in several steps of RNA metabolism, including end maturation, RNA stability, RNA splicing, and translation (Schmitz-Linneweber and Small, 2008; Shikanai and Fujii,

2013). The PLS-type PPR (PLS PPR) proteins possess E or DYW domains at their C-terminus, which are involved in RNA editing in plant organelles (Chateigner-Boutin et al., 2013; Okuda et al., 2009). Systematic comparison of the protein sequence of all the identified PPRs and the RNA sequence of their targets enabled the establishment of a so-called “PPR code” (Barkan et al., 2012). This code can be used to predict or to change the binding specificity of a PPR protein (Rojas et al., 2019).

Examples of confirmed regulators of chloroplast translation are rare, the majority of which possess a PPR motif (Zoschke and Bock, 2018). One of the best-studied examples is PPR10, which binds upstream of the *atpH* start codon and activates translation by preventing the formation of an RNA structure that masks the ribosome binding site (Prikryl et al., 2011). PPR10 stimulates not only *atpH* translation but is also involved in the *atpH* transcript stabilization. Some other PPRs were shown or suggested to be involved in translational regulation in a similar way (Zoschke and Bock (2018). Recently, a PPR protein in *Arabidopsis*, LPE1, was suggested to affect *PsbA* synthesis (Jin et al., 2018). However, this finding was later shown to be a secondary effect of the translational regulation of *psbJ* (Williams-Carrier et al., 2019). In addition, few of the RBPs promoting translation are not PPRs. HCF173 and HCF244 possess an atypical short-chain dehydrogenase/reductase (SDR) domain and are involved in light-dependent translational regulation of *psbA* (Chotewutmontri et al., 2020; Link et al., 2012; Schult et al., 2007; Williams-Carrier et al., 2019). HCF107 is a HAT protein involved in the translational activation of *psbH* (Felder et al., 2001; Hammani et al., 2012; Williams-Carrier et al., 2019).

1.7 Purification of chloroplast ribonucleoproteins (RNP) complexes

Given the small number of transcripts retained in the chloroplast versus the large number of RNA-binding proteins imported from the nucleus, it is assumed that chloroplast translation itself adapted to a transcript-specific regulation. In this regard, each chloroplast transcript is believed to be bound by a set of factors few of which were shown to be translation activators (Barkan, 2011; Barkan and Small, 2014). In order to selectively identify the factors involved in translational regulation in general and the translation feedback regulation of some of the candidates that emanated in this study as potential CES subunits, a selective purification of specific chloroplast transcripts and their bound proteome is needed.

1.7.1 Classical methods for RNP purification

Small, genetically introduced protein affinity tags (epitopes) have been used for decades to produce recombinant proteins in order to isolate defined protein complexes. The widespread use of protein affinity tags led to the development of similar tags for nucleic acids. Different methods have been developed to tag RNAs either for localization purposes or for affinity purification. Over the past decade, RNA immunoprecipitation (RIP) -approaches have been developed to identify RNAs that bind to a specific protein (Gagliardi and Matarazzo, 2016; Schmitz-Linneweber et al., 2005). A modified version of RIP, CLIP, including a UV cross-linking step has also been used (Sugimoto et al., 2012). In both techniques, the RNA-binding protein is purified and the bound RNAs are identified by microarray

hybridization or deep-sequencing. In plants, these techniques have been employed to study several protein-RNA interactions (Barkan, 2009). Recently, RIP-seq has been used to immunoprecipitate an artificial PPR protein customized to target an endogenous RNA *in vivo* and its RNP complex (McDermott et al., 2019). One of the major limitations of RIP is that a known RNA-binding protein and an antibody directed against the endogenous or the epitope-tagged proteins are needed. Although engineered RBP might be, in theory, a gateway to target any RNA, an off-targeting effect cannot be completely ruled out. Another approach to study protein-RNA interactions has been developed in yeast (Lapointe et al., 2015). In this technique, the RNA-binding protein is fused to a *Caenorhabditis elegans* poly (U) polymerase. Upon binding of the RBP to the RNA, the polymerase tags the RNA with 3' terminal uridines, which enables the identification of the tagged RNA by RNA sequencing. Although in this approach an antibody is not needed for immunoprecipitation, a known RBP is still required to target the polymerase to the RNA. Instead of tagging and targeting the protein counterpart to identify protein-RNA interactions, a tagging approach in which the RNA is the anchor for the purification has advantages. Biotinylated antisense oligonucleotides have been used to purify small nuclear ribonucleoproteins (snRNPs) in human cells (Blencowe et al., 1989). This technique has also been successfully applied in chloroplasts to co-purify the *psbA* mRNA with its RNA-binding proteome (Watkins et al., 2019). A major drawback of this approach is that the antisense oligonucleotides might cover an RBP binding site, which might affect either the trafficking or the loading of the RNA into an RNP complex.

1.7.2 Aptamer-based affinity purification

Aptamer tagging of RNAs is another method that has been used to affinity purify RNP complexes (Said et al., 2009; Walker et al., 2008). The term “aptamer” has been coined by Andrew D. Ellington and derives from Latin “aptus” and “meros” meaning to fit and part, respectively (Ellington and Szostak, 1990). Aptamers are short, oligonucleotide or peptide molecules that bind to a specific target, including proteins, peptides, carbohydrates, small molecules, toxins, and even living cells. Nucleic acid aptamers (referred to as aptamers hereafter) are single-stranded DNA or RNA (ssDNA or ssRNA) molecules that bind with high specificity to their target (Ellington and Szostak, 1992). Aptamers tend to form helices and single-stranded loops and bind to their targets via hydrophobic and van der Waals interactions. Aptamers with affinity for a desired target can be synthetically produced or occur naturally. Synthetic aptamers were selected through a process called “Systematic Evolution of Ligands by Exponential Enrichment” (SELEX) invented in 1990 (Ellington and Szostak, 1990; Tuerk and Gold, 1990). The desired aptamer is selected from a large random sequence pool by cycles of selection and amplification. In these cycles, the target molecule is incubated with a library of sequences. Only the aptamers with the highest affinity bind to the target molecule and are therefore selected and amplified. Using SELEX, D8 and S1 aptamers that bind to Sephadex and streptavidin, respectively, were identified and shown to be eligible for use as RNA affinity tags (Srisawat and Engelke, 2001; Srisawat et al., 2001). The D8

Sephadex-binding RNA aptamer (referred to as the Sephadex-binding aptamer hereafter) possesses a motif of 33 nt that binds specifically to Sephadex resins (Srisawat et al., 2001), made by crosslinking dextran B12 with epichlorohydrin. The S1 streptavidin-binding RNA aptamer (referred to as the streptavidin-binding aptamer hereafter) has a motif of 44 nt and was selected to bind streptavidin (Srisawat and Engelke, 2001), a protein from the bacterium *Streptomyces avidinii*. Both of these aptamers are characterized by a low background affinity, have compact structures, and have been previously used to purify the ribonuclease P enzyme (RNase P) from *Saccharomyces cerevisiae* under native conditions (Srisawat and Engelke, 2001; Srisawat et al., 2001). Additionally, the streptavidin-binding aptamer has been employed to isolate RNP complexes in *E.coli* (Leonov et al., 2003) and human cells (Li and Altman, 2002).

The term “aptamers” also encompasses naturally evolving RNA elements that bind to their targets with high specificity. Some examples include the 25 nt hairpin binding to the PP7 bacteriophage coat protein (Larson et al., 2011), a 29 nt hairpin that binds to the BglG protein (Chen et al., 2009; Gulati and Mahadevan, 2001), and a 21 nt RNA fragment that binds to the splicing protein U1Ap (Chung and Takizawa, 2010). The most frequently applied natural RNA-protein interaction is the phage MS2 RNA element (referred to as MS2 aptamer hereafter) that binds with high specificity to the MS2 bacteriophage coat protein (Bardwell and Wickens, 1990; Peabody, 1993). The MS2 aptamer consists of a short sequence that folds into a hairpin with a stem of seven base pairs containing a protruded adenine (Zhou et al., 2002). The MS2 sequence is usually added in multiple tandem copies to the RNA to be tagged. MS2-based affinity purification has been widely used to effectively purify a variety of ribonucleoprotein particles (RNPs) of small non-coding RNAs (sRNAs) (Said et al., 2009) or to purify the human spliceosome (Jurica et al., 2002; Zhou et al., 2002). More recently, the MS2 aptamer was used to purify the postcatalytic P complex (Fica et al., 2019).

Despite the wide use of aptamers to purify RNPs in yeast, *E.coli*, and human cells, neither synthetic nor natural aptamers were reported to be used in plants. In this work, an aptamer-based affinity purification protocol was adapted to isolate chloroplast RNPs. Three different RNA affinity tags (MS2, Sephadex-, and streptavidin-binding aptamers) were selected to tag the 3' UTR of chloroplast transcripts. Sephadex- and streptavidin-binding aptamers were previously inserted into the 3' UTR of *psbA* and *rbcL* in tobacco chloroplast (Reimo Zoschke (MPIMP), unpublished). In this work, the MS2 aptamer was additionally used to tag *psbA* and *rbcL*, and affinity purification was optimized for all three aptamers. *psbA* and *rbcL* transcripts were chosen as targets because *rbcL* is a known CES subunit and *psbA* was found in the present work to be feedback regulated by factors, which were unknown when this study was initiated. In addition, these two transcripts possess a stable stem-loop structure in their 3' UTR. The stem loop does not only stabilize the transcript (Stern and Gruissem, 1987) but also protects the aptamer sequence by preventing its removal by nucleases or RNA processing events (Srisawat and Engelke, 2002). Additionally, these two transcripts are both monocistronic and very abundant in the

chloroplast with *psbA* being the most abundant transcript with an estimated 14,000 molecules per chloroplast (Nakamura et al., 2001), which facilitates the enrichment, detection, and verification of translation factors.

1.8 Aptamer-based RNA tracking

Targeted mRNA localization and localized translation were proposed to be important for site-specific protein deposition and assembly into complexes (Slobodin and Gerst, 2010). It is also thought to affect different cellular processes such as cell division (Du et al., 2007), motility, and responses to external triggers (Du et al., 2007; Elson et al., 2009; Yoon et al., 2009). Okita and Choi (2002) have demonstrated that also plants can localize RNA transcripts in order to target protein production to a specific compartment and therefore control cell fate and growth. Also in *Chlamydomonas*, several studies have demonstrated mRNA-based mechanisms in the targeting of specific proteins in the chloroplast (e.g., the D1 subunit) (Uniacke and Zerges, 2009; Weis et al., 2013). Most recently, Ouyang et al. (2020) showed that liquid-liquid phase separation could account for cargo sorting in the chloroplast. Many methods have been employed to track individual RNAs and to study their intracellular localization. Among these techniques are fluorescent *in situ* hybridization (FISH) and chimeric RNAs tagged with traceable elements like RNA aptamers. The MS2 aptamer has been widely used to tag the 3' UTR of endogenous RNAs. In one study, the MS2 coat protein was fused to the fluorescent protein GFP and the hybrid protein was used to visualize the RNA tagged with the MS2 aptamer (Wang et al., 2012). This MS2 system has been applied, for example, to visualize the localization of RNA decay intermediates in cytoplasmic processing bodies in yeast (Sheth and Parker, 2003), to confirm that Argonaute and miRNAs suppress the mRNA translation in processing bodies in mammalian cells (Liu et al., 2005), and to study the mRNA transport in oocytes of *Drosophila* (Forrest and Gavis, 2003). More recently, Morisaki et al. (2016) used the MS2 system in combination with tagging of the nascent polypeptide with a FLAG tag to track *in vivo* the translation of a single transcript. Despite the frequent usage of the MS2 aptamer in different organisms to track RNA, no application in plant cells nor chloroplast has been reported.

1.9 Aim and strategies

1.9.1 Search for feedback regulation in the expression of the subunits of photosynthetic complexes in land plants

In chloroplasts, about half of the subunits of photosynthetic complexes are chloroplast-encoded most of which are core subunits essential for the assembly of the respective complexes. In *Chlamydomonas*, the stoichiometric accumulation of the photosynthetic complexes is coordinated by proteolysis of unassembled subunits (Adam, 2007) and by assembly-dependent translation feedback regulation known as CES (Choquet and Wollman, 2009). Many studies have shown that CES is a major feature of chloroplast gene expression in *Chlamydomonas*. Conversely, in embryophytes, the only identified CES

regulation is that of the large subunit of Rubisco (Wostrikoff and Stern, 2007). The aim of this work was to identify whether the CES process is also the main regulator of protein synthesis of photosynthetic subunits in embryophytes or if the stoichiometric accumulation is always realized by proteolytic degradation of the unassembled subunits. To comprehensively answer this question, a combination of microarray-based and sequencing-based ribosome profiling was used to monitor regulation at translation or transcript level in mutants with assembly defects in each of the photosynthetic complexes. Potential CES loops were identified in PSII and the Cyt *b₆f* complex. Subsequent pulse labeling experiments confirmed the observed CES regulation. For the photosynthetic complexes where no CES regulation was identified, mutants of core subunits of these complexes were crossed with inducible knockdown mutants of the chloroplast proteases Clp and FtsH (Moreno et al., 2018). These double mutants shall enable the analysis of the contribution of proteolytic degradation of unassembled photosynthetic subunits during complex assembly. Altogether, this work aimed to achieve a deeper understanding of the regulatory mechanisms that coordinate the expression and assembly of the photosynthesis machinery in embryophytes.

1.9.2 Identification of chloroplast translation factors using an aptamer-based affinity purification approach

To unravel the molecular mechanisms of the identified translation feedback regulation, potentially involved transcript-specific translation factors need to be identified. Aptamer tagging has been used for the purification of RNPs (Walker et al., 2008). Transplastomic plants in which *psbA* and *rbcL* mRNAs were tagged with Sephadex- and streptavidin- binding aptamers were previously created. Furthermore, transplastomic tobacco plants with MS2-tagged *psbA* and *rbcL* mRNAs were created by chloroplast transformation in this study. This work was expected to adapt and optimize the purification of *psbA* and *rbcL* RNPs using these aptamers. Both transcripts were efficiently and specifically purified with the MS2 aptamer and only *psbA* mRNA was enriched using the streptavidin-binding aptamer. Follow up work encompasses identification of the co-purified proteins, including translation factors, by mass spectrometry. Functional analysis of specific translation factors is a long-term aim. Additionally, the aim of this work was to examine the suborganellar localization of *psbA* and *rbcL* mRNAs *in vivo*, which shall shed light on the spatial localization of translation. This pioneering approach is expected to be used as a model to analyze the RNA-binding proteomes and to localize any transcript in the chloroplast in the future.

2 Material and methods

2.1 Material

2.1.1 Suppliers of chemicals and consumables

Most of the general and specialized chemicals were of molecular biology or analytical grade and were obtained from the following suppliers: Bio-Rad Laboratories GmbH (Hercules, CA, USA), Carl Roth GmbH (Karlsruhe, Germany), Eppendorf (Hamburg, Germany), Fisher Scientific (Schwerte, Germany), Merck KGaA (Darmstadt, Germany), Sarstedt AG & Co. KG (Nümbrecht, Germany), New England Biolabs GmbH (Frankfurt am Main, Germany), Serva (Heidelberg, Germany), Promega GmbH (Mannheim, Germany), Sigma-Aldrich Corporation (St. Louis, MO, USA), Thermo Fisher Scientific Inc. (Waltham, MA, USA), Unigloves GmbH (Troisdorf, Germany), VWR International (Darmstadt, Germany).

2.1.2 Specialized chemicals

Chemical	Order number	Supplier
2-Mercaptoethanol	63689-100ML-F	Sigma-Aldrich Corporation (St. Louis, MO, USA)
[α - ³² P]-dCTP	SRP-205	Hartmann Analytic (Braunschweig, Germany)
Ammonium Persulfate	V3131	Promega Corporation (Fitchburg, WI, USA)
ATP 100 mM Solution WI/DI	GE27-2056-01	Sigma-Aldrich Corporation (St. Louis, MO, USA)
Bis-acrylamide 19:1, 40 % Solution	1300-500ML	Merck Chemicals GmbH (Darmstadt, Germany)
Certified™ Molecular Biology Agarose	1613101EDU	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
Chloramphenicol	C0378-5G	Sigma-Aldrich Corporation (St. Louis, MO, USA)
Chloroform–isoamyl alcohol mixture	25666-100ML	Sigma-Aldrich Corporation (St. Louis, MO, USA)
cOmplete™ Protease Inhibitor Cocktail	5056489001	Sigma-Aldrich Corporation (St. Louis, MO, USA)
Coomassie Brilliant Blue G250	140739	Serva (Heidelberg, Germany)
Coomassie Brilliant Blue R250	150671	Serva (Heidelberg, Germany)
Cycloheximide 100 mg/mL	C4859-1ML	Sigma-Aldrich Corporation (St. Louis, MO, USA)
EasyTag™ EXPRESS 35S Protein Labeling Mix, [35S]	NEG772002MC	PerkinElmer (Waltham, Massachusetts, USA)
Ethidium bromide 1 % (10 mg/mL)	1239-45-8	Carl Roth GmbH (Karlsruhe, Germany)
Formaldehyde solution	F8775-500ML	Sigma-Aldrich Corporation (St. Louis, MO, USA)
Formaldehyde solution	1.04003.1000	Merck Chemicals GmbH (Darmstadt, Germany)

Deionized formamide	P040.1	Carl Roth GmbH (Karlsruhe, Germany)
GlycoBlue™ Coprecipitant	AM9516	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Hakaphos®	-	Compo (Münster, Germany)
IPTG	088M4049V	Sigma-Aldrich Corporation (St. Louis, MO, USA)
Methylene blue	61-73-4	Serva (Heidelberg, Germany)
Plus One Bromophenol Blue	L784172345	GE Healthcare (Little Chalfont, UK)
Polyoxyethylene (10) tridecyl ether	P2393-500G	Sigma-Aldrich Corporation (St. Louis, MO, USA)
Ponceau S	5938.2	Carl Roth GmbH (Karlsruhe, Germany)
Recombinant RNasin® RNase Inhibitor	N2511	Promega Corporation (Fitchburg, WI, USA)
Restore™ Western Blot Stripping Buffer	46430	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Roti®-Phenol	38.1	Carl Roth GmbH (Karlsruhe, Germany)
Roti®-Phenol/Chloroform/Isoamyl alcohol	A156.2	Carl Roth GmbH (Karlsruhe, Germany)
Rotiphorese® Gel A	3037.1	Carl Roth GmbH (Karlsruhe, Germany)
Rotiphorese® Gel B	3039.1	Carl Roth GmbH (Karlsruhe, Germany)
Sodium azide	71289-5G	Sigma-Aldrich Corporation (St. Louis, MO, USA)
TEMED	T7024-25ML	Sigma-Aldrich Corporation (St. Louis, MO, USA)
TRIzol Reagent-200 mL	15596018	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Tween 20	P9416-50ML	Sigma-Aldrich Corporation (St. Louis, MO, USA)
Urea molecular biology reagent	U5378-1KG	Sigma-Aldrich Corporation (St. Louis, MO, USA)
Xylencyanol FF	3132	Riedel-De Haen AG (Seelze, Germany)

2.1.3 Consumables

Consumable	Order number	Supplier
TGX™ Precast Protein Gels	4561096	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
Amersham Protran 0.2 NC Membrane	10600001	GE Healthcare (Little Chalfont, UK)
Amylose Resin	E8021S	New England Biolabs GmbH
<i>Arabidopsis</i> Microarray	-	Arbor Biosciences (Ann Arbor, MI, USA)
Corning® Sterile Vacuum Filter Unit	430758	Corning Inc. (Corning, NY, USA)
Bio-Spin Chromatography Columns	732-6008	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
Dualfilter tips 0.1-10 µL	0030 077 504	Eppendorf (Hamburg, Germany)
Dualfilter tips 2-100 µL	0030 077 547	Eppendorf (Hamburg, Germany)
Dualfilter tips 50-1000 µL	0030 077 571	Eppendorf (Hamburg, Germany)

Dynabeads MyOne™ Streptavidin T1	65601	Life Technologies GmbH (Darmstadt, Germany)
HiTrap™ Heparin HP	17-0406-01	GE Healthcare (Little Chalfont, UK)
Goat anti Rabbit IgG 15 nm gold conjugate	EM.GAR15	BBI solutions (Crumlin, UK)
Hybond™-N+ Membrane	RPN303B	GE Healthcare (Little Chalfont, UK)
Hybond™-N Membrane	RPN303N	GE Healthcare (Little Chalfont, UK)
LRWhite Resin	-	London Resin Company (Berkshire, UK)
MBPTrap™ HP	28-9187-79	GE Healthcare (Little Chalfont, UK)
Millex-GS 0.22µm 33mm sterile	SLGS033SB	Merck Millipore (Darmstadt Germany)
Pierce™ Protein A/G Magnetic Beads	88803	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Sephadex G-200	-	Pharmacia (provided by AG Dobbek, Humboldt University of Berlin)
SW 55Ti centrifuge tubes	326819	Beckman Coulter (Brea, CA, USA)
Tobacco Microarray	-	Arbor Biosciences (Ann Arbor, MI, USA)
Tube 14mL, 105x16,8mm, PP	55.538	Sarstedt AG & Co. KG (Nümbrecht, Germany)
VWR® Disposable Transfer Pipets	414004-038	VWR International (Darmstadt, Germany)

2.1.4 Reaction kits

Reaction kit	Order number	Supplier
Agilent 2100 Small RNA	5067-1549	Agilent Technologies (Santa Clara, CA, USA)
Agilent High Sensitivity DNA Kit	5067-4626	Agilent Technologies (Santa Clara, CA, USA)
Agilent RNA 6000 Nano Kit	5067-1511	Agilent Technologies (Santa Clara, CA, USA)
ECL Plus Western Blotting Detection	RPN2133	GE Healthcare (Little Chalfont, UK)
EXTRACT-N-AMP	XNAP-1KT	Sigma-Aldrich Corporation (St. Louis, MO, USA)
Megaprime DNA Labeling System	RPN1607	GE Healthcare (Little Chalfont, UK)
NEXTflex™ Small RNA-Seq Kit v3	5132-06	Bioo Scientific (Austin, TX, USA)
NucleoSpin® Gel and PCR Clean-up Kit	740609.240C	Macherey-Nagel GmbH (Düren, Germany)
NucleoSpin® Plasmid	740588.250	Macherey-Nagel GmbH (Düren, Germany)
NucleoBond® Xtra Midi Plus	740412.10	Macherey-Nagel GmbH (Düren, Germany)
Pierce™ BCA Protein Assay Kit	23227	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Qubit dsDNA HS Assay Kit	Q32851	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Qubit microRNA Assay Kit	Q32880	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Qubit RNA HS Assay Kit	Q32852	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
RiboMinus™ Plant Kit for RNA-Seq	A1083808	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
ULS aRNA Labeling Kit	EA-006	Kreatech (Amsterdam, The Netherlands)

2.1.5 Common enzymes

Enzyme	Order number	Supplier
Ambion™ RNase I, cloned	AM2295	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Antarctic phosphatase	M0289L	New England Biolabs GmbH (Frankfurt am Main, Germany)
DreamTaq polymerase	EP0701	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
EcoRV-HF	R3195S (20,000 units/mL)	New England Biolabs GmbH (Frankfurt am Main, Germany)
FastAP	EF0654	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
MfeI	R05895	New England Biolabs GmbH (Frankfurt am Main, Germany)
NUCLEASE S7	10107921001	Carl Roth GmbH (Karlsruhe, Germany)
Phusion High-Fidelity polymerase	F530L	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
T4 DNA Ligase	15224041	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
T4 DNA Ligase	M180B	Promega Corporation (Fitchburg, WI, USA)
T4 Polynucleotide Kinase	EK0031	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
TURBO™ DNase	AM2238	Thermo Fisher Scientific Inc. (Waltham, MA, USA)

2.1.6 Molecular weight markers

Marker	Order number	Supplier
DNA marker		
Gene Ruler 100 bp DNA Ladder	SM0241	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Gene Ruler 1kb DNA Ladder	SM0311	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
RNA marker		
DynaMarker® Prestain Marker for Small RNA Plus	DM253	Biodynamics Laboratory Inc. (Bunkyo-ku, Tokyo, Japan)
Century™-Plus RNA Markers	AM7145	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Millennium™ RNA Markers	AM7150	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Protein marker		
Precision Plus Protein™ Dual Xtra Prestained Protein Standards	1610377	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
Seeblue™ Plus2 Pre-stained Protein Standards	LC5925	Thermo Fisher Scientific Inc. (Waltham, MA, USA)

2.1.7 Equipment

Equipment	Supplier
Agilent 2100 Bioanalyzer	Agilent Technologies (Santa Clara, CA, USA)
Allegra™ 25R Centrifuge	Beckman Coulter (Brea, CA, USA)
Äkta explorer	Amersham Biosciences (Buckinghamshire, UK)
Biolistic® PDS-1000/He Particle Delivery system	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
Centrifuge 5427 R	Eppendorf (Hamburg, Germany)
Centrifuge MiniSpin®	Eppendorf (Hamburg, Germany)
Desiccator	neoLab (Heidelberg, Germany)
Digital Sonifier® W-250D	Brandon Ultrasonic (Danbury, USA)
DUAL-PAM	Heinz Walz GmbH (Effeltrich, Germany)
DynaMag™- Spin Magnet	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
ECX-F20.M UV transilluminator	Vilber Lourmat (Marne La Vallee, France)
Eppendorf Multipette M4 Starter Kit	Eppendorf (Hamburg, Germany)
FujiFilm BAS storage phosphor screen cassette 2040	GE Healthcare (Little Chalfont, UK)
GenePix 4000B Microarray-Scanner	Molecular Devices (San Jose, CA, USA)
Grinding mill Retsch® MM301	Retsch (Haan, Germany)
High Pressure Freezing Machine HPM 100	Leica (Wetzlar, Germany)
Image eraser 810-UNV	Amersham Biosciences (Buckinghamshire, UK)
Labor-pH-Meter Lab 850	Carl Roth GmbH (Karlsruhe, Germany)
LS 6500 Multi-Purpose Scintillation Counter	Beckman Coulter (Brea, CA, USA)
Mastercycler EPGradient PCR Thermocycler	Eppendorf (Hamburg, Germany)
Microplate reader CLARIOstar®	BMG Labtech (Ortenberg, Germany)
Mini centrifuge with slide rotor	Carl Roth GmbH (Karlsruhe, Germany)
Mini Trans-Blot® Electrophoretic Transfer Cell	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
Mini-Multi-Rotator	Kisker Biotech GmbH (Steinfurt, Germany)
Mini-PROTEAN® Vertical Electrophoresis Cell	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
MobyLux GroBank BrightBoy XL 5	CLF Plant Climatics (Wertingen Germany)
MODEL 583 GEL DRYER	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
NanoDrop™ One	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
NeoLab Horizontal Gel Electrophoresis System	NeoLab Migge (Heidelberg, Germany)
Optima™ L-80 XP Ultracentrifuge	Beckman Coulter (Brea, CA, USA)
Orbital Shaker 3017	GFL GmbH (Burgwedel, Germany)
Peqlab Horizontal Gel Electrophoresis System	Peqlab Biotechnologie GmbH (Erlangen, Germany)
Power Supply EV233	Carl Roth GmbH (Karlsruhe, Germany)
Quantum CX5 Gel Documentation System	Vilber Lourmat (Marne La Vallee, France)
Qubit 4 Fluorometer	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Quick-Count QC-2000	BioScan (Washington, USA)

Sorvall® RC6 Centrifuge	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
SpeedVac™ System	Eppendorf (Hamburg, Germany)
Storage Phosphor Screens	GE Healthcare (Little Chalfont, UK)
SW55-Ti Rotor	Beckman Coulter (Brea, CA, USA)
Syngene G:BOX Chemi XT4	SynOptics (Santa Clara, CA, USA)
T100™ Thermal Cycler	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
Thermomixer comfort	Eppendorf (Hamburg, Germany)
TEM 912 Omega	ZEISS (Oberkochen, Germany)
Typhoon™ TRIO+ Variable Mode Imager	Amersham Biosciences (Little Chalfont, UK)
Ultracut UCT Ultramicrotome	Leica Microsystems (Wetzlar, Germany)
Ultrospec® 3100 pro	Amersham Biosciences (Buckinghamshire, UK)
UV-crosslinker BLX-254	Vilber Lourmat (Marne La Vallee, France)
Vortex-Genie®	VWR International (Darmstadt, Germany)
VWR Thermal Shake lite	VWR International (Darmstadt, Germany)
Water bath FBC 620	Fisher Scientific, Schwerte, Germany

2.1.8 Antibodies

The antibodies used in this work are listed in Table 1.

Table 1: Overview of antibodies used in this work

IgG: Immunoglobulin G

Antibody against	Target size (kDa)	Working dilution	Source/ Supplier
AtpB	54	1: 5000	Agrisera, Vännäs, Sweden (AS05085)
PetB	24	1:5000	Agrisera Vännäs, Sweden (AS03034)
PsaD	20	1: 1000	Agrisera, Vännäs, Sweden (AS09461)
PsbA	38	1: 10,000	Agrisera, Vännäs, Sweden (AS10704)
PsbD	39.5	1: 10,000	Agrisera, Vännäs, Sweden (AS06146)
NdhH	45	1: 5000	Prof. Dr. Peter Westhoff, HHU, Düsseldorf, Germany
RbcL	50	1: 3750	Agrisera (Vännäs, Sweden) (AS03037)
Actin	41.6	1: 5000	Sigma-Aldrich Corporation (St. Louis, MO, USA) (A0480)
Rabbit IgG	n/a	1: 10,000	Bio-Rad Laboratories GmbH (Hercules, CA, USA) (170-6515)

Mouse IgG	n/a	1: 40,000	Sigma-Aldrich Corporation (St. Louis, MO, USA) (A9044)
MBP	n/a	-	Abcam (Cambridge, UK) (ab9084)

2.1.9 Oligonucleotides

Oligonucleotides were designed using SeqBuilder (DNASTAR Lasergene 10 Core Suite) and synthesized by Metabion International AG or Sigma-Aldrich (see Table 2).

Table 2: List of oligonucleotides used in this work

Primer	Sequence (5'→3')	Description
rbcL-Tag-span-fw	GGAGAAAGAACTCAAGTAATTATCC	Genotyping of aptamer-tagged lines
rbcL-Tag-span-rev	TCTTGTATATCTAGGTAAGTATATAC	
3' rbcL-fw	AAAGGATTGAGCCGAATACAAC	
psbA-Tag-span-fw new	CGAAGCTCCATCTACAAATGG	
psbA-Tag-span-rev	CGGGAAATCCCTATGTAAAC	
psbA-3' UTR-fw	GAGGGTGCTATTGCTCCTTTC	
aadAoutrev	ACTGCGGAGCCGTACAAATG	
f_RNAi-ClpC	CGGGATCCAGTGTGATAGAGAAAAGGAGG	Genotyping of proteases RNAi mutants
f_RNAi-FtsH	GCGGATCCCTATATGATTGCCITTTACCCC	
r_NOS	ATGATAATCATCGCAAGACCCGGCAACAGG	
f_RNAi-ClpT1	CGGGATCCAAAGCTCTTTTAATGGGAATC	
RNAiClpC-FR	CCTCCCTTCTCTATCACACT	
RNAiFtsH-FR	GGGGTAAAGGCAATCATATA	
RNAiClpT1T2-FR	CCTTCGACCAAGATTCCCAT	
f_pA1cA(1)	GGATTGGATGCATGGGAACC	
f_pA1cA(2)	AGAGACGGAGCACITTTCTGG	
r_asRNAi-ClpC	GAGTTGAAACAGTACTTCAGGCC	
r_asRNAi-FtsH	CTGCTCTCTCCAACTTCTTTGG	
r_asRNAi-ClpT1T2	CTTACTGAAACCAGCTCAAAGGG	
r_FtsHshort	CAAATCTTTGTCTCTCCTCCC	
r_ClpCshort	CCGTCAGCTCCTAAAGTTAGAG	
r_ClpT1T2short	GCACAAAAGGCTCTTGATTGG	

f_linkerprotease	CATTTGGAGAGGACGACCTG	
r_Intron	GAATCTTGCAGTCAGCTGAAAC	
f_psbA probe	CCTTGACTGTCAACTACAGATTGG	Generation of <i>psbA</i> northern probe
r_psbA probe	TACTTGGCGTAGCTTGTACATGG	
f_rbcL probe	GCATTATGTTAAAACCTTTCCAAGG	Generation of <i>rbcL</i> northern probe
r_rbcL probe	ATATGATCTCCACCAGACATACG	
f_atpH probe	TTAAACA AAAAGGATTCGCA	Generation of <i>atpH</i> northern probe
atatr_atpH probe	ATGAATCCACTGATTTCTG	
f_atpF probe	CTTTCATTTGTTCCCAACATGC	Generation of <i>atpF</i> northern probe
r_atpF probe	GAAATTCAGAAGAAGACTGCGTG	
f_psbA-MS2_southern	GAAGGAGAGGTTAATTTCTTGCATTT	Generation of <i>psbA</i> Southern probe
r_psbA-MS2_southern	GATTCACAAATCCACTGCCTTGAT	
f_rbcL-MS2_southern	GATTGGTATAATCTTTTCTATCC	Generation of <i>rbcL</i> Southern probe
r_rbcL-MS2_southern	ATCTTTCCATAGAGTTTCTTGC	
MS2-1	AAAAATCAATTGGATATCCGTACACCATCAGGGTACGAGCTAGCCCATGGCGTACA CCATCAGGGTACGA	Synthesis of MS2 aptamer sequence
MS2-2	AAAAATCAATTGGAAATTCGGTACCCTGATGGTGTACGAGATCTACTAGTCGTACCC TGATGGGTACCG	
psbA-fw-PstI	TAATATCTGCAGTCCCTATTCAGTGCTATGCATGG	Amplification of genomic DNA fragments for cloning of the MS2 aptamer
trnH-rev-PstI	TAATATCTGCAGATGGTGGATTCAACAATCCACTGC	
rbcLfw-PstI	TAATATCTGCAGATTCACCTGGTACCCGTAGTAGG	
accD-PstI rev	TAATATCTGCAGCCATCTTTCCATAGAGTTTCTTGG	
f_pBluescript	CTGCAAGGCCGATTAAGTTGGG	Sequencing of the cloning vector
r_pBluescript	CGCCAAGCTCGAAATTAACC	
f_pBluescript 1	CACCTAAAATTGTAAGCGTT	
f_pBluescript 2	GTTGTAAAACGACGGCCAGTGAA	
f_pBluescript 4	TTGATCAACGACCTTTTGGAAAC	
f_pBluescript 5	CAGGCATGCAAGCTTGACTCA	
f_pBluescript 7	GAAATGTTATCCGCTCAC	
r_pBluescript 4	AGCCAGATCAATGTCGATCGTGG	
r_pBluescript 5	GATGTTTATGCTAATTCACATAAACATC	

r_pBluescript 7	AGCGCCCAATACGCAAACC	
f_MS2	GTACGAGCTAGCCCATGGCG	Check the orientation of the MS2 aptamer
r_MS2	CGCCATGGGCTAGCTCGTAC	
f_MBP-MS2 (1)	ACCAGTAAGGCAACCCCGCC	Sequencing of the expression vector
f_MBP-MS2 (2)	ATGTTCCGGCGTTATTCTTIG	
f_MBP-MS2 (3)	TGGCACGACAGGTTTCGCC	
f_MBP-MS2 (4)	AATCTGGCCTGTTGGCTGAAATC	
f_MBP-MS2 (5)	CTCGAAAACATATCTGCTGAC	
r_MBP-MS2 (1)	CGCCATTCAGGCTGCGCAAC	
f_hcf244	TCCTGAAGAAAACCTTCTCAGTGCAG	Genotyping of <i>hcf244</i> T-DNA insertion line
LB-hcf244	ATAATAACGCTGCGGACATCTACATTTT	
r_hcf244	CTTTTAAACTTTTGACACCTCTTGC	
f_hcf173	GAGATTTGAGGTTACCGTTTAGGC	Genotyping of <i>hcf173-2</i> T-DNA insertion line
LB-hcf173	ATAATAACGCTGCGGACATCTACATTTT	
r_hcf173	CTGAAAATGGCACCCCTTACC	
R-B-1_GABI-kat	CGCCAGGTTTTTCCCAGTCAACGACG	
R-B-2_GABI-kat	GTGGATTGATGTGATATCTCC	
f_psd1-1	CCAAACAACCTCTAGCTTCTCATTTCCG	Genotyping of <i>psad1-1</i> T-DNA insertion line
r_psd1-1	TAAACACTGCTCTTTCCCTCGGTAGC	
3'dSpm	TACGAATAAGAGCGTCCATTTAGAGTGA	
f_hcf107-2 new	GTAAAGCAAGGAATCTTCTAGC	Genotyping of <i>hcf107-2</i> T-DNA insertion line
r_hcf107-2 new	GTGACTGATGAGGAAAATACAG	
aadA-outprm fw	GAACTCCGGCGGAATACGA	Genotyping of <i>psbD</i> translation knockdown mutants
aadA-outpsbA rev	TCCTTCTTTTCAAAAACCTCCTA	
psbD-ATG fw new	AGGAGAGGTTATTTCTTTC	
psbD-aadA-span fw	ATCGCTTGGTCCCTTGAAGAG	

2.2 Methods

2.2.1 Growth, cultivation, and transformation of plants and bacteria

2.2.1.1 Cultivation of *Escherichia coli* (*E.coli*)

The *E. coli* strain DH5-Alpha (genotype *dlacZ* Delta M15 Delta(*lacZYA-argF*) U169 *recA1 endA1 hsdR17(rK-mK+)* *supE44 thi-1 gyrA96 relA1*) was used for standard cloning experiments. Rosetta (DE3) *E. coli* cells (genotype *F⁻ ompT hsdS_B (r_B⁻ m_B⁻) gal dcm (DE3) pRARE (Cam^R)*) were used for heterologous protein expression.

In all experiments, bacteria were grown overnight at 37 °C either in liquid LB medium (Bertani, 1951) (180 rpm; Orbital Shaker 3017), or on agar supplemented LB medium (15 g/L agar), containing the appropriate concentration of antibiotics for selection.

LB medium: 1 % (w/v) tryptophane, 1 % (w/v) NaCl, 0.5 % (w/v) yeast extract, 1.5 % (w/v) micro-agar.

2.2.1.2 Preparation of heat-shock competent *E.coli*

The same protocol was used for both DH5-Alpha and Rosetta (DE3) strains: a single colony of *E. coli* cells was inoculated into 10 mL LB medium. The culture was incubated overnight at 37 °C with continuous shaking and subsequently inoculated into 500 mL LB medium. Following growth to an OD₆₀₀ of 0.4 (measured with Ultrospec[®] 3100), the cell culture was collected by centrifugation for 10 min at 5000 g and 4 °C in an Allegra[™] 25R centrifuge. The pellet was gently resuspended in 300 ml of sterile-filtered ice-cold CCMB80 buffer and rested for 20 min on ice before centrifugation for 10 min at 4 °C. The pellet was resuspended in 40 ml of sterile-filtered ice-cold CCMB80 buffer, 100 µl aliquots were snap-frozen (liquid N₂). Cells were stored at -80 °C before use.

CCMB80 buffer: 10 mM KOAc pH 7.0, 80 mM CaCl₂·2H₂O, 20 mM MnCl₂·4H₂O, 10 mM MgCl₂·6H₂O, 10% (v/v) glycerol, pH 6.4 adjusted with HCl.

2.2.1.3 Heat-shock transformation of *E.coli*

An aliquot of heat-shock competent cells (50 µL) was thawed on ice, mixed with 5 to 10 µL of the ligation reaction or 100 pg of the plasmid DNA. Following 30 min resting on ice, the heat shock was applied for 90 s at 42 °C before 2 min further resting on ice. The cells were recovered in 1 mL LB medium for 1 h at 37 °C and 400 rpm. Subsequently, 100 µL of cells were plated on LB medium supplemented with 100 µg/mL ampicillin for growth overnight at 37 °C.

2.2.1.4 Plant material

Nicotiana tabacum cv. Petit Havana and *Arabidopsis thaliana* Columbia accession (Col-0) (referred to throughout as tobacco and *Arabidopsis*) were used for all experiments.

2.2.1.5 Sterilization of tobacco seeds

Approximately 200 μL tobacco seeds were mixed with 1 mL of 70 % ethanol and 20 μl of Tween 20 for 2 min with constant shaking. After removal of the supernatant, 1 mL of 6 % NaOCl was added followed by further shaking for 5 min. The supernatant was thereafter immediately removed and the seeds were washed 5 to 6 times with sterile MQ-H₂O. Subsequently, the seeds were distributed on MS medium (Murashige and Skoog, 1962) supplemented with the corresponding antibiotic, allowed to imbibe overnight at 4 °C and grown under grow-bank conditions (section 2.2.1.4).

MS medium: 0.44 % (w/v) MS elements M0222, pH 5.8 adjusted with KOH, 0.56 % (w/v) agar, without or with sucrose (2 % or 3 % (w/v) sucrose).

2.2.1.6 Sterilization of *Arabidopsis* seeds

Roughly, 20 mg of *Arabidopsis* seeds were surface-sterilized by mixing with 2 mL of 70 % ethanol and 0.5 % SDS for 10 min followed by an additional 10 min incubation with 100 % ethanol. After ethanol removal, the seeds were dried for at least 1 h under the clean bench followed by mixing with 1 mL sterile 0.15 % agarose and plating on MS medium supplemented with 2 % sucrose medium (Murashige and Skoog, 1962). After stratification at 4 °C for 48 h, the plates were transferred to controlled conditions ($\sim 100 \mu\text{mol m}^{-2} \text{s}^{-1}$ for 12/12 h light/dark, 21 °C).

2.2.1.7 Gas sterilization of tobacco seeds

Gas sterilization was used for a large number of samples. To this end, tobacco seeds were surface sterilized by incubation for 5 h in the presence of chlorine gas, produced by the addition of 4 mL 37 % [v/v] HCl to 50 mL of 12 % NaOCl. Sowing and growth were performed as described in section 2.2.1.5.

2.2.1.8 Growth conditions

Detailed growth conditions for each mutant are listed below:

Mutant	Medium	Growth conditions
$\Delta ycf4$	On sterile media	50 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for 16/8 h light/dark, 25 °/22 °C (grow-bank conditions)
$\Delta psbN$	At four leaves stage transferred to soil	
KD- <i>psaA</i>	On soil	70 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for 16/8 h light/dark, 22 °/18 °C
<i>psbD</i> -GTG		150 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for 16/8 h light/dark, 22 °/18 °C, 70 %

<i>psbD</i> -TTG		then transferred to 350 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for 16/8 h light/dark, 22 °/18 °C, 75/70 % humidity (standard conditions)
KD- <i>psbD</i>		
<i>atpB</i> -GTG		
as- <i>AtpC</i>		
as- <i>RBCS</i>		
Δ <i>psaI</i>		
Δ <i>ycf10</i>		
<i>mrl1-1</i>	On soil	~ 120 $\mu\text{mol m}^{-2} \text{s}^{-1}$, 16/8 h light/dark, 21 °C (greenhouse)
<i>rbcsl1a3b-1</i>		
<i>rbcsl1a3b-1</i>		
<i>hcf111-1</i>	On sterile media	~ 50 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for 16/8 h light/dark, 23 °/18 °C, 70 %
<i>hcf173-2</i>	On sterile media	~ 25 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for 16/8 h light/dark, 22 °/18 °C
Δ <i>petL</i>	On soil	150 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for 16/8 h light/dark, 22 °/18 °C, 70 %
<i>ndhC/K/J</i>		
<i>ndhA/H/I</i>		
Δ <i>psbD/C</i>	In magenta boxes	~ 5-10 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for 16/8 h light/dark, 22 °/18 °C
Δ <i>psbB</i>		
Δ <i>atpB</i>		
<i>psad1-1</i>	On soil	~ 120 $\mu\text{mol m}^{-2} \text{s}^{-1}$, 16/8 h light/dark, 25 °/20 °C

2.2.1.8.1 Tobacco cultivation on soil

Seeds were germinated on compost moisturized with water supplemented with 0.15 % [v/v] Previcur fungicide and grown under 150 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for 16/8 h light/dark, 22 °/18 °C, 70 %. Plants were watered with tap water supplemented with 1 g/L Hakaphos[®] fertilizer. The first nine days the seedlings were covered by a plastic cultivation dome to maintain humidity. Eight days after sowing, seedlings were transplanted to individual pots (6 cm diameter) and kept for an additional two days under the same conditions before they were transferred to standard conditions (section 2.2.1.8). After three to four weeks, at a developmental stage with four true leaves, the aerial part of the plant was harvested and immediately flash-frozen in liquid nitrogen. In all cultures, the plants were harvested 30 min into the light cycle because chloroplast gene expression peaks early after the start of illumination (Chotewutmontri and Barkan, 2016).

2.2.1.8.2 Cultivation of tobacco on vermiculite

Seeds were sown directly on a nylon net placed on vermiculite and grown for 7 days under standard conditions (section 2.2.1.8). The seeds were covered with a plastic dome to keep the humidity and the actual measured light intensity was ~ 280 $\mu\text{mol m}^{-2} \text{s}^{-1}$.

2.2.1.8.3 Cultivation on sterile media

Sterilized seeds were grown on plates containing MS medium (Murashige and Skoog, 1962) supplemented with 3 % sucrose and the appropriate antibiotic. After one night stratification at 4 °C, the seeds were grown under grow-bank conditions (section 2.2.1.8). For tobacco transplastomic lines, the antibiotic concentration used for selection was either 500 µg/mL for spectinomycin or 200 µg/mL for kanamycin. For further growth, 10 to 14 days after sowing, individual seedlings were transplanted to Magenta boxes containing MS medium supplemented with 3 % sucrose (Murashige and Skoog, 1962).

2.2.1.8.4 Cultivation of *Arabidopsis* on soil

Arabidopsis was grown for two weeks after sowing on MS medium (Murashige and Skoog, 1962) supplemented with 2 % sucrose under long-day conditions (16 h light/8 h dark, ~ 100 µmol m⁻² s⁻¹, 21 °/16 °C, 75 % humidity). Afterward, the seedlings were transplanted into soil and kept in the same condition for one week before being moved to the greenhouse (2.2.1.8).

2.2.1.9 Plastid transformation and regeneration of transplastomic plants

Chloroplast transformation was performed using biolistic bombardment as described in Ruf and Bock (2011). Briefly, young leaves from aseptically grown wild-type tobacco plants in Magenta boxes under grow-bank conditions were bombarded with gold particles (0.6 µm diameter, Bio-Rad) coated with 20 µg plasmid DNA using a PDS-100/He Biolistic gun (Bio-Rad) with the Hepta adapter setup. Bombarded leaves are cut into pieces and placed on RMOP medium supplemented with 500 µg/mL spectinomycin for selection. Transplastomic resistant shoots start to appear after three to six weeks incubation under 25 µmol m⁻² s⁻¹, 16 h light/8 h dark cycle. In general, one to two regeneration rounds were needed to reach the homoplastomic state.

RMOP medium: 0.44 % (w/v) MS, 3 % (w/v) sucrose, 0.01 % (w/v) NAA 1 mg/mL (in 0.1 M NaOH), 0.1 % (w/v) BAP 1 mg/mL (in 0.1 M HCl), pH 5.8 adjusted with KOH, 0.54 % (w/v) agar.

2.2.2 Nucleic acid analysis

2.2.2.1 Plasmids

The plasmids generated or used in this work are listed in Table 3.

Table 3: Overview of plasmids used or generated in this work

Plasmid	Purpose	Source	Resistance in bacteria	Resistance in plants
pRG1	Tagging <i>rbcL</i> mRNA with the D8 aptamer and used as a template for pRG2	This work	Ampicillin	Spectinomycin, streptomycin

pRG2	Tagging <i>rbcL</i> mRNA with MS2 aptamer	This work	Ampicillin	Spectinomycin, streptomycin
pRG3	Tagging <i>psbA</i> mRNA with D8 aptamer and used as a template for pRG4	This work	Ampicillin	Spectinomycin, streptomycin
pRG4	Tagging <i>psbA</i> mRNA with MS2 aptamer	This work	Ampicillin	Spectinomycin, streptomycin
pMBP-MS2	Expression of MS2 coat protein fused to the maltose-binding protein (MBP)	Addgene plasmid #65104	Ampicillin	n/a

2.2.2.2 Vector design and cloning techniques

In silico restriction cuttings and alignments were performed using SeqBuilder and SeqMan Pro, DNASTAR Lasergene Core Suite (Burland, 2000), respectively. All the plasmids generated in this work have pBluescript II SK(+) as a backbone.

For tagging of *rbcL* and *psbA* transcripts, corresponding fragments and promoters were cloned as follows (ptDNA: plastid DNA, *Cr*: *Chlamydomonas reinhardtii*, *Nt*: *Nicotiana tabaccum*, P: promoter, f/fw: forward, r/rev: reverse):

Vector	Fragment	Template	Primer
pRG1	D8 aptamer, <i>aadA</i> cassette with <i>NtPrn</i> , <i>CrTrbcL</i> and flanking regions	tobacco ptDNA of transplastomic line <i>rbcL</i> -seph	<i>rbcL</i> fw-PstI accD-PstI rev
pRG2	<i>aadA</i> cassette with <i>NtPrn</i> , <i>CrTrbcL</i> and flanking regions MS2 aptamer	pRG1 synthetic sequence	- MS2-1 MS2-2
pRG3	D8 aptamer, <i>aadA</i> cassette with <i>NtPrn</i> , <i>CrTrbcL</i> and flanking regions	tobacco ptDNA of transplastomic line <i>psbA</i> -seph	<i>psbA</i> -fw-PstI trnH-rev-PstI
pRG4	<i>aadA</i> cassette with <i>NtPrn</i> , <i>CrTrbcL</i> and flanking regions MS2 aptamer	pRG3 synthetic sequence	- MS2-1 MS2-2

The pRG1 and pRG3 vectors were derived by amplifying fragments from the genomic DNA (gDNA) of the unpublished transplastomic *psbA* and *rbcL* Sephadex-tagged plants. The latter lines contain the Sephadex-binding aptamer sequence inserted into the 3' untranslated region (UTR) of either *psbA* or *rbcL* transcripts followed by an *aadA* cassette that confers spectinomycin resistance. The PCR fragments were then ligated into the EcoRV-linearized pBluescript II SK(+) vector. Both vectors

contained at least 400 bp of native tobacco plastid gDNA at both sides of the transgenic region to ensure homologous recombination into the plastid genome. The *aadA* cassette contained the strong, constitutive *Prrn* promoter of the tobacco plastid rRNA operon and 3' UTR of plastid *rbcl* from *Chlamydomonas*. The plastid transformation vectors pRG2 and pRG4 were derived from pRG1 and pRG3, respectively. For this, pRG1 and pRG3 vectors were linearized with MfeI cutting to remove the Sephadex-binding aptamer followed by dephosphorylation. Subsequently, the amplified MS2 aptamer (generated by synthetic oligos MS2-1 and MS2-2; Table 2) was cut with MfeI and then ligated into pRG1 and pRG3. Vectors maps are shown in Figure 3.25.

2.2.2.3 Isolation of plasmid DNA

Small (< 50 µg) and big (> 50 µg) plasmid amounts were isolated with the NucleoSpin® Plasmid kit (Macherey-Nagel) and the NucleoBond® Xtra Midi Plus kit (Macherey-Nagel) following the manufacturer's instructions.

2.2.2.4 Isolation of plant genomic DNA

Plant genomic DNA was isolated by the CTAB-based method (Doyle and Doyle, 1990). In brief, 100 to 200 g of frozen plant tissue was homogenized in liquid nitrogen and thawed in 1 mL CTAB extraction buffer followed by 30 min incubation at 60 °C with 500 rpm shaking. The plant extract was then mixed with 400 µL of chloroform: isoamyl alcohol (24:1) followed by centrifugation for 10 min at 12,000 g at 4 °C. This step was repeated for the recovered aqueous phase from the first centrifugation. The DNA was precipitated by adding 0.7 volumes of isopropanol and subsequent incubation for 5 min at room temperature (RT) and centrifugation for 30 min at 18,000 g and 4 °C. The DNA pellet was washed with 500 µL of 70 % pre-cooled ethanol, air-dried at RT for 20 min and resuspended in 50 µL of MQ-H₂O. Quick DNA extraction was performed using the EXTRACT-N-AMP kit (following the manufacturer's instructions) for PCR genotyping of big sample sizes.

CTAB extraction buffer: 2 % CTAB, 1.4 mM NaCl, 20 mM EDTA pH 8.0, 100 mM Tris-HCl pH 8.0, 100 mM 2-Mercaptoethanol (added immediately before use).

2.2.2.5 Isolation of total plant RNA

RNA was isolated using the TRIzol® Reagent following the manufacturer's instructions (Thermo Fisher Scientific). Briefly, 1 mL of TRIzol® Reagent was added to the frozen plant lysates and mixed immediately by vortexing followed by 10 min rotation. After 5 min centrifugation (Eppendorf Centrifuge 5417 R) at 18,000 g at 4 °C, the supernatant was mixed with 200 µL chloroform/isoamyl alcohol (24:1) followed by 15 min centrifugation at 18,000 rpm at 4 °C. The RNA was precipitated with 0.5 mL 100 % isopropanol at -20 °C overnight and collected by centrifugation for 40 min at 4 °C at

18,000 g. The RNA pellet was washed with 1 mL pre-cooled 75 % ethanol, air-dried for 10 min, resuspended in 20 μ L filtered MQ-H₂O and stored at -20 °C.

2.2.2.6 Assessment of nucleic acid concentration and purity

Nucleic acid concentrations were quantified based on the optical density measurements at 260 nm using NanoDrop™ One (Thermo Fisher Scientific). An absorbance of $A_{260\text{nm}} = 1$ is equivalent to 50 μ g dsDNA/mL and 40 μ g RNA/mL. The absorbance at 280 nm reflects the protein contamination and that at 230 nm reflects the contamination with organic compounds. The purity of nucleic acids in solution was determined based on the $A_{260\text{nm}}/A_{280\text{nm}}$ and $A_{260\text{nm}}/A_{230\text{nm}}$ ratios. High purity DNA and RNA have $A_{260\text{nm}}/A_{280\text{nm}}$ ratios of 1.8 to 2.0 and 1.9 to 2.1, respectively, and $A_{260\text{nm}}/A_{230\text{nm}}$ in the range of 2.0 to 2.2.

2.2.2.7 Polymerase chain reaction (PCR)

For standard genotyping PCR was performed using the DreamTaq DNA polymerase (Thermo Fisher Scientific). For cloning purposes and other products requiring a low error rate, the Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific) was used. The reactions and the programs used are described in Table 4. PCR products were purified using the PCR Clean-up Kit (Macherey-Nagel) according to the manufacturer's instructions.

Table 4: PCR basic reactions mix and program

T_m: melting temperature. The primers used are shown in Table 2.

DreamTaq PCR	1x DreamTaq buffer, 0.2 mM of each dNTP, 0.5 μ M of each primer, 1.25 U/50 μ l reaction of the DreamTaq DNA polymerase, 100 ng gDNA or 10 pg plasmid DNA. Program: 2 min 95 °C; 30-35 cycles of 30 sec 95 °C, 30 sec T _m -2 °C or T _m -5 °C, 1 min/kilobase (kb) 72 °C; 10 min 72 °C.
Phusion PCR	1x Phusion High-Fidelity, 0.2 mM of each dNTP, 0.5 μ M of each primer, 1 U/50 μ l reaction of the Phusion High-Fidelity DNA polymerase, 100 ng gDNA or 10 pg of plasmid DNA. Program: 30 sec 98 °C, 30-35 cycles of 10 sec 98 °C, 30 sec 45-72 °C, 15 sec/kb 72 °C; 10 min 72 °C.

2.2.2.8 Size-based separation of DNA fragments by agarose gel electrophoresis

DNA was separated with 0.8 to 3 % (w/v) agarose gels in 1x TAE buffer. To enable the visualization of DNA under UV light, ethidium bromide (EtBr) was added to a final concentration of 0.1 mg/L. Samples were mixed with 0.2 volume 10x loading buffer and run in parallel with either 100-bp or 1-kb

DNA marker (Gene Ruler DNA Ladder, Thermo Fisher Scientific) at 5 to 10 V/cm until adequately separated using the Peqlab Horizontal Gel Electrophoresis System (PeqLab). The gels were visualized using the Quantum CX5 Gel Documentation System (Vilber Lourmat).

For cloning and sequencing purposes, DNA fragments were recovered from excised gel pieces under UV light. DNA fragments were then purified using NucleoSpin® Gel (Machery-Nagel) according to the manufacturer's instructions.

10x TAE buffer: 400 mM Tris-acetic acid pH 8.0, 10 mM EDTA pH 8.0.

10x loading buffer: Bromophenol blue, xylene cyanol, 50 % (v/v) glycerin, 100 mM EDTA pH 8.0.

2.2.2.9 Preparation of hybridization probes for Southern and northern analyses

Hybridization probes were produced by amplifying target sequences from gDNA using gene-specific primers. PCR products were then labeled with [α - 32 P] dCTP using the Megaprime DNA labeling System (GE Healthcare) according to the manufacturer's instructions. Briefly, 100 ng DNA was mixed with 5 μ L random nonamer primers in 33 μ L reaction volume, denatured for 5 min at 99 °C, followed by the addition of 10 μ L labeling buffer, 2 U Klenow polymerase enzyme and 50 μ Ci [α - 32 P] dCTP. The reaction was incubated for 20 min at 37 °C and 300 rpm followed by enzyme inactivation at 95 °C for 5 min. The unincorporated nucleotides were then removed using Pierce® G-50 columns (Thermo Fisher Scientific). The labeling efficiency was estimated by measuring the counts per minute (cpm) using Quick-Count QC-2000 gamma particle counter (BioScan). Subsequently, the radiolabeled probe was denatured for 10 min at 95 °C then added to the membrane (see sections 2.2.2.10 and 2.2.2.11).

2.2.2.10 Southern blot analysis

For restriction fragment length polymorphism (RFLP) analysis, 3 μ g of gDNA was cleaved with MfeI restriction enzyme at 37 °C overnight and size-separated by gel electrophoresis in 0.8 % agarose gel in 1x TAE buffer (section 2.2.2.8) with the following regime: 30 min at 1.5 V/cm, 2 h at 3.5 V/cm, 1 h at 5 V/cm. The DNA was deputed in the gel in solution I for 15 min, denatured for 30 min in solution II and solution III, each, before the pH of the gel was neutralized by incubation in solution IV for 15 min. All the steps were done at RT under rocking incubation with rinses of MQ-H₂O between each incubation. Subsequently, DNA was transferred onto a Hybond™-N+ membrane (GE Healthcare) by overnight capillary blotting in 10x SSC buffer. After blotting, DNA was covalently cross-linked to the membrane by UV treatment with 0.120 Joules/cm² in a BLX-254 Crosslinker (Vilber Lourmat) and, to block the membrane, pre-incubated in a hybridization tube by rotation for at least 1 h at 60 °C in 10 mL Church buffer. Subsequently, the membrane was incubated overnight at 60 °C with the appropriate radioactive probe in 8 mL of fresh Church buffer followed by two washes for 10 min each with 50 mL wash buffer I and by 10 min wash with 50 mL wash buffer II, all at 60 °C and under rotation. Afterward,

a Storage Phosphor Screen (GE healthcare) was exposed to the membrane for 2 to 24 h, depending on signal intensity, and the signal was visualized using the TyphoonTM TRIO+ Variable Mode Imager (Amersham Biosciences).

Solution I: 0.25 M HCl.

Solution II: 0.5 M NaOH.

Solution III: 0.5 M NaOH, 1.5 M NaCl.

Solution IV: 1 M Tris-HCl pH 6.5, 3M NaCl.

20x SSC buffer: 3 M NaCl, 0.3 M Na-citrate, pH 7.0 adjusted with HCl.

Church buffer: 0.5 M sodium phosphate buffer pH 7.2 [1M Na₂HPO₄, 1 M NaH₂PO₄], 1 mM EDTA pH 8.0, 7 % SDS.

Wash buffer I: 2x SSC buffer, 0.1 % SDS.

Wash buffer II: 0.5x SSC buffer, 0.1 % SDS.

2.2.2.11 Northern blot analysis

RNA samples were vacuum-dried and denatured in 10 µL of Northern sample buffer for 5 min at 65 °C. Subsequently, the samples were placed on ice for 10 min and RNAs were size-separated on a 1.2 % agarose gel in 1x MOPS gel buffer (including 6 % formaldehyde; Sigma-Aldrich). The gel was run for ~ 1.5 h at 6 V/cm in 1x MOPS running buffer containing 3.7 % formaldehyde (Merck). The gel was then rinsed twice with MQ-H₂O to remove the formaldehyde. Afterward, the gel was washed for 5 min with 5x SSC buffer. The RNAs were transferred overnight with 5x SSC buffer by capillary blotting onto an Amersham HybondTM- N membrane (GE Healthcare). The RNAs were crosslinked to the membrane as described in section 2.2.2.10 and rRNAs were visualized by methylene blue staining to check the integrity of the RNA and blotting efficiency. The probe labeling, hybridization, membrane washings, and visualization were performed as described in section 2.2.2.10 with one minor modification: after hybridization, the membrane was washed twice with 50 mL 1x SSC, 0.5 % SDS buffer followed by one wash with 50 mL 0.5x SSC, 0.1 % SDS buffer.

Northern sample buffer: 62.5 % (v/v) deionized formamide, 21 % (v/v) formaldehyde, 12.5 % (v/v) MOPS pH 8.0, 0.02 % (v/v) bromophenol blue, 0.02 % (v/v) xylene cyanol.

10x MOPS running buffer: 200 mM MOPS, 80 mM NaAc, 10 mM EDTA, pH 7.0.

10x MOPS gel buffer: 200 mM MOPS, 80 mM NaAc, 10 mM EDTA, pH 8.0.

2.2.2.12 DNA sequencing

Purified PCR products and plasmid DNA were sequenced by Eurofins MWG Operon (Ebersberg, Germany). Sequence alignments were performed using Seqman (DNASTAR Lasergene 10 Core Suite).

2.2.3 Microarray-based ribosome profiling

2.2.3.1 Ribosome footprint preparation

Ribosome footprints and total RNA were isolated as described in (Schuster et al., 2020; Trösch et al., 2018; Zoschke et al., 2013). Concisely, 500 mg of frozen plant tissue were homogenized in liquid nitrogen with mortar and pestle and thawed with 5 mL of fresh ribosome extraction buffer. For subsequent total RNA isolation and microarray-based transcript profiling, a 0.5 mL aliquot of the lysate was flash-frozen in liquid nitrogen. The remaining lysate was filtered through glass wool and centrifuged for 10 min at 15,000 g at 4 °C in an SS-34 rotor to remove cell debris. Thereafter, 4 mL of the supernatant was incubated with 600 U of Micrococcal nuclease (MNase, Roche) and 20 µL of 1M CaCl₂ for 1 h at RT with slow rotation. MNase-treated supernatant was then gently loaded onto a 1 mL sucrose cushion and ultracentrifuged for 1.5 h at 303,800 g and 4 °C in an SW55-Ti rotor using Optima™ L-80 XP Ultracentrifuge (Beckman Coulter). The resulting monosome pellet was resuspended with 0.5 mL footprint isolation buffer. Ribosome footprints and total RNA were isolated with TRIzol reagent following the manufacturer's instructions. The RNA was resuspended in 50 µL of filtered MQ-H₂O and the concentration and purity were determined by NanoDrop™ One (Thermo Fisher Scientific).

Ribosome footprints were purified by electrophoresis on a 12 % denaturing polyacrylamide gel (19:1, acrylamide: bisacrylamide) prepared in 1x TBE buffer containing 8 M urea. 30 µg RNA from the monosome fraction were vacuum-dried, resuspended in 40 µL of ribosome footprint loading buffer and incubated for 10 min at 45 °C with 600 rpm and mixed by pipetting before denaturation for 10 min at 70 °C with 450 rpm. In parallel to the denatured RNA, 4 µL of prestained RNA marker (Biodynamics Laboratory) was loaded. The gel was run in 1x TBE buffer with a constant power of 30 W while cooling the gel chamber to 12 °C until adequate separation.

Gel slices including RNA from 23 to 45 nt were excised and incubated in 4 mL of TESS buffer overnight at 4 °C with slow rotation to elute the RNA. The ribosome footprints were then recovered by mixing with an equal volume of phenol:chloroform: isoamyl alcohol (25:24:1) followed by centrifugation for 5 min at 4,000 rpm at RT using the Allegra™ 25R Centrifuge (Beckman Coulter). The RNA was precipitated from the supernatant by adding 2.5 volumes of ethanol and incubation at -20 °C overnight. Following centrifugation for 1 h at 15,000 g at 4 °C, the ribosome footprint pellet was suspended in 500 µL of filtered MQ-H₂O with 0.1 M NaCl and 2.5 µL of GlycoBlue (Thermo Fisher Scientific) and then subjected to a second round of phenol:chloroform:isoamyl alcohol (25:24:1) extraction followed by chloroform:isoamylalcohol (24:1) extraction and ethanol precipitation. After precipitation, the ribosome footprint pellet was washed with 900 µL of 75 % ethanol, air-dried for 10 min at RT, resuspended in 20 µL MQ-H₂O, and stored at -80 °C. The concentration and purity of ribosome footprints were determined by NanoDrop™ One (Thermo Fisher Scientific).

Ribosome extraction buffer:	0.2 M sucrose, 0.2 M KCl, 40 mM Tris-OAc pH 8.0, 10 mM MgCl ₂ , 10 mM 2-Mercaptoethanol, 2 % (v/v) polyoxyethylene (10) tridecyl ether, 1 % (v/v) Triton X-100, 100 µg/mL chloramphenicol, 100 µg/mL cycloheximide.
Sucrose cushion:	30 % (w/v) sucrose, 0.1 M KCl, 40 mM Tris-OAc pH 8.0, 15 mM MgCl ₂ , 5 mM 2-Mercaptoethanol, 100 µg/mL chloramphenicol, 100 µg/mL cycloheximide.
Footprint isolation buffer:	10 mM Tris pH 8.0, 1 mM EDTA pH 8.0, 100 mM NaCl, 1 % (w/v) SDS, 0.1 M EGTA pH 8.0.
10x TBE buffer:	0.89 M Tris, 0.89 M boric acid, 2 mM EDTA pH 8.0.
Ribosome footprint loading buffer:	90 % (v/v) deionized formamide, 20 mM Tris-HCl pH 7.5, 20 mM EDTA pH 8.0, 0.04 % (w/v) bromophenol blue, and 0.04 % (w/v) xylene cyanol.
TESS buffer:	10 mM Tris pH 8.0, 1 mM EDTA pH 8.0, 0.1 M NaCl, 0.2 % (w/v) SDS.

2.2.3.2 Total RNA fragmentation

12 µg total RNA were chemically fragmented by incubation with 2.5 µL of RNA fragmentation buffer in a final volume of 25 µL for 12.5 min at 85 °C. The reaction was stopped by adding 225 µL TESS buffer (section 2.2.3.1) containing 1 mM EDTA pH 8.0. The fragmented total RNA was extracted with 250 µL of phenol:chloroform:isoamylalcohol (25:24:1) and chloroform:isoamylalcohol (24:1), respectively, before ethanol precipitation (all as described in section 2.2.3.1).

RNA fragmentation buffer: 0.4 M Tris-OAc pH 8.3, 1 M KOAc, 0.3 M Mg(OAc)₂.

2.2.3.3 RNA labeling and hybridization

RNA labeling was performed according to (Trösch et al., 2018; Zoschke et al., 2013). Briefly, 4 µg of ribosome footprints and 3.5 µg of fragmented total RNA isolated from control and mutant plants were differentially labeled with Cy5 and Cy3 (ULS aRNA Labeling Kit, Kretech), respectively, following the manufacturer's instructions. Labeled RNA was vacuum-concentrated to ~1 µL at RT, resuspended in 110 µL of hybridization buffer and denatured for 10 min at 70 °C. After denaturation, RNA was hybridized to a custom tiling microarray (Arbor Biosciences) that cover all the open reading frames (ORFs) in the tobacco or *Arabidopsis* chloroplast genome with approximately 30-nt resolution (Trösch et al., 2018) Microarrays were scanned with a GenePix 4000B microarray scanner (Molecular Devices).

Hybridization buffer: 2.25 M NaCl, 15 mM EDTA pH 8.0, 89 mM NaH₂PO₄, 61 mM Na₂HPO₄, 10 % (v/v) deionized formamide, 0.01 mg/mL acetylated BSA, 0.01 % (v/v) Tween 20.

2.2.3.4 Data processing and analysis

Data analysis was performed as described previously (Schuster et al., 2020; Trösch et al., 2018), with some modifications. Briefly, the data were analyzed with GenePix Pro 7.2 software (Molecular Devices). Low-quality spots on the microarray were manually removed after visual inspection. Only probes for which at least two out of four technical replicate spots showed sufficient quality were considered in the analysis and are represented in the figures. The background was subtracted using the local subtraction method. Afterward, the median value from technical replicates was calculated for each probe for both channels (F635-B635 or F532-B532) with values ≤ 100 considered below background and removed from the analysis. All median values of probes covering the protein-coding regions in ribosome and transcriptome profiling were normalized to a constant value (5000) in order to exclude biases caused by technical variations such as labeling and hybridization efficiencies. Subsequently, the average value of normalized probes signals in each ORF was calculated. In order to obtain the relative expression levels (RNA or ribosome footprint) for every single ORF compared to the average of all chloroplast ORFs in the mutant and the control, the relative abundance of ribosome footprints and total RNA were calculated by normalizing the average signal of each ORF to the average signal of all ORFs in a logarithmic scale. To relatively compare the changes in the transcript accumulation and protein synthesis level between the mutant and the control, the log-transformed relative abundance values of the ribosome footprints and the total RNA in the control were subtracted from the corresponding relative abundance values in the mutant for each replicate. Translation efficiencies were calculated for each ORF by subtracting the log-transformed relative value of the total RNA from that of the ribosome footprints in each replicate. The average and of relative changes of ribosome footprints, total mRNA, and translation efficiencies were calculated for each ORF from two or three biological replicates. The standard deviation was calculated for datasets with three biological replicates.

Differential distribution of elongating ribosomes was evaluated as described in (Chotewutmontri and Barkan, 2018; Schuster et al., 2020). Shortly, the ribosome footprint signal of each probe located in an ORF was normalized to the sum of the signals of all the probes in the same ORF followed by the calculation of the ratio of these relative ribosome occupancies between the mutant and the control. This enables a local assessment of changes of footprints abundances on probe level regardless of the overall change of the whole ORF. Significance of differential distribution of ribosomes was assessed using the empirical Bayes method in limma package (Smyth, 2004) and the *P* values were adjusted according to the False Discovery Rate procedure (FDR) (Benjamini and Hochberg, 1995).

2.2.4 Ribo-seq

2.2.4.1 Ribosome footprint preparation

For the sequencing approach, the ribosome footprints were isolated similarly to the microarray-based ribosome profiling (section 2.2.3.1) with some minor modifications. 250 U/ml RNase I (Thermo Fisher Scientific) was used for the digestion of unprotected ribosome-free RNA. Following the nuclease treatment, the lysate was layered on a 2 mL sucrose cushion to ensure better separation of monosomes from smaller ribonucleoprotein (RNP) particles. Additionally, ribosome footprints from 20 nt to ~ 39 nt were recovered from the polyacrylamide gel (section 2.2.3.1). To obtain higher accuracy, the concentration of purified ribosome footprints was determined by Qubit with the microRNA assay kit (Thermo Fisher Scientific) and the size distribution was assessed with the 2100 Bioanalyzer Instrument (Agilent Technologies) with the small RNA assay kit (Agilent) before proceeding to the library preparation.

2.2.4.2 Total RNA preparation

For total RNA samples, gDNA was removed using TURBO™ DNase (Thermo Fisher Scientific) according to the manufacturer's instructions. After precipitation, 6 µg of purified total RNA were subjected to rRNA depletion using a RiboMinus™ Plant Kit for RNA-Seq (Thermo Fisher Scientific) following the manufacturer's instructions.

2.2.4.3 Library Construction and sequencing

Before library preparation, the 5' and 3' ends of the ribosome footprints were phosphorylated and dephosphorylated, respectively, using T4 Polynucleotide Kinase (Thermo Fisher Scientific). To this end, 100 ng of ribosome footprints were denatured in a volume of 7 µL by heating to 65 °C for 5 min, followed by addition of 2 µL of the kinase master mix: 1 µl of 10x T4 Polynucleotide Kinase buffer, 0.5 µl RNasin® RNase Inhibitor (40 U/µL), and 0.5 µl T4 Polynucleotide Kinase. The reaction was incubated for 10 min at 37 °C after which 1 µL of 10 mM ATP (Sigma-Aldrich) was added. Subsequently, the RNA samples were incubated for 30 min at 37 °C followed by 20 min at 65 °C to deactivate the enzyme.

70 ng of end-fixed ribosome footprints (7 µL) were used as input for the NEXTflex™ Small RNA-Seq Kit v3 (Bioo Scientific) following the manufacturer's instructions. The concentration of the cDNA library was measured using Qubit with dsDNA HS assay kit (Thermo Fisher Scientific) and the size distribution was assessed with the 2100 Bioanalyzer Instrument (Agilent Technologies) with the High Sensitivity DNA assay kit (Agilent Technologies). The libraries were barcoded according to the NEXTflex™ Small RNA-Seq Kit v3 (Bioo Scientific) and nine times multiplexed for single-end 75-bp sequencing on Illumina NextSeq 500. Next-generation RNA sequencing was performed by the Sequencing Core Facility of the Max Planck Institute for Molecular Genetics.

2.2.4.4 Ribo-seq data analysis

Ribo-seq data were processed by Michael Ting, MPIMP. The quality of the obtained footprint reads was checked with FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Adapters were removed using cutadapt (Martin, 2011) followed by a custom python script to clip the unique molecular indexes and preserve their identity in the read name. All alignments were performed with STAR aligner v2.7.1a (Dobin et al., 2013), using a sequential mapping approach where unmapped reads were used as input for the proceeding alignment. The alignment order is as follows: 1) rRNA and tRNA contaminants; 2) chloroplast genome; 3) mitochondria genome; and 4) nuclear genome (Edwards et al., 2017). The following sequences were obtained from NCBI (as well as their respective annotation format): chloroplast genome (Z00044.2), mitochondria genome (NC_006581.1), 5.8S rRNA (AJ300215.1), 5S rRNA (AJ222659.1), 18S rRNA (AJ236016.1), and 26S rRNA (AF479172.1). Sequences for the nuclear genome and tRNA were obtained from solgenomics (ftp://ftp.solgenomics.net/genomes/Nicotiana_tabacum/). Alignments were processed through UMI-tools (Smith et al., 2017) to remove PCR duplicates. Reads mapping to the CDS of genes were summarized using featureCounts (Liao et al., 2014). Additional manipulations of alignment data were done using samtools (Li et al., 2009) and bedtools (Quinlan and Hall, 2010).

2.2.5 Protein analysis

2.2.5.1 Phenol based total protein extraction

Total protein was extracted according to Cahoon et al. (1992 with some modifications. 200 mg of plant tissue was homogenized in liquid nitrogen, thawed with 500 μ L protein isolation buffer, and then mixed with 500 μ L of phenol using a Vortex-Genie[®] followed by centrifugation for 10 min at 18,000 g at RT. Subsequently, 200 μ L of the aqueous phase was transferred to 1 mL of 0.1 M NH₄OAc in methanol for overnight precipitation at -20 °C. The protein pellet was recovered by centrifugation for 5 min at 18,000 g at 4 °C followed by washing with 500 μ L of 1 M NH₄OAc in methanol, air-drying for 20 min at RT and resuspension in 100 μ L of 1 % (w/v) SDS.

Protein isolation buffer: 0.7 M sucrose, 0.5 M Tris-HCl, 50 mM EDTA pH 8.0, 0.1 M KCl adjusted to final pH of 9.4 with KOH. Before usage 2 % (v/v) 2-Mercaptoethanol and 2 % (v/v) cOmplete[™] EDTA-free protease inhibitor was freshly added.

2.2.5.2 Soluble and membrane protein extraction

Quick isolation of soluble and membrane proteins from plant tissue was carried out according to Barkan (1998 with minor modifications: the homogenate from plant tissue was mixed with soluble protein extraction buffer. The thylakoid membrane was pelleted by centrifugation for 15 min, at 13,000 g.

Subsequently, the supernatant containing the soluble proteins was transferred to a new tube and the pellet was suspended with carbonate buffer.

Soluble protein extraction buffer: 100 mM Tris pH 8, 5 mM EGTA pH 8, 5 mM EDTA pH 8, 150 mM NaCl. Before usage 2 % (v/v) 2-Mercaptoethanol, 2 % (v/v) cOmplete™ EDTA-free protease inhibitor was freshly added.

Carbonate protein extraction buffer: 100 mM Na₂CO₃, 10% (w/v) sucrose. Before usage 2 % (v/v) 2-Mercaptoethanol, 2 % (v/v) cOmplete™ EDTA-free protease inhibitor was freshly added.

2.2.5.3 Quantification of protein concentration

PIERCE™ BCA Protein Assay Kit (Thermo Fisher Scientific) was used to quantify the protein concentration according to the manufacturer's instruction. The photometric measurements were performed with the microplate reader CLARIOstar® (BMG Labtech).

2.2.5.4 Separation of proteins by SDS-PAGE

Depending on the follow-up experiment, the separation of the proteins by SDS-PAGE (SDS-polyacrylamide gel electrophoresis) was performed according to either Laemmli (1970) or Schagger (2006). For western blot purposes, 4 µg proteins were mixed with 2x SDS protein sample buffer, denatured for 10 min at 70°C and loaded onto a 0.75 mm thick 4 to 20 % (w/v) precast polyacrylamide gradient gels in the Mini-PROTEAN® vertical electrophoresis cell system (Bio-Rad). The separation was performed in 1x Laemmli buffer at 4 °C with 5 V/cm for 30 min followed by 8 V/cm for 60 min. For *in vivo* protein labeling experiments, self-made tricine gels were used according to Schagger (2006). Protein samples were mixed with 2x tricine sample buffer, denatured for 10 min at 70 °C and separated on a 1 mm thick 4 to 10 % tris-tricine gradient gel with 1x Anode buffer and 1x cathode buffer in a Mini-PROTEAN® vertical electrophoresis cell system (Bio-Rad). The following running regime was used: 4 V/cm for 40 min (till the proteins pass through the stacking gel), 6 V/cm for 80 min, and 12 V/cm until the dye reached the gel base. In both experiments, the protein separation was performed next to 4 µL of prestained protein marker (Precision Plus Protein™ Dual Xtra Prestained Protein Standards, Bio-Rad).

2x SDS protein sample buffer: 125 mM Tris-HCl pH 6.8, 4 % (w/v) SDS, 20 % (v/v) glycerol, 25 mM EDTA pH 8.0, 0.04 % (w/v) bromophenol blue, 2 % (v/v) 2-Mercaptoethanol.

10x LaemmLi buffer: 0.25 M Tris, 1.92 M glycine, 1 % (w/v) SDS.

2x Tricine sample buffer:	100 mM Tris-HCl pH 6.8, 24% glycerine (v/v), 12.5% SDS (w/v), 0.02 % Coomassie Brilliant Blue G250 (w/v), 2 % (v/v) 2-Mercaptoethanol.
10x Anode buffer:	2 M Tris-HCl pH 8.9.
10x Cathode buffer:	1M Tris, 1M Tricine, 1% SDS, pH 8.25.
4 % Tricine stacking gel:	13.3 % Rotiphorese Gel A, 6 % Rotiphorese Gel B, 760 mM Tris-HCl pH 8.45, 0.07 % SDS, 5.3 mM TEMED, 0.08 % APS.
10 % Tricine running gel:	32.6 % Rotiphorese Gel A, 15 % Rotiphorese Gel B, 1 M Tris-HCl pH 8.45, 0.1 % SDS, 10.6 % Glycerine, 2.38 mM TEMED, 0.036 % APS.

2.2.5.5 Protein transfer and staining

Gel separated proteins were transferred to a nitrocellulose membrane (AmershamTM ProtranTM 0.2 µm NC, GE Healthcare) in the Mini Trans-Blot[®] Electrophoretic Transfer Cell system (Bio-Rad) in 1x Transfer buffer. The transfer was done overnight at 4 °C with a constant low voltage (20 V). For quality control, abundant proteins were stained with Ponceau S solution for 5 min followed by destaining with MQ-H₂O and subsequent scanning using EPSON Perfection V700 Photo. The membrane was either directly used for immunodetection or dried and stored at RT until use.

The tricine gels were visualized with Coomassie blue staining as described in Schägger (2006 with small modifications: the gel was incubated in the fixing solution for 20 min, stained with blue brilliant Coomassie staining buffer for 10 min and de-stained with three consecutive incubations in the destaining buffer for 10 min each. All incubations were done at RT with gentle shaking. For higher sensitivity, in some cases, colloidal Coomassie (G250) was used. To this end, the gel was stained for 30 min with colloidal Coomassie buffer followed by 10 min incubation with the colloidal destaining buffer.

1x Transfer buffer:	1x LaemmLi buffer, 20 % ethanol.
Ponceau S solution:	0.1 % (w/v) Ponceau S, 5 % acetic acid.
Fixing solution:	50 % methanol, 10 % acetic acid.
Blue brilliant Coomassie staining buffer:	0.025 % Coomassie R250, 10 % acetic acid.
Destaining buffer:	40 % methanol, 7 % acetic acid.
Colloidal Coomassie buffer:	0.02 % Coomassie G250, 5 % aluminium sulfate-(14-18)-hydrate, 10 % ethanol, 2 % orthophosphoric acid.
Colloidal destaining buffer:	10 % ethanol, 2 % orthophosphoric acid.

2.2.5.6 Immunodetection

Protein membranes were incubated with the blocking buffer for 1 h at RT followed by incubation with the primary antibody solution and the secondary antibody solution for 1 h each at RT. After each incubation, three washes were performed with 1x TBST buffer for 10 min each at RT. All steps were performed under rocking incubation. The signals were visualized using the ECL Plus™ detection system (GE Healthcare) according to the manufacturer's instructions and documented with Syngene G:BOX Chemi XT4 (SynOptics). An incubation with Restore PLUS Western Blot Stripping Buffer (Thermo Fisher Scientific) was performed for 20 to 30 min at RT with gentle shaking before re-immunoblotting.

10x TBS buffer:	0.2 M Tris-HCl pH 7.5, 1.5 M NaCl.
1x TBST buffer:	10 % (v/v) 10x TBST buffer, 0.1 % Tween 20.
Blocking buffer:	4 % (w/v) milk powder, 15 mM EDTA pH 8.0, in 1x TBST buffer.
Primary antibody solution:	2 % (w/v) milk powder, 15 mM EDTA pH 8.0, the dilution of primary antibody given in Table 1 in 1x TBST buffer.
Secondary antibody solution:	2 % (w/v) milk powder, 15 mM EDTA pH 8.0, anti-rabbit secondary antibody 1:10,000 or anti-mouse 1:40,000 in 1x TBST buffer.

2.2.5.7 *In vivo* labeling of chloroplast proteins

Radioactive pulse labeling of newly synthesized chloroplast proteins was done in tobacco plants at two different developmental stages: three- to four-week-old seedlings and eight to 12-day-old seedlings (cotyledon stage). In both cases, the experiment was performed as follow: 12 leaf discs of 0.5 cm diameter from primary leaves or 25 seedlings of the cotyledon stage were soaked in 400 μ L of labeling buffer supplemented with \sim 440 μ Ci of EasyTag™ EXPRESS³⁵S Protein Labeling Mix (PerkinElmer: 11 mCi/mL, $>$ 1000Ci/mmol; ³⁵S -methionine and cysteine) with 20 μ g/mL cycloheximide to block cytosolic translation. In order to enable the delivery of the radiolabeled amino acids into the cells, the samples were vacuum infiltrated three times for 20 sec using a desiccator (neoLab) followed by incubation for 20 min under \sim 80 μ mol m⁻² s⁻¹ light intensity with care taken that the leaf discs or seedlings do not shade each other. After labeling, the leaf discs and seedlings were washed twice in 1 mL labeling buffer without cycloheximide, dabbed shortly on Whatman paper for drying, flash-frozen in liquid nitrogen, and homogenized as described in section 2.2.5.2. An equal amount of 1,000,000 cpm for soluble proteins and 100,000 cpm for thylakoid proteins, measured with LS 6500 Multi-Purpose Scintillation Counter (Beckman Coulter), were separated on tricine gels and visualized as described in section 2.2.5.5. The gel was then incubated for 1 h in the gel drying solution and vacuum-dried for 45 min at 80 °C using MODEL 583 GEL DRYER, Bio-Rad. FujiFilm phosphor screens (GE healthcare)

were exposed to the dried gel and radiolabeled proteins were visualized by autoradiography using TyphoonTM TRIO+ Variable Mode Imager (Amersham Biosciences).

Gel drying solution: 40 % methanol, 10 % acetic acid, 5 % glycerol.

2.2.5.8 Immunoprecipitation assay of radiolabeled proteins

To visualize newly synthesized core subunits of the cytochrome *b₆f* (cyt *b₆f*) complex, an immunoprecipitation approach was applied. The assay was adapted from Lennartz et al. (2001) and Xiao et al. (2012). In brief, 14 tobacco leaf discs were labeled for 20 min in labeling buffer supplemented with 20 µg/mL cycloheximide and 4 µCi/µL EasyTagTM EXPRESS³⁵S Protein Labeling Mix as described in section 2.2.5.7. After labeling, the leaf discs were homogenized in the frozen state using the Grinding mill Retsch[®] MM301 (Retsch), thawed in 150 µL of soluble protein extraction buffer and centrifuged for 15 min at 13,000 g. The pellet containing the thylakoid membranes was resuspended with 150 µL of soluble protein extraction buffer supplemented with 2 % SDS final concentration. Subsequently, the thylakoid membranes were solubilized by 10 min incubation at 25 °C followed by 30 sec at 70 °C and then quickly centrifuged to remove the tissue debris. For immunoprecipitation, an aliquot of 200,000 cpm of labeled thylakoid proteins was 10 fold diluted with immunoprecipitation buffer and incubated for 2 h with a mix of antibodies containing ~ 26 µg of each. Meanwhile, 90 µL of PierceTM Protein A/G Magnetic Beads (Thermo Scientific Scientific) were washed three times with 300 µL immunoprecipitation buffer. The incubation was continued overnight at 4 °C after which the beads were washed four times with 300 µL of the immunoprecipitation buffer and the immunoprecipitated proteins were eluted by incubation with 2x SDS protein sample buffer at 70 °C for 10 min. The proteins were then size-separated by electrophoresis on a 4 to 20 % (w/v) precast polyacrylamide gradient gel according to Laemmli (1970) as described in section 2.2.5.4. The radiolabeled proteins were visualized by autoradiography as described in section 2.2.5.7.

Immunoprecipitation buffer: 100 mM Tris pH 8, 5 mM EGTA pH 8, 5 mM EDTA pH 8, 150 mM NaCl. Before usage 2 % (v/v) 2-Mercaptoethanol, 2 % (v/v) cOmpleteTM EDTA-free protease inhibitor and 1 % Nonidet P-40 were freshly added.

2.2.6 Expression and purification of the affinity-tagged adapter protein MS2-MBP

2.2.6.1 Bacterial transformation and induction

pMBP-MS2 was a gift from Josep Vilardell (Addgene plasmid # 65104; <http://n2t.net/addgene:65104>; RRID: Addgene_65104). This plasmid was transformed into competent *E.coli* Rosetta (DE3) cells as described in section 2.2.1.3. A single colony was used to inoculate 5 mL LB medium supplemented with 100 µg/mL ampicillin. The culture was grown at 37 °C overnight with shaking and subsequently

used to inoculate 1 L of LB medium (Bertani, 1951) supplemented with 2 % glucose and 100 µg/mL ampicillin till reaching an OD₆₀₀ of 0.5 (Ultrospec[®] 3100). The expression of the MS2-MBP protein was induced for 3 h by adding 1 mL of 1 M IPTG. Subsequently, the cell culture was collected by centrifugation at 5,510 g for 10 min at 4 °C in a Sorvall[®] RC6 centrifuge (rotor: Fiberlite F14-6x 250y). The supernatant was removed and the cell pellet was stored at -20 °C or directly used for fast protein liquid chromatography (FPLC) purification.

2.2.6.2 Protein purification

For purification of the affinity-tag adapter protein MS2-MBP, 1 g of IPTG-induced cell pellet was thawed and resuspended in 10 mL pre-cold AB1 buffer supplemented with 200 µL of cOmplete[™] EDTA-free protease inhibitor. The cells were sonicated on ice with 20 % amplitude for 90 sec using Digital Sonifier[®] W-250D (Brandon Ultrasonic) with the following regime: 3 sec ON, 6 sec OFF. Afterward, the cell debris was removed by 30 min centrifugation at 15,000 g at 4 °C in an Allegra[™] 25R centrifuge (Beckman Coulter). The supernatant, containing the MS2-MBP protein, was then loaded at 0.3 mL/min flow speed on a 5 mL AB1-equilibrated MBPTrap[™] HP amylose column (GE healthcare). The column was then washed with 40 mL AB1 buffer at 5 mL/min followed by 10 mL of AB2 at 2.5 mL/min; this was intended to lower the concentration in preparation for the heparin chromatography. Thereafter the MS2-MBP fusion protein was eluted with 20 mL of ABE buffer at 1.5 mL/min. Based on OD₂₈₀ of the eluted fractions, the peak fractions were pooled, concentrated to about 3.5 mL in a 30K Amicon[®] Ultra-Y 4 mL centrifugal filter (Millipore, Merck). To remove the nucleic acid contaminants, the concentrate was subsequently purified with 1 mL HiTrap[™] Heparin HP column (GE healthcare) previously equilibrated with a mixture of HB1 and HB2 to 20 mM KCl at 1 mL/min. The MS2-MBP recombinant protein was eluted with a 20 to 400 mM KCl gradient, concentrated with a 30K Amicon[®] Ultra-Y 4 mL centrifugal filter, glycerol (10% final concentration) was added and 200 µl aliquots were stored at -80 °C. The FPLC was performed at 4 °C using the ÄKTA explorer (Amersham Biosciences).

AB1 buffer:	20 mM HEPES pH 7.9, 200 mM KCl, 1 mM EDTA pH 8.0.
AB2 buffer:	20 mM HEPES pH 7.9, 20 mM KCl, 1 mM EDTA pH 8.0.
ABE buffer:	20 mM HEPES pH 7.9, 20 mM KCl, 1 mM EDTA pH 8.0, 10 mM maltose.
HB1 buffer:	20 mM HEPES pH 7.9, 1 mM EDTA pH 8.0.
HB2 buffer:	20 mM HEPES pH 7.9, 1M KCl, 1 mM EDTA pH 8.0.

2.2.7 Aptamer-based affinity purification

2.2.7.1 Affinity purification with S1 (streptavidin-binding) aptamer

To remove RNases and other contaminations prior to the experiment, 200 μ L of Dynabeads MyOne™ Streptavidin T1 (Life Technologies) per sample were washed by adding 0.5 mL of the pulldown extraction buffer, mixing, magnetizing and discarding the supernatant for five times. The protocol was adapted from Walker et al. (2008 and Leppek and Stoecklin (2014. Leaves from one transplastomic tobacco seedling expressing tagged *psbA* or *rbcL* with the streptavidin-binding aptamer as well as from control lines (~ 200 mg fresh weight each) were homogenized in liquid nitrogen with mortar and pestle and thawed in 2 mL of pulldown extraction buffer. A 0.1 mL aliquot was removed and flash-frozen for subsequent total RNA isolation. The remainder of the suspension was filtered through glass wool and centrifuged for 5 min at 18,000g and 4 °C to remove cell debris. For RNA and protein follow-up experiments and quality control, 0.2 mL aliquots were kept as input fractions. The remaining lysate was incubated with pre-washed beads for 40 min with slow rotation at 4 °C. Afterward, the beads were washed three times with pulldown extraction buffer, and then the streptavidin-binding RNAs were eluted by incubation with 0.5 mL pulldown extraction buffer supplemented with 5 mM D-biotin (1.22 mg/mL; Sigma-Aldrich) for 10 min at 4 °C with slow rotation. All the steps were carried out on ice and different fractions were collected: input, flow-through, all washings steps, and the elution fractions, were flash-frozen after splitting each into two sub-fractions dedicated for RNA quality/efficiency check and liquid chromatography-mass spectrometry (LC-MS/MS).

Pulldown extraction buffer: 200 mM sucrose, 200 mM KCl, 40 mM Tris-acetate, pH 8.0, 10 mM MgCl₂, 10 mM 2-Mercaptoethanol, 2 % polyoxyethylene tridecyl ether, 1 % Triton X-100, 100 μ g/mL chloramphenicol, and 100 μ g/mL cycloheximide, 2 mg/mL heparin, 2 % (v/v) cOmplete™ EDTA-free protease inhibitor, 0.1 U/ μ L RNasin® RNase Inhibitor.

2.2.7.2 Affinity purification with D8 (Sephadex-binding) aptamer

Sephadex G-200 resins were prepared by swelling 0.2 g Sephadex G-200 in 10 mL of HEPES buffer for two days at RT, resulting in approximately 6 mL of resins. After 5 min centrifugation at 4000 g at 4 °C, the supernatant was discarded and the resin pellet was washed four times with 6 mL of HEPES buffer each. The resins were mixed with HEPES buffer to get a 50 % suspension. The resins were then stored at 4 °C after adding sodium azide to 0.02 % final concentration. Before the pulldown, affinity columns were prepared as follows: 1 mL of the 50 % Sephadex suspension was applied to Bio-Spin® disposable chromatography columns (Bio-Rad), washed seven times each with 1 mL pulldown extraction buffer to remove RNases and other contaminations (section 2.2.7.1). The steps are similar to those of streptavidin pulldown (section 2.2.7.1) with minor differences: after the binding step, the column was washed for five times followed by 10 min incubation with 1 mL pulldown extraction buffer supplemented with 100 mg/mL enzymatically synthesized dextran (Sigma-Aldrich) at 4 °C with slow

rotation. The dextran binds with a high affinity to the Sephadex slurry and thus enables the elution of the Sephadex-bound RNA species. In all steps, the supernatant was drained by gravity flow.

Sephadex swelling buffer: 100 mM HEPES-KOH pH 7.4, 10 mM MgCl₂, 100 mM NaCl, 1 mM DTT, 0.1 % Triton X-100, 10 % glycerol.

2.2.7.3 Affinity purification with the MS2 aptamer

The MS2 affinity purification was adapted from Said et al. (2009 with minor modifications required for the work with tobacco. The affinity columns were prepared as follows: 50 µl of amylose resin (New England Biolabs) was added to Bio-Spin[®] disposable chromatography columns (Bio-Rad) and washed twice with 2 mL of MS2 extraction buffer. Approximately 100 pmol of the purified MS2-MBP fusion protein, diluted in 1 mL of MS2 extraction buffer, was added to the column followed by two washes with 2 mL of MS2 extraction buffer. Plant tissue was homogenized as described in section 2.2.7.1 and the cleared lysate was applied directly to the prepared amylose column, non-covalently coupled to MS2-MBP. Upon drainage of the flow-through and binding of the MS2-tagged RNAs, the column was washed three times with 2 mL of MS2 extraction buffer. Finally, the MS2-tagged RNAs were eluted with MS2 extraction buffer supplemented with 12 mM maltose.

MS2 extraction buffer: 0.2 M sucrose, 0.1 M KCl, 40 mM Tris-OAc pH 8.0, 10 mM MgCl₂, 100 mM 2-Mercaptoethanol, 2 % (v/v) polyoxyethylene (10) tridecyl ether, 1 % (v/v) Triton X-100, 100 µg/mL chloramphenicol, 100 µg/mL cycloheximide, 2 % (v/v) cOmplete[™] EDTA-free protease inhibitor, 0.1 U/µL RNasin[®] RNase Inhibitor.

2.2.8 Mass spectrometry

2.2.8.1 Sample preparation and processing

Mass spectrometry was performed in collaboration with Dr. Frederick Sommer from the University of Kaiserslautern. Protein precipitation, in solution tryptic digest and desalting of the peptides, was adapted and performed as described by (Hammel et al., 2018). Peptide analysis via LC-MS/MS (Eksigent nano-LC 425 coupled to TripleTOF 6600; ABSciex) was performed in information-dependent acquisition (IDA) mode. High-performance liquid chromatography (HPLC) separation was performed in trap-elution mode using ReprosilPur 120 C18-AQ (5 µm particle, 0.15 × 10 mm for trapping and a self-packed 3 µm particle, 75 µm × 200 mm column for separation; Dr Maisch, Germany). A constant flow of 300 nl/min was used and the gradient was ramped within 35 min from 2 % to 35 % of HPLC buffer B, then within 4 min to 50% HPLC buffer B, followed by washing and equilibration steps. The mass spectrometer was run in IDA mode recording one survey scan (250 ms, 350–1250 *m/z*) and fragment spectra (100–1600 *m/z*) of the 30 most intense parent ions (50 ms, charge state > 2, intensity > 100 cps,

rolling collision energy), resulting in a 1.8 sec cycle time. After MS/MS analysis, the selected precursors were excluded for 10 sec from the analysis.

HPLC buffer B 90 % acetonitrile, 0.1 % formic acid.

2.2.8.2 Data analysis

Protein identification and quantification of the raw MS/MS spectra were performed using MaxQuant software (Cox and Mann, 2008) with default settings. Details about the settings used are shown in Supplemental Dataset 1. The peptide fragment spectra were searched against a target database for tobacco (POTanno4) developed and kindly provided by Dr. Michael Tillich (formerly AG Bock, MPIMP). The following parameters were used in the analysis: Protein and peptide false discovery rate (FDR) was set to 0.01 and peptides with at least seven amino acids were considered. The protein intensity output from the MaxQuant analysis was further analyzed using Perseus software (Tyanova et al., 2016). Statistical analysis in Perseus was performed as follows: first, the intensity values were filtered by excluding the proteins identified only by site, matching to the reverse database and the contaminants. After \log_2 transformation, the data from the tagged lines and the controls were grouped separately and the missing values were imputed considering a distribution with a factor of 0.3 (width), shifted down by 1.8 standard deviations. A two-sample t-test was performed with 0.05 FDR and 2 fold change as thresholds.

2.2.9 Transmission electron microscopy (TEM)

TEM was used to examine the suborganellar localization of MS2-tagged *psbA* and *rbcL* transcripts. Leaves from two-week-old seedlings grown under standard conditions (section 2.2.1.4) were used. Sample preparation and microscopy studies were performed by Anja Froehlich and Dr. Arun Sampathkumar (both MPIMP) following the protocol in McFarlane et al. (2008) with minor modifications. In brief, samples for TEM were processed following five steps: cryofixation, freeze substitution, resin embedding, sectioning and post-staining, and imaging. After dissection, the samples were high-pressure frozen in copper type 'B' sample carriers filled with hexadecane using a High Pressure Freezing Machine HPM 100 (Leica Microsystems), followed by incubation with the freeze-substitution solution. The samples were then embedded with LR White Resin (London Resin Company) and sectioned using a Leica Ultracut UCT Ultramicrotome (Leica Microsystems) with a Diatome diamond knife. Following sectioning, the samples were placed on nickel grids and incubated in blocking solution for 30 min at RT followed by three washings with TEM buffer. Immunolabelling was performed by the following incubations: (1) 1 h with the MS2-MBP fusion protein (1 $\mu\text{g}/\mu\text{l}$) 1:5 diluted in TEM buffer supplemented with 1 % BSA, (2) 1 h with the primary antibody (Anti-MBP antibody, ab9084, abcam) 1:10 diluted in TEM buffer, 1 % BSA, (3) 1 h with gold conjugate (Goat anti-Rabbit

IgG 15nm Gold, BBI solutions) diluted in TEM buffer, 1 % BSA to 1:100 ratio. All incubations were done at RT and each step was followed by thorough washings with TEM buffer. Subsequently, the grids were post-stained in post-staining buffer for 10 min followed by 4 min incubation with lead citrate. Finally, the samples were examined using TEM 912 Omega (ZEISS) at 120 kV.

Freeze substitution solution: 0.25 % glutaraldehyde, 0.1 % uranyl acetate, and 8 % dimethoxypropane in anhydrous acetone.

TEM buffer: 0.02 M Tris, 0.5 M NaCl, 0.2 % Tween 20.

Blocking solution: 5 % BSA in TEM buffer.

Post-staining buffer: 2 % uranyl acetate in 50 % ethanol.

3 Results

3.1 Identification of assembly-dependent translational feedback regulation in photosynthetic complexes of embryophytes

3.1.1 Selection of mutants with assembly defects in photosynthetic complexes

In order to reveal potential translational feedback regulation in the assembly of photosynthetic complexes in embryophytes, several existing *Arabidopsis* and tobacco (*Nicotiana tabacum*) mutants with assembly defects were selected for each complex: photosystem II (PSII), cytochrome *b₆f* (Cyt *b₆f*), photosystem I (PSI), ATP synthase, and NAD(P)H dehydrogenase-like (NDH) complexes as well as ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) (Table 5). In order to avoid secondary effects caused by the knockout of photosynthesis, whenever possible, weak alleles of photosynthetic mutants that are able to grow photoautotrophically were selected. At least one weak mutant allele for each of the four core photosynthetic complexes (PSII, Cyt *b₆f*, PSI, and ATP synthase) was analyzed in this work. Only three mutants that knockout photosynthesis and the ability for autotrophic growth were included. The mutants employed in this study belong to two different groups: (i) mutants of the non-essential assembly factors PsbN and Ycf4, that cause defects in the assembly of PSII and PSI, respectively, were selected (Krech et al., 2013; Krech et al., 2012); (ii) mutants of specific photosynthetic subunits which cause assembly defects to different degrees (Table 5). Most of the chloroplast mutants were created by insertion of an *aadA* cassette either within or upstream of the ORF of the target gene. The *aadA* gene encodes the enzyme aminoglycoside 3'-adenylyltransferase, which confers resistance to spectinomycin and streptomycin and as such serves as a selectable marker in transformed chloroplasts (Goldschmidt-Clermont, 1991; Svab and Maliga, 1993).

Chloroplast gene expression peaks early in plant development and diurnally after the onset of the light phase (Chotewutmontri and Barkan, 2016). Hence, plant material used in this study was harvested shortly after the start of the illumination period. All the tobacco weak mutant alleles were harvested at the three-week-old developmental stage of their corresponding control (four-leaf stage). On the other hand, mutants that knockout photosynthesis were grown heterotrophically *in vitro* till they reached the size of a four-leaf old tobacco seedling. *Arabidopsis* mutants were harvested just before bolting. In specific cases, I analyzed in addition some of the tobacco mutants at cotyledon stage (before the emergence of the true leaves).

To analyze chloroplast translation in the selected tobacco and *Arabidopsis* mutants, a chloroplast-targeted ribosome profiling approach was used (Zoschke et al., 2013; Trösch et al., 2018). This targeted approach enabled a fast, quantitative, chloroplast genome-wide study of translation. Sequencing-based ribosome profiling (Ingolia et al., 2009) was used to analyze translation in the mutants that knockout photosynthesis due to the small amount of ribosome footprints that could be

recovered in these mutants. As a confirmatory approach of the CES interaction identified by ribosome profiling, *in vivo* pulse experiments were used.

Table 5: Analyzed tobacco and *Arabidopsis* mutants

Mutants of subunits and assembly factors of the photosynthetic apparatus in embryophytes and their essentiality for autotrophic growth (nucleus-encoded subunits are underlined).

Complex	Mutant	Type of mutation	Gene essential for autotrophic growth	Reference
PSI	$\Delta ycf4$	insertion knockout	-	(Krech et al., 2012)
	KD- <i>psaA</i>	Shine-Dalgarno mutation (knockdown)	+	Unpublished
	<i>psad1-1</i>	T-DNA insertion	-	(Ihnatowicz et al., 2004)
PSII	$\Delta psbN$	frameshift (knockout)	-	(Krech et al., 2013)
	<i>psbD</i> -GTG	start codon exchange	+	(Fu, 2012)
	<i>psbD</i> -TTG	start codon exchange	+	(Fu, 2012)
	KD- <i>psbD</i>	5' UTR-insertion transcriptional knockdown	+	(Fu, 2012)
	<u><i>hcf111-1</i></u>	EMS mutation	-	(Météignier et al., 2020)
	<u><i>hcf173-2</i></u>	T-DNA insertion	+	(Schult et al., 2007)
	$\Delta psbD/C$	Insertion knockout	+	(Hager, 2002)
Cyt <i>b₆</i>	$\Delta petL$	insertion knockout	-	(Schöttler et al., 2007)
	$\Delta psal$	insertion knockout	-	(Schöttler et al., 2017)
	$\Delta ycf10$	insertion knockout	-	Unpublished
	$\Delta psbB$ operon	Insertion knockout	+	(Hager, 2002)
ATP synthase	<i>atpB</i> -GTG	start codon exchange	+	(Rott et al., 2011)
	<u>as-<i>AtpC</i></u>	antisense- RNA-induced	+	(Rott et al., 2011)
	$\Delta atpB$	insertion knockout	+	(Hager, 2002)
	<u>as-<i>RBCS</i></u>	<i>RBCS</i> antisense-RNA	+	(Rodermeil et al., 1988)

Rubisco	<i>mrl1-1</i>	T-DNA insertion	-	(Johnson et al., 2010)
	<i>rbcsl1a3b-1</i>	T-DNA insertion	+	(Izumi et al., 2012)
NDH	<i>ndhC/K/J</i>	insertion knockout	-	(Hager, 2002)
	<i>ndhA/H/I</i>	insertion knockout	-	(Kofer et al., 1998)

3.1.2 Analysis of feedback regulation in chloroplast gene expression

Chloroplast genome-wide translation of each mutant and its corresponding control (Table 5) was profiled. Given that each ribosome footprint is generated by one translating ribosome which, ultimately, produces one protein, the footprint abundance per ORF (translation output) provides an estimation of protein synthesis levels. The signal intensities of the probes covering the protein-coding regions were normalized as described in section 2.2.3.4. The reproducibility between the biological replicates on probe level is shown in Supplemental Table 4 (Pearson's correlation R values of $\sim 70\%$ of all the replicates ≥ 0.93). The average of the probe signals was calculated for each chloroplast ORF to represent the translation output per gene (section 2.2.3.4) and showed high reproducibility between the biological replicates (Pearson's correlation R values of $\sim 90\%$ of all the replicates ≥ 0.93 ; Supplemental Table 2). The relative chloroplast translation output of each ORF was compared between mutants and corresponding controls. This enabled us to detect genes whose translation output is substantially affected by the mutation (i.e., more than twofold change in translation output). The statistical significance of changes in the translation output, transcript accumulation, and ribosome redistribution was assessed using the empirical Bayes methods in the limma package (section 2.2.3.4) (Smyth, 2004). False discovery rate (FDR) method was used for P value adjustments with 0.05 as a threshold to define significance (Benjamini and Hochberg, 1995).

The ribosome footprint data was also used to investigate how the assembly defect affects the translation elongation of the chloroplast genes. Ribosomes redistribution in chloroplast reading frames was examined for all the mutants with three biological replicates (as described in section 2.2.3.4). Only the data of the mutants that exhibited a significant alterations (according to the thresholds defined) in ribosome occupancy is shown (see below).

Translation output is not only determined by the translational activity but is also affected by the transcript abundance. To monitor the changes on transcript level and assess the extent to which the transcript changes influence the changes in the translation output, the transcript abundance was profiled using the same plant material and microarrays used for ribosome profiling (section 2.2.3.2). The data were processed as described above for ribosome footprint data. Biological replicates showed very high reproducibility for the normalized probes signal intensities (Pearson's correlation R values of $\sim 80\%$

of all the replicates ≥ 0.93 ; Supplemental Table 3) and the average of ORF transcript abundance (Pearson's correlation R values of 95 % of all the replicates ≥ 0.94 ; Supplemental Table 1).

3.1.3 Confirmation of the only known CES in embryophytes: translation of RbcL is regulated by the availability of its assembly partner, RbcS

Carbon fixation in photosynthetic organisms depends on the first enzyme in the Calvin-Benson-Bassham cycle, Ribulose-1,5-Bisphosphate Carboxylase Oxygenase (Rubisco). In embryophytes, Rubisco is a hexadecameric complex (Andersson and Backlund, 2008) with eight large subunits, encoded by the chloroplast *rbcL* gene, and eight small subunits, encoded by the nuclear *RBCS* gene family. It was previously reported for tobacco that the large subunit of Rubisco obeys the classical CES paradigm in that its translation decreases in the absence of the small subunit of Rubisco (Rodermel et al., 1996; Wostrikoff and Stern, 2007).

In order to validate the ability of the targeted ribosome profiling to identify translational feedback regulation, transgenic tobacco plants expressing antisense RNA for the small subunit of Rubisco (as-*RBCS*) were used (Rodermel et al., 1988) (Figure 3.1A). The mutant was grown side by side to control plants (Petit Havana SR1 cultivar) in standard conditions (section 2.2.1.8) (Figure 3.1A). The aerial part of the control was harvested after three weeks of growth. Since the mutant displayed growth retardation, it was further grown for three days until it reached the same developmental stage as the control. It was reported for this antisense mutant that the *rbcS* mRNA level is tenfold decreased (down to 12% to that in the SR1 control) (Rodermel et al., 1988). Even though the mRNA level of *rbcL* did not change, the protein level of RbcL decreased to about 38% of the control (Rodermel et al., 1988). Such a discrepancy between the mRNA and protein changes suggested that the translation might be the step that regulates the accumulation of RbcL in regard to its assembly state. In this work, I revisited the impact of reduced *RBCS* accumulation on the transcript and footprint abundance of *rbcL* and protein level of RbcL. Indeed, the immunoblot analysis showed a strong decrease in the protein accumulation of RbcL (Figure 3.1B) validating the previously observed defect (Rodermel et al., 1988). The accumulation of photosynthetic complexes involved in the light reactions was examined by assessing the protein level of one core subunit per complex. None of these complexes was found to be substantially altered (Figure 3.1B). Transcript and ribosome profiling was performed for the mutant and its control and the transcript abundance and translation output (section 3.1.2) of each chloroplast-encoded ORF were calculated. The translation output of *rbcL* decreased significantly by about 11-fold (Figure 3.1C). In contrast, the transcript level of *rbcL* was decreased only by approximately threefold (Figure 3.1C), which indicates that the transcript change cannot account for the massive reduction observed in translation output and protein accumulation of RbcL. This finding substantiates what was previously highlighted by Wostrikoff and Stern (2007). With this result, I confirmed the only known CES regulation in land plant chloroplasts and, importantly, demonstrated that the targeted ribosome profiling is well suited for a

screen for CES feedback regulation. In addition to *rbcL*, transcript accumulation of *psbA* and *psbL* decreased significantly while that of *ndhI* showed a significant increase (Figure 3.1C). Furthermore, the translation output of *ndhJ* and *psbN* significantly increased (Figure 3.1C). It is possible that these changes (Figure 3.1C) are caused by off-target effects of the antisense RNA in these plants or represent secondary effects due to the strong defects of Rubisco.

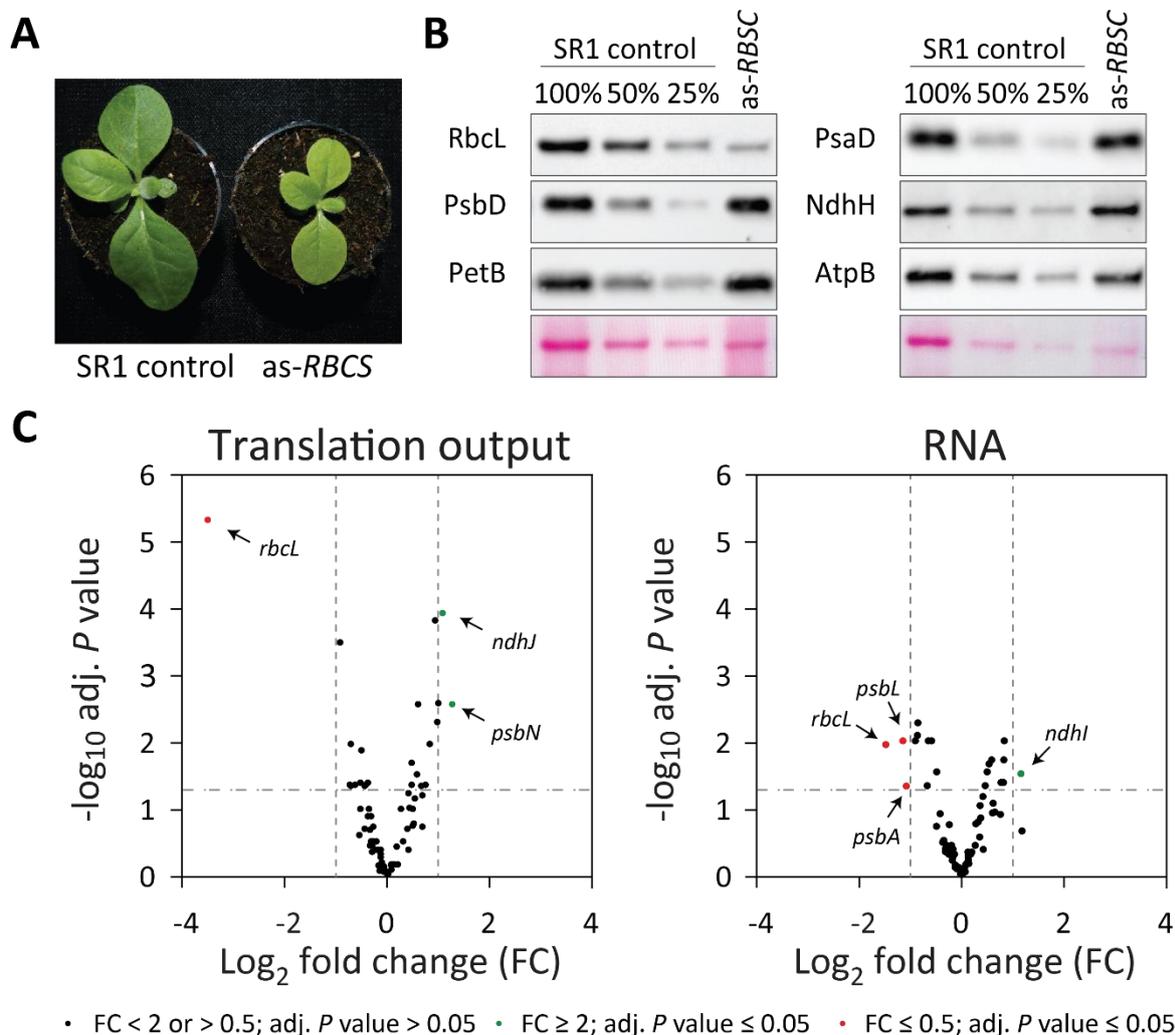


Figure 3.1: Reduced levels of RBCS provoke a decrease in the transcript and ribosome footprint abundances of *rbcL* and a defect in the protein accumulation of RbcL.

A. as-RBCS mutant was created by transforming the diploid *Nicotiana tabacum* Petit Havana, SR1 plants (Maliga et al., 1973) with the antisense DNA construct. SR1 plants were used as control and were grown for three weeks under standard conditions (section 2.2.1.8) next to the as-RBCS mutant. The aerial part was harvested and used for total RNA, ribosome footprint, and protein isolations. **B.** Total proteins from as-RBCS and SR1 control were separated by SDS-PAGE. Representative immunoblot analysis of photosynthetic core subunits (labeled on the left) was performed to assess the accumulation of photosynthetic complexes (three replicate immunoblots showed similar results). Note the strong reduction of RbcL in the as-RBCS mutant in the absence of substantial alterations for the other tested complexes. The results were obtained from three biological replicates, one representative replicate is shown. **C.** Changes in the translation output (left) and transcript abundance (right) of all the chloroplast ORFs represented in volcano plots. The average ribosome footprint and transcript abundances per ORF was calculated. Fold change ratios of mutant compared to control values were log₂-transformed and

plotted against the adjusted P values. Genes whose expression is at least twofold downregulated with an adjusted P value ≤ 0.05 are labeled in red. Genes whose expression is at least twofold upregulated with an adjusted P value ≤ 0.05 are labeled in green. The vertical dashed lines represent the fold change cutoff (twofold). The horizontal dashed line represents the adjusted P value threshold (0.05).

3.1.4 The translational feedback regulation of *rbcL* occurs at the initiation level

Ribosome profiling provides a snapshot of ribosome abundances and distribution on transcripts. However, the method cannot distinguish between actively translating and pausing or stalling ribosomes. To acquire further confirmation that translation is the post-transcriptional event that regulates the accumulation of RbcL, and to distinguish initiation and elongation regulation, an analysis of *de-novo* synthesized chloroplast proteins was performed. To this end, proteins were labeled *in vivo* (in leaf discs) by a 20 minute “pulse” with ^{35}S -radiolabeled amino acids (methionine and cysteine) (section 2.2.5.7). The cytoplasmic translation was blocked with cycloheximide to simplify the interpretation of the pattern of labeled proteins. Consistent with the ribosome profiling results (Figure 3.1D), the *as-RBCS* mutant showed a clear reduction in the newly synthesized RbcL compared to the SR1 control (Figure 3.2A). This confirms that ribosome footprint levels provide a good indicator of protein synthesis levels for the vast majority of genes whose translation output is regulated at the level of translation initiation. As explained in section 2.2.5.7, equally labeled soluble proteins (1,000,000 cpm) were separated on tricine gels. However, it should be noted that the *as-RBCS* mutant has a defect in the synthesis of RbcL, the most abundant and highly expressed protein in plants (Ellis, 1979). Hence, if RbcL synthesis is reduced, equal loading based on labeled protein is difficult, which becomes apparent by the higher background in the mutant lane (indicating overloading of all labeled proteins) (Figure 3.2A). Consequently, the RbcL signal was quantified and normalized to the background signal. Following this strategy, the synthesis of RbcL in the *as-RBCS* mutant was reduced to 40 % of that in the SR1 control (Figure 3.2B), validating the CES regulation of RbcL.

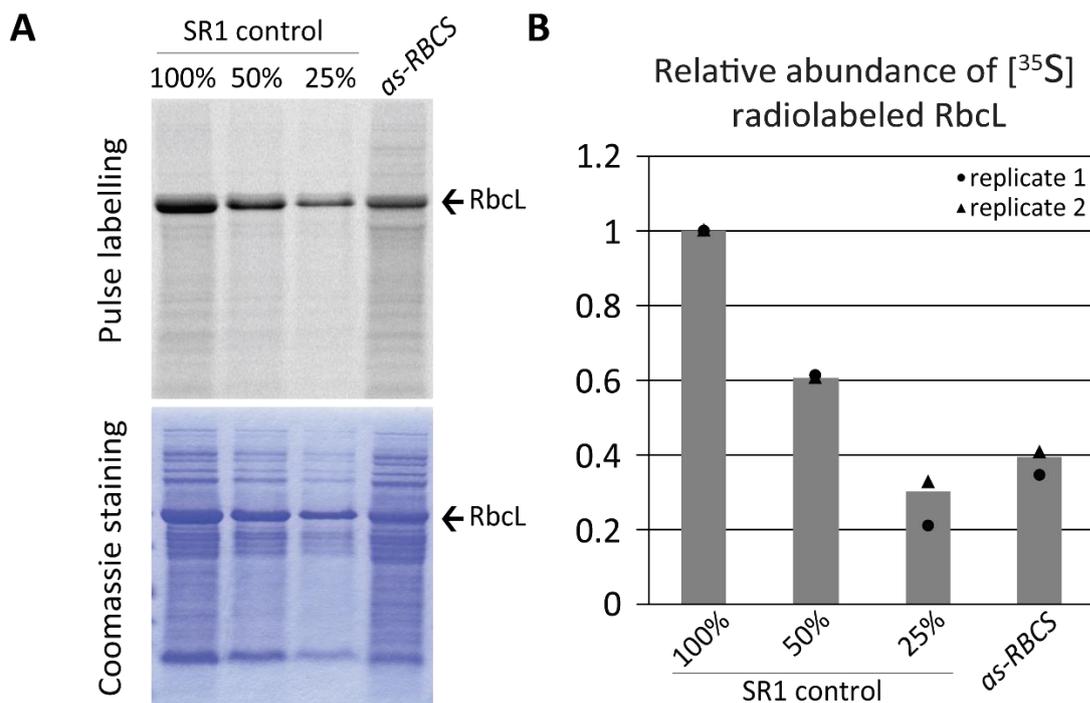


Figure 3.2: Pulse labeling of as-RBCS mutant validates the CES regulation of *rbcL*.

A. Leaf punches were pulse-labeled for 20 min with a ³⁵S labeling mix (section 2.2.5.7). Total soluble proteins were isolated and separated on a tricine gel. The top image shows a representative autoradiograph of the newly synthesized proteins, including RbcL, the bottom image shows the respective Coomassie staining of all proteins. RbcL is marked by an arrow. **B.** The band intensities of RbcL in the autoradiograph in A were quantified using Image Lab (Bio-Rad) and normalized to the background of the lane. The results were obtained from two biological replicates.

3.1.5 MRL1 is a potential translation factor involved in the CES regulation of *rbcL*

The main finding inferred by Wostrikoff and Stern (2007) was that the translation of *rbcL* is not transactivated by its assembly partner (RBCS) but rather *rbcL* undergoes autoregulation of translation, which could be either direct (i.e., by *rbcL* mRNA-binding of RbcL as suggested by (Wostrikoff and Stern, 2007)) or mediated by another factor. It has been shown that in *Chlamydomonas*, MRL1 binds the 5' untranslated region (UTR) of *rbcL* in a high molecular mass complex located in the stroma (Johnson et al., 2010). Based on that, I hypothesized that MRL1, a nuclear-encoded protein that stabilizes the *rbcL* mRNA in *Chlamydomonas* and *Arabidopsis* (Johnson et al., 2010) could be a candidate factor that is involved in *rbcL* feedback regulation. To test this hypothesis, I collaborated with Dr. Hannes Ruwe from Humboldt University, Berlin. A double T-DNA insertion mutant of RBCS1A and RBCS3B subunits (designated as *rbcs1a3b-1*) (Izumi et al., 2012) was crossed with the T-DNA insertion mutant *mrl1-1* (Johnson et al., 2010) (Figure 3.3A). In the double RBCS mutant (*rbcs1a3b-1*), two out of the four *Arabidopsis* RBCS genes were knocked out which resulted in a reduction of the RBCS mRNA to 23 % of the wild-type level (Izumi et al., 2012). A preliminary ribosome profiling analysis of chloroplast translation was performed for the single *mrl1-1* mutant, the double *rbcs1a3b-1* mutant, and the triple *rbcs1a3b-1* x *mrl1-1* mutant. For all three mutants, wild-type Col-0 was used as

control. The translation output of *rbcL* decreased about fourfold in *rbcS1a3b-1* (Figure 3.3B), which is consistent with the published extent of *RBCS* knockdown in this mutant and substantiates the CES regulation of *rbcL*. In accordance with our hypothesis, the translation output of *rbcL* decreased by ~3.5-fold in *mrl1-1* indicating that MRL1 is a translational activator of *rbcL*. Remarkably, in the triple *rbcS1a3b-1* x *mrl1-1* mutant, the translation defect of *rbcL* was not additive, and with a 2.5-fold reduction compared to the wild type, even less pronounced than in *mrl1-1* or *rbcS1a3b-1* alone (Figure 3.3B). These preliminary results suggest a model in which MRL1 acts as a translational enhancer that is repressed by the unassembled RbcL. In this case, in the absence of its assembly partner RbcS, the interaction of unassembled RbcL with MRL1 would repress the activation of *rbcL* translation (Figure 3.3C). However, due to time-consuming crossings to obtain a homozygous triple mutant, these results were obtained at the end of my Ph.D. from a single ribosome profiling experiment and consequently do not allow firm conclusions. More replicates (including transcript profiles and pulse labeling experiments) are needed to validate this model.

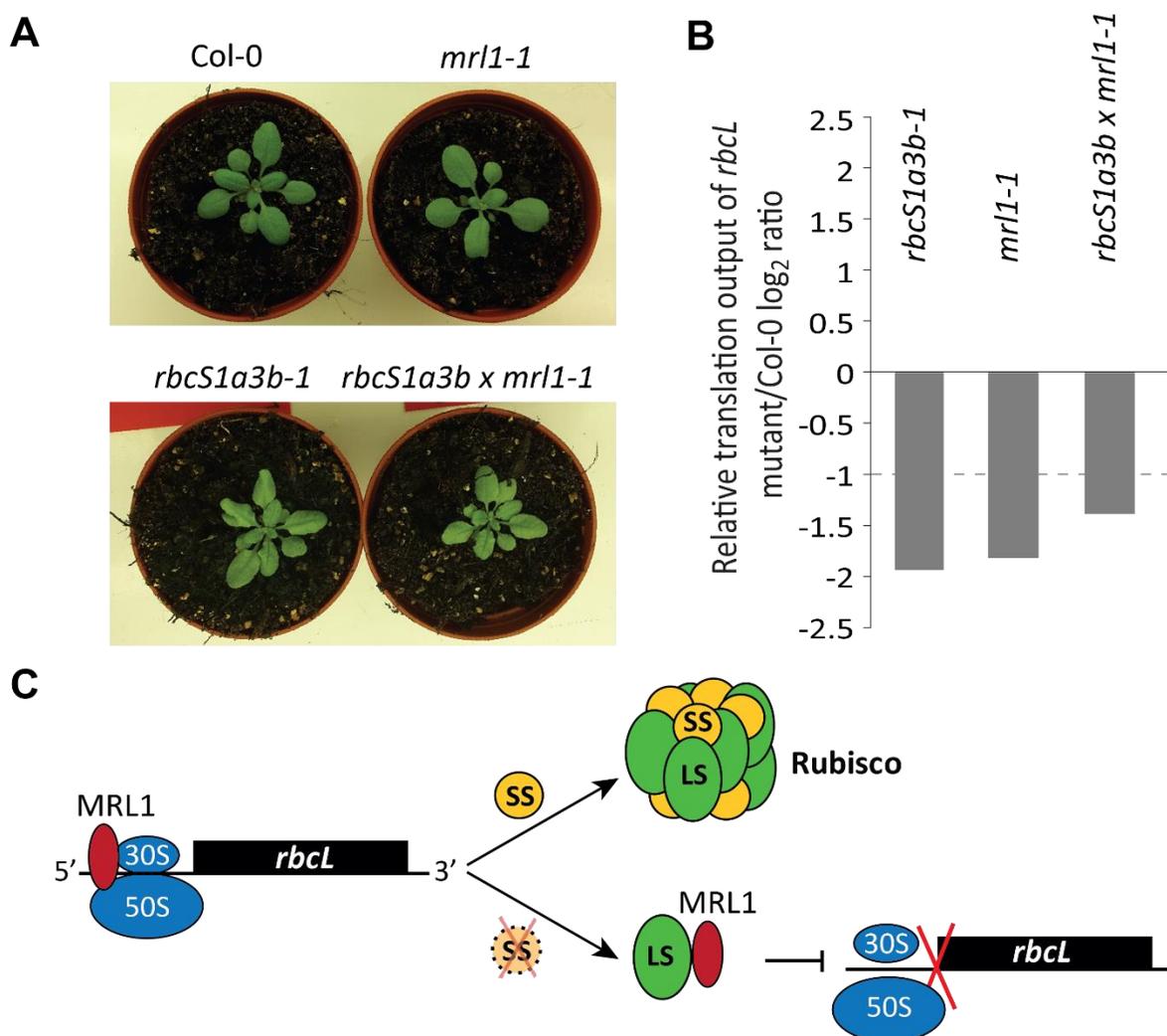


Figure 3.3: MRL1 is a potential mediator of *rbcL* CES regulation.

A. Phenotype of the harvested material from the wild type (Col-0) and *mrl1-1*, *rbcS1a3b-1*, and *rbcS1a3b-1 x mrl1-1* mutants and the aerial part used for ribosome profiling. **B.** Bar graph plots of the translation output of *rbcL* in each of the three mutants in comparison to Col-0. The results were obtained from one biological replicate. **C.** Hypothetic model for the CES regulation of *rbcL* based on data shown in B. In the wild type, MRL1 enables the efficient translation of *rbcL*. In plants with reduced levels of RBCS, Rubisco assembly is impaired and unassembled RbcL interacts with MRL1. This interaction blocks the MRL1-mediated translation activation of *rbcL*, thereby mediating an assembly-dependent translational feedback regulation.

3.1.6 Identification of potential CES regulation in PSII

PSII is a giant complex of numerous nucleus- and plastid-encoded proteins, cofactors, and lipids whose assembly is highly ordered and coordinated (Nickelsen and Rengstl, 2013). In *Chlamydomonas*, the PSII assembly is regulated by a CES cascade in which the lack of PsbD causes a reduction of the PsbA synthesis that leads to the reduction of PsbB synthesis (Choquet and Wollman, 2009; Minai et al., 2006) (Figure 1.3). PSII core subunits and their mode of assembly are highly conserved between cyanobacteria, algae, and embryophytes (Nickelsen and Rengstl, 2013) raising the question if regulatory CES mechanisms exist in embryophytes as well. To address the question if CES-like mechanisms regulate the synthesis of PSII subunits in embryophytes, I selected different tobacco and *Arabidopsis* mutants of PSII core subunits and associated factors that interrupt PSII assembly at different stages (see Table 5). Mutants that affect the synthesis of PSbE (*hcf111-1*), PsbD (*psbD-GTG*, *psbD-TTG*, *KD-psbD*, Δ *psbD/C*), PsbA (*hcf173-2*) and PsbB/H (Δ *psbB* operon) were analyzed in this work.

PsbD is placed at the origin of PSII assembly (Nickelsen and Rengstl, 2013). Hence, I chose mutants with a defect in *psbD* translation for the analysis (kindly provided by Dr. Mark Aurel Schöttler, MPIMP). In these mutants, the start codon of *psbD* was mutated from ATG into either GTG or TTG. In addition, an *aadA* selection marker cassette was inserted upstream of *psbD* (Figure 3.4A, C). Surprisingly, both mutant lines were phenotypically indistinguishable from wild-type plants (data not shown). Ribosome and transcript profiling was performed to investigate the effects of *psbD* start codon mutation on the expression and stoichiometric synthesis of the plastid-encoded PSII subunits. Transcriptome analyses revealed in both mutant lines a drastic increase in the transcript levels of *psbD*, *psbC*, and *psbZ* (Figure 3.4B, D), most likely as a consequence of read-through transcription from the upstream *aadA* cassette, a phenomenon that has previously been reported (Krech et al., 2012). Interestingly, in the *psbD-GTG* line, the increase in *psbD* transcript accumulation caused an increase in translation output (Figure 3.4D) thus the translation efficiency remained unchanged, which indicates that GTG can be efficiently used as a start codon in the *psbD* context. In contrast, in the *psbD-TTG* line, the translation output of *psbD* was reduced leading to a fourfold decrease in translation efficiency (ratio of translation output to transcript abundance). This result implies that *psbD* translation cannot efficiently initiate at a TTG start codon. However, the increase in transcript level partially compensated the defect in the translation of *psbD* (Figure 3.4B). Interestingly, the translation defect in the *psbD-TTG* mutant

did not cause an impaired translation of the downstream *psbC* reading frame which overlaps with *psbD* (Figure 3.4B). This finding suggests that the translation of the downstream *psbC* reading frame is, if at all, only partially coupled to the translation of the upstream *psbD* reading frame, a conclusion which may be not surprising taking into consideration that *psbC* is translated from a dicistronic and a monocistronic transcript isoform (Figure 3.4A, C). Interestingly, *psbK*, encoding a small PSII subunit, showed a significant decrease in footprint level (~ threefold in *psbD*-GTG and twofold in *psbD*-TTG). Although the function of PsbK has not been characterized in embryophytes, it was shown that it is non-essential in cyanobacteria and its deletion has only minor effects on growth and PSII function (Shi and Schröder, 2004; Zhang et al., 1993), whereas, in *Chlamydomonas*, PsbK is essential for autotrophic growth (Shi and Schröder, 2004; Takahashi et al., 1994). This divergence between *Chlamydomonas* and cyanobacteria indicates the different stability requirements of PSII in these two organisms. Yet, it is difficult to explain these drastic differences, especially that in both algae and cyanobacteria PsbK has a similar location next to CP43 (Nickelsen and Rengstl, 2013). Consequences of the altered PsbK translation on PSII accumulation and function cannot be directly addressed given that no *psbK* mutant has been characterized in land plants and no antibody is available. Taken together the impaired translation initiation of *psbD* in lines with mutated start codons was compensated by overexpression on RNA level. Hence, these lines did not provide an adequate study subject to analyze the potential PsbD-dependent CES mechanism in PSII.

To comprehensively investigate the effect of manipulating the translation initiation of *psbD* on the other subunits of PSII, the *psbD*-TTG mutant was compared against a control line where an *aadA* cassette was inserted exactly at the same position upstream of *psbD* while the start codon of *psbD* was kept unchanged (ATG) (referred to as *s-aadA-psbD* control). Indeed, such comparison excluded the overexpression on the transcript level that was caused by the read-through of *aadA* and enabled us to assess exclusively the effect of the start codon mutation. As expected, TTG cannot serve as an efficient start codon as reflected by the observed decrease of the translation of *psbD* (Figure 3.4E). Surprisingly, the translation output of *psbA* and *psbB* decreased by approximately 3.5-fold and twofold, respectively (Figure 3.4E, F). This may indicate that PsbA and PsbB are potential CES subunits whose translation depends on the availability of PsbD. In addition, the translation output of *psbC* decreased about twofold, which is much weaker than the effect on *psbD* and points to a partial translation coupling between the overlapping reading frames of *psbD* and *psbC* *in vivo*, a conclusion that has previously been drawn based on *in vitro* translation experiments (Adachi et al., 2012).

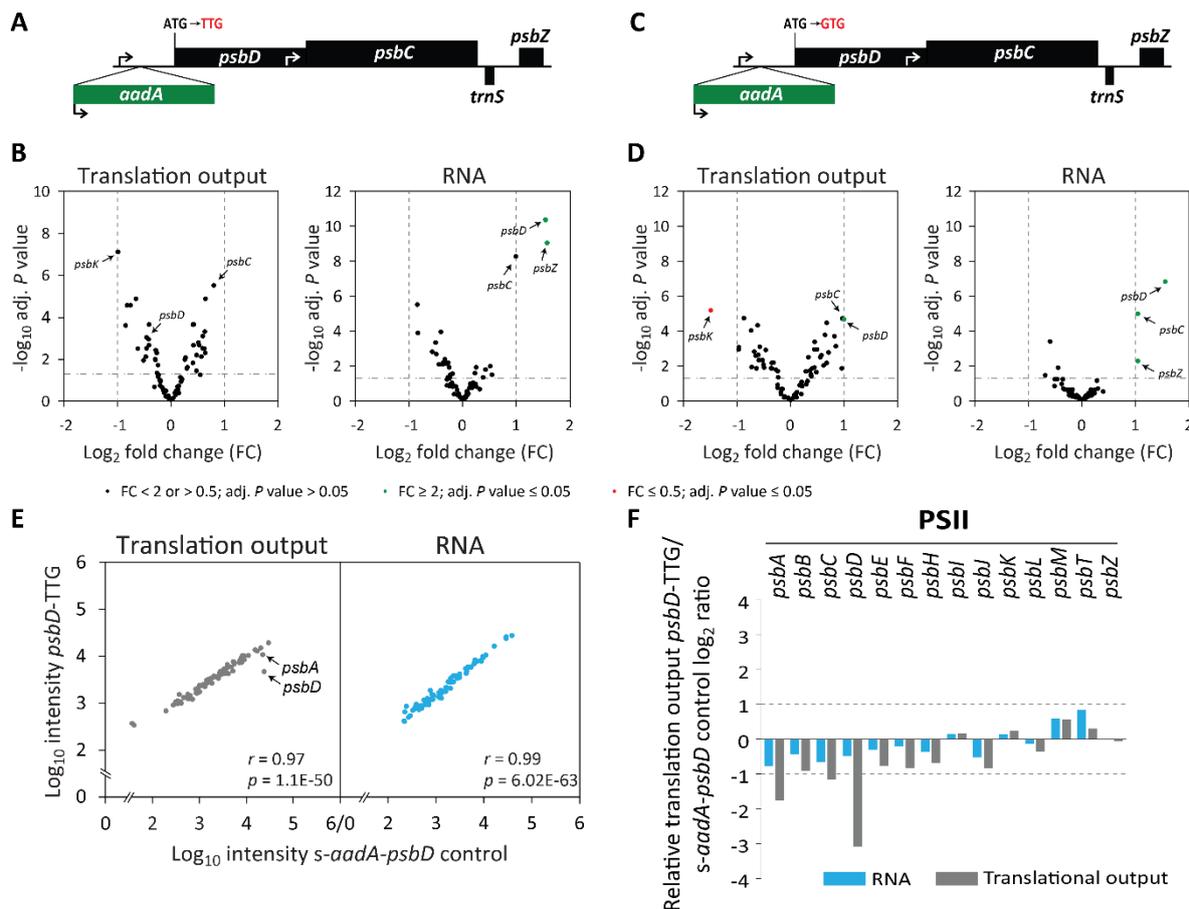


Figure 3.4: Transcriptome and ribosome footprint profiling of *psbD* translation mutants reveals potential CES regulation of *psbA*.

A, C. Physical gene map of the mutants containing the selection marker inserted upstream of *psbD* start codon. Black boxes: chloroplast protein-coding regions; green box: *aadA* cassette; black and white arrows: promoters of the dicistronic and monocistronic transcripts, respectively. The start codon of *psbD* was mutated from ATG to TTG (**A**) or GTG (**C**). **B, D.** Changes in the translation output (left) and transcript abundance (right) of all the chloroplast ORFs in *psbD*-TTG mutant (**B**) and *psbD*-GTG mutant (**D**). Wild type was used as control. Data is represented in volcano plots. Refer to Figure 3.1D for labeling details. Data was obtained from three biological replicates. **E.** Comparison of the translation output and transcript abundance for each chloroplast gene in *psbD*-TTG in comparison to *s-aadA-psbD* control. Results were obtained from one biological replicate. The average values per ORF are plotted. For better visualization, the x- and y- axes are broken. Pearson's r value and Anova's P value are shown for each plot. **F.** Ratio of the translation output and the transcript abundance in *psbD*-TTG mutant relative to *s-aadA-psbD* control for chloroplast-encoded PSII subunits.

3.1.7 Evidence for a PSII CES cascade in embryophytes

3.1.7.1 The translation of *psbB* is downregulated when PsbD/C synthesis is reduced

To further investigate the presence of CES regulation that might tune the stoichiometry of PSII subunits, a knockdown mutant of *psbD* (referred to as *KD-psbD*) was included in this work. In this mutant the *aadA* cassette was inserted in antisense orientation between the *psbD* promoter and start codon, thereby disrupting transcription and/or transcript accumulation (Figure 3.5A). The *KD-psbD* mutant displayed

a pale phenotype with thin leaves and retarded growth in comparison to the control (Figure 3.5B). To take the burden of expressing a transgene into consideration, plants containing the *aadA* resistance marker at a neutral position between the *petA* and *psbE* genes were used as control (line pRB8-S5 in Bock et al. (1994) referred to in this work as pRB8 control). By transcript and ribosome profiling, the transcript level and translation output of the chloroplast-encoded genes were compared in KD-*psbD* to those of the pRB8 control (Figure 3.5C). The transcript profiling showed that the *aadA* insertion caused a 2.5-fold decrease in the transcript accumulation of *psbD* that, at least partially, explains the reduced PsbD translation output (Figure 3.5C). Interestingly, the *psbD* translation output of the residual transcript in the KD-*psbD* line is even stronger reduced than expected from its transcript accumulation, leading to a ~ 3.5-fold reduction of translation efficiency. Hence, it can be concluded that the *aadA* insertion also interferes with *psbD* translation initiation in the reduced *psbD* transcript pool. Additionally, the transcript accumulation of *psbC* was affected, although to a lesser extent than that of *psbD*. This can be explained by co-transcription and accumulation of a dicistronic *psbD/C* transcript. The attenuated effect on *psbC* transcript accumulation is likely to be caused by the internal promoter in the *psbD*-coding region that produces monocistronic *psbC* transcripts (Yao et al., 1989). The defect in the transcript accumulation of both genes was accompanied by an even stronger reduction in their footprint levels (Figure 3.5C), which was explained for *psbD* above and maybe caused for *psbC* by partial translational coupling in the dicistronic transcript (section 3.1.6). In addition, in PSII, the footprint level of *psbB* was decreased more than twofold (Figure 3.5C). This result may indicate that PsbB, similar to the situation in *Chlamydomonas*, is a PSII CES subunit whose synthesis is controlled by the availability of PsbD (Minai et al., 2006; Trösch et al., 2018). However, a very modest effect was observed on the footprint abundance of *psbA*. Based on the *Chlamydomonas* PSII CES, a stronger reduction of PsbA synthesis would be expected upon *psbD* knockdown, since PsbA is recruited upstream of PsbB but downstream of PsbD in the hierarchical assembly of PSII (Nickelsen and Rengstl, 2013) (Figure 1.3). Besides changes in the expression of PSII subunits, the translation output of the photosystem I (PSI) components (*psaA* and *psaB*) was significantly lower which is likely attributable to the defect in PSII as previously reported for other mutants (Swiatek et al., 2003; Trösch et al., 2018). The ribosome profiling data also showed that the translation output of *rps14* and *atpH* is significantly compromised. The defect in *rps14* translation output could be either a direct regulatory effect or due to the fact that it resides in a polycistronic transcription unit with *psaA* and *psaB*, indicating that the reduced ribosome loading of the upstream *psaA/B* reading frames causes a reduced translation of *rps14*, presumably by translational coupling (Rex et al., 1994). The reduction in *atpH* translation can neither be explained by a direct effect nor does it have substantial consequences for the accumulation of the ATP synthase complex (communication with Dr. Mark Aurel Schöttler, MPIMP). Furthermore, the translation output of *ycf10*, *matK*, and *petB* increased significantly, which probably represent the indirect effects of the strong defect in PSII.

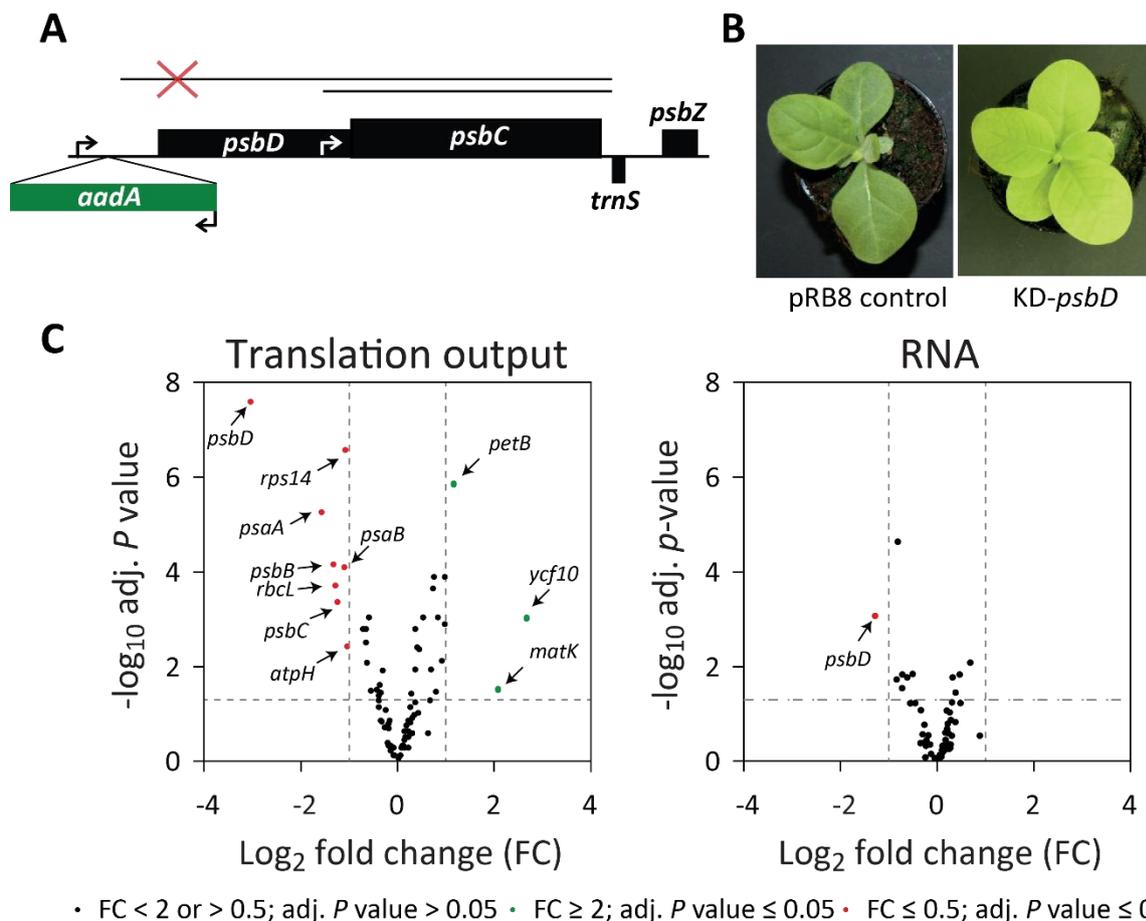


Figure 3.5: Transcriptome and ribosome footprint analyses reveals potential CES regulation of *psbB* in KD-*psbD* mutant.

A. Physical gene map of the mutant. The selection marker, *aadA*, was inserted upstream of *psbD* in the antisense direction. The accumulation of a dicistronic transcript produced by the promoter upstream of *psbD* (black arrow) is disrupted by the *aadA* insertion. An internal promoter (white arrow) produces a monocistronic *psbC* transcript (labeling details are given in Figure 3.4). **B.** The KD-*psbD* mutant was grown to the same developmental stage of three-week-old pRB8 control plants under standard conditions (section 2.2.1.8). The aerial part was used for total RNA and ribosome footprint isolations. **C.** Comparison of the ribosome footprint abundance (left) and transcript abundance (right) of all the chloroplast ORFs represented in volcano plots. Results from three biological replicates were collected and ratios of \log_2 fold change in the mutant compared to the control were plotted against the adjusted *P* values (labeling details are given in Figure 3.1D).

3.1.7.2 *psbA* and *psbH* translation is reduced in the KD-*psbD* mutant at cotyledon stage

We speculated that a potential CES regulation might be particularly important during the thylakoid membrane biogenesis at a very early developmental stage (cotyledon stage) where the synthesis of the PSII subunits is highly active. Hence, CES regulation may dominate other regulations (e.g., by light) during chloroplast biogenesis, which could simplify its detection. To address this hypothesis and to further disentangle the translational regulation during the assembly of PSII, I performed transcriptome, and ribosome profiling analyses on eight-day-old tobacco seedlings of the KD-*psbD* mutant (Figure 3.6A). The transcriptome-wide analysis revealed a significant defect in the abundance of *psbD* transcripts, and the ribosome profiling showed an even stronger decrease in its ribosome footprints,

similar to the result observed in the four-leaf stage for the same mutant (section 3.1.7.1). Also comparable to the four-leaf stage was the observed effect for *psbC* whose translation output is significantly decreased (Figure 3.6B). Furthermore, at this developmental stage, the KD-*psbD* mutant exhibited a significant decrease in the translation output of *psbB* and *psbZ* and more than twofold decrease for *psbA* and *psbH* (Figure 3.6B). A possible interpretation of the data is that due to the strong knockdown of *psbD*, the assembly of PSII is impaired, which causes a downregulation in the translation output of other subunits in PSII in order to maintain the stoichiometry. Overall, most of the changes observed were similar to those described in *Chlamydomonas* (Trösch et al., 2018). A comparison of the translation output of the tobacco KD-*psbD* mutant and the *Chlamydomonas dU* mutant, which is also strongly impaired in *psbD* expression (Trösch et al., 2018) showed a clear similarity between the two organisms in regard to the effects observed for *psbA*, *psbB* and *psbH* from PSII and *psaA*, *psaB* and *psaC* from PSI (Figure 3.6C). Consequently, this result might lead to the assumption that a cascade of feedback regulation controls the assembly of PSII where PsbA, PsbB, and PsbH are potential CES subunits downstream of PsbD (Figure 3.6D). However, the hierarchy of the cascade in embryophytes remains to be disentangled. In addition, there seems to be feedback regulation from PSII assembly to the expression of core subunits of PSI.

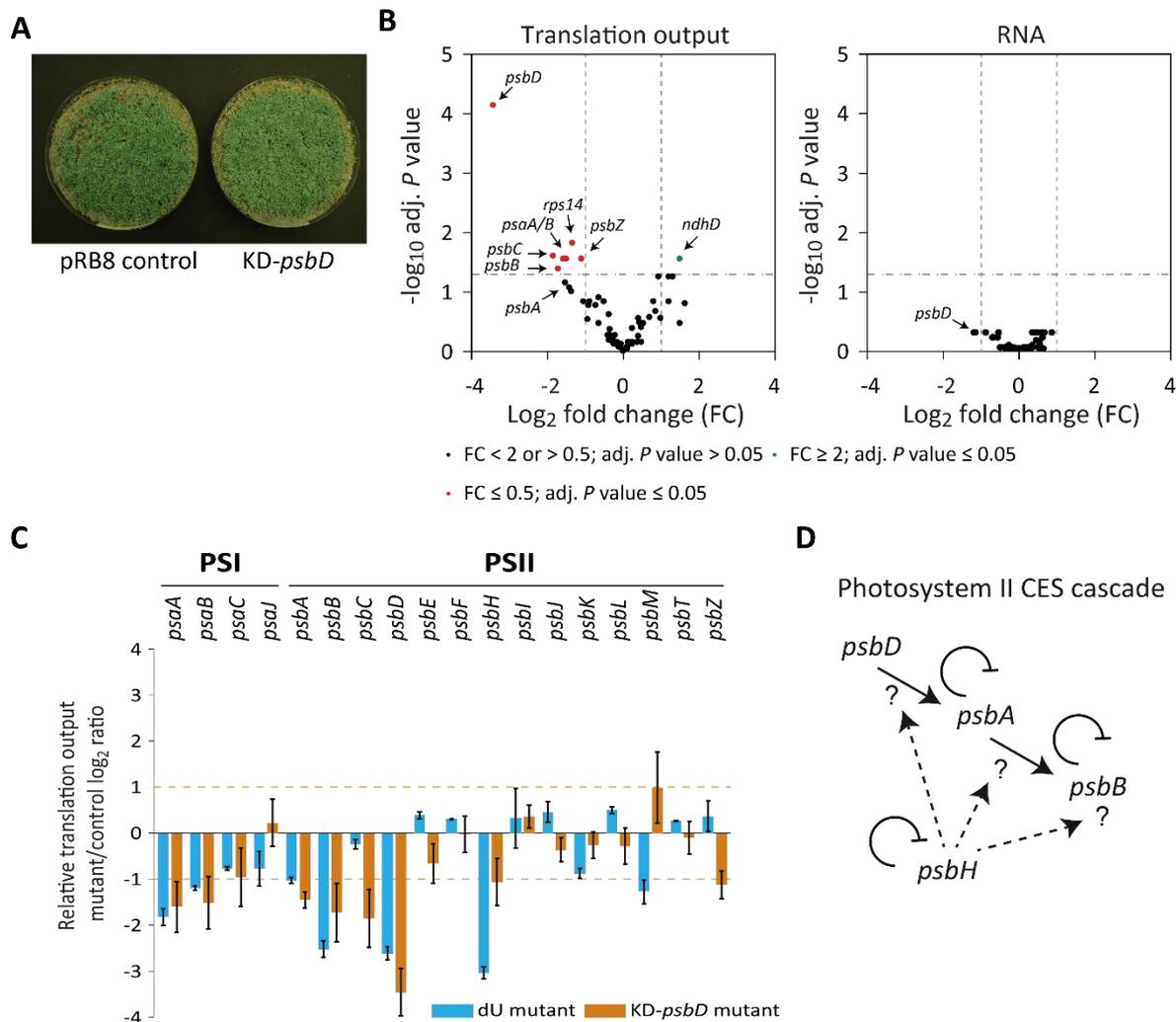


Figure 3.6: Knockdown of *psbD* results in similar translation regulatory effects in embryophytes in comparison to *Chlamydomonas*.

A. KD-*psbD* mutant and pRB8 control were grown on vermiculite under standard conditions (section 2.2.1.8). Seedlings were harvested eight days after sowing at the cotyledon stage (section 2.2.1.8.2). **B.** Ribosome profiling (left) and transcriptome profiling (right) in KD-*psbD* at the cotyledon stage compared to pRB8 control (for labeling details see Figure 3.1D). **C.** Relative translation output of PSI and PSII subunits in the KD-*psbD* mutant in comparison to pRB8 control. Comparison between data from tobacco (orange) and *Chlamydomonas* dU mutant grown under $80 \mu\text{mol m}^{-2} \text{s}^{-1}$ (blue). Error bars represent standard deviations of three biological replicates. **D.** Cartoon of the proposed PSII CES cascade in *Chlamydomonas*. Effects of the reduced PsbD synthesis on the translation of *psbA*, *psbB*, and potentially *psbH*.

3.1.7.3 Confirmation of translational downregulation of *psbA* in the KD-*psbD* mutant by pulse labeling

Further confirmation of the translation feedback regulation of *psbA* was achieved by inspecting the synthesis rate of newly synthesized PsbA with an *in vivo* protein labeling assay. Extracts from the soluble fraction and thylakoid membranes were visualized following an *in vivo* protein labeling with

^{35}S -radiolabeled amino acids (section 2.2.5.7). In order to have a simplified pattern of labeled proteins, cytoplasmic translation was blocked. A comparison of the pRB8 control and the mutant revealed less incorporation of ^{35}S -labelled amino acid into the newly synthesized PsbA (Figure 3.7A). As discussed previously (section 3.1.4), the mutant lane displayed a stronger background than the control. This is likely due to the fact that PsbA dominates the thylakoid membrane fraction in the chloroplast, hence loading an equal amount of freshly synthesized proteins when PsbA synthesis is reduced causes overloading of all the labeled proteins. Quantification of *psbA* band signal and normalization to the background revealed a reduction of about 3.5-fold in comparison to the pRB8 control (Figure 3.7B), which suggests that the synthesis rate and/or stability of the D1 subunit is reduced when PsbD accumulation is compromised. This result further hints at a potential CES regulation of *psbA* in PSII.

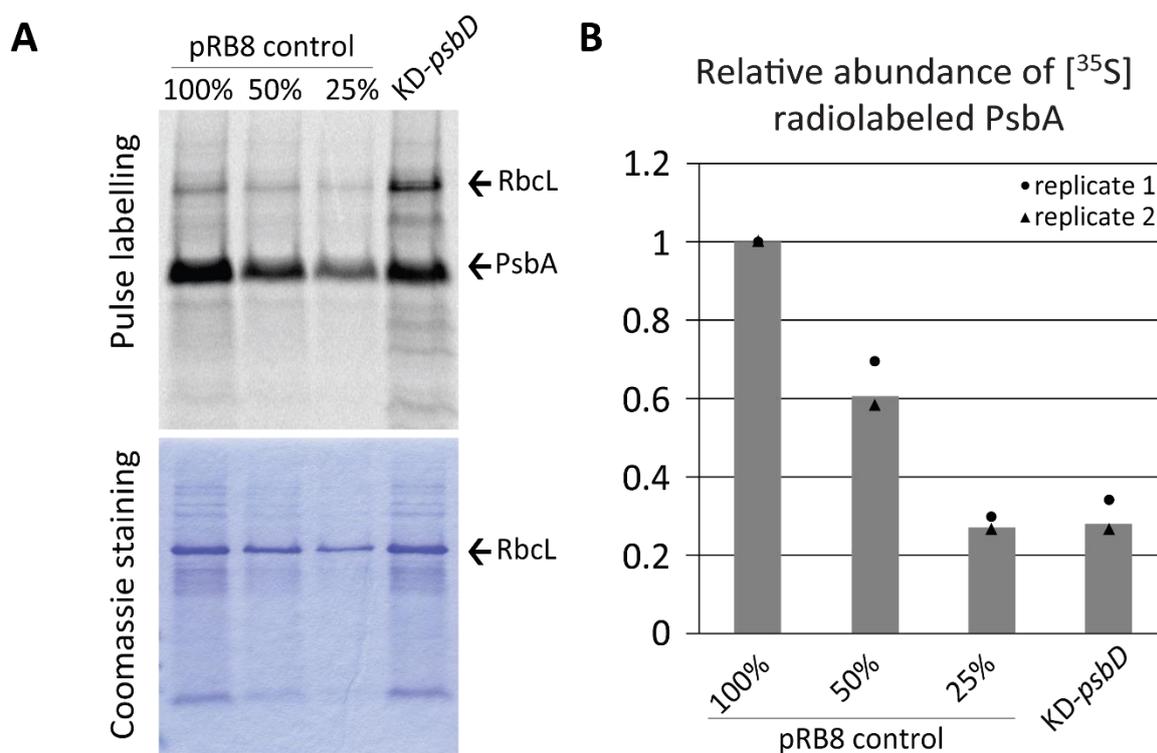


Figure 3.7: PsbA synthesis is downregulated when PsbD protein level is reduced.

A. Chloroplast proteins were radiolabeled with ^{35}S -methionine and cysteine and extracted. Membrane proteins equivalent to 100,000 incorporated cpm from KD-*psbD* mutant and pRB8 control were separated on tricine gels. Autoradiograph and Coomassie staining of a representative gel are shown. **B.** The signal intensity of the band corresponding to PsbA was quantified using Image Lab software (Bio-Rad) and normalized to the background excluding the RbcL band. Results were obtained from two biological replicates.

3.1.8 Reduced PsbD levels induce significant alterations of chloroplast ribosome occupancy on PSII and PSI transcripts

The translation is predominantly regulated at the initiation level, which is detectable by ribosome profiling as a change of the overall footprints abundance per ORF (Ingolia et al., 2018). In addition, translation elongation rate can significantly differ within a given reading frame (e.g., Chotewutmontri and Barkan (2018)). Changes at the level of elongation can cause a local increase or decrease of ribosome occupancies within an ORF. To check whether reduced *psbD* expression causes a redistribution of ribosomes in chloroplast reading frames, I examined the ribosome distribution for each ORF in the chloroplast. To do so, the signal intensity of each probe in protein-coding regions was normalized to the sum of the signal of the respective ORF (section 2.2.3.4) to obtain ribosome occupancy values. Next, I calculated the ratio of these ribosome occupancies between the mutant and the control. With such an analysis, local alteration of footprint distribution regardless of the overall change of the ORF translation output could be identified. Note that with the targeted ribosome profiling approach, footprints can be detected in a resolution of approximately 30 nucleotides (nt) and, hence, only strong pausing can be observed in this analysis. The significance of changes was assessed using the empirical Bayes methods in the limma package (section 2.2.3.4) (Smyth, 2004). FDR was used for *P* value adjustments with 0.05 set as a threshold to define significance. To make conclusions reliable and only consider exceptional strong alterations in ribosome pausing, only regions with two consecutive probes with \geq twofold change and adjusted *P* value \leq 0.05 were considered as sites with potentially altered ribosome pausing.

Inspection of altered ribosome distribution in KD-*psbD* mutant (Figure 3.5B) showed 57 probes with \geq twofold significant change in ribosome occupancy compared to the pRB8 control. Notably, 12 of these significant changes appeared in two consecutive probes in a total of four genes (labeled in red in Figure 3.8A). Three potential pausing sites were detected in *psbA* transcript with increased local ribosome footprint coverage (Figure 3.8B). Likewise, two pausing sites were identified in *psaC* transcript. Interestingly, a substantial decrease in ribosome coverage was observed at two locations, in *psaB* and in the middle of *psbC* ORF (Figure 3.8B). As a control, the same analysis was done for the transcript abundance data and no sites with significant alterations were observed (Supplemental Figure 1) reflecting that the local changes detected in the ribosome footprint data are rather caused by ribosome redistribution. Overall, these findings suggest that the translational regulation of the subunits of PSII and PSI triggered by the absence of *psbD* could be an interplay between regulation of the initiation and elongation of these genes.

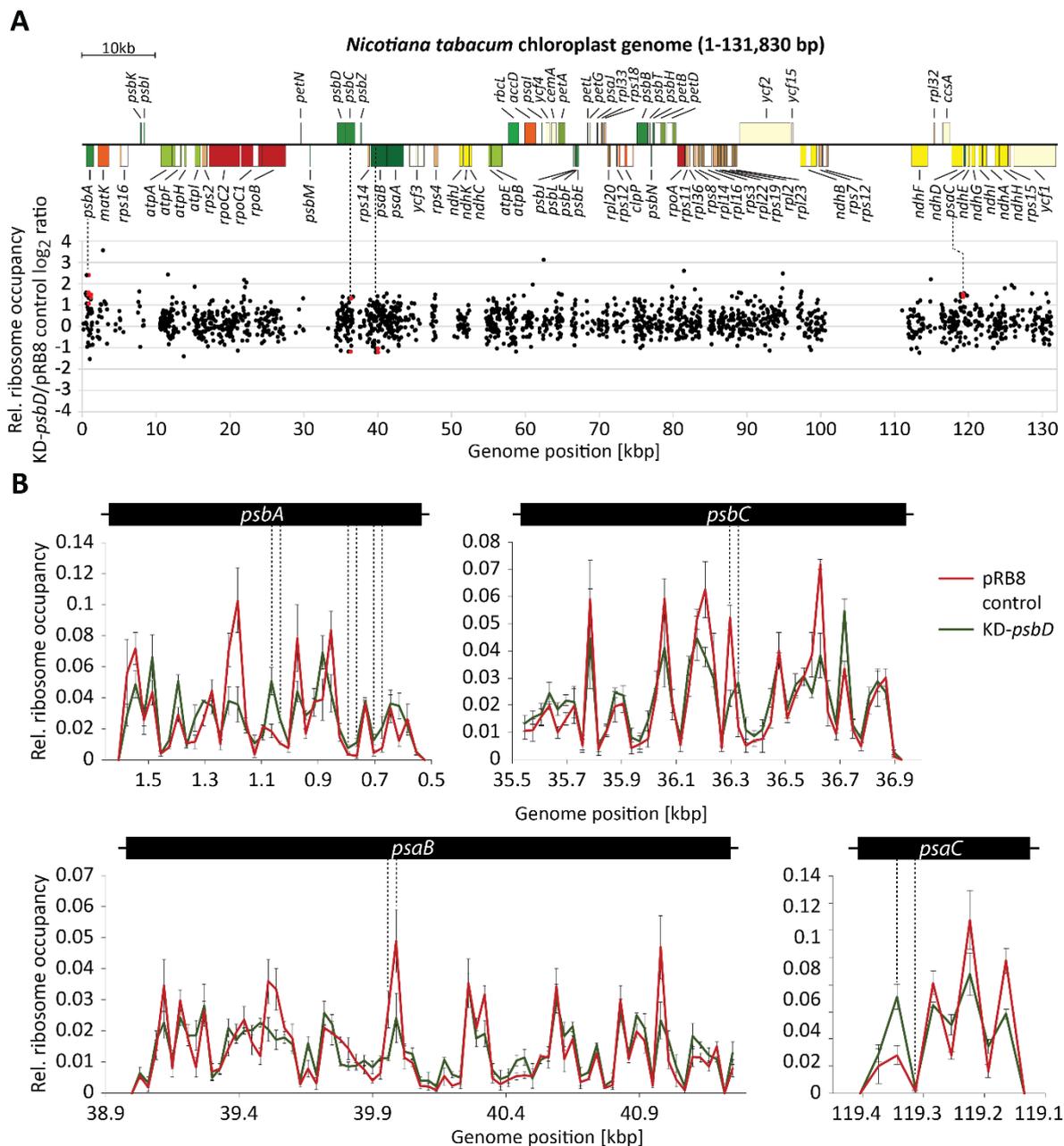


Figure 3.8: Transcriptome-wide analysis of ribosome occupancy in KD-*psbD* mutant reveals significant alterations in PSII and PSI transcripts.

A. Chloroplast genome-wide comparison of the relative local ribosome occupancies in KD-*psbD* mutant compared to pRB8 control in Figure 3.5B. For each probe located in an ORF, the footprint signal was normalized to the sum of the signals of all the probes located in the same ORF to calculate the contribution of each probe region to the ribosome occupancy of the whole reading frame. Ratios of these ribosome occupancy values in KD-*psbD* in comparison to pRB8 control were calculated and plotted according to the position in the chloroplast genome. Results were obtained from three biological replicates. Probes in regions with consecutive probes with \geq twofold change and P values ≤ 0.05 are labeled in red. The physical map of the chloroplast genome was generated from the NCBI reference sequence Z00044.2 using OGDRAW (Greiner et al., 2019). The map shows the first of the two inverted repeats in the chloroplast genome and illustrates only the protein-coding genes. **B.** Detailed magnification of the ORFs containing probes with significantly altered ribosome coverage. The signal fraction of each probe to the sum of the ORF signal is plotted. The dashed lines mark the sites with

significantly altered ribosome occupancy. Error bars represent the standard deviation. The black boxes depict the protein-coding regions.

3.1.9 Impaired *psbA* expression does not cause alterations in the expression of other PSII subunits

The high similarity in the regulation of PSII subunits between *Chlamydomonas* and embryophytes may argue for the possibility of a conserved CES cascade. In order to assess the effect on *psbB* translation when PsbA is not produced, the *hcf173-2* T-DNA insertion mutant was analyzed. HCF173 has been shown to stabilize the *psbA* transcript and has been assumed to activate its translation (Schult et al., 2007; Williams-Carrier et al., 2019). In both studies, the PSII deficiency in the *hcf173-2* T-DNA mutant was attributed to a defect in *psbA* translation and only negligently to a defect in the mRNA level. However, the mRNA hybridization assays performed in both studies left the magnitude of the effect on the transcript abundance unclear. I revisited the PSII deficiency in *hcf173-2* T-DNA mutant using microarray-based transcriptome and ribosome profiling of chloroplast genes. Surprisingly, the transcriptome and translatoome analyses demonstrated that the *hcf173-2* mutant has a severe decrease of the *psbA* mRNA level (~ fourfold), as well as an approximately fivefold decrease in the translation output (Figure 3.9A, B). These results indicate that HCF173 acts predominantly as mRNA stabilization factor and plays if at all, only a minor role as a translational activator of *psbA*. However, even though the translation output of *psbA* is strongly reduced, only a marginal effect was observed on the footprint accumulation of *psbB* or other PSII-coding genes, which is in line with the findings reported in Williams-Carrier et al. (2019). These results put in doubt the presence of a CES cascade in PSII (Figure 3.6F) in which PsbA is required for the efficient translation of *psbB*.

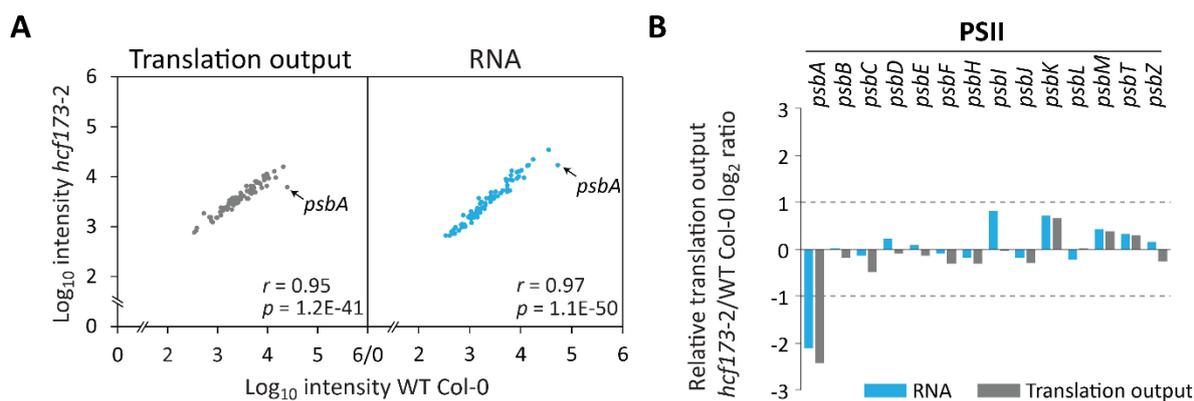


Figure 3.9: Defect in the expression of *psbA* has no effect on the translation of *psbB* or other transcripts encoding PSII subunits.

A. Transcript accumulation and translation output of one biological replicate were calculated in *hcf173-2* T-DNA mutant in comparison to WT Col-0 (data plotted as in Figure 3.4E). **B.** Ratio of the translation output and transcript accumulation of chloroplast ORFs encoding PSII subunits calculated from data shown in A.

3.1.10 Impaired cytochrome *b₅₅₉* expression induces reduced *psbA* translation

It has been shown in various studies that cytochrome *b₅₅₉* (cyt *b₅₅₉*) is at the very top of the PSII assembly cascade (Morais et al., 1998; Pakrasi et al., 1989; Swiatek et al., 2003). Cyt *b₅₅₉* consists of the alpha (α) and beta (β) apoproteins, encoded by the *psbE* and *psbF* plastid genes, respectively. In a collaboration with Dr. Karin Meierhoff (EMP-Düsseldorf University) and Dr. Kamel Hammani (IBMP-CNRS) the *mda1* mutant allele *hcf111-1* was analyzed. MDA1 (mTERF defective in *Arabidopsis* 1) (Robles et al., 2012) is a mitochondrial transcription termination factor (mTERF) protein, also known as mTERF5, which showed a clear defect in PSII activity and synthesis based on chlorophyll fluorescence measurements and pulse labeling experiments (Métégnier et al., 2020). In parallel to my studies, Ding et al. (2019) reported that MDA1 is a positive regulator of *psbE/F/L/J* transcription. Also, my plastome-wide analysis of transcript abundance and translation output revealed a major defect on the expression of the *psbE/F/L/J* operon in the *hcf111-1* mutant (four to ninefold decrease on transcript abundance and translation output levels; Figure 3.10), which is sufficient to explain its PSII defect. Interestingly, in the *hcf111-1* mutant also the translation level of *psbA* decreased more than threefold, which may suggest that *psbA* is a potential CES subunit regulated by its state of assembly. Interestingly, an increase of the translation level was observed for several genes that are located on the complement strand directly upstream of the *psbE/F/L/J* operon (e.g., translation output of *petL*, *petG* and *rpl33* increased more than twofold). If their overexpression is a direct effect of the MDA1 mutation or rather an indirect consequence needs to be elucidated. In addition to the defects in the expression of PSII genes, an effect on *ndhA* and *ndhI* mRNA accumulation was observed as well, which suggests that these transcripts might be additional targets of the HCF111 transcription activator.

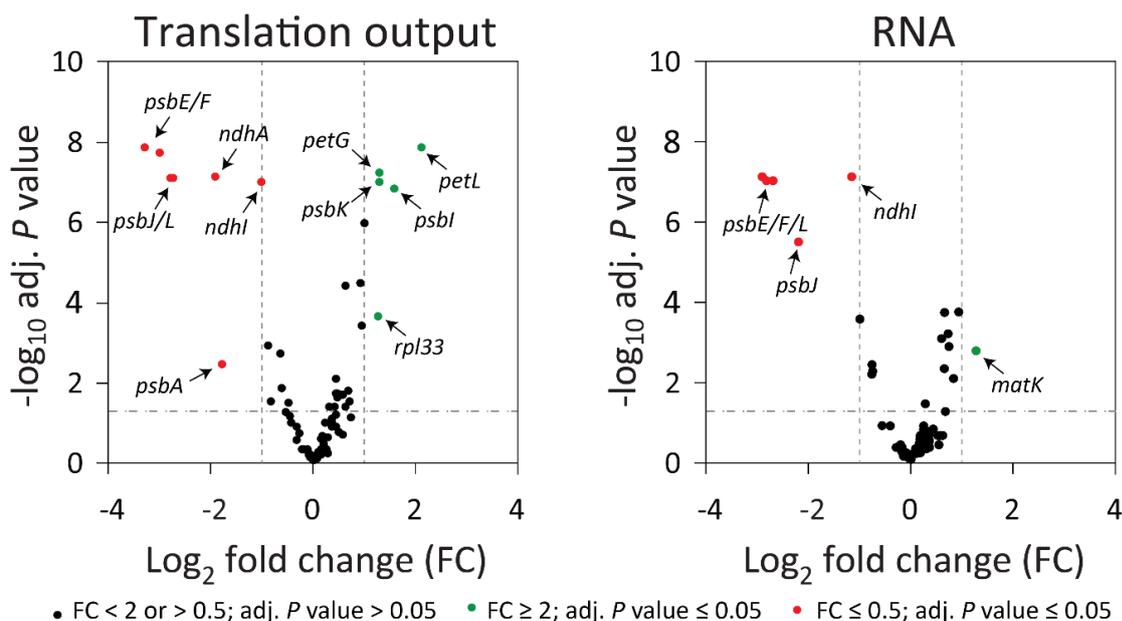


Figure 3.10: Reduced expression of cyt *b₅₅₉* results in the decrease of the translation output of *psbA*.

Footprint (left) and transcript (right) data was collected from three biological replicates (labeling as in Figure 3.1D).

3.1.11 The knockout of a PSII assembly factor causes a decrease in the translation output of *psbB*

Besides the abovementioned mutants of PSII core subunits, a mutant of the non-essential plastid-encoded PSII assembly factor PsbN was analyzed (Krech et al., 2013). This line contains a frameshift mutation in the *psbN* reading frame causing a premature stop codon close to the start codon. The *psbN* gene is located on the complement strand of the pentacistronic *psbB* operon-like transcription unit (Krech et al., 2013; Stoppel and Meurer, 2013; Zghidi-Abouzid et al., 2011) (Figure 3.11A). It was previously reported that the expression of *psbN* affects the cleavage in the intergenic *psbT/psbH* region (Chevalier et al., 2015). However, it was shown in Krech et al. (2013) that the mutant displayed unaltered RNA processing of the *psbB* operon. RB70 plants (Ruf et al., 2001) in which the *aadA* cassette was targeted to the same intergenic region between two tRNA genes as in the $\Delta psbN$ mutant were used as control. $\Delta psbN$ mutant and RB70 control were grown side by side under controlled conditions (section 2.2.1.8) (Figure 3.11B). Transcriptome profiling showed no significant change in comparison to RB70 control (Figure 3.11C). Despite the pronounced PSII assembly defect (Krech et al., 2013), only the translation output of *psbB* in PSII decreased significantly by threefold (Figure 3.11C). This result further supports the hypothesis that in embryophytes PsbB is a CES subunit whose synthesis is dependent on the assembly status of PSII. In addition, the translation output of *atpH* decreased significantly by more than twofold, which is in line with the lower ATP synthase activity and lower AtpB accumulation described in this mutant (Krech et al., 2013). Furthermore, the translation output of *rbcL*, *ycf10*, and *matK* changed significantly by more than twofold. These observed variations could be secondary effects of the strong defect in the accumulation of PSII as they were observed in other mutants (e.g., Figure 3.5).

The ribosome distribution in this mutant was checked as described in section 3.1.8. Interestingly, 12 *psbA*-localized probes and 13 *psbB*-localized probes showed more than twofold change of the relative ribosome occupancy with a P value ≤ 0.05 (data not shown). However, following the multiple testing, the adjusted P values rise above the threshold (0.05). This alteration, though not significant, in the ribosome occupancies within genes encoding PSII subunits (PsbA, PsbB) hints to a potential regulation on the elongation level upon defects in the assembly of the complex.

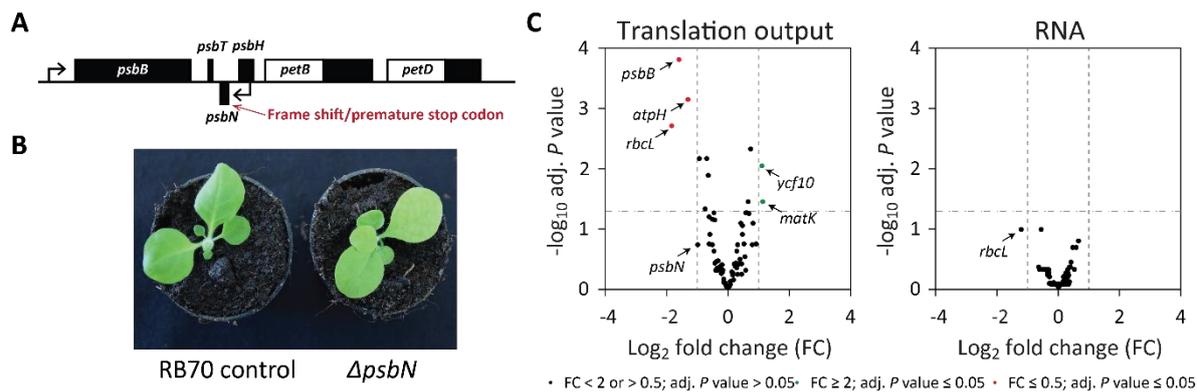


Figure 3.11: Decrease of the translation output of *psbB* in $\Delta psbN$ mutant.

A. Physical gene map of the $\Delta psbN$ mutant. The *psbB* operon comprises *psbB*, *psbT*, *psbH*, *petB*, and *petD* genes, which are transcribed in the sense direction. The *psbN* gene is located on the reverse strand and transcribed in the opposite direction. Black box: protein-coding region; open box: introns; black arrows: transcription start site and direction of the transcription. **B.** The *psbN* knockout mutant ($\Delta psbN$) and RB70 control were grown *in vitro* for two weeks (section 2.2.1.8.3) before transfer to soil for three weeks under controlled conditions (section 2.2.1.8). The aerial part was used for total RNA and ribosome footprints isolation. **C.** Ribosome footprint (left) and RNA (right) levels in $\Delta psbN$ in comparison to RB70 control. Refer to Figure 3.1D for labeling details.

3.1.12 The knockout of a PSII assembly factor evokes significant alteration of ribosome distribution on *psbA*, *psbB*, and *rbcL* transcripts at the cotyledon stage

$\Delta psbN$ mutant was also analyzed at the cotyledon stage (Supplemental Figure 2A) because the synthesis and assembly of PSII subunits are highly active at early developmental stages and thus an assembly defect might be more pronounced as observed for KD-*psbD* (section 3.1.7.2). The changes in transcript accumulation and translation output were highly similar to those observed in older seedlings (section 3.1.11) (Supplemental Figure 2B). In addition to these effects, the translation output of *psaA* decreased significantly (Supplemental Figure 2B) which may be an indirect consequence of the defect of PSII as observed in other mutants with PSII defects (section 3.1.7.1). The relative footprint distribution for each probe across the chloroplast ORFs was assessed as described in sections 2.2.3.4 and 3.1.8. Altogether, 53 probes showed \geq twofold significant change (adj. *P* value \leq 0.05). 16 of these changes appeared in at least two consecutive probes (colored with red in Figure 3.12A) from three genes, *psbA*, *psbB*, and *rbcL*. Strikingly, most of these sites with local constricted alterations were located in *psbA* (12 sites out of 16) and only two in *psbB* and *rbcL* (marked with red and black dashed lines in Figure 3.12B). The substantially altered pausing in *psbA* suggests that its translation is regulated on the elongation level upon a defect in PSII assembly. For this gene, an interesting phenomenon of ribosome redistribution was observed: A cluster five close sites showed a significant increase in local ribosome footprint coverage and was followed by a downstream site with a decrease in ribosome coverage (marked with blue dashed lines). The same analysis was performed for the transcript abundance data as a control. No

Chlamydomonas studies (Wostrikoff et al., 2004), PsaB initiates a translation feedback regulation cascade, in which *psaA* translation depends on the availability of PsaB and PsaA is required for the efficient translation of *psaC*. To check whether CES regulation is involved in the assembly of PSI in embryophytes, a knockdown mutant of *psaA* (KD-*psaA*) was used and the effects on the RNA accumulation and translation level of the other PSI subunits were assessed. This mutant was kindly provided by Dr. Mark Aurel Schöttler (MPIMP). In the KD-*psaA* mutant, the Shine-Dalgarno (SD) sequence was point-mutated (Figure 3.13A), which causes a reduction of the translation initiation efficiency of the *psaA* gene (D. Bednarczyk, R. Bock and M.A. Scöttler, unpublished data). In order to control for the transcriptional read-through and other effects caused by the *aadA* expression, a line that harbors the *aadA* resistance marker gene in the same position but lacks the SD mutation was used as a control. The control line is referred to as *psaA* control and was grown next to the KD-*psaA* mutant (Figure 3.13B). Indeed, the RNA level of all the chloroplast-encoded transcripts remained unchanged (Figure 3.13C). The translation output of *psaB* and *rps14*, which reside in the same gene cluster as *psaA* was not altered (Figure 3.13C). Interestingly, this observation argues against a translational coupling between *psaA* and *psaB* or *rps14*. Furthermore, no alteration in the translation output of *psaC* and other PSI subunits was observed (Figure 3.13C), making a translation feedback regulation downstream of *psaA* unlikely. In fact, apart from the primary defect in *psaA* translation, only the translation output of *rbcL* was altered (Figure 3.13C). This defect of *rbcL* translation was observed in some of the analyzed photosynthesis mutants to different extents (e.g., in the *psbN* mutant (section 3.1.11)) (Figure 3.11), which may be explained as a general, indirect regulation by photosynthetic activity (e.g., by redox signals).

To test the hypothesis that the absence of a nucleus-encoded subunit might be the trigger of CES regulation of the chloroplast-encoded subunits of PSI as it is the case in Rubisco, I selected the T-DNA insertion mutant *psad1-1* that has reduced levels of PsaD (Ihnatowicz et al., 2004). Wild-type Columbia (Col-0) was used as control (Figure 3.13D). PsaD is an essential stromal subunit of PSI (Haldrup et al., 2003) that forms the ferredoxin binding site together with PsaC and PsaE. It was shown that PsaD stabilizes the binding of PsaC to the heterodimer (Antonkine et al., 2003) and is required for the stabilization and accumulation of PSI (Haldrup et al., 2003; Ihnatowicz et al., 2004). Although the decrease of PsaD caused a major decrease in the protein abundance of the other PSI subunits (Ihnatowicz et al., 2004), no substantial alteration was observed on the level of transcript accumulation or translation output (Figure 3.13E). Altogether, these results suggest that the stoichiometric accumulation of PSI is mainly regulated on the level of proteolysis.

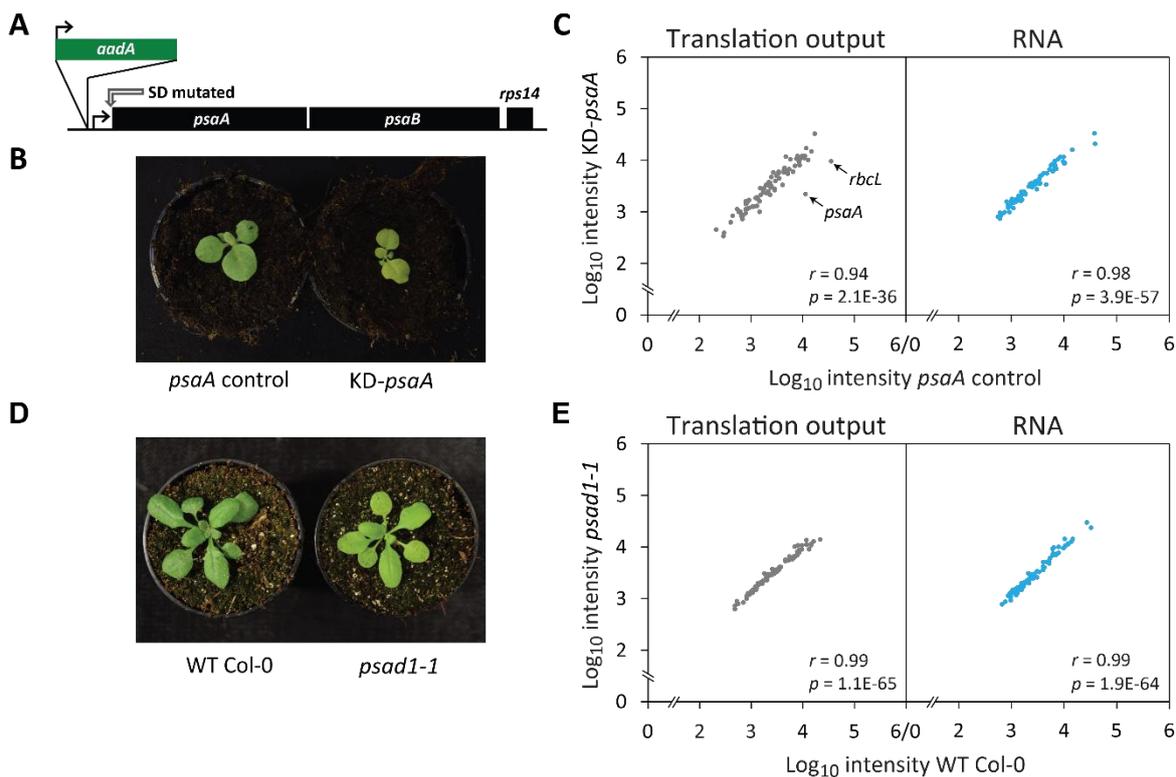


Figure 3.13: Reduced expression of *psaA* and *psaD* does not trigger CES regulation within PSI.

A. Physical map of the KD-*psaA* mutant. The selection marker (*aadA* cassette) was inserted upstream of the start codon of *psaA*. The SD sequence was mutated causing a defect in the translation of *psaA*. **B.** The KD-*psaA* mutant was grown under standard conditions (section 2.2.1.8) and displayed a pale green phenotype in comparison to the *psaA* control. **C.** The average of ribosome footprint and transcript abundances for each ORF in KD-*psaA* are compared to the *psaA* control. The results were obtained from two biological replicates (labeling details as in Figure 3.4E). **D.** Four weeks old seedlings grown under greenhouse conditions (section 2.2.1.8). The T-DNA mutant *psad1-1* had a pale green phenotype in comparison to WT Col-0. **E.** Ribosome footprint and RNA abundance from two biological replicates. Labeling details are given in Figure 3.4E.

3.1.14 Knockout of a PSI assembly factor drives a potential CES regulation of *psaC* and *psaI*

Similar to PSII, I analyzed a knockout mutant of a plastid-encoded PSI assembly factor, Ycf4 (for hypothetical chloroplast reading frame no. 4) (Boudreau et al., 1997; Krech et al., 2012). The *ycf4* gene is part of a gene cluster in the tobacco plastid genome, which comprises additional three genes: *psaI* (encoding a small non-essential subunit of PSI), *ycf10* (encoding a nonessential envelope membrane protein) (Rolland et al., 1997), and *petA* (encoding the cytochrome *f* subunit (PetA) of the Cyt *b₆f* complex) (Figure 3.14A). In this mutant, most of the *ycf4* reading frame was deleted and replaced with an *aadA* cassette integrated in sense direction (Figure 3.14A) (Krech et al., 2012). In contrast to *Chlamydomonas* (Boudreau et al., 1997), the tobacco Δ *ycf4* mutant can grow photoautotrophically (Krech et al., 2012). The severely retarded mutant was grown side-by-side with the pRB8 control (Figure 3.14B). The aerial part was harvested and used for transcript and ribosome profiling. In

accordance with Krech et al. (2012), this mutant displayed an increase in the transcript accumulation of *petA* and *ycf10* (Figure 3.14C), which is due to the read-through transcription from the upstream *aadA* gene. Moreover, the transcript accumulation of *ycf4* increased (Figure 3.14C), which is probably caused by the production of a tetracistronic unprocessed precursor from *psaI* to *petA* including the *aadA* cassette, which was shown to accumulate in the mutant (Krech et al., 2012). This tetracistronic precursor containing the remaining part of the *ycf4* ORF could probably hybridize to the *ycf4* probes on the microarray. On the footprint level, the translation output of *petA* and *ycf10* increased significantly to the same extent as the transcript level (Figure 3.14C). In *Chlamydomonas*, *petA* translation is feedback regulated according to its assembly status (Choquet et al., 1998; Choquet et al., 2003). If that was the case in tobacco one would expect that an overaccumulation of the *petA* transcript is counterbalanced by a translational downregulation to achieve an unaltered translation output of *petA*, which was not observed in the $\Delta ycf4$ mutant. A potential interpretation of this result is that the *Chlamydomonas* CES regulation of *petA* is not conserved in tobacco. Furthermore, the translation output of *petB* increased significantly (~ 2.5-fold) (Figure 3.14C). This increase reflects a potential regulatory link between *petA* and *petB* and could be interpreted as a potential positive feedback regulation within the Cyt *b₆f* complex. Among the subunits of PSI, only *psaI* and *psaC* showed a decrease in their translation output. While the effect on *psaI* might be caused by the *aadA* insertion in the same transcription unit, the effect on *psaC* could be explained as a potential negative feedback regulation within PSI in response to its assembly defect. However, the translation defect was not specific to PSI since genes from the other photosynthetic complexes (*psbB*, *atpH*, *rbcl*, *petG*, *psbA*), plastid-encoded plastid RNA polymerase (*rpoC2*), *matK* and *ycf2* exhibited also significant alterations of footprint abundances (Figure 3.14C). The high number of changes in this mutant may be related to the exceptionally severe defect of its photosynthetic performance (Krech et al., 2012). Thus, the potential CES regulation of *psaC* and the potential positive feedback regulation in the Cyt *b₆f* complex remain highly speculative given the broad range of secondary effects observed.

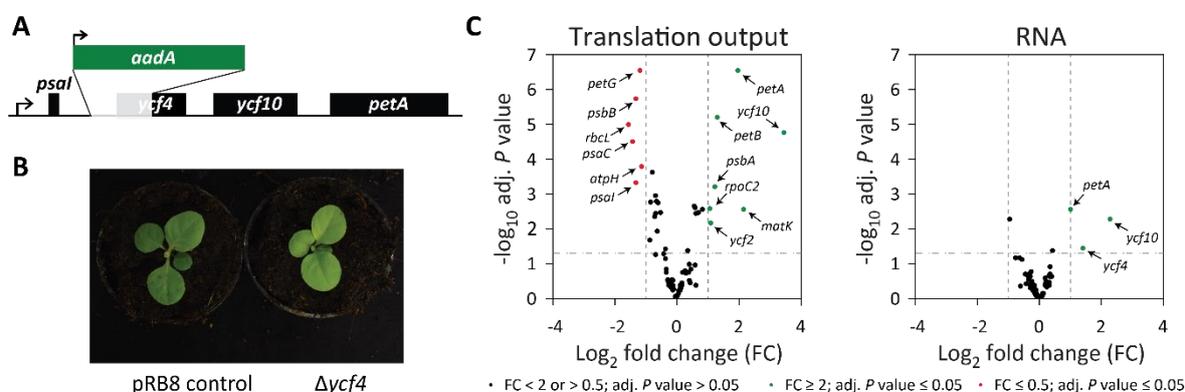


Figure 3.14: *ycf4*-dependent assembly defect of PSI provokes numerous alterations on translation level.

A. Physical map of the *ycf4*-containing gene cluster in the chloroplast genome in $\Delta ycf4$ mutant. For labeling details, refer to Figure 3.4A. **B.** The $\Delta ycf4$ mutant was grown under low light conditions (section 2.2.1.8) and displayed a retarded growth and pale green phenotype in comparison to the pRB8 control. **C.** Volcano plots of relative changes in translation output and transcript abundance in $\Delta ycf4$. Results were obtained from three biological replicates. Refer to Figure 3.4B for details.

3.1.15 Potential positive feedback regulation between *petA* and *petB* in cytochrome *b₆f* complex

The cytochrome *b₆f* complex (Cyt *b₆f*) is one of the two major sites that control photosynthetic fluxes together with the ATP synthase complex. It is the smallest complex in the photosynthetic electron transport apparatus with eight subunits, two of them are nucleus-encoded: PetC and PetM (Baniulis et al., 2009; Cramer and Zhang, 2006; Schöttler et al., 2015). Although largely unknown, it is assumed that the assembly of the Cyt *b₆f* complex in land plants follows a similar pattern to that described in bacteria and yeast mitochondria (reviewed by Zara et al. (2009); Smith et al. (2012)). It is speculated that the assembly starts with cytochrome *b₆* (Pet B) and subunit IV (Pet D) assembling first, followed by the insertion of PetA and the Rieske protein (PetC) (Schöttler et al., 2015). It has been shown in *Chlamydomonas* that *petA* is a CES subunit that autoregulates its synthesis depending on the availability of PetD or PetB (Kuras and Wollman, 1994; Lemaire et al., 1986). A prominent study evoked the possibility of a CES regulation that governs the assembly of the Cyt *b₆f* complex also in tobacco (Monde et al., 2000). In order to verify a potential CES regulation between PetA and PetB suggested by our data (section 3.1.14), and since the *petB* and *petD* mutants used in Monde et al. (2000) were not able to grow autotrophically, I decided to analyze an independent *petA* overexpressor mutant. This overexpression of *petA* is indirectly caused by the knockout of *psaI* created by *aadA* insertion within the *psaI* ORF (Schöttler et al., 2017) (Figure 3.15A). *PsaI* is a non-essential subunit of PSI, and its knockout does not cause any phenotype under standard growth conditions (Figure 3.15A) (in contrast to the *ycf4* mutation, which causes a severe growth defect due to its disturbed PSI assembly). Spectroscopic analyses have shown that the content of the Cyt *b₆f* complex is significantly higher in *psaI* mutants grown in standard conditions (section 2.2.1.8) (Schöttler et al., 2017). Similar to the $\Delta ycf4$ mutant, the insertion of the *aadA* cassette caused overexpression of *petA*, *ycf4* and *ycf10* on transcript level. In addition, an increase in the transcript abundance of *psaI* was observed, which is likely due to the overaccumulation of a fused transcript that contains *aadA* and the 3' region of *psaI* that hybridizes to the microarray (Schöttler et al., 2017). As seen in the $\Delta ycf4$ mutant (Figure 3.14), the overaccumulation of the transcripts downstream of *aadA* caused an increase in *ycf10* and *petA* translation output. As expected from the position of the *aadA* cassette, also the translation output of *ycf4* was increased (Figure 3.15B). In contrast, the ribosome footprint level of *psaI* decreased significantly due to the disruption of the reading frame by the *aadA* cassette (Figure 3.15B). Interestingly, PetB synthesis increased by ~ 2.5-fold (Figure 3.15B) thereby

substantiating the result obtained with the $\Delta ycf4$ mutant arguing for a regulatory link between PetA and PetB.

Both $\Delta ycf4$ and $\Delta psal$ are PSI related mutants. To exclude the possibility that a PSI-related defect might be causing the increase in the translation output of PetB a third, PSI-unrelated overexpressor of *petA* was investigated. This overexpressor mutant was created by insertion of an *aadA* cassette within the ORF of *ycf10*, a non-essential gene, in the sense direction. The $\Delta ycf10$ mutant was kindly provided by Dr. Mark Schöttler (MPIMP) and grows autotrophically (Figure 3.15C). The transcript profiling revealed a ~ 3.5 -fold increase in the transcript accumulation of *petA*, which is certainly due to the read-through from the upstream *aadA* cassette. Consequently, the translation output of *petA* showed a fourfold increase. In addition, the translation output of *petB* increased significantly twofold despite no change in the transcript abundance. These results indicate that the increase of *petB* ribosome footprints in the mutant is entirely due to the enhanced translation of *petA*. This finding does not only support the hypothesis of a potential positive feedback regulation between PetA and PetB, but also contradicts the order of regulation reported in *Chlamydomonas* where *petA* is upstream of *petB*. Nevertheless, this data does not mutually exclude the possibility of a potential negative feedback loop similar to that in *Chlamydomonas*. It is important to mention that *ycf10* is a lowly expressed gene and therefore it was technically challenging to detect the defect of its expression using the microarray as only very few probes were hybridized.

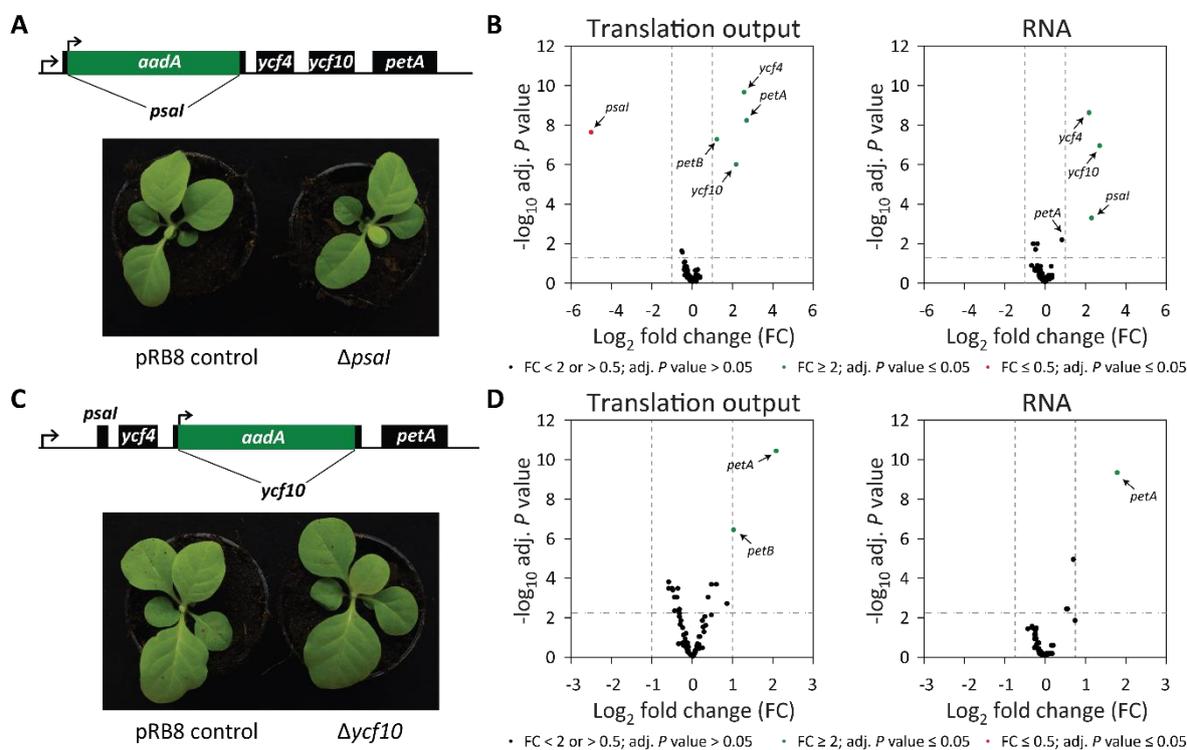


Figure 3.15: Enhanced translation output of *petA* triggers the translation of *petB* in $\Delta psal$ and $\Delta ycf10$ mutants.

A, C. Physical gene map of $\Delta psal$ and $\Delta ycf10$ knockout mutants both created by inserting an *aadA* cassette in sense direction to disrupt the *psal* and *ycf10* genes, respectively. Refer to Figure 3.4A for labeling details. **B, D.** Changes in translation output and transcript accumulation are shown in volcano plots. Results were obtained from three biological replicates. Refer to Figure 3.4B for details.

3.1.16 Confirmation of enhanced synthesis of PetB upon overexpression of *petA* by pulse labeling

To independently confirm the positive feedback regulation of PetB, an *in vivo* pulse labeling experiment was performed in which leaf punches from $\Delta psal$ and pRB8 control plants were radio-labeled with ^{35}S -methionine and cysteine (section 2.2.5.7). After pulse labeling, the soluble and membrane-bound proteins were separated as described in section 2.2.5.7. Due to the exceptionally high synthesis levels of PsbA and other PSII subunits, the subunits of the Cyt *b₆f* complex are usually not detectable in such approaches. Hence, immunoprecipitation experiments were performed on solubilized thylakoid proteins with a mixture of three antibodies against PetB, PetA and AtpB (section 2.2.5.8). The translation of *atpB* was virtually unchanged in $\Delta psal$ based on ribosome profiling data (Figure 3.15B), thus it was used as a control for the immunoprecipitation efficiency and the gel loading. As shown in Figure 3.16, the normalized protein synthesis level of PetB increased while that of PetA was slightly down. The quantification and normalization of the signals to the AtpB signal placed the enhancement of PetB synthesis at ~ threefold, similar to what is detected with the ribosome profiling experiment. This result, together with the ribosome profiling data, suggests a positive feedback regulation of *petB* based on the synthesis level of PetA. Since both *petA* and *petB* showed an increased translation output (Figure 3.15B) but only PetB synthesis level is increased in pulse labeling, one potential explanation is that PetB protein is more stable or has a slower turnover than that of PetA. Still speculative, this hypothesis could explain the negligible accumulation of Cyt *b₆f* (Schöttler et al., 2017) despite the increase in the synthesis of two of its core subunits.

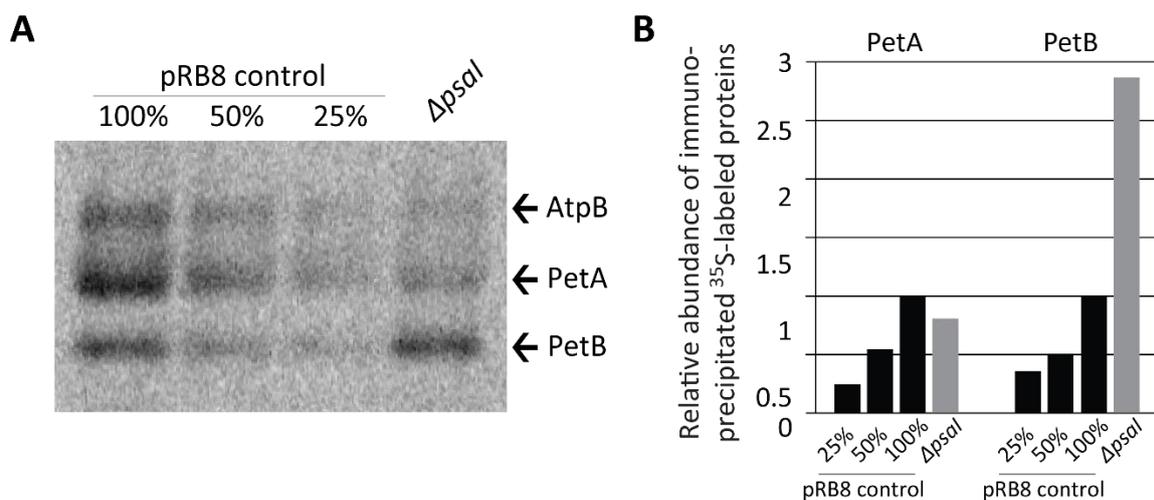


Figure 3.16: PetB synthesis is upregulated in $\Delta psal$.

A. Chloroplast proteins were radiolabeled with ^{35}S -methionine and -cysteine for 20 min and thylakoid proteins were extracted. An equivalent to 200,000 incorporated cpm was used for immunoprecipitation with antibodies against AtpB, PetA and PetB subunits (indicated on the right). Precipitated proteins were separated by SDS-PAGE and detected by autoradiography. **B.** Signals of the detected bands were quantified using Image Lab software (Bio-Rad). PetA and PetB signals in ΔpsaI were normalized to the ratio of AtpB signal in the pRB8 control to that in ΔpsaI mutant.

3.1.17 Enhancement of *petB* translation output in ΔpetL mutant

PetL is a non-essential chloroplast-encoded subunit of the Cyt *b₆f* complex located at the periphery of the complex (Hasan et al., 2013; Schöttler et al., 2007). PetL was shown to play a role in the stabilization of the Cyt *b₆f* complex in mature and old leaves but is not required for the proper function and accumulation of the complex (Schöttler et al., 2007). The transcript abundance and the translation of the plastid-encoded genes were investigated in a *petL* knockout mutant (Schöttler et al., 2007). Interestingly, this mutant exhibits a strong overexpression of *petG* likely due to the read-through from the upstream *aadA* cassette inserted in sense direction (Figure 3.17A). PetG is an essential subunit, indispensable for the stability and accumulation of the Cyt *b₆f* complex (Schwenkert et al., 2007). In addition, the transcript profiling showed overexpression of *psaJ*, *rps18* and *rpl33*, all of which are located downstream of the *aadA* insertion site (Figure 3.17B). Consequently, the footprint abundance of the aforementioned genes increased by more than twofold. Surprisingly, *petL* was overexpressed on transcript level probably due to the production of a fused transcript containing *aadA* and the 3' region of *petL* that can partially anneal to the *petL* probes on the microarray. As expected by the interruption of the *petL* reading frame, the translation output of *petL* decreased drastically by ~ 3.5-fold (Figure 3.17B). Remarkable, beside these direct consequences of the genomic location and expression of the *aadA* resistance cassette, the translation output of *petB* increased by ~ 2.3-fold and that of *petA* by 1.7-fold (Figure 3.17B), which raises the question whether the overexpression of an essential subunit triggers the overexpression of other subunits of the Cyt *b₆f* complex through positive feedback regulation. However, this increase in the translation output of the Cyt *b₆f* complex subunits is not reflected at the protein accumulation level (Schöttler et al., 2007; Schwenkert et al., 2007), which suggests that the observed overexpression might be an artifact. Potential concerns regarding the normalization to the pRB8 control are discussed below (section 4.1.8).

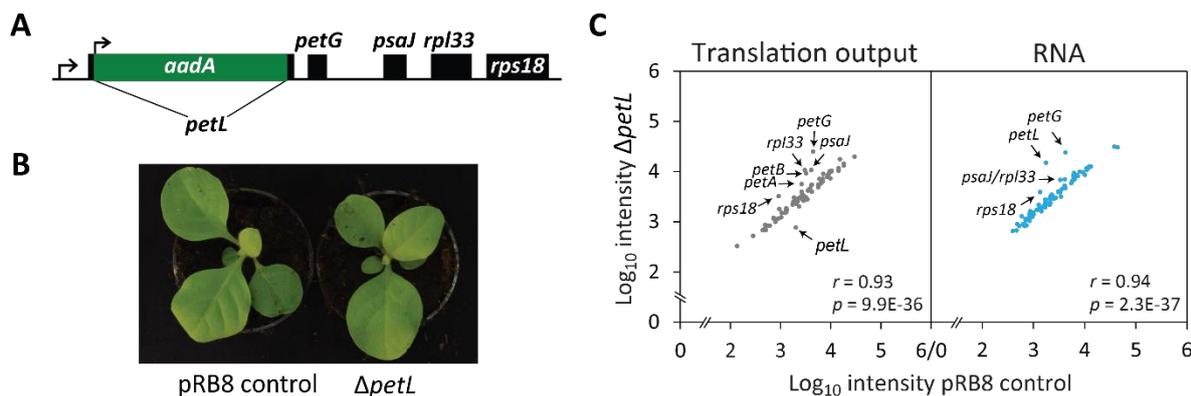


Figure 3.17: Overexpression of *petG* enhances *petB* translation.

A. Physical gene map of the *petL* knockout line that induces overexpression of the downstream located *petG* (Schöttler et al., 2007). Refer to Figure 3.4A for labeling details. **B.** The $\Delta petL$ mutant was grown beside the pRB8 control under controlled conditions (section 2.2.1.8). $\Delta petL$ had no visible phenotype. **C.** Chloroplast ORF average of translation output and RNA abundance was calculated and plotted. Results were obtained from two biological replicates. For labeling details, refer to Figure 3.4E.

3.1.18 Lack of CES interaction between AtpB and AtpA and potential side effects upon reduction of AtpC

In plants, the thylakoid CF0-CF1 ATP synthase complex catalyzes the ATP synthesis fueled by the translocation of protons from the lumen to the stroma (Seelert et al., 2003). A special feature of the ATP synthase complex that differs from all the other photosynthetic complexes is that its subunits are not found in a 1:1 ratio (Schöttler et al., 2015). The chloroplast-encoded subunits of the ATP synthase complex (AtpA, AtpB, AtpE, AtpF, AtpH, AtpI) have a 3: 3: 1: 1: 14: 1 ratio (Groth and Pohl, 2001; Vollmar et al., 2009). It was previously reported that the translation output but not the transcript abundance of these subunits follows their stoichiometric ratio (Chotewutmontri and Barkan, 2016; Trösch et al., 2018). This finding accentuates the relevance of the translation regulation in the synthesis of the ATP synthase subunits. Although the assembly of the ATP synthase complex in embryophytes is not yet understood in detail, a CES cascade that fine-tunes the stoichiometric production of the ATP synthase complex was described in *Chlamydomonas*. Drapier et al. (2007) have shown that AtpC is required for the translation of *atpB* which, in turn, is required for the production of AtpA. Based on the *Chlamydomonas* CES cascade, AtpB and AtpC mutants were analyzed in this work to investigate whether a similar regulation occurs in embryophytes.

The gene expression in *atpB* knock-down mutants with a mutated start codon (ATG to GTG) (Rott et al., 2011) was previously investigated, and it was shown that despite a fivefold decrease in the footprint accumulation of *atpB*, the translation of the other subunits in ATP synthase complex was virtually unchanged (Trösch et al., 2018). However, based on our results in PSII assembly regulation (section 3.1.7.2), earlier developmental stages may exhibit stronger CES-like regulation. Hence, I tested if the *atpB* mutant shows a defect in the translation of other ATP synthase subunits at an early developmental

stage. As a control, plants with only the *aadA* resistance marker inserted at the same position were used. Apart from the primary defect of *atpB*, only the translation output of *psbA* was altered (Supplemental Figure 4) which is likely an indirect effect of the decreased ATP synthase activity in this mutant (Rott et al., 2011; Zoschke et al., 2013). This result further confirmed the translation uncoupling between *atpB* and its downstream, partially overlapping gene, *atpE*, as no defect at the footprint level was observed for *atpE* (Supplemental Figure 4B). Furthermore, the translation output of *atpA* remained unchanged (Supplemental Figure 4B) which confirmed the absence of a translation feedback regulation between *atpB* and *atpA*.

An additional mutant that knocks down the nucleus-encoded *AtpC*, the first subunit of the *Chlamydomonas* CES cascade in the ATP synthase complex, was included in this work (Rott et al., 2011). The *AtpC* gene encodes the essential γ -subunit whose availability was shown to control the biogenesis of the ATP synthase complex in *Chlamydomonas* (Drapier et al., 2007). This mutant was obtained from a collection of antisense lines in tobacco (Lein et al., 2008; Rott et al., 2011). Wild-type plants (cultivar SNN) were used as control (Figure 3.18A). The immunoblot analysis showed ~ fourfold decrease in the abundance of the AtpB subunit, which reflects a defect in the accumulation of the whole ATP synthase complex (Figure 3.18B). Ribosome and transcriptome profiling revealed a pronounced defect in the transcript abundance of *atpH*, *atpF*, and *psbN* in the as-*AtpC* mutant (Figure 3.18C). The transcript abundance of *atpF* and *atpH* was further investigated by RNA hybridization analysis (Figure 3.18D). *atpF* and *atpH* genes are part of a polycistronic transcription unit, which produces a highly complex band pattern. Northern blot analyses with *atpF*-specific probe and *atpH*-specific probe showed an overall decrease in the expression of all the transcripts produced from this transcription unit (Figure 3.18D). Interestingly, a band of ~ 1.2 kb was observed in both blots, which probably correspond to a dicistronic *atpH*-spliced *atpF* transcript (Figure 3.18D). The ribosome profiling analysis showed a clear defect in the translation of *atpA* (fourfold) which hints to a potential CES regulation of *atpA* without the involvement of AtpB. Additionally, a major decrease in the footprint accumulation was also observed for *atpH* (5.6-fold) and *atpF* (3.5-fold) (Figure 3.18C). However, the alterations in translation output were not restricted to ATP synthase components: while most of the chloroplast-encoded genes showed a decrease in their translation output, a clear increase in the synthesis of the subunits of NDH complex, the RNA polymerase subunits, *ycf10*, and *matK* was observed (Supplemental Figure 5). Based on this, the defects observed for *atpA* and *atpH* translation output do not necessarily result from translation feedback regulation in response to the *AtpC* knockdown as they could be secondary effects of the energy depletion caused by the lack of the ATP synthase complex.

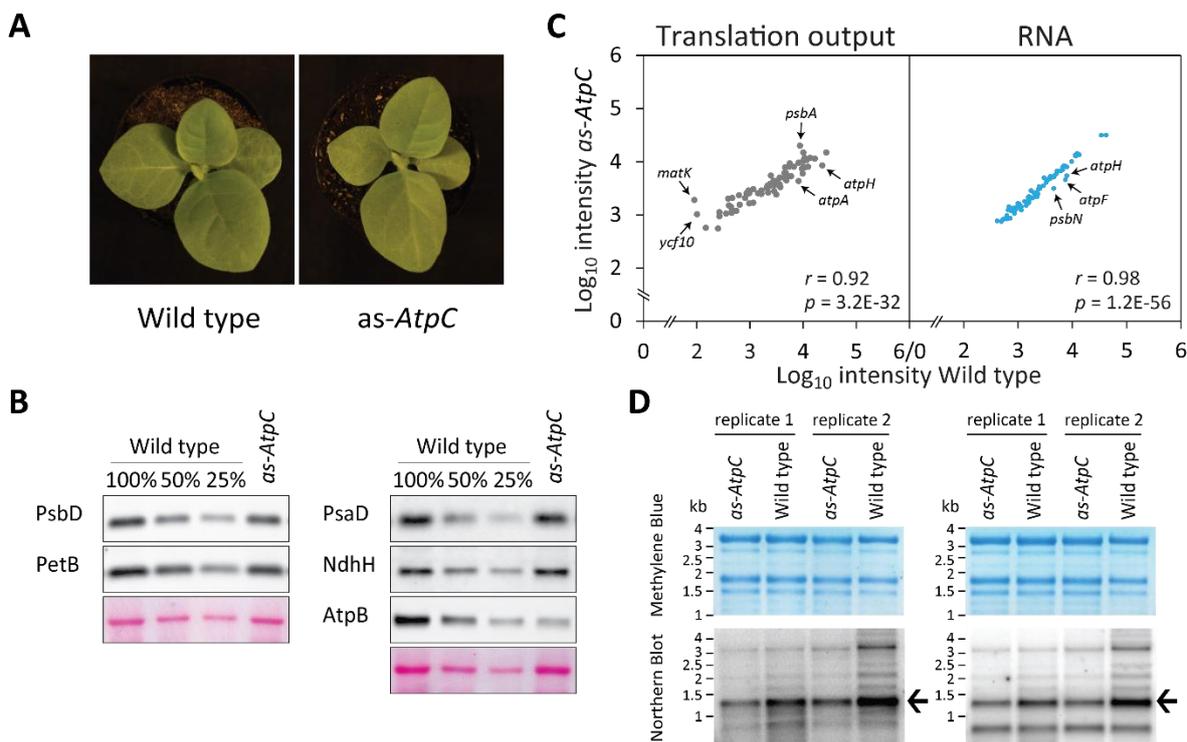


Figure 3.18: Reduced AtpC levels causes alterations in the expression of several chloroplast-encoded genes.

A. The *as-AtpC* mutant was grown together with wild type cultivar SNN under standard conditions (section 2.2.1.8). The mutant displayed a pale green phenotype and retarded growth. **B.** Immunoblot analysis of the core subunits of the photosynthetic complexes in *as-AtpC* and the wild type. Total proteins were extracted and separated by SDS-PAGE. Results were obtained from three biological replicates, one representative replicate is shown. **C.** ORF average of footprint abundance and RNA level for the chloroplast-encoded genes were calculated and plotted as described in Figure 3.4E. Results were obtained from two biological replicates. **D.** Northern blot experiment utilizing a probe specific for the *atpH* CDS (left) and *atpF* CDS (right). The methylene blue staining is shown as a loading control. RNA sizes are labeled on the left side of the gel. The black arrow indicates the *atpH-atpF* dicistronic transcript.

3.1.19 Study of gene expression in NDH dehydrogenase mutants does not reveal any substantial feedback regulation

In embryophytes, the nonessential NAD(P)H dehydrogenase-like (NDH) complex forms a supercomplex with PSI and thereby mediates the cyclic electron transport among other suggested functions (Endo et al., 2008). By doing so, the NDH complex enables lowering the photodamage that can result from stromal over-reduction (Burrows et al., 1998; Shikanai et al., 1998). The chloroplast NDH complex consists of 11 subunits, which are homologous to the genes of the respiratory complex I (Ohyama et al., 1986; Shinozaki et al., 1986). The fact that *ndh* genes exist in all photosynthetic embryophytes whereas they are absent in algae including *Chlamydomonas* (Martín et al., 2015), suggests a role of the NDH complex in the land adaptation of photosynthesis. The NDH complex

subunits are distributed in five subcomplexes: A, B, L, EDB, and M (Shikanai, 2016). To investigate whether the stoichiometry of the NDH complex is coordinated by CES feedback regulation, two knockout mutants were analyzed in this work. First, the triple knockout mutant *ndhC/K/J* was selected (Hager, 2002). Genes encoding the NdhC subunit from the subcomplex M and NdhK and NdhJ subunits from subcomplex A are located in the same polycistronic unit, enabling the generation of a triple mutant by insertion of one *aadA* cassette (Figure 3.19A). This insertion resulted in the deletion of the *ndhK* sequence and deletions in the C-terminal and N-terminal coding sequences of *ndhC* and *ndhJ*, respectively (Figure 3.19A). pRB8 plants were used as control and were grown next to the mutant under controlled conditions (section 2.2.1.8) (Figure 3.19B). Transcript accumulation of *ndhC* in the triple mutant was unaffected reflecting that an *ndhC*-containing transcript (lacking part of the 3' end of *ndhC*) is produced (Figure 3.19C). Given the low expression level of *ndhK*, the assessment of its transcript abundance using the microarray was technically challenging. For that reason, and despite the complete knockout of *ndhK* gene, only a mild decrease of the transcript abundance of *ndhK* was observed. Conversely, *ndhJ* was overexpressed on RNA level, most likely due to read-through from the *aadA* cassette, overexpressing the 3' end of the reading frame (Figure 3.19C). The translation output of *ndhK* decreased by 17.5-fold and that of *ndhJ* by ~ twofold (Figure 3.19C). However, *ndhC* was only mildly affected on footprint level (Figure 3.19C) suggesting that the truncated *ndhC* transcript is translated as evidenced by the footprints signal detected from the remaining region (data not shown). Overall, apart from the primary defect, no effects were observed for any of the other NDH complex subunits or other chloroplast-encoded proteins. This result suggests that the reduced translation output of *ndhK* and *ndhJ* does not trigger any CES regulation within the NDH complex.

The second mutant that was used to assess a potential CES regulation in the NDH complex is the triple knockout mutant *ndhA/H/I* (Kofer et al., 1998) (Figure 3.19E). In this mutant, NdhA subunit from the subcomplex M and NdhI and NdhH subunits from subcomplex A were deleted by insertion of an *aadA* cassette (Figure 3.19D). The insertion of the *aadA* cassette caused the complete deletion of *ndhI* and *ndhA* and the removal of the C-terminal of *ndhH* (Figure 3.19D), which caused the expected pronounced reduction in RNA abundances (Figure 3.19F). Additionally, the mutant showed a decrease in the transcript abundance of *rps15*, which is possibly caused by the insertion of the *aadA* cassette leading to an overexpression of the antisense RNA of *rps15* (Figure 3.19F). In addition to the primary effects, a decrease in the translation outputs of *ndhE* and *ndhG* of two and threefold, respectively, was observed. While these reduced translation outputs could hint to a possible CES effect caused by the reduced synthesis level of either NdhI, NdhA, NdhH or Rps15, it is also possible that it is a direct effect of the *aadA* insertion, given that *ndhE* and *ndhG* are located in close proximity to the *aadA* insertion site (in the same transcription unit) (Figure 3.19D). A decrease was also observed in the transcript abundance of *ndhG* and *ndhE*, which further supports the possibility that the effect observed on the footprint

abundance of these two genes is a direct effect of *aadA* insertion. Future northern blot experiments are required to validate this possibility.

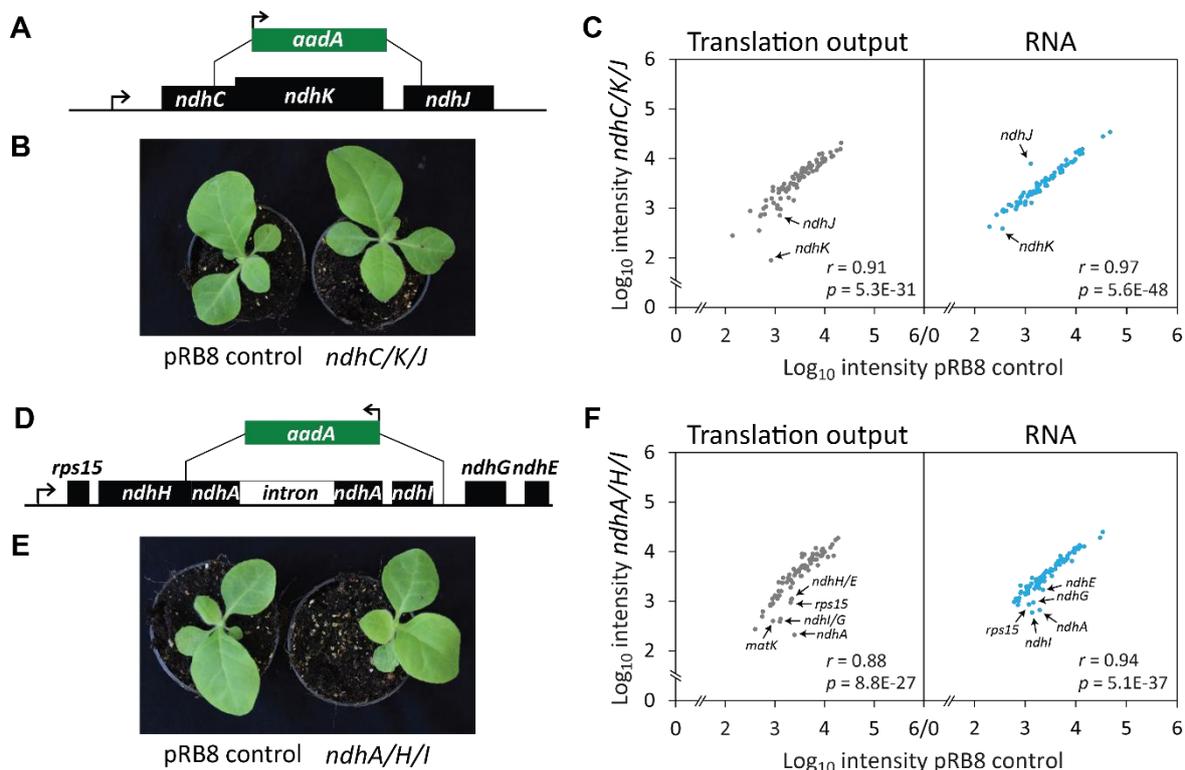


Figure 3.19: Lack of substantial CES interactions in *ndhC/K/J* and *ndhA/H/I* mutants.

A, D. Physical map of the *ndhC/K/J* mutant (**A**) and *ndhA/H/I* mutant (**D**) showing the insertion site of the *aadA* cassette. For labeling details see Figure 3.4A. **B, E.** The *ndhC/K/J* (**B**) and *ndhA/H/I* (**E**) mutants were grown next to the pRB8 control under controlled conditions (section 2.2.1.8). The aerial part was used for total RNA and ribosome footprints isolation. **C, F.** Comparison of the translation output and the RNA abundance in the *ndhC/K/J* mutant (**C**) and *ndhA/H/I* mutant (**F**) to the pRB8 control. Results were collected from one biological replicate for the *ndhC/K/J* mutant and two biological replicates for *ndhA/H/I* mutant. Labeling details are given in Figure 3.4E.

3.1.20 Ribo-seq enables analysis of translation in knockout mutants of photosynthetic subunits

In all instances reported in *Chlamydomonas*, CES regulation was identified in heterotrophically grown knockout mutants of photosynthetic subunits. In order to rule out that mild effects are overlooked in the autotrophically grown knockdown mutants analyzed above, I examined potential CES regulation in three knockout mutants of PSII, Cyt *b₆f*, and ATP synthase complexes, namely $\Delta psbD/C$, $\Delta psbB$ operon and $\Delta atpB$ (Figure 3.20). These mutants were previously created using *aadA* cassette insertion (see following sections for details) and were grown heterotrophically on MS medium supplemented with 3 % sucrose (section 2.2.1.8.3) under low light (5 to 10 $\mu\text{mol m}^{-2} \text{s}^{-1}$). pRB8 plants grown under the same conditions were used as control. Since these mutants grow very slowly and have very thin

chlorotic leaves, it was not practicable to isolate an amount of ribosome footprints that is sufficient for the microarray approach. For this reason, next-generation sequencing-based ribosome profiling, which requires much less tissue, was used to assess translational regulation in these mutants. Single-end 75-bp sequencing on Illumina NextSeq 500 was applied and datasets from two biological replicates were collected for each line. The numbers of raw counts and mapping statistics of each dataset are listed in Supplemental Table 5. It is important to note that I optimized Ribo-seq of chlorotic tissue until the end of my Ph.D. and therefore the results presented here are still preliminary and only comprise data for chloroplast genes. Further in-depth and thorough data analysis will be performed in the future. In addition, RNA-seq analysis still needs to be performed to determine the transcript accumulation in these mutants and eventually assess changes in translation efficiency.

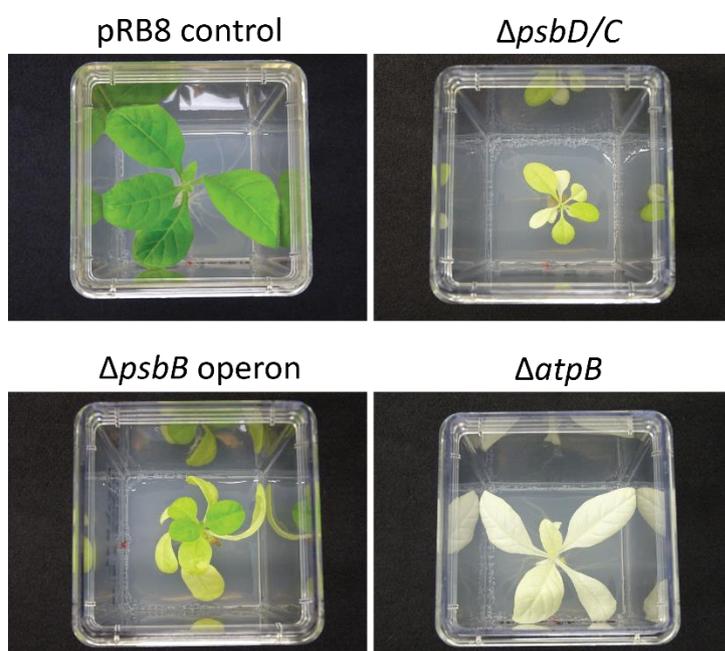


Figure 3.20: Phenotype of the knockout mutants.

The mutants and the corresponding pRB8 control were grown under heterotrophic conditions (section 2.2.1.8.3). Leaves were harvested 30 min after the onset of the light phase and were used for ribosome footprint isolation.

3.1.20.1 Deletion of the C-terminus of PsbD induces the CES regulation of PsbA but not of PsbB

The knockout *psbD/C* mutant ($\Delta psbD/C$) was created by inserting an *aadA* cassette in the ORFs of *psbD* and *psbC* thus replacing the last 315 bp of the *psbD* ORF and first 948 bp of the *psbC* ORF (Figure 3.21A). The ribosome footprint isolation and libraries were prepared as described in sections 2.2.4.1 and 2.2.4.3, respectively. Reads per kilobase per million (RPKM) values of the chloroplast ORFs were calculated and represented the translation output of the respective ORF (section 2.2.4.4). The relative changes in the translation output in $\Delta psbD/C$ for each chloroplast ORF in comparison to the pRB8 control were calculated as described before (section 2.2.4.4). Both replicate experiments showed a reduction of *psbC* translation output to background levels, confirming the complete lack of *psbC* translation due to the removal of the start codon (Figure 3.21B, C). In contrast, for *psbD*, footprints mapping to the 5' end of the ORF were detected, suggesting that a truncated PsbD protein is synthesized

(Supplemental Figure 6). Interestingly, the mutant exhibited a specific sevenfold decrease in the translation output of *psbA* (Figure 3.21B). In order, to clearly visualize the effect on the ribosome footprint abundance of the other PSII subunits and PSI subunits, the ratios of the relative RPKM changes in the $\Delta psbD/C$ mutant in comparison to the pRB8 control were plotted (Figure 3.21C). In addition to *psbA*, the translation output of *psbE* decreased more than twofold (Figure 3.21C). Comparison of these results with KD-*psbD* mutant (section 3.1.7.2) revealed a discrepancy regarding the PSII subunits whose translation output is affected. The knockdown of the full-length *psbD* transcript caused a decrease in the translation output of *psbA* and *psbB* (Figure 3.6) whereas the truncation of the C-terminus of PsbD didn't affect the footprint abundance of *psbB* (Figure 3.21C). This finding substantiates PsbA as a CES subunit, however, it argues against the possibility that PsbA is required for the efficient synthesis of PsbB. Furthermore, this result questions whether PsbD, specifically its C-terminus, is required for the translational activation of *psbB*. Based on these findings, a new model of the intricate CES regulation of PSII in land plants is discussed below (section 4.1.2).

Ribosome profiling captures primarily defects in translation initiation (Ingolia et al., 2018). To rule out the possibility that the observed defect on *psbA* translation is caused by altered translation elongation, the general distribution of the ribosome footprints along the *psbA* mRNA was inspected. No substantial alteration in the ribosome occupancy along the *psbA* reading frame was observed (Figure 3.21D), which implies that mainly a defect in translation initiation accounts for the overall decrease in the footprint level of *psbA* in the $\Delta psbD/C$ mutant.

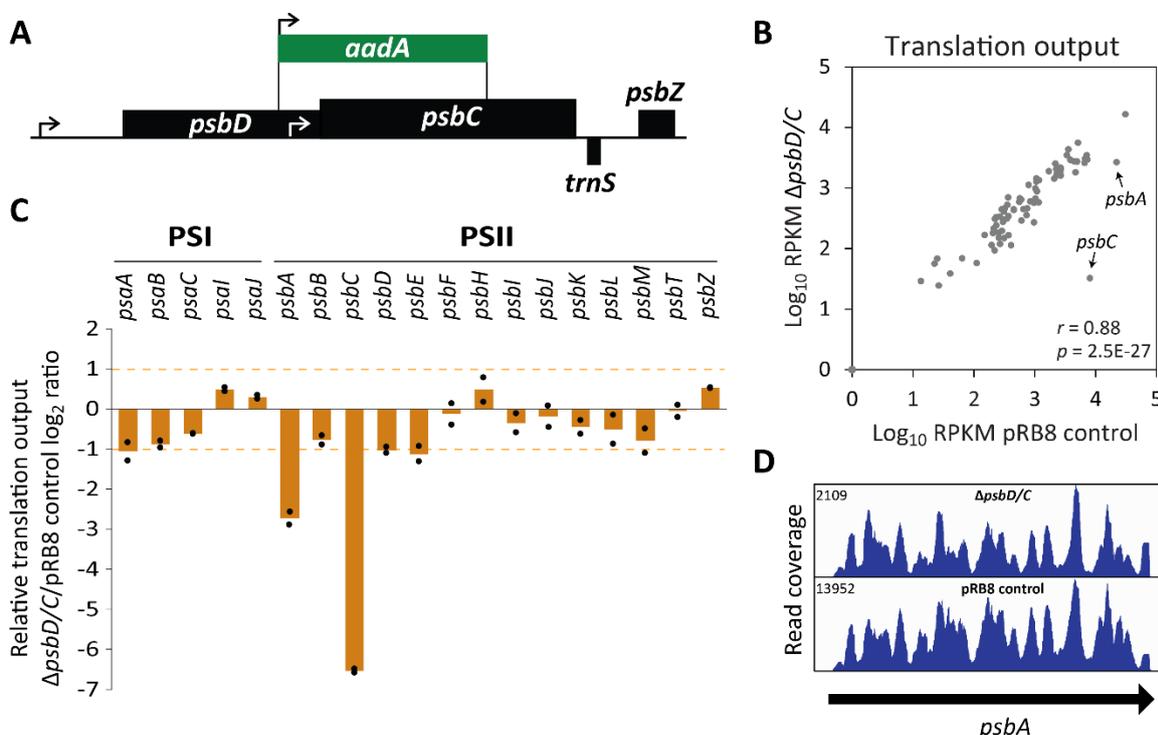


Figure 3.21: Truncation of the C-terminus of PsbD induces a decrease in the translation output of *psbA*.

A. Physical gene map of the mutant showing the insertion site of the *aadA* selection marker. Labeling details are given in Figure 3.4A. **B.** Comparison of the ribosome footprint abundance of all chloroplast ORFs in $\Delta psbD/C$ to that in pRB8 control. Data are collected from two biological replicates and the RPKM values of each chloroplast ORF are plotted. Pearson's *r* value and Anova's *P* value are shown. **C.** Ratio of the Ribo-seq reads in $\Delta psbD/C$ mutant relative to pRB8 control for chloroplast-encoded PSI and PSII subunits. Values from individual biological replicates are shown as black dots. **D.** Screenshot from the Integrated Genome Viewer (IGV) showing the ribosome footprints distribution along the *psbA* ORF in $\Delta psbD/C$ mutant and pRB8 control. The y-axes represent the number of reads and were adjusted to facilitate the comparison. The maximum y-axis values are shown in the upper left corner.

3.1.20.2 The translation of *petA* does not require the availability of PetD and PetB

Results presented in this work (sections 3.1.15 and 3.1.16) propose a new translational regulatory mechanism to fine-tune the stoichiometry within the Cyt *b₆f* complex that has not been shown in *Chlamydomonas*. However, the presence of positive feedback between PetA and PetB does not preclude the possibility of an additional negative feedback regulation reminiscent to that reported in *Chlamydomonas* (Boulouis et al., 2011; Choquet et al., 1998). To address whether PetA is a CES subunit whose synthesis depends on the availability of PetD or PetB, a previously created knockout mutant of the *psbB* operon was investigated. In this mutant, a large region of the chloroplast genome that includes the genes *psbB*, *psbT*, *psbN*, *psbH* from PSII and *petB*, and the first 286 bp of *petD* from the Cyt *b₆f* complex was replaced with an *aadA* cassette (Figure 3.22A). This mutant displayed a strong growth defect (even if grown heterotrophically) and, consequently, was examined by Ribo-seq. Chloroplast translation in $\Delta psbB$ operon was compared to that of the pRB8 control. As expected, the translation output of all the knocked out genes in the *psbB* operon decreased (Figure 3.22B). The translation output of *clpP* displayed a pronounced decrease, which is most likely due to the read-through of the *aadA* cassette producing an antisense transcript of *clpP*. In addition to the primary defect, the mutant displayed a pronounced decrease in the translation output of *psbA*. Previous reports have shown that the *psbH* mutant (*hcf107* T-DNA insertion mutant) which lacks PSII showed a reduction of *psbA* translation (Felder et al., 2001; Williams-Carrier et al., 2019). Also, in the present work, I showed that the $\Delta psbN$ mutant caused an alteration in the ribosome distribution along *psbA* mRNA at early developmental stages (section 3.1.12). Taken together, the trigger of the translation defect of *psbA* could be an assembly defect of PSII caused by the absence of the following PSII subunits: PsbB, PsbH, and PsbN. Regarding the Cyt *b₆f* complex, no effect was observed on *petA* ribosome footprint abundance and distribution (Figure 3.22B, C) despite the knockout of *petB* and *petD*, which suggests that its translation is independent of the presence of PetB or PetD. This result further supports that PetA is not a CES subunit whose synthesis is dependent on its assembly status. Unfortunately, this mutant could not be used to further investigate the potential translational co-regulation of *psbB* and *psbH* that was described in *Arabidopsis* (Felder et al., 2001; Levey et al., 2014) as it lacks both subunits.

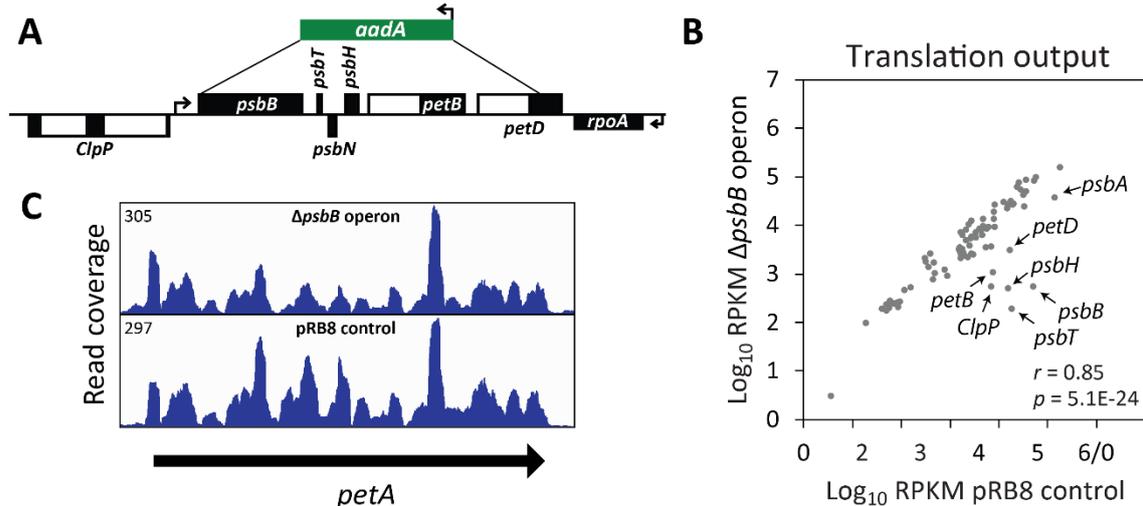


Figure 3.22: Translation of *psbA* is compromised by the PSII assembly defect in $\Delta psbB$ operon mutant whereas *petA* translation is independent of PetB or PetD.

A. Physical map of the mutant. The *aadA* cassette replaced *psbN* and multiple genes in the *psbB* operon including *psbB*, *psbT*, *psbH*, *petB*, and the 5' end of the *petD* ORF. Labeling details are given in Figure 3.4. **B.** Comparison of the ribosome footprint abundance in the $\Delta psbB$ operon mutant to that in the pRB8 control. Results were obtained from two biological replicates. **C.** IGV screenshot showing the ribosome footprints distribution along the *petA* ORF in $\Delta psbB$ operon mutant and pRB8 control. Labeling details are explained in Figure 3.21D.

3.1.20.3 Knockout of *atpB* causes a general translation defect in the chloroplast

In order to further check for CES regulation in the ATP synthase complex, I included a knockout mutant where *atpB* is replaced with an *aadA* cassette. The replaced fragment begins 56 bp downstream of the *atpB* start codon and ends 170 bp downstream of the start codon of *atpE* (Figure 3.23A). This mutant was completely white and had to be grown heterotrophically (Figure 3.20). Ribo-seq was performed to study translation. As expected, the translation output of *atpB* decreased substantially (Figure 3.23B, C). The inspection of the ribosome footprints distribution along *atpB-atpE* transcript showed a complete lack of *atpB* translation (Figure 3.23D). In contrast, ribosome footprints mapped along the *atpE* ORF downstream of the *aadA* cassette insertion site (Figure 3.23D). This implies that a short *atpE* isoform could be produced, however less efficiently. This finding is in line with the presence of an internal SD sequence within *atpE* mRNA reflecting that translation could initiate within *atpE* (Hirose and Sugiura, 2004a). On the other hand, the translation output of all the other subunits of the ATP synthase complex decreased by more than twofold, with *atpF* being the most affected (after *atpE*) (~ fivefold) (Figure 3.23C). In addition to the effects on subunits from the ATP synthase complex, and in agreement with the data from the as-*AtpC* mutant (section 3.1.18), effects on other genes were also observed. The translation output of almost all photosynthetic subunits decreased, except *ndhB* whose translation was enhanced (Supplemental Figure 7). Furthermore, the Ribo-seq reads of the *rpo* genes and some ribosomal proteins genes increased (Supplemental Figure 7). Such expression pattern has been observed

in other mutants with defective chloroplast gene expression (unpublished data, personal communication Dr. Reimo Zoschke, MPIMP). Hence, it is challenging to assess whether the defects observed on the other subunits of the ATP synthase complex are secondary effects or caused by translation feedback regulation.

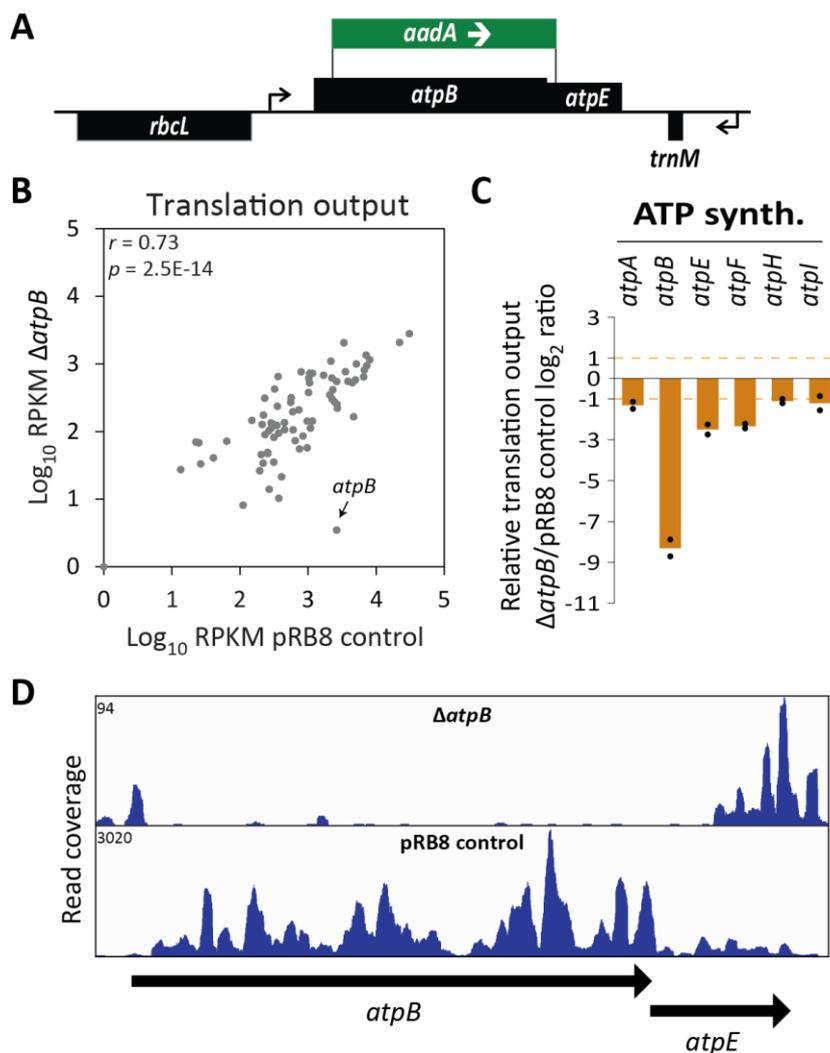


Figure 3.23: The knockout of the ATP synthase complex results in global alteration of translation in the chloroplast.

A. Physical map of the mutant showing the insertion site of the *aadA* cassette (labeling details as in Figure 3.4A). **B.** Comparison of the translation output in $\Delta atpB$ mutant to that in pRB8 control. Results were obtained from two biological replicates. **C.** Ratio of the Ribo-seq reads in $\Delta atpB$ mutant relative to pRB8 control for chloroplast-encoded ATP synthase subunits. Values from individual biological replicates are shown as black dots.

3.2 Identification and characterization of *psbA* and *rbcL* translation factors using an aptamer-based affinity purification

In order to gain insights into the molecular mechanism of the translation feedback interactions identified in this work, attempts to identify the involved factors has been made. Different approaches have been employed to unravel the RNA-binding proteomes of specific mRNAs in the chloroplast. Most recently RIP-seq of engineered PPR proteins that target specific transcripts in the chloroplast was used (McDermott et al., 2019). Furthermore, biotinylated antisense oligonucleotides were also employed to purify the factors that bind to *psbA* mRNA (Watkins et al., 2019). Here, an aptamer-based affinity purification approach was adapted for chloroplast transcripts. Three RNA affinity tags, the D8 aptamer (referred to as Sephadex-binding aptamer hereafter), the S1 aptamer (referred to as streptavidin-binding

aptamer hereafter), and the MS2 RNA element (referred to as MS2 aptamer hereafter) were selected to tag the chloroplast *psbA* and *rbcL* mRNAs. Most of the identified translation factors bind to the 5' UTR of their target (Zoschke and Bock, 2018), hence the aptamers were inserted into the 3' UTR of the transcripts to avoid disrupting the binding sites of these factors. Transplastomic plants with the Sephadex- and streptavidin-binding aptamers were previously created by Dr. Reimo Zoschke (MPIMP). MS2 tagged plants were created in this work and affinity purification with all aptamers was optimized.

3.2.1 Establishment of an MS2 aptamer-based affinity purification in chloroplast

3.2.1.1 Insertion of aptamers does not disrupt the conformation of the 3' end of *psbA* and *rbcL* mRNAs *in silico*

Given that RbcL is a CES subunit and PsbA was found to be a potential CES subunit in this work (sections 3.1.7 and 3.1.20.1), *rbcL* and *psbA* mRNAs were chosen to be tagged with MS2 aptamer. Both of these transcripts are monocistronic and abundant in the chloroplast, which facilitates their enrichment and verification of the co-purified proteins. In addition, they possess a stem-loop in their 3' UTR (Figure 3.24), which protects the inserted aptamer from being cleaved off. A major concern in the design of the chloroplast transformation vector was not to disrupt the expression of the targeted genes, *psbA* and *rbcL*, which might be deleterious or even lethal given that these proteins are both essential for autotrophic growth. Prior to the design of the transformation vectors, the Mfold server (Zuker, 2003) was used to check how the insertion of the MS2 aptamer would affect the folding structure of the 3' UTR of *psbA* and *rbcL*. The Sephadex- and streptavidin-binding aptamers were used as controls. As shown in Figure 3.24, the *in silico* analysis revealed that the insertion of either of the aptamers into the designated location results in a larger loop structure placed at the end of a longer stem. Thus, this secondary structure prediction indicated that the MS2 aptamer would disrupt the conformation of the 3' end of *psbA* or *rbcL* *in silico*. In addition, plants with Sephadex- and streptavidin-binding aptamers inserted into the 3' end of either *psbA* or *rbcL* did not display a visible phenotype. Hence, I expected that the insertion of the MS2 aptamer in the same position would not be deleterious.

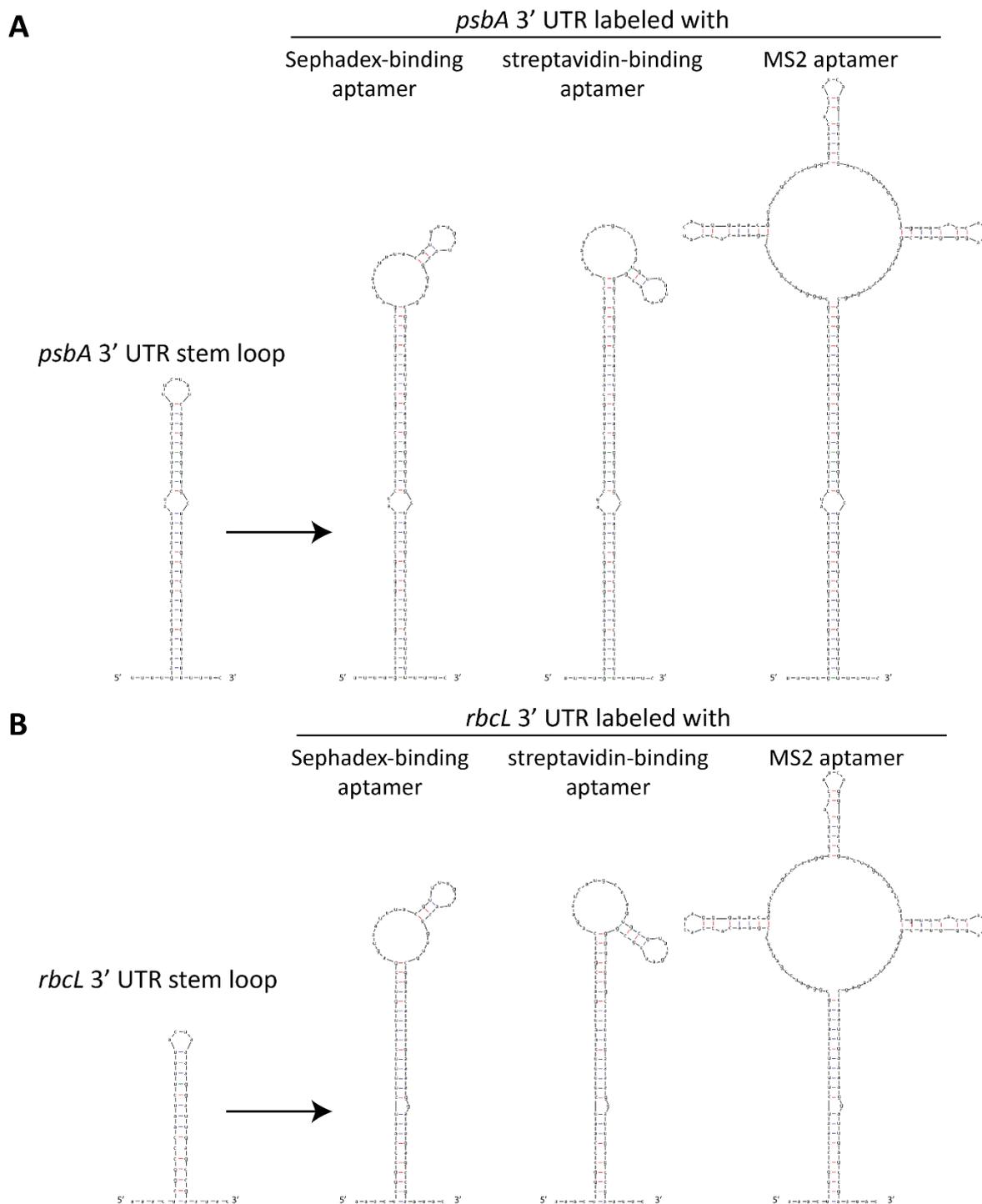


Figure 3.24: Insertion of the selected aptamers does not disrupt the structures of the 3' UTR of *psbA* and *rbcL* *in silico*.

A. From left to right: Mfold-predicted structures of the *psbA* 3' UTR and the *psbA* 3' UTR including inserted Sephadex-binding, streptavidin-binding, and MS2 aptamer sequences (Zuker, 2003). **B.** From left to right: Mfold-predicted structures of the *rbcL* 3' UTR and the *rbcL* 3' UTR including inserted Sephadex-binding, streptavidin-binding, and MS2 aptamer sequences (Zuker, 2003).

3.2.1.2 Tagging of the *psbA* and *rbcL* transcripts with the MS2 aptamer

The MS2 aptamer is one of the most frequently used natural aptamer for affinity purification of RNPs in other systems (e.g., yeast, humans) (Fica et al., 2019; Jurica et al., 2002; Said et al., 2009; Zhou et al., 2002). The MS2 system was widely exploited to examine the localization of RNAs *in vivo* (Forrest and Gavis, 2003; Liu et al., 2005; Sheth and Parker, 2003; Wang et al., 2012) and even to track the translation of a single transcript (Morisaki et al., 2016), however, its use was not extended to chloroplasts. In this work, the MS2 aptamer was inserted into the 3' UTR of *psbA* and *rbcL* by chloroplast transformation. Map of the chloroplast genome of the aptamer-tagged *psbA* and *rbcL* lines is represented in Figure 3.25A, bottom. For MS2 cloning, genomic DNA (gDNA) fragments amplified from the Sephadex-tagged plants were used (section 2.2.2.2). The amplified fragments contained the sequence of the *aadA* gene that confers resistance against spectinomycin placed 254 bp and 378 bp downstream of the stop codon of *psbA* and *rbcL*, respectively. In addition, the Sephadex-binding aptamer sequence and ~ 400-500 bp flanking sequences identical to the gDNA were included in the amplified fragments for insertion of the transgenes by homologous recombination (Figure 3.25). The Sephadex-binding aptamer was cleaved off by MFeI restriction enzyme and replaced with a PCR-amplified synthetic sequence harboring three copies of the MS2 aptamer. These constructs were then introduced into wild-type tobacco plastids by biolistic transformation (section 2.2.1.9). Correct integration was ensured by homologous recombination between the flanking regions in the transformation vectors and the plastid genome. To test for potential effects of the expression of the spectinomycin resistance gene in the respective genomic region, plants with only *aadA* inserted downstream of *psbA* or *rbcL* were used (Figure 3.25B). These *aadA*-control plants were previously created by Dr. Reimo Zoschke (MPI-MP) and were used as a control for all three corresponding tagged lines.

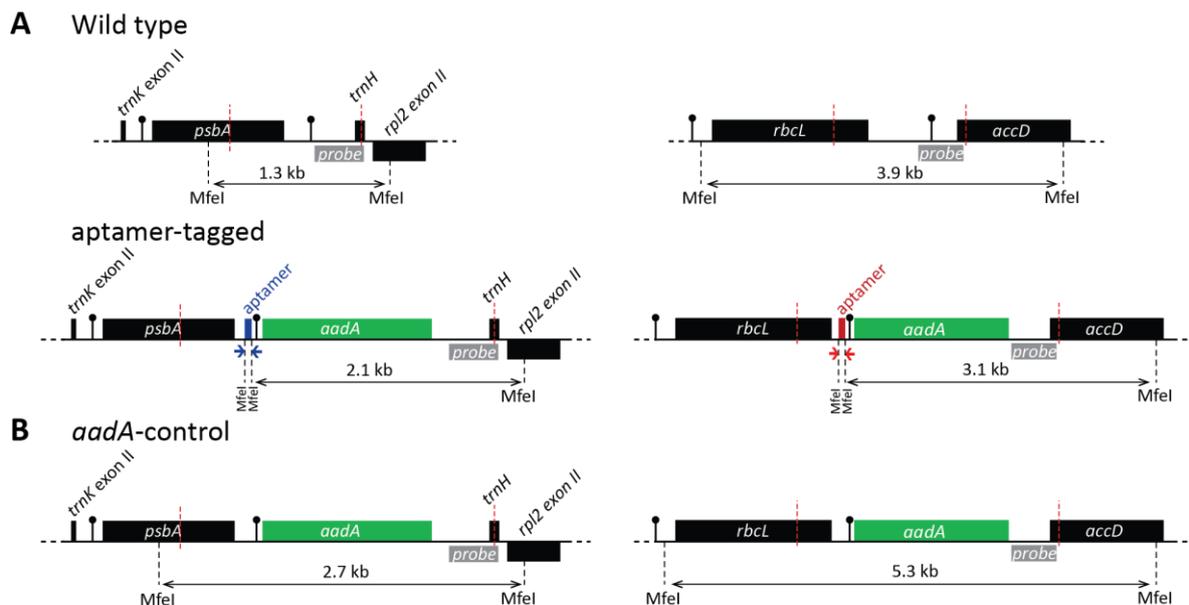


Figure 3.25: Physical map of the chloroplast genome in wild type and in *psbA* and *rbcL* transplastomic lines.

A. Physical map of the wild-type plastid genome where the aptamers and the *aadA* cassette were inserted (upper panel) and map of the transplastome containing the aptamer sequence and the selection marker inserted into the 3' UTR of *psbA* (left) or *rbcL* (right) (lower panel). **B.** Physical map of the *aadA*-control lines transformed only with the selection marker integrated downstream *psbA* (left) or *rbcL* (right). **A, B.** The restriction sites of *MfeI* used for RFLP, the location of the probe, and the expected sizes of the fragments in Southern blot analysis are indicated. Note that the *aadA* cassette contains the tobacco *P_{rrn}* promoter and the *Chlamydomonas rbcL* terminator. The black circles mark the 5' and 3' ends of *psbA* and *rbcL* transcripts. The red dashed lines mark the borders of the flanking regions. Black boxes: chloroplast genes; green box: *aadA* cassette; blue/red boxes: aptamer sequence; grey box: probe-binding region; blue/red arrows: primers for spanning PCR.

For *psbA* MS2-tagged lines (referred to as *psbA*-MS2 hereafter) and *rbcL* MS2-tagged lines (referred to as *rbcL*-MS2 hereafter), several independent lines per transformation were verified. In total, we obtained five *psbA*-MS2 and 12 *rbcL*-MS2 green homoplastomic plants. The high number of green transplastomic lines indicated already that the expression of *psbA* and *rbcL* was not impaired by the MS2 aptamer or the *aadA* insertions. To select for homoplastomy, the primary transformants were regenerated on spectinomycin-containing medium (500 $\mu\text{g}/\text{mL}$) and the emerging resistant shoots were first screened using a tag-spanning PCR (primers indicated in Figure 3.25A) to check for the integration of the MS2 sequence (data not shown). The homoplastomy and the correct integration of the MS2 aptamer and the *aadA* cassette were then confirmed by RFLP Southern blot analysis (section 2.2.2.10). For *psbA*-MS2, three independent lines were confirmed to be homoplastomic and showed a single band at 2.1 kilobases (kb) without any wild-type plastome at 1.3 kb (Figure 3.26A). Likewise, for *rbcL*-MS2 five lines were analyzed by Southern blot and showed a strong 3.1 kb band in comparison to a 3.9 kb band for the wild type (Figure 3.26B). However, a faint signal that corresponds to the wild-type DNA was observed (Figure 3.26B). These signals are often seen in RFLP analyzes of transplastomic plants

(e.g., Krech et al. (2012)) and are normally considered to derive from chloroplast DNA fragments that integrated into the nuclear genome (promiscuous DNA) (Bock and Timmis, 2008; Hager et al., 1999; Ruf et al., 2000). Based on that, the *rbcL*-MS2 lines were considered to be homoplasmic.

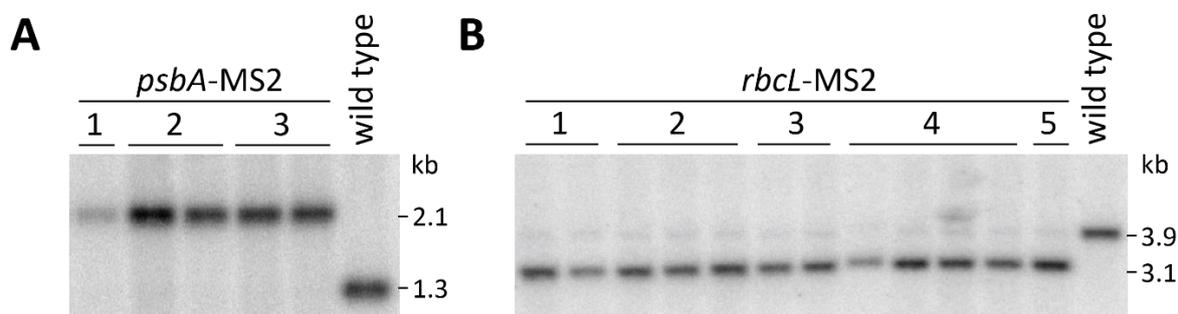


Figure 3.26: RFLP analysis of the *psbA*-MS2 and *rbcL*-MS2 primary transformants.

A. Genomic DNA from five plants deriving from three independent *psbA*-MS2 resistant lines (labeled with 1, 2 and 3) and **B.** 12 plants from five independent *rbcL*-MS2 resistant lines (labeled with 1, 2, 3, 4 and 5) was fragmented with *Mfe*I, separated by gel electrophoresis, blotted to nylon membranes, and hybridized to the probes indicated in Figure 3.25. Bands with expected sizes (Figure 3.25) are observed for *psbA* and *rbcL*.

The homoplastomy of the progeny of two independent lines for each construct was checked. To compare all tagged lines (Sephadex, streptavidin, and MS2), two representative independent lines for each aptamer tag and two lines of the corresponding *aadA*-control were included in the analysis.

For *psbA*-tagged plants, as illustrated in Figure 3.27A, the insertion of the Sephadex-, streptavidin-, and MS2-binding aptamers into the 3' end of *psbA*, together with the insertion of the *aadA* resistance cassette caused a shift of the fragment length from 1.3 kb (wild type) to 2.1 kb. The *aadA*-control lines showed a band at 2.7 kb. To further confirm the homoplastomy of all the lines, a tag-spanning PCR was performed. All the tagged lines displayed a single larger PCR product compared to the wild-type and *aadA*-control lines (according to the size of the corresponding tag, Figure 3.27A, lower panel): 286 bp for the MS2 aptamer, 223 bp for the Sephadex-binding aptamer, 234 bp for the streptavidin-binding aptamer. The *aadA*-control and the wild type showed a band at 190 bp. In addition, the transplastomic lines were tested for the segregation of the resistance marker by germination on spectinomycin (500 µg/mL). All tested *psbA* lines were uniformly resistant to spectinomycin, which confirms the stable homoplasmic state of these lines (Figure 3.27B). The phenotype of the *psbA*-tagged homoplasmic lines and the kinetic of growth at different developmental stages were assessed. Plants were transferred and grown under greenhouse conditions. All lines sustained autotrophic growth and showed a comparable phenotype to the wild type (Figure 3.28C). All the lines reached the reproductive developmental stage and produced flowers and seeds similar to the wild type.

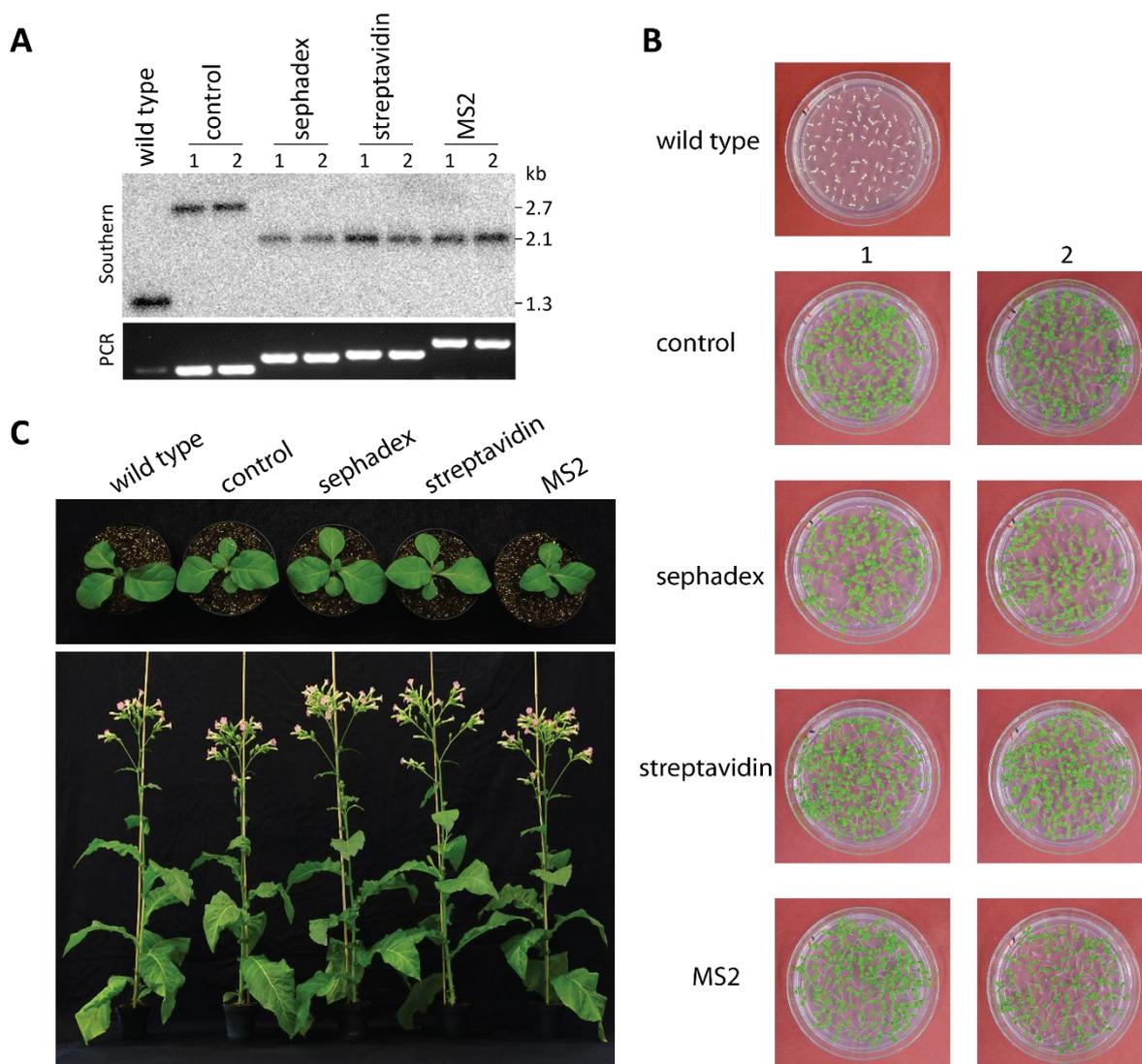


Figure 3.27: Tagging of *psbA* with three different aptamers by stable transformation of the plastome.

A. PCR and RFLP analysis of the offspring of two independent lines for each *psbA* construct. Upper panel: Southern blot analysis, gDNA was fragmented with *MfeI* restriction enzyme, separated by gel electrophoresis, immobilized on a nylon membrane, and hybridized to a radiolabeled probe indicated in Figure 3.25A. Lower panel: gDNA was used for spanning PCR with primers indicated in Figure 3.25A. **B.** Growth on spectinomycin-containing media demonstrated the uniform resistance of the offspring. All the lines are homoplastomic with no seedlings displaying wild-type bleached phenotype. **C.** Phenotype of representative *psbA* tagged lines. Four-week-old plants (top) and eight-week-old plants (bottom) grown under greenhouse conditions (section 2.2.1.8). All lines display a wild-type phenotype.

Similarly, for *rbcL*, RFLP analysis was performed, and based on the restriction sites, the tagged lines showed a band at 3.1 kb whereas the *aadA*-control and the wild type showed a band at 5.3 kb and 3.9 kb, respectively (Figure 3.28A). Furthermore, the spanning PCR resulted in the expected band sizes: 253 bp for the MS2 aptamer, 190 bp for the Sephadex-binding aptamer, 201 bp for the streptavidin-binding aptamer and 151 bp for the *aadA*-control and the wild type (Figure 3.28A). These results confirmed the homoplastomy of the T1 progeny of the *rbcL*-tagged lines. Additionally, seeds from these

lines were grown on spectinomycin-containing medium. In contrast to the wild type, all the lines were uniformly resistant to the antibiotic, further indicating their homoplasmic state (Figure 3.28B). Homoplasmic lines were transferred to soil and grown under greenhouse conditions to investigate their phenotype. All lines sustained autotrophic growth and showed wild-type-like growth phenotypes at different developmental stages (Figure 3.28C).

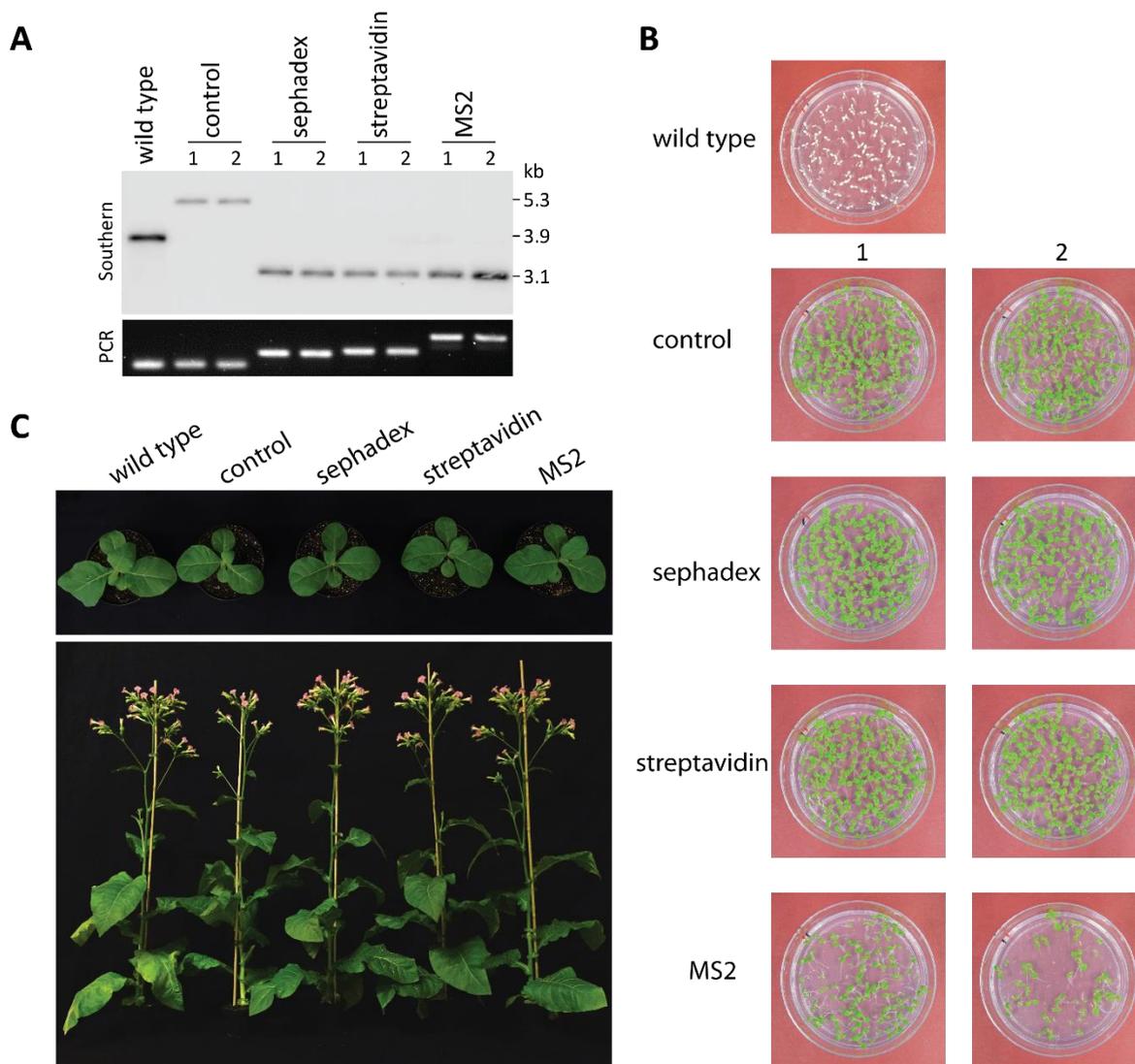


Figure 3.28: Tagging of *rbcL* with three different aptamers by stable transformation of the plastome.

A. PCR and RFLP analysis of the offspring of two independent lines for each *rbcL* construct. Upper panel: Southern blot analysis, gDNA was fragmented with *MfeI* restriction enzyme, separated by gel electrophoresis, immobilized on a nylon membrane, and hybridized to the probe indicated in Figure 3.25A. Lower panel: gDNA was used for spanning PCR with primers indicated in Figure 3.25A. **B.** The uniform resistance of the offspring was confirmed by growth on spectinomycin. All the lines are homoplasmic without any bleached seedling. **C.** Phenotype of representative *rbcL* tagged lines. Four-week-old plants (top) and eight-week-old plants (bottom) grown under greenhouse conditions (section 2.2.1.8). All lines resemble the wild type.

It is important to mention that Sephadex-tagged *psbA* plants (Figure 3.27C) and *rbcL aadA*-control plants (Figure 3.28C) were male-sterile and thereby had to be pollinated with wild-type plants for propagation. It should be noted that *in vitro* propagation of tobacco regularly causes male sterility (communication with Dr. Stephanie Ruf, MPIMP). Hence, taken together with the absence of similar phenotypes in the other tagged and *aadA*-control lines, it is unlikely that the sterility results from the aptamer or the *aadA* insertions. Overall, the integration of an aptamer ~ 100 bp apart from the translation stop codon of *psbA* and *rbcL* did not substantially affect the physiology of the plants and homoplastomic transformants were obtained.

3.2.1.3 Expression and purification of an active MS2-MBP fusion protein

Despite the wide usage of the MS2 aptamer for affinity purification approaches, no commercial affinity columns to which this aptamer can bind are available. Consequently, to exploit the high-affinity interaction between the MS2 aptamer and the MS2 coat protein from the bacteriophage capsid ($K_d = 3 \times 10^{-9}$ M) (Lim and Peabody, 1994), I aimed to use a fusion of the MS2 coat protein to a maltose-binding protein (MS2-MBP), which enables the purification of the mRNA and its bound proteome using amylose resins (Figure 3.29A) (Said et al., 2009). In the fusion protein, the maltose-binding protein is located at the N-terminus of the MS2 coat protein that carries a double mutation (V75Q and A81G) (Macias et al., 2008) to prevent the protein oligomerization (LeCuyer et al., 1995). The pMS2-MBP plasmid (addgene #6501) was used to express the MS2-MBP fusion protein and the transcription of the protein-coding sequence was induced by IPTG (section 2.2.6.1, Figure 3.29B). The IPTG induction was assessed by SDS-PAGE followed by Coomassie Brilliant staining. A 60 kDa band that corresponds to the size of the MS2-MBP fusion protein was detected in the bacterial lysate only after induction (Figure 3.29B). Several tests were performed to check the efficiency of each purification step and the purity of the MS2-MBP fusion protein. First, each step of the FPLC chromatography was evaluated by SDS-PAGE and a subsequent Coomassie colloidal staining as described in section 2.2.5.5 (Figure 3.29C). The purification over the amylose column yielded a single band (~ 60 kDa) on the Coomassie-stained gel, which was not detected in the flow through (Figure 3.29C). This indicates the strong affinity of the MBP to the amylose matrix however, the major pitfall was the significant contamination with nucleic acid as shown with the UV light excitation of the agarose gel (Figure 3.29D). This contamination accounts for the high affinity of the MS2 coat protein to nucleic acid and was previously shown to cause binding and trapping of *E. coli* nucleic acid and to affect the protein stability and the RNA-binding efficiency of the fusion protein (Jurica et al., 2002). Therefore, heparin chromatography, as a second purification step, was necessary to eliminate this contaminant by taking advantage of the fact that heparin mimics the polyanionic structure of nucleic acids and thus acts as an

affinity ligand competitor of DNA/RNA-binding proteins such as the MS2 coat protein (Xiong et al., 2008).

Indeed, the double purified MS2-MBP fusion protein did not contain any substantial nucleic acid contamination (Figure 3.29D). In addition, it had an A_{280}/A_{260} ratio of 0.57 that is within the range of optical density (OD) of a pure protein (OD = 0.59). According to Jurica et al. (2002), the MS2-MBP protein elutes at ~ 60 mM KCl. Hence, a gradient from 20 to 400 mM KCl (section 2.2.6.2) was used to elute the MS2-MBP protein from the heparin column. Unexpectedly, a clear band at ~ 60 kDa was detected over the whole gradient (Figure 3.29C) indicating the saturation of the heparin column. Consequently, all the peak fractions were checked for nucleic acid contamination (Figure 3.29D), pooled, and concentrated (section 2.2.6.2). The MS2-MBP protein was later used for affinity purification and to localize the MS2-tagged transcripts (see below).

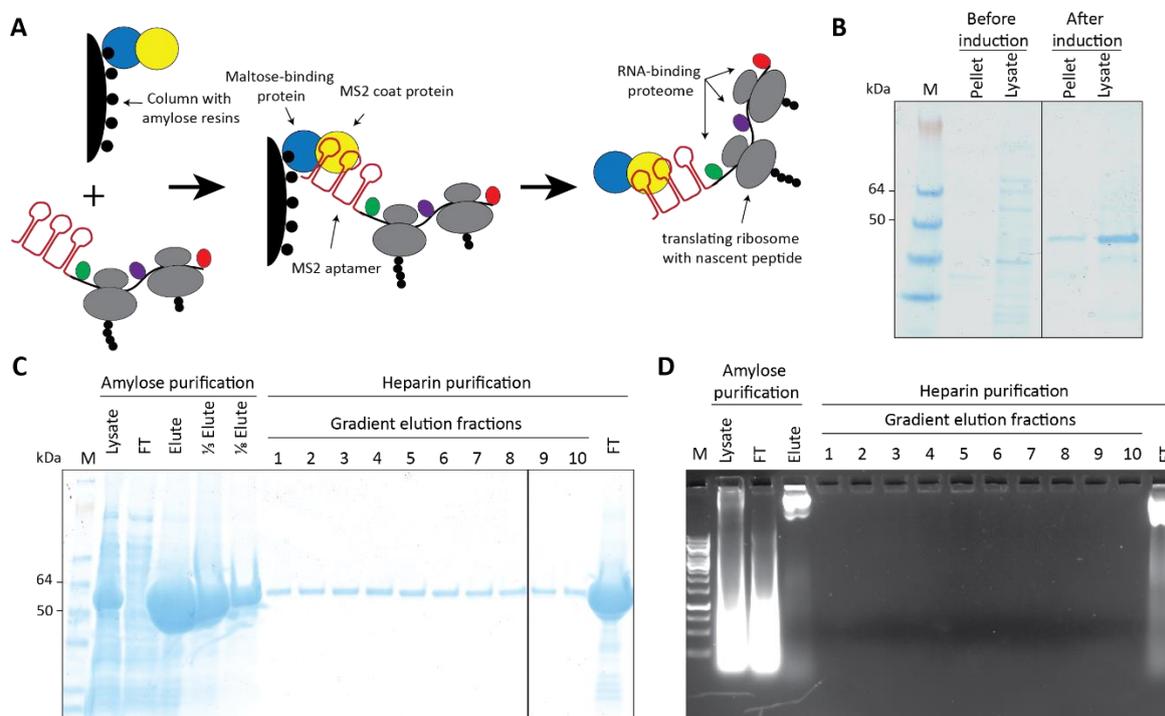


Figure 3.29: Double maltose and heparin column purification yields a pure MS2-MBP recombinant protein.

A. Schematic representation of the experimental strategy of the MS2-based affinity purification adapted from Said et al. (2009). The red loops denote the MS2 aptamer, the yellow and blue circles represent the MS2 coat protein and the maltose-binding protein, respectively. The latter binds to the subunits of the amylose column (black circles). Translating ribosomes are shown in gray with the nascent peptide represented in small black circles. The colored circles depict the RNA-binding proteins. **B.** Bacterial lysate and cell debris (pellet), before and after induction of MS2-MBP expression, separated by SDS-PAGE. The gel was stained with colloidal Coomassie blue (section 2.2.5.5). A band of 60 kDa, corresponding to the MS2-MBP recombinant protein was observed. M: protein marker. **C.** Protein products in the fractions of the amylose purification (the lysate, flow-through (FT), the eluate with different dilutions) and all 10 fractions of the heparin chromatography were separated by SDS-PAGE and visualized with colloidal Coomassie staining. The black arrow indicates the MS2-MBP recombinant protein, which appeared to be efficiently purified with the amylose column and was

detected in all the elution fractions of the heparin column. **D.** Agarose gel electrophoresis of the different fractions from the amylose and the heparin chromatography. A nucleic acid signal was observed in the elution fraction of the amylose purification. In contrast, none of the elution fractions of the heparin column was showing the nucleic acid signal.

3.2.1.4 MS2 aptamer enables specific purification of *psbA* and *rbcL* mRNAs

MS2-tagged plants and the corresponding *aadA*-control were grown in standard conditions for three weeks (section 2.2.1.8). The aerial part was harvested 30 min after the onset of light given that the chloroplast gene expression peaks early after the start of the illumination. The MS2-tagged plants did not exhibit any visible phenotype compared to the *aadA*-control plants (Figure 3.30A and B).

The affinity purification was performed using plant lysates of the MS2-tagged plants in parallel to the corresponding *aadA*-control as described in section 2.2.7.3. The efficiency of the pulldown was first investigated at the RNA level using northern blot and microarray hybridization analyses (sections 2.2.2.11 and 2.2.3.3). RNA was isolated from total plant lysate, the input, the flow-through, the washing steps, and the elution fractions from MS2-tagged and *aadA*-control lines.

The northern blot analysis (with *psbA*-specific probe) of the *psbA*-MS2 pulldown showed a specific band at the expected size of the *psbA* transcript only in the MS2-tagged line (Figure 3.30B, elution lane). Although *psbA* is a monocistronic transcript of 1.2 kb size (Shinozaki et al., 1986), an additional band was detected that migrated slower into the gel (at ~2.5 kb) (Figure 3.30B). I assumed that the insertion of the *aadA* cassette downstream of the MS2 tag causes an accumulation of a co-transcribed dicistronic *psbA-aadA* hybrid transcript. To confirm this hypothesis, another hybridization using an *aadA*-specific probe was performed on the same membrane (Figure 3.30B). Two bands were observed: a large band (~ 2.5 kb), which based on the size represents the dicistronic *psbA-aadA*, and a small band that corresponds to the *aadA* transcript (~ 1.3 kb). Altogether, this shows that the *psbA* transcript can be efficiently purified using the MS2 aptamer however, this does not prove the specificity of the purification. To reveal potential contaminating chloroplast transcripts that were co-purified with the MS2 aptamer, I hybridized the RNA from the elution fraction on our custom tobacco chloroplast microarrays (section 2.2.3.3). Several optimization steps were needed before a final RAP-Chip (RNA aptamer-based purification) protocol was established. Especially, the fragmentation of the purified RNA improved the signal intensity on the microarrays considerably (see section 2.2.3.2). To normalize my results to the general input, 3.5 µg of flow-through RNA was fragmented and hybridized in addition. After background subtraction, probes with negative values were set to zero. All background-subtracted signals in the elution and flow-through of the MS2-tagged line and the corresponding *aadA*-control were normalized to the average signal of all these four datasets to remove alterations caused by technical variations (e.g., labeling or hybridization efficiencies). Subsequently, the average values of the normalized signals were calculated in a moving window of 270 nt (moving by 30 nt). In order to

preclude differences in the input, the ratio of the flow-through average values in the MS2-tagged line to those in the *aadA*-control line was calculated. This ratio was then used to normalize the values obtained in the elution of the MS2-tagged line. Results showed a substantial enrichment of *psbA* in the MS2-tagged line compared to the *aadA*-control of about tenfold (Figure 3.30E).

For *rbcL* pulldown with the MS2 aptamer, the efficiency was first checked using northern blot (Figure 3.30D). The RNA was extracted from the same fractions as described for the *psbA*-MS2 pulldown. A band of 1.6 kb, the expected size of *rbcL*, was observed in the MS2-tagged line (Figure 3.30D) and was not detected in the *aadA*-control. An additional larger band of 3 kb (Figure 3.30D) was also observed, which by considering the size, corresponds to a dicistronic *rbcL-aadA* hybrid transcript. RNA hybridization with an *aadA*-specific probe showed two bands with the larger one corresponding by size to *rbcL-aadA* dicistronic transcript and the small one to the *aadA* transcript. Moreover, the specificity of the *rbcL* purification with the MS2 aptamer was checked by microarray hybridization. *rbcL* was specifically enriched (~ fivefold). A signal from rRNAs was observed, which most likely comes from the translating ribosomes bound to the mRNA giving that the plastid translation was blocked by chloramphenicol during isolation (section 2.2.7.3). In sum, I could efficiently and specifically purify *rbcL* using the MS2 aptamer.

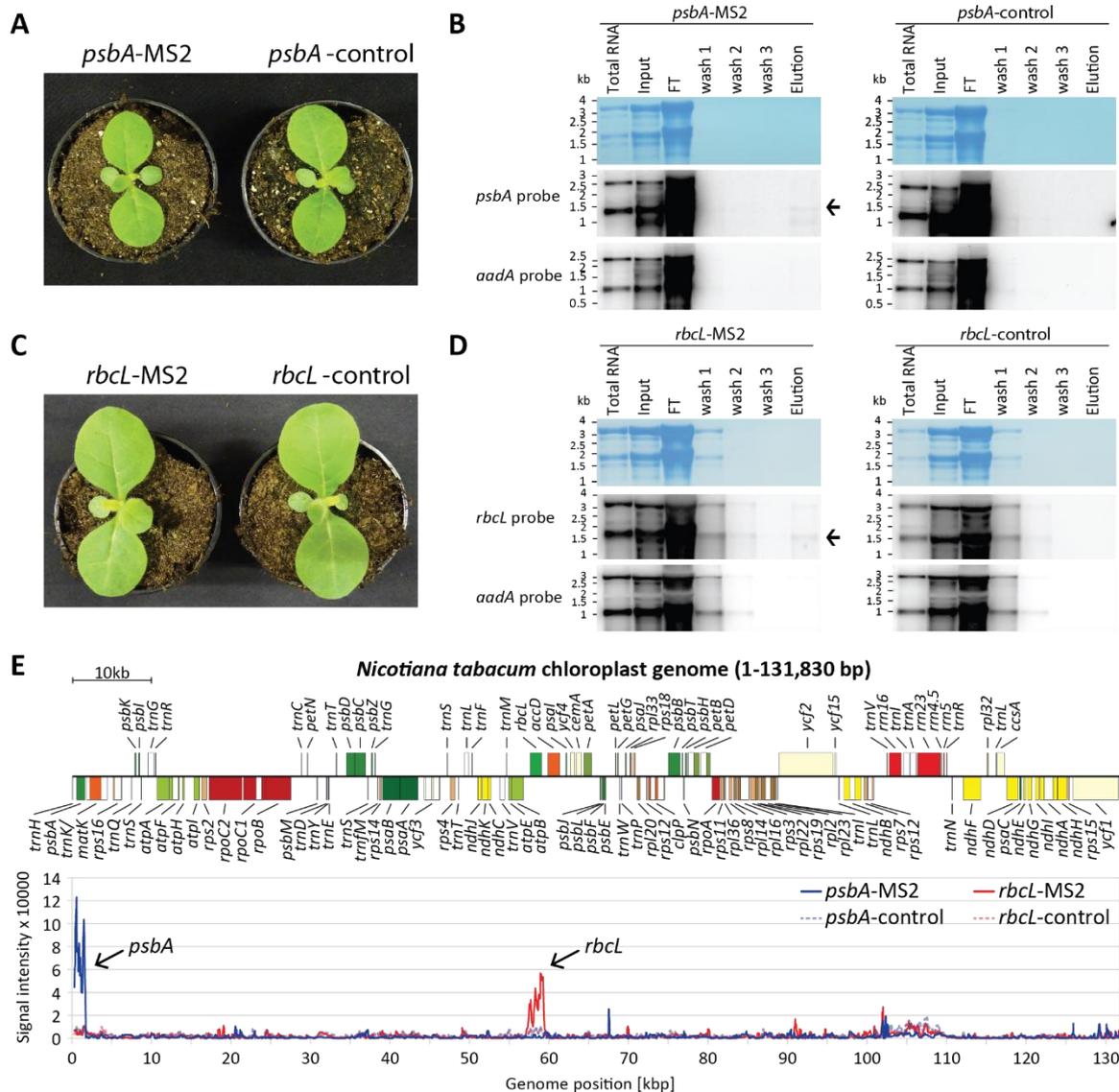


Figure 3.30: *psbA* and *rbcL* mRNAs are efficiently and specifically purified with the MS2 aptamer. **A, C.** Tagged and control plants were grown for three weeks under standard conditions (section 2.2.1.8). The aerial part of the plants was harvested and total plant lysates were used for the MS2-based purification (section 2.2.7.3). **B.** Northern blot analysis of *psbA* purification with the MS2 aptamer: one-fifth of the volume of each fraction was separated on a 1.2 % agarose gel, blotted to a nylon membrane, and hybridized with a probe complementary to the coding sequence of *psbA* (upper panel) and *aadA* (lower panel). RNA sizes in kb are labeled on the left side of the membrane. The black arrow indicates the detected *psbA* transcript in the elution fraction. Methylene blue staining is shown as RNA integrity control. FT: flow-through. **D.** Northern blot analysis for *rbcL*-MS2 purification with *rbcL* probe (upper panel) and *aadA* probe (lower panel). **E.** RNA from the elution and flow-through fractions was fragmented and hybridized to our tobacco chloroplast microarrays (sections 2.2.3.2 and 2.2.3.3). The microarray signal intensities from the elution fractions were plotted according to the position in the tobacco chloroplast genome. These intensities were background-subtracted and normalized to the signals in the input fraction (see section 3.2.1.4 for details). Blue and red solid lines represent the signal intensities from the *psbA*-MS2 and *rbcL*-MS2 elution fractions, respectively. The dashed blue and red lines depict the intensities from the *psbA*-control and *rbcL*-control elution fraction of the MS2 purification, respectively. The physical map of the chloroplast genome was generated from the NCBI reference sequence Z00044.2 using OGDRAW (Greiner et al., 2019). The map shows one of the two inverted repeats in the chloroplast genome and illustrates all the genes.

3.2.1.5 The MS2 aptamer enables the examination of the suborganellar localization of *psbA* and *rbcL*

Aptamer-tagging of RNAs was also used to study RNA localization (Forrest and Gavis, 2003; Morisaki et al., 2016; Sheth and Parker, 2003; Wang et al., 2012; Xiao et al., 2005). Hence, I asked, if the aptamer-tagged chloroplast transcripts can be utilized to examine the localization of *psbA* and *rbcL* using immunogold labeling and transmission electron microscopy (TEM) (collaboration with Dr. Arun Sampathkumar (MPIMP)). Leaf samples were taken from two weeks old MS2-tagged plants grown under standard conditions (section 2.2.1.4) and fixed as described in section 2.2.9. Heterologously expressed and purified MS2-MBP fusion protein (section 3.2.1.3) was added to the fixed tissue followed by incubation with an anti-MBP antibody and a secondary antibody with gold conjugate. Control samples were only incubated with primary and secondary antibodies. TEM enabled the visualization of the chloroplast ultrastructure with massive starch grains and stacked grana (Figure 3.31). The first trials showed black particles in the chloroplast of the MS2-tagged lines but not in the negative controls (Figure 3.31). The black particles, which correspond to *psbA*-MS2 transcripts, were detected in the stroma and close to the thylakoid membranes (Figure 3.31A). This finding supports the dual localization of the *psbA* transcript with one fraction cotranslationally targeted to the thylakoid membrane and an untranslated fraction in the stroma (Zoschke and Bock, 2018). In contrast, in *rbcL*-MS2 lines, *rbcL* transcripts were only localized in the stroma (Figure 3.31B). Some particles were detected outside of the chloroplast, which suggests that more or extended washings are needed. This result proves that the MS2 aptamer can be used to examine transcript localization in chloroplasts. However, further methodological refinements such as different blocking and washing steps are needed to optimize the sensitivity.

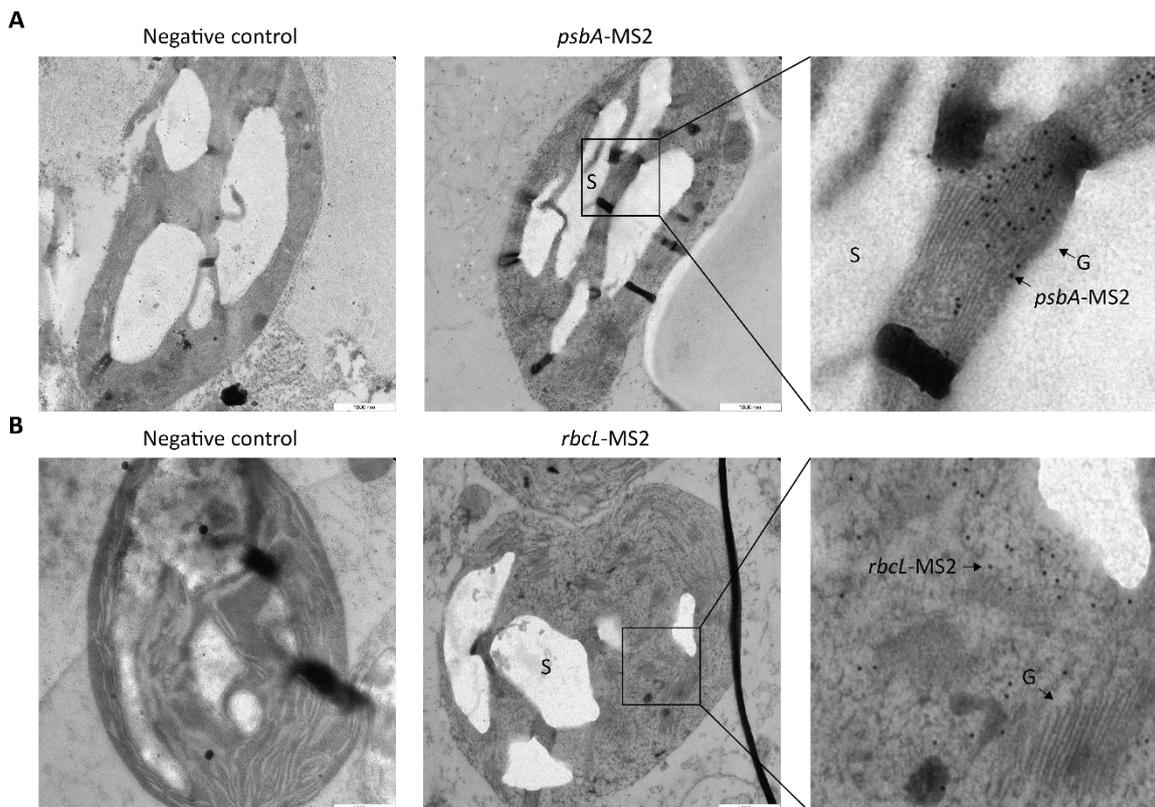


Figure 3.31: Aptamer-based measurements of suborganellar localization of *psbA* and *rbcL* transcripts.

A, B. TEM pictures of chloroplasts from MS2-tagged seedlings grown for two weeks under standard conditions (section 2.2.1.8). The left side shows the negative controls for *psbA*-MS2 (**A**) and *rbcL*-MS2 (**B**). The middle shows the chloroplasts from *psbA*-MS2 (**A**) and *rbcL*-MS2 (**B**) incubated with MS2-MBP fusion protein prior to incubation with antibodies and the gold conjugate (for details see section 2.2.9). The black squares mark the magnified regions shown on the right side. The white bar indicates the scale, 1 μ m. S: Starch granules; G: Stacked grana.

3.2.2 Streptavidin-binding aptamer enables the specific affinity purification of *psbA* but not *rbcL*

Plants tagged with streptavidin-binding aptamer were grown next to the control plants and harvested as described in section 3.2.1.4. The affinity purification was performed from plant lysates of the tagged *psbA* and control lines (section 2.2.7.1). In the course of optimization, different trials were done to maximize the enrichment of the targeted transcript while preventing RNA degradation and background contamination from other transcripts. Affinity purification without elution in which the bound RNAs and proteins are isolated from beads was performed. Additionally, affinity purification experiments including 10 min-elution step or 30-min elution step were performed. The efficiency and specificity of the pulldown of *psbA* were tested with microarray hybridization of the RNA purified from the beads (in case of no elution) or the elution fraction. With the elution step omitted, the data showed a clear enrichment of *psbA* (Figure 3.32A). Interestingly, other abundant RNAs (e.g., rRNAs, *rbcL*, *psbB*,

psaA/B) were co-purified with the streptavidin beads. Highly abundant RNAs were also purified (Figure 3.32A). Together, this indicates that the streptavidin beads unspecifically bind RNA (or RNA-binding proteins), a property that has previously been observed (Theil et al., 2019). The average of the signal intensity of *psbA* probes was calculated and normalized to the ratio of the average of *psbA* signal intensity in the flow-through in the tagged line to that in the control. Comparison of the average of the normalized signal intensity of *psbA* in the streptavidin-tagged line to that in the control revealed a clear enrichment of *psbA* of 2.5-fold. Overall, I was able to enrich for *psbA* using the streptavidin-binding aptamer, however, the purification was not completely specific. To reduce the contamination with other transcripts, a 30-min elution step with D-biotin (Sigma-Aldrich), which has a strong competing affinity to streptavidin, was included. Subsequently, the RNA in the elution fraction was extracted, fragmented, and hybridized to the microarray. The contamination level with abundant RNAs was slightly reduced (Figure 3.32B). However, the *psbA* enrichment (compared to other RNAs) decreased. This may be caused by RNA degradation with RNases introduced by the D-biotin. To overcome the RNA degradation problem, I shortened the elution time to 10 min, which was sufficient given the extraordinarily high affinity of streptavidin to biotin. Indeed, the enrichment was higher relative to the purification with 30-min elution or without elution while keeping lower contamination with abundant RNAs (Figure 3.32C). In conclusion, I found after these optimization steps that intact *psbA* can be efficiently purified using the streptavidin-binding aptamer while maintaining a low level of contamination.

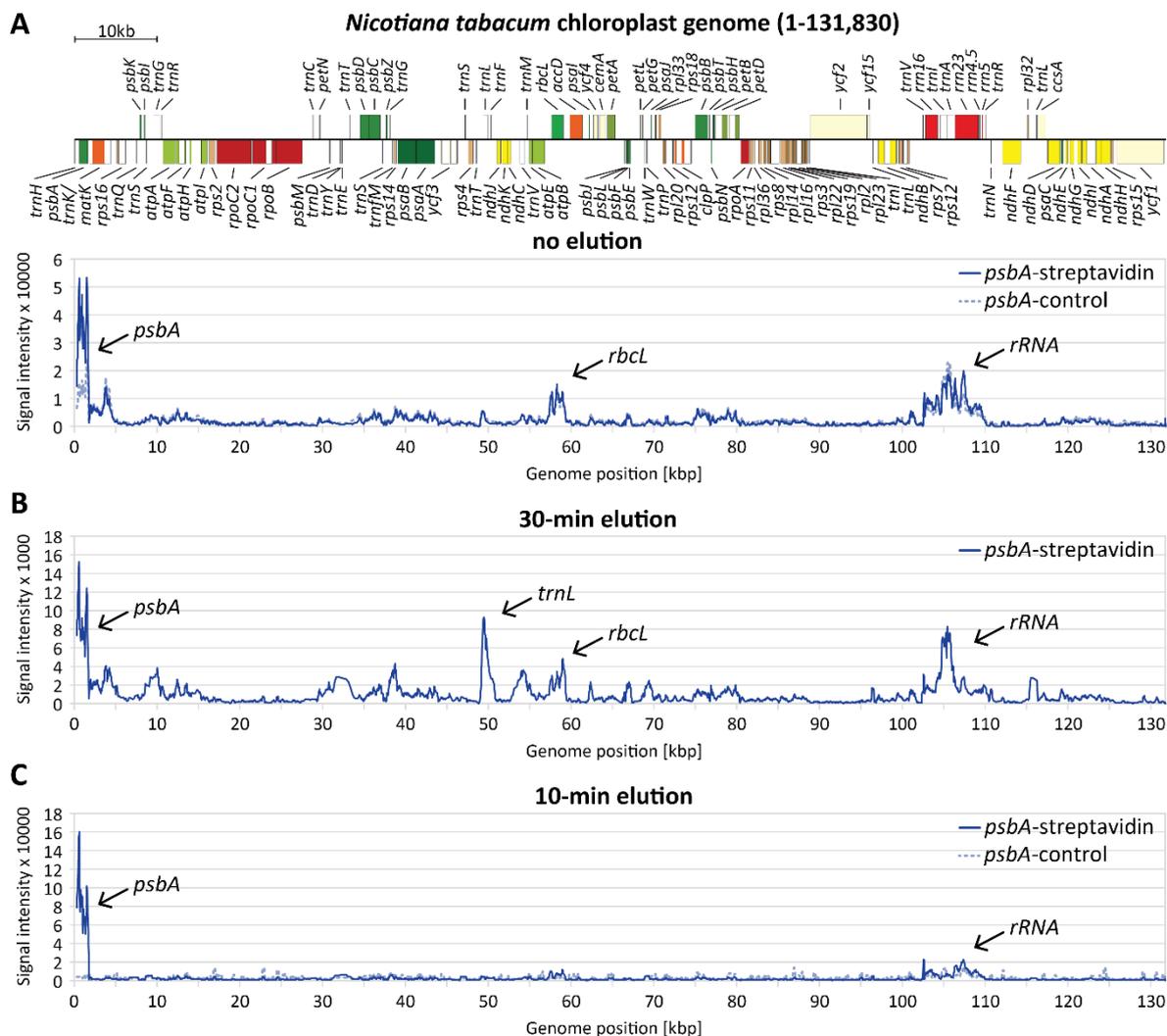


Figure 3.32: Purification of *psbA* mRNA with the streptavidin-binding aptamer.

A. RNA from the elution or the beads and flow-through fractions was fragmented and hybridized to our tobacco chloroplast microarrays (section 2.2.3.3). Microarray signal intensities from the streptavidin purification without elution step, **B.** with 30-min elution, **C.** with 10-min elution was analyzed and plotted as described in Figure 3.30E. The blue solid line indicates the signal from the *psbA*-streptavidin purification while the dashed blue line indicates the RNA signal intensity of the *psbA*-control purification (chloroplast genome map as in Figure 3.30E). The controls gave reproducibly low background signals and therefore were not always analyzed. The black arrows indicate the most abundant RNA species detected.

Several attempts were performed to purify *rbcL* using the streptavidin-binding aptamer. However, northern blot analysis showed that the *rbcL* transcript did not co-precipitate to detectable amounts with streptavidin beads (data not shown). One possibility could be that the streptavidin-binding aptamer inserted into the 3' end of *rbcL* has a conformation that hinders efficient interaction with streptavidin. The context-dependency of the aptamer's conformation remains puzzling.

3.2.3 *rbcL* and *psbA* can be purified using the Sephadex-binding aptamer

Sephadex-tagged plants were grown next to their corresponding controls and harvested as described in section 3.2.1.4. Extracts from *rbcL*-Sephadex line and *rbcL*-control were subjected to affinity purification with Sephadex G200 resins (section 2.2.7.2). RAP-Chip analysis was used to evaluate the efficiency and the specificity of the purification, as well as the integrity of the purified RNA. Data from the affinity purification of *rbcL* without an elution step showed a minor enrichment of *rbcL*, with a major degradation of the mRNA noted by the peak of the signals of the probes at the 3' end of the transcript (Figure 3.33A). Mostly highly abundant RNAs such as rRNA, tRNA, and *psbA* appeared to be co-purified. Additionally, the same contaminant RNAs were detected in the untagged control highlighting the unspecificity of binding to the Sephadex G200. To overcome the unspecific binding to the Sephadex resins, I decided to include a 30-min elution with enzymatically synthesized dextran that competes with the Sephadex resins for binding to the Sephadex-tagged *rbcL*. Indeed, fewer contaminants were detected in the pulldown (Figure 3.33B). However, the degradation, in this case, was more pronounced as the length of the experiment got extended. In addition, the dextran might have introduced some RNases making the RNA more prone to degradation. To reduce the degradation while compromising the contamination with other RNAs, the elution length was reduced to 10 min. Furthermore, the Sephadex resins were thoroughly pre-washed prior to purification. These adjustments yielded higher enrichment of *rbcL* (Figure 3.33C). Nonetheless, the purified *rbcL* was still partially degraded.

In parallel, I performed affinity purification without elution to purify *psbA* using the Sephadex-binding aptamer. The subsequent RAP-Chip analysis revealed a minor enrichment of *psbA* with mild degradation. Other abundant RNAs were also co-purified (data not shown). However, no further optimization was applied as I was able to purify *psbA* with two independent aptamers, MS2, and streptavidin-binding aptamer.

Taken together, I demonstrated that the Sephadex-binding aptamer could be used to purify both *rbcL* and *psbA* however, the integrity of the purified RNA in both cases was not ideal in comparison to the other two aptamers used in this work.

Note that Sephadex G100 was also tested and it appeared to be inefficient to purify Sephadex-tagged transcripts (data not shown).

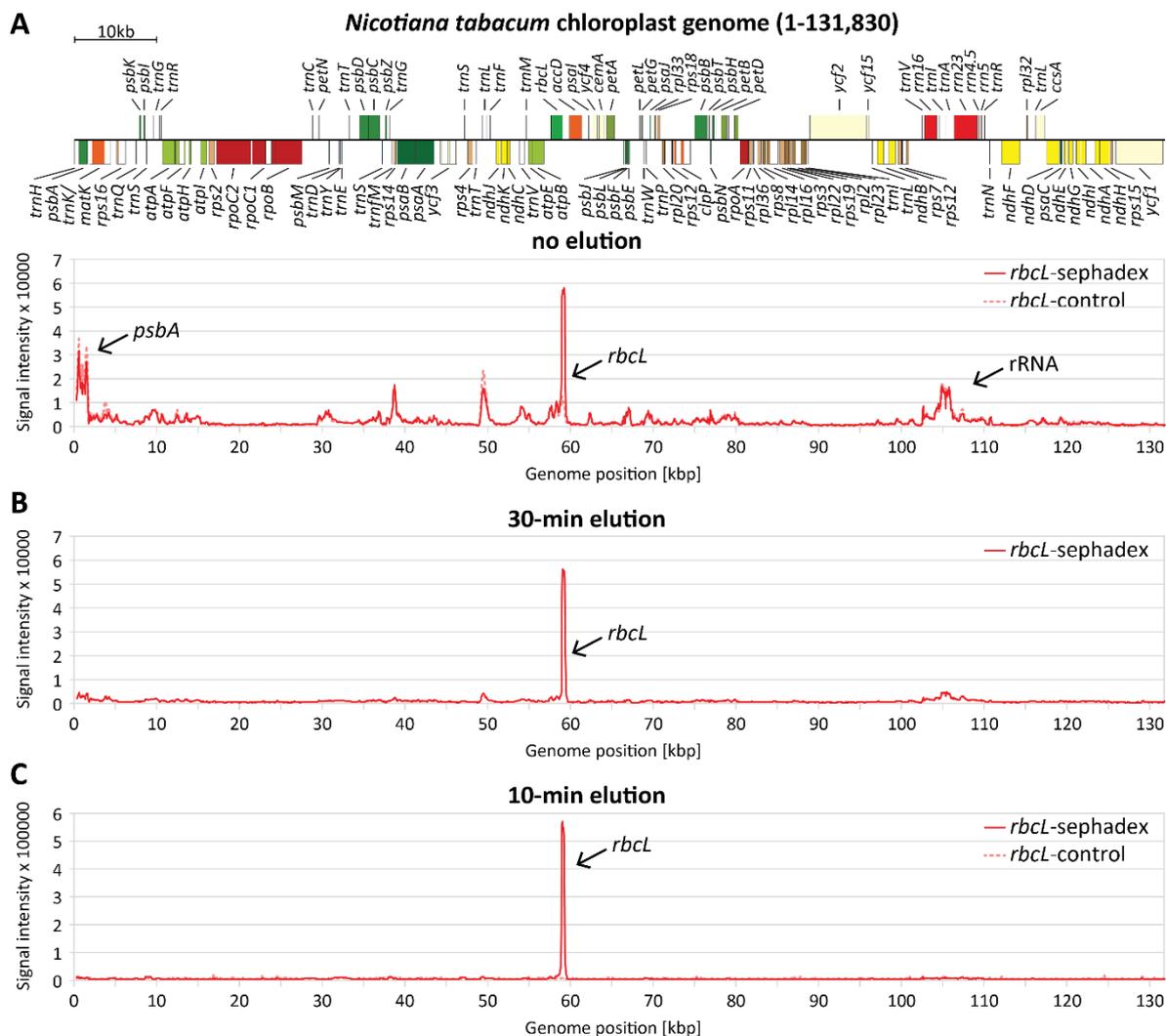


Figure 3.33: Sphadex-binding aptamer enables the purification of *rbcL* mRNA.

A. RNA from the elution or Sephadex resins and the flow-through fractions were fragmented and hybridized to the custom tobacco chloroplast microarray as described in section 2.2.3.3. Microarray signal intensities from the Sephadex purification without elution step, **B.** with 30-min elution, **C.** with 10-min elution were analyzed and plotted as described in Figure 3.30E. The red solid line indicates the signal from the *rbcL*-Sephadex purification while the dashed red line indicates the *rbcL*-control purification. See Figure 3.30E for details about the physical map. The controls were not always analyzed since they showed consistently low background signals. The black arrows indicate the most abundant RNA species detected.

3.2.4 Comparison of the MS2, streptavidin- and Sephadex-binding aptamers in regard to the enrichment efficiency and integrity of the purified *psbA* and *rbcL* mRNAs

In order to assess the best performing aptamer in purifying *psbA* mRNA, the RAP-Chip analysis of all the *psbA* pulldown experiments was compared. The microarray data of only the tagged lines from all the *psbA* pulldown experiments were analyzed together. Following the subtraction of the background signals, all the probes with negative values were set to zero. All the background-subtracted values were

then normalized to the average of the signals in all the datasets. This was intended to remove biases caused by technical variations such as the hybridization and labeling efficiencies. In order to normalize the differences in the input used in each pulldown experiment, the ratio of the flow-through probe signals in each dataset to that in the MS2 pulldown was calculated. The probe signals in the elution or beads fractions were then normalized to this ratio. In order to obtain the enrichment level of the purified *psbA*, the average value of normalized probes signals in the ORF was calculated. The assessment of the efficiency of the pulldown was based on three criteria: (i) the integrity of the purified RNA, (ii) the enrichment level, and (iii) the contamination with other abundant RNAs. In terms of enrichment, a similar level of *psbA* was purified with the MS2 aptamer and the streptavidin-binding aptamer with 10-min elution (Figure 3.34A). On the other hand, the enrichment of *psbA* in the Sephadex-binding aptamer-mediated purification was 1.7- and 1.6-fold lower than that with the MS2 aptamer and streptavidin-binding aptamer with 10-min elution, respectively. As mentioned in section 3.2.2, *psbA* is less efficiently purified with the streptavidin-binding aptamer when no elution or long elution is implemented (Figure 3.34A). RNA degradation could account for this decrease, especially with such a long experimental protocol.

The integrity of the purified mRNA is an essential prerequisite to identify new translation factors as most of the studied ones bind to the 5' UTR of the transcript (Zoschke and Bock, 2018). In the case of an intact aptamer-tagged mRNA, a signal should be detected across the whole transcript. Given that in our current experimental design the aptamer sequence was placed in the 3' UTR of the transcript in question, a potential degradation is noted by an enrichment of the probes located in the 3' end of the mRNA with a lower signal of the probes in the 5' end and the middle of the ORF. In the *psbA*-MS2 pulldown, the probe signal was comparable across the whole transcript with no indication of degradation (Figure 3.34C). Similarly, in all three purification experiments with the streptavidin-binding aptamer, no pronounced degradation was observed, however, a slight decrease of signal intensities of the probes at the 5' end of *psbA* can be observed in the experiment with 10-min elution step. Although the purification with the streptavidin-binding aptamer without elution and with 30-min elution yielded two and 3.5-fold less enrichment compared to the purification with the MS2 aptamer (Figure 3.34A), the signal distribution across the transcript was uniform with a similar pattern to that obtained in the MS2 aptamer-mediated purification. In contrast, a major degradation of *psbA* was observed in the purification with the Sephadex-binding aptamer emphasized by the higher signal intensities of the probes at the 3' end of *psbA* where the aptamer is inserted (Figure 3.34C).

Furthermore, regarding contamination, the purification with the MS2 aptamer showed fewer signals from contaminant RNAs in comparison to the purifications with the Sephadex-binding aptamer and the streptavidin-binding aptamer (all three conditions) (Figure 3.30E and Figure 3.32). This cross-comparison of the three aptamers used to purify *psbA* showed that the MS2 and streptavidin-binding aptamers yield a higher and more specific enrichment than the Sephadex-binding aptamer.

A similar comparison was performed for *rbcl* purification experiments to assess the best aptamer strategy for further work. The *rbcl*-Sephadex affinity purification with 10-min elution yielded the highest enrichment, approximately 2.6-fold higher than the purification with the MS2 aptamer (Figure 3.34B). However, the main pitfall for the Sephadex-binding aptamer was the massive degradation of *rbcl* noted by the increased signals toward the 3' end (Figure 3.34D). In the MS2-based purification, a comparable signal intensity level along the *rbcl* transcript was observed with only slight drift toward the 3' end (Figure 3.34D). On the contrary, in all the purification experiments with the Sephadex-binding aptamer, *rbcl* was partially degraded. The addition of a 30-min elution step increased the degradation probably by extended exposure to RNases (e.g., from plant material). By shortening the elution step to 10 min, more signals were detected in the body of the *rbcl* gene. Surprisingly, however, signals at the 3' end of *rbcl* were eightfold higher than that with the MS2 aptamer. Taken together, these results indicate that the MS2 aptamer is the most suitable aptamer to specifically purify intact *psbA* and *rbcl*.

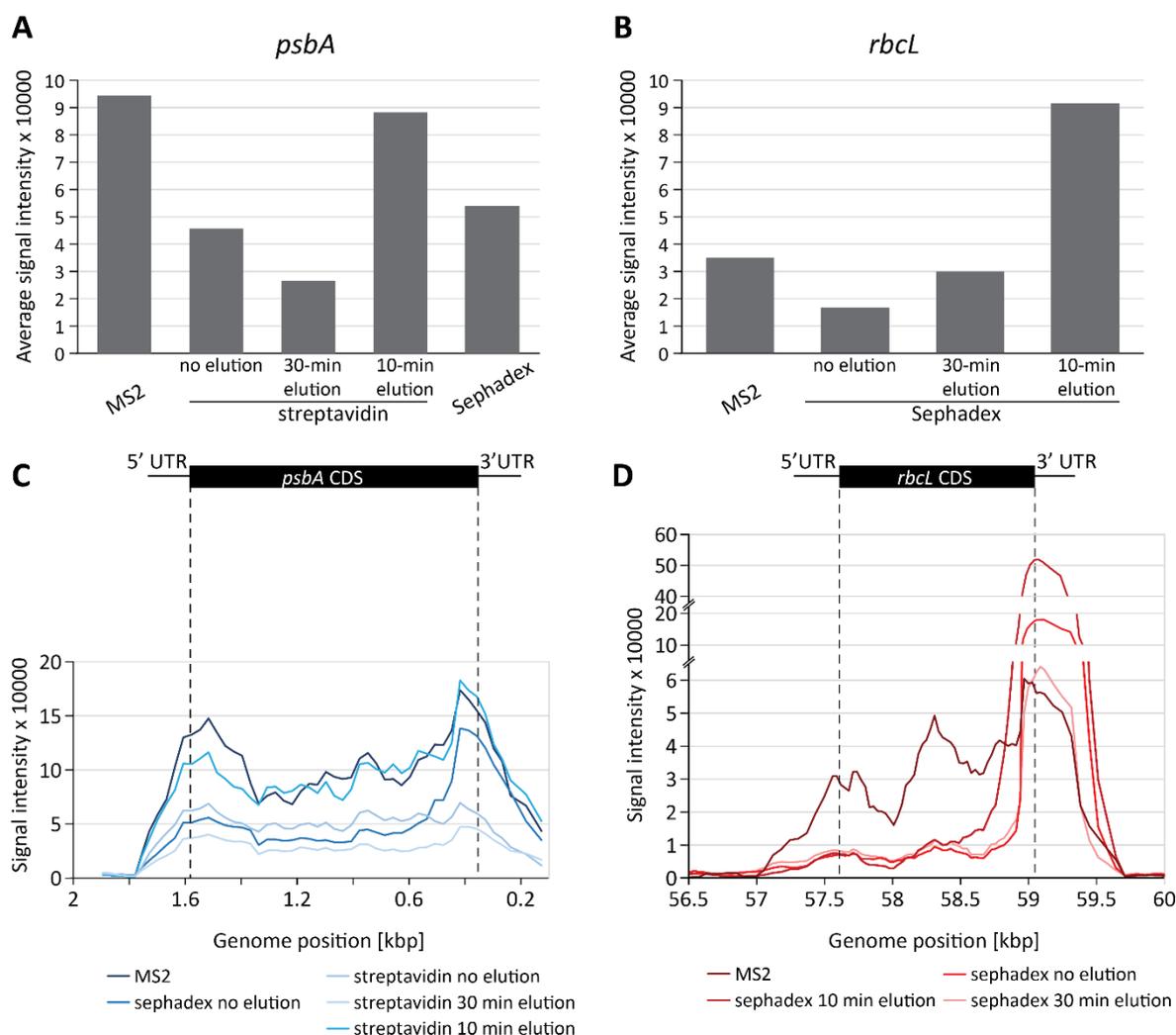


Figure 3.34: Comparison of the purification of *psbA* and *rbcl* with different aptamers.

A. Average signal intensity of *psbA* probes in the purification experiment with the MS2 aptamer, the Sephadex-binding aptamer, and the streptavidin-binding aptamer with different elution steps. **B.** Average signal intensity of *rbcL* probes in the purification experiment with the MS2 aptamer and the Sephadex aptamer with different elution steps. **A, B.** The average values are relative to the flow-through *psbA* and *rbcL* signal intensity in the corresponding MS2 purification experiment. **C, D.** Magnification of the signal distribution along the **C.** *psbA* and **D.** *rbcL* ORF in all the purification experiments.

3.2.5 Mass spectrometry analysis of the RNA-binding proteome of *psbA*

To identify the protein composition of the RNA-binding proteome associated with *psbA*, the co-purified proteins were identified by LC-MS/MS in collaboration with Dr. Frederick Sommer from the University of Kaiserslautern. The data obtained were processed as described in section 2.2.8. A twofold change was applied as a rule of thumb for the final identification of *psbA*-associated proteins. The uncharacterized proteins were blasted against the *Arabidopsis thaliana* proteome using NCBI Blastp. Out of 48 enriched proteins, only five were plastid-localized (Supplemental Table 6), none of which have an RNA-binding domain. Interestingly, a nucleus-encoded tetratricopeptide repeat (TPR) -like superfamily protein (AT2G33680) was detected. This protein is located in the chloroplast and is involved in RNA modification. TPR superfamily consists of helical repeat proteins with regulatory functions (Bohne et al., 2016). Many studies have shown that these proteins are involved in different steps of the thylakoid membrane biogenesis including the protein import into the chloroplast (Qbadou et al., 2007), the chloroplast gene expression (Boudreau et al., 2000; Trösch et al., 2018), assembly/stability of photosynthetic complexes (Bhuiyan et al., 2015) and chlorophyll synthesis (Kauss et al., 2012).

I tried to combine the MS data from three purification experiments: *psbA*-MS2, *psbA*-streptavidin without elution, and *psbA*-streptavidin with 30-min elution. This was intended to have a rough idea of the proteins that can be enriched with *psbA* and to get a proxy whether a cross-linking step is needed or not. For that, the three control datasets were grouped and compared to the combination of the three tagged-*psbA* datasets. The intensity data output from MaxQuant was loaded for subsequent analysis in Perseus (Tyanova et al., 2016). Subsequent filtering and processing were undertaken as described in section 2.2.8. A two-sample t-test followed by an FDR correction did not result in any significant enrichment of any interactor with *psbA*, which may be not surprising given that the datasets compared are no true replicates.

Supplemental Table 7 summarizes the proteins that were identified with more than twofold enrichment. Out of 13 proteins, two proteins are plastid-localized, two are uncharacterized proteins and the rest have random locations in the plant cell. Based on these preliminary results, no conclusive statement about potential *psbA*-binding proteins can be drawn. It is important to mention that the detection score was low which reflects a necessity to enrich more for the proteins. These low levels of co-purified proteins may explain the failure to detect HCF173, the only known *psbA*-binding protein so far (Link et al., 2012; McDermott et al., 2019; Schult et al., 2007; Watkins et al., 2019). Strategies to overcome these

counter boundaries are discussed below (section 4.2.2). Details about all the identified proteins, intensities, number of unique peptides and sequence coverage are included in Supplemental dataset 1.

3.2.6 Mass spectrometry analysis of the RNA-binding proteome of *rbcL*

Data from one replicate *rbcL*-MS2 pulldown was analyzed using MaxQuant. As only one pulldown experiment was performed for *rbcL*, the output of MaxQuant was no further processed in Perseus. A threshold of twofold was set to filter potential RNA regulators of *rbcL*. The data are summarized in Supplemental Table 8. The data were analyzed as described in section 2.2.8. Most of the identified proteins were plastid ribosomal proteins, which is in agreement with the fact that a translation elongation inhibitor was added into the purification buffer. Additionally, three plastid-localized proteins were enriched in the tagged *rbcL* plants compared to the untagged control: AT2G22450, AT5G42650, and AT5G26570 functioning, respectively, in riboflavin biosynthesis, allene oxide synthesis and as carbohydrate kinase. All of the aforementioned proteins do not have an RNA binding capacity and have no direct relation to Rubisco. Nonetheless, this does not exclude a potential regulatory function of these proteins. However, given that this result is based only on one replicate further confirmation is needed with additional independent biological replicates and, ideally, by independent purification of the same protein with differentially tagged *rbcL* mRNAs (e.g., Sephadex-binding aptamer). Further optimization is needed to identify the full set of proteins that control *psbA* and *rbcL* expression (discussed in section 4.2.2).

4 Discussion and outlook

4.1 Identification of assembly-dependent translational feedback regulation in photosynthetic complexes of embryophytes

Several studies of *Chlamydomonas* mutants have shown that the availability of one photosynthetic subunit triggers the synthesis of another subunit in the same protein complex (Choquet and Wollman, 2009). Those subunits whose synthesis rate is assembly-dependent were designated as CES subunits (controlled by epistasy of synthesis). CES cascades were found to be widespread in *Chlamydomonas*. In many cases, these cascades are believed to enable the nucleus to control the rate of expression of the most upstream assembly partner in a biogenesis pathway to ensure the stoichiometric production of the downstream subunits (Figure 1.3). In embryophytes, a few observations supported the idea that the CES process is involved in the biogenesis of chloroplast proteins (Monde et al., 2000; Rodermel et al., 1996; Wostrikoff and Stern, 2007). However, it remained unclear to which extent CES regulation is conserved in the green lineage. The work presented here provides an extensive investigation of potential CES regulation in the chloroplasts of embryophytes.

4.1.1 Ribosome profiling readily detects CES in the assembly of Rubisco in embryophytes

Ribosome profiling has revolutionized the study of translation and has been employed in many biological systems to assess the proportional synthesis of subunits in protein complexes (Dephoure et al., 2014; Taggart and Li, 2018). As a fast method to survey chloroplast translation, a targeted ribosome profiling approach has been established for several photosynthetic organisms including tobacco and *Arabidopsis* (Trösch et al., 2018). Using this approach, I was able to confirm the previously known CES regulation of RbcL in *RBCS* knockdown mutants in tobacco and *Arabidopsis* (sections 3.1.3 and 3.1.5) (Rodermel et al., 1996; Wostrikoff and Stern, 2007). The observed CES effect on *rbcL* was more pronounced in tobacco than in *Arabidopsis* (11.3-fold vs 3.8-fold), which may be due to the stronger suppression of *RBCS* expression in the tobacco mutant compared to the *Arabidopsis* mutant (transcript levels reduced to 12 % and 23 %, respectively). In addition, species-specific differences in CES regulation between tobacco and *Arabidopsis* may account for this variation. In sum, the efficient detection of CES in the *RBCS* mutants demonstrates that ribosome profiling is well suited to identify CES regulation in the chloroplasts of embryophytes.

The regulation of Rubisco synthesis and assembly is of high physiological relevance given that Rubisco is the first rate-limiting enzyme in the CBB cycle and its expression and assembly is regulated in response to different environmental changes (Cavanagh and Kubien, 2014). Therefore, it would be highly interesting to identify the molecular mechanism of this regulation, including the involved *cis*-elements and potential *trans*-factors. Little is known about the factors involved in the translational

regulation of *rbcL*. MRL1 is a PPR protein that was shown to be involved in *rbcL* transcript stabilization in *Chlamydomonas* and *Arabidopsis* and its conservation in other green photosynthetic eukaryotes (Johnson et al., 2010) makes it a candidate factor for the *rbcL* CES. It was shown that MRL1 is targeted in an RNA-mediated fashion to a high molecular mass complex in *Chlamydomonas*, which might suggest a potential association with polysomes (Johnson et al., 2010). A systematic characterization of the molecular function of MRL1 in *Arabidopsis* is lacking. Given the RNA-binding capacity of MRL1 to the 5' end of *rbcL* (Johnson et al., 2010), I investigated whether it is involved in the CES regulation of *rbcL*. Indeed, ribosome profiling analysis of the *mrl1-1* T-DNA insertion mutant showed a defect of the translation of *rbcL* in addition to its known transcript accumulation defect (Figure 3.3C). This result implies that MRL1 is a potential translational activator of *rbcL* in *Arabidopsis*. The triple *rbcS1a3b-1* x *mrl1-1* mutant exhibited a similar defect in *rbcL* translation as the double *rbcS1a3b-1* and single *mrl1-1* mutants (Figure 3.3C). The lack of additive effects suggests that the accumulating unassembled RbcL in the *rbcS1a3b-1* mutant might cooperatively regulate *rbcL* translation together with MRL1. These preliminary results together with other described CES interactions in *Chlamydomonas* (e.g., MCA1-mediated CES regulation of *petA*), may hint to a model in which unassembled RbcL directly or indirectly blocks the MRL1-mediated translational activation of the *rbcL* mRNA (Figure 3.3B). This could happen either by direct interaction between RbcL and MRL1 or the unassembled RbcL might outcompete MRL1 from binding to the 5' end of *rbcL* mRNA, which is supported by the RNA binding capacity of RbcL (Yosef et al., 2004). To test either of these scenarios, further work is planned in collaboration with Dr. Hannes Ruwe (Humboldt University, Berlin). Creation of tagged lines of MRL1 followed by western blot and polysome analyzes to assess the turnover of MRL1 and its association to polysomes in wild-type and *rbcS1a3b-1* backgrounds would clarify its involvement in the translational regulation of *rbcL*. In parallel, strategies to unravel the *rbcL* mRNA-binding proteome were established in this work (see below).

4.1.2 Evidence for a PSII CES network in embryophytes

One of the most intriguing CES regulation reported in *Chlamydomonas* is that of PSII that endorses a cascade initiated by PsbD and followed sequentially by the two CES subunits PsbA and PsbB (Figure 1.3). The CES cascade in PSII of *Chlamydomonas* has been initially discovered by pulse labeling experiments and was recently validated by targeted ribosome profiling (Minai et al., 2006; Trösch et al., 2018). To investigate PSII CES regulation in embryophytes, several PSII mutants were analyzed in this work. The knockdown mutant of *psbD* (KD-*psbD*) displayed a pronounced reduction in the translation output of *psbD*, *psbB*, *psbA*, and *psbH* (Figure 3.5E and Figure 3.6D); although, a pronounced decrease in *psbA* and *psbH* translation was only seen at early developmental stages (Figure 3.6D). Comparison of the results obtained in this work for KD-*psbD* mutant at the cotyledon stage with a *Chlamydomonas psbD* mutant (dU mutant; Trösch et al. (2018)) showed a high similarity in the translation dynamics of PSII subunits (Figure 3.6E). Consequently, my data support a model in which

PsbA, PsbB, and PsbH are CES subunits downstream of PsbD also in embryophytes. It is important to note that the effect on *psbA* was not significant, which could be due to the rather high-light intensity used for growth ($350 \mu\text{mol m}^{-2} \text{s}^{-1}$) that may cause an increase in the turnover of the D1 subunit. The increase in the *psbA* translation needed for PSII repair might mask the CES effect caused by the knockdown of *psbD*, a conclusion that has previously been drawn based on data in *Chlamydomonas* (Trösch et al., 2018). To disentangle *de novo* translation and PSII repair, the KD-*psbD* mutant could be analyzed under lower-light intensity.

In addition to the effects observed for PSII subunits, the KD-*psbD* mutant displayed also a decrease in the translation output of PSI subunits similar to what was observed in *Chlamydomonas* (Trösch et al., 2018) (Figure 3.6E). This effect points to a regulatory connection of PSII and PSI biogenesis. Interestingly, the KD-*psbD* mutant showed also alterations in ribosome pausing behavior within the reading frames of some of PSII and PSI subunits (*psbA*, *psbC*, *psaC*, and *psbB*) (Figure 3.8) suggesting that proper co-translational assembly of PSII subunits may be the checkpoint for proceeded elongation at specific positions. Further support for this hypothesis comes from similar patterns of altered ribosome pausing that were observed in the Δ *psbN* mutant, mainly in the *psbA* transcript (Figure 3.12) (although the overall effect on the translation output of *psbA* was rather mild). In contrast, the translation output of *psbB* was significantly affected in the Δ *psbN* mutant (Figure 3.11), which further points to PsbB as a CES subunit.

To further investigate the presence of a CES cascade in PSII, a T-DNA mutant of the *psbA* translation activator HCF173 was analyzed. Despite the strong defect in the translation output of *psbA*, no effect was observed on the translation of *psbB* (Figure 3.9). Similar results were obtained by Williams-Carrier et al. (2019) in a Ribo-seq analysis of the same mutant. Furthermore, the analysis of the *hcf244* T-DNA mutant (Chotewutmontri et al., 2020), which is deficient in another factor involved in the translational activation of *psbA* (Link et al., 2012), revealed only a mild defect on *psbB* ribosome footprint abundance. Taken together, these results question the existence of a CES cascade in PSII synthesis (Figure 3.6F) and raise the question of whether the translational feedback regulation of *psbB* is rather directly triggered by the reduced accumulation of PsbD instead of being caused by a cascade including PsbA. In addition, the effect on *psbH* expression appeared to be linked to the availability of PsbB since in mutants lacking an effect on *psbB* (for example *hcf173-2* (section 3.1.9) and Δ *psbD/C* (section 3.1.20.1)), no effect on *psbH* was observed. This observation is consistent with previous speculations of translational co-regulation of *psbB* and *psbH* (Levey et al., 2014; Trösch et al., 2018). Interestingly, the ribosome profiling analysis of *hcf173-2* mutant showed a drastic decrease in the transcript level of *psbA*, which can explain most of the defect observed on the level of translation output (Figure 3.9). This mutant has been analyzed in two studies in which a strong decrease in *psbA* transcript abundance was assessed by RNA hybridization (Schult et al., 2007; Williams-Carrier et al., 2019). However, due to the missing RNA quantification in these studies, the contribution of the defect in transcript accumulation

to the decreased *psbA* translation output was not quantitatively assessed in these studies. Based on the result obtained in this work, HCF173 seems to act mainly as a factor of *psbA* mRNA stabilization.

Furthermore, the knockdown of PsbE/F (cytochrome *b₅₅₉*), which is at the very top of the PSII assembly cascade (Komenda et al., 2004; Nickelsen and Rengstl, 2013), caused a decrease in the translation of *psbA* (Figure 3.10) (Météignier et al., 2020), which points to the speculation that this defect might be caused by an impairment of the early steps of PSII assembly. However, no effect was observed on *psbB* translation, which further contradicts the presence of an epistatic relationship between *psbA* and *psbB*. Altogether, based on these data, I came to the conclusion that PsbD might be the upstream subunit triggering the CES regulation of both PsbA and PsbB. This model is supported by the structural location of PsbD in between PsbA and PsbB within PSII (van Bezouwen et al., 2017; Wei et al., 2016) (Figure 4.1A). Furthermore, Wei et al. (2016) have shown that the C-termini of PsbA and PsbD are in close proximity and are both stabilized by a loop of PsbP that contacts both C-terminal tails (Figure 4.1B). This suggests that PsbA probably interacts with PsbD via its C-terminus (Figure 4.1C).

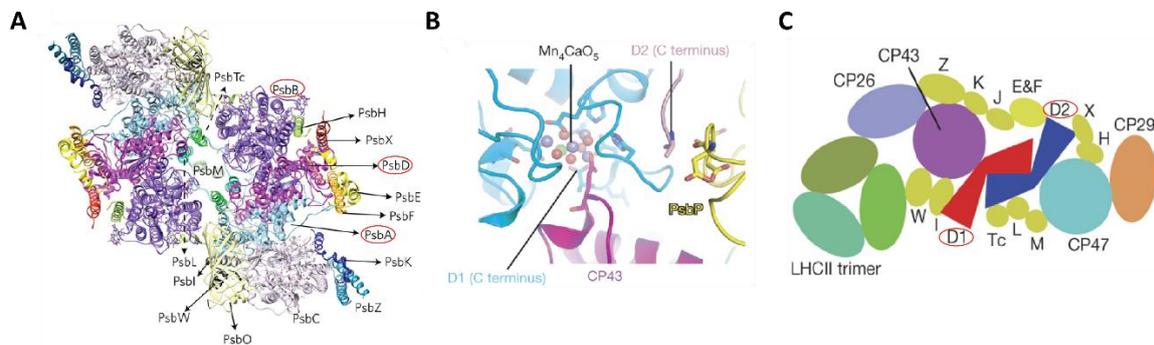


Figure 4.1: Structure of PSII complex

A. Structure of the PSII complex in *Arabidopsis*; PsbA, PsbB, and PsbD are marked with red circles. The figure was modified from van Bezouwen et al. (2017). **B.** Structure showing the PsbP loop involved in the stabilization of the C-terminus of PsbA (D1) and PsbD (D2). The figure was taken from Wei et al. (2016). **C.** Cartoon showing the arrangement of the spinach PSII-LHCII supercomplex; D1 and D2 are marked by red circles. The figure was modified from Wei et al. (2016).

To further elaborate this model and to deepen our understanding of assembly-dependent translation feedback regulation, a truncated PsbD mutant ($\Delta psbD/C$) was analyzed. Surprisingly, the truncation of the PsbD C-terminus caused a decrease in the translation output of *psbA* (Figure 3.21) but not of *psbB*, in contrast to the KD-*psbD* mutant (Figure 3.5 and Figure 3.6). These data suggest that the C-terminus of PsbD triggers the CES regulation of *psbA* whereas the N-terminus is the trigger for *psbB* CES regulation. In order to verify this model, further experiments are required such as immunoprecipitation of polysome-associated truncated PsbD followed by microarray hybridization of coprecipitated transcripts. Co-immunoprecipitation of *psbB* mRNA and the absence of *psbA* mRNA could support that the co-translational interaction with the C-terminus of PsbD is required for D1 synthesis. Additionally,

blue native PAGE experiments of the analyzed mutants shall reveal the alterations in the assembly intermediates of PSII and whether PsbA subunits can assemble into PSII with truncated PsbD.

Overall, PSII biogenesis in embryophytes seems to be tightly regulated by translation feedback regulation, however, at least somewhat differently than in *Chlamydomonas*. Based on the results presented here, I propose a CES network rather than a cascade for PSII in embryophytes (Figure 4.2). Further experiments are needed to elaborate on this model.

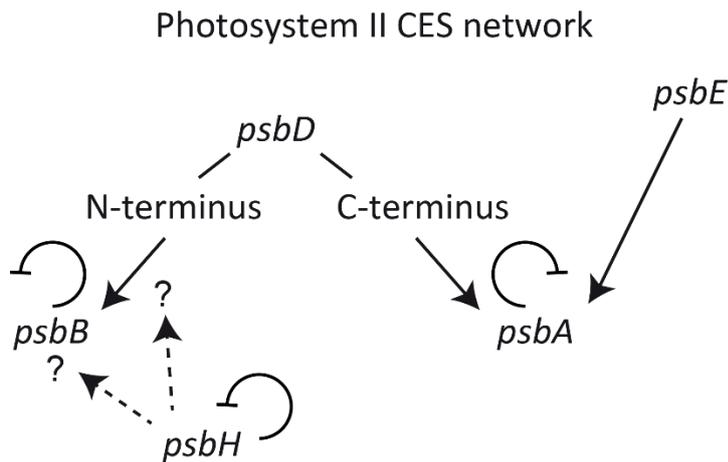


Figure 4.2: PSII CES network in embryophytes

Cartoon showing the epistatic relationships between the subunits of PSII based on data presented in this work.

Just before the end of my Ph.D., a working model of the negative feedback regulation of *psbA* was described by Chotewutmontri and Barkan (2020). In this model, the HCF244 complex consisting of HCF244, OHP1, and OHP2 is located in the thylakoid membrane and was proposed to activate *psbA* translation by impacting the activity of HCF173. Yet, the factor that fulfills the connection between the HCF173 and the HCF244 complex still to be identified. HCF136 facilitates the incorporation of PsbA into the HCF244 complex (Komenda et al., 2008; Plücker et al., 2002), which is thought to block the activation of *psbA* translation. This model was proposed to be the underlying mechanism controlling *de novo* PSII assembly and repair. To investigate whether this model holds for the PsbD-dependent regulation of *psbA* translation, I propose to create a double mutant of *psbD* and *hcf136*. If this model is the sole underlying mechanism of the CES regulation of *psbA*, it would be expected that in this double mutant unassembled PsbA cannot incorporate into the HCF244 complex due to the absence of HCF136 and thus cannot deactivate *psbA* translation. In this case, the absence of HCF136 would repress the PsbD-dependent CES regulation of *psbA*.

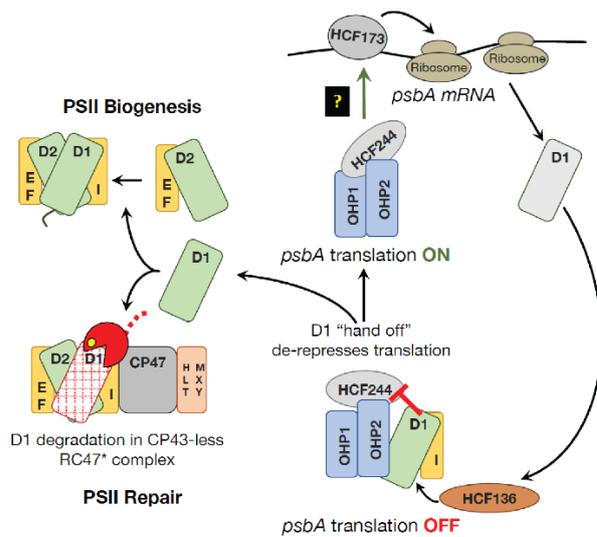


Figure 4.3: Proposed model for translational feedback regulation of *psbA*

The HCF244/OHP1/OHP2 complex mediates the translation activation of *psbA*, which is blocked by the association of the D1 subunit into the complex. This suppression is alleviated when D1 is released to bind to a D1-less PSII upon PSII biogenesis or repair. The Figure was taken from Chotewutmontri and Barkan (2020).

4.1.3 Identification of potential positive feedback regulation in Cyt *b₆f* complex

For the Cyt *b₆f* complex, the analysis of three independent overexpressors of *petA* ($\Delta ycf4$, $\Delta psal$, and $\Delta ycf10$) showed a potential feedback regulation that links the production of the two major subunits, PetA and PetB. The similarity of the observed effects in these mutants suggests that they are a consequence of the overexpression of *petA* and not a direct result of the knockout or co-overexpression of *ycf4*, *psal* or *ycf10*, respectively. Interestingly, the observed translational regulation of *petB* is different from that in *Chlamydomonas*, where PetB is required for PetA production but not vice versa. This finding further highlights the conclusion that feedback regulation of chloroplast translation evolved differently in embryophytes than in *Chlamydomonas*, which was drawn already based on the coordinated expression of PSII subunits. Further inspection of the data of these overexpressors suggests that *petA* is not a CES subunit since its increased translation output was triggered by its overexpression on RNA level and no negative feedback regulation on translation was observed as it would be the case for a CES subunit. This argument is partially weakened by the fact that the Cyt *b₆f* level significantly increased in $\Delta psal$ under standard conditions (15-20 % increase) (Schöttler et al., 2017). A similar increase in Cyt *b₆f* level was also observed in $\Delta ycf10$ mutant (unpublished data, personal communication Dr. Mark Schöttler, MPIMP). The higher accumulation of the complex in these mutants demonstrates that the bulk of excess PetA does not accumulate unassembled but rather assembles into the Cyt *b₆f* complex. It is important to mention that the extent of upregulation on translation output was higher than the complex accumulation level reflecting that PetA is not the only rate-limiting subunit of the Cyt *b₆f* complex. Taken this into consideration, the possibility that PetA is a CES subunit whose synthesis is dependent on the availability of its assembly partners PetD or PetB cannot be fully excluded. It was previously shown that the polysome association of *petA* decreased in $\Delta petB$ and $\Delta petD$ mutants in tobacco (Monde et al., 2000), suggesting that *petA* could be a potential CES subunit. The previously created Cyt *b₆f* knockout mutants in tobacco (Monde et al., 2000) were unable to grow autotrophically

and were not maintained in tissue culture (personal communication David Stern and Francis-Andre Wollman). Therefore, it was not possible to analyze these mutants with ribosome profiling. As an alternative, I analyzed a multisubunit knockout mutant, $\Delta psbB$ in which a cluster of genes is deleted including *petB* and the 5' end of *petD* (Figure 3.22). Ribo-seq analysis of $\Delta psbB$ confirmed that *petA* translation was not affected in the absence of PetD and PetB, which strongly argues against *petA* as a PetD- or PetB-dependent CES subunit. The partially contrasting effects observed in this work compared to previous studies shows the limitations of polysome and pulse labeling analyzes when used as only methods to study chloroplast translation. This was previously discussed in several studies (Chotewutmontri et al., 2020; Williams-Carrier et al., 2019; Zoschke et al., 2013) where the authors show that classical methods to study translation might be misleading as was the case for OHP1/OHP2 (Li et al., 2019) and LPE1 (Jin et al., 2018). Combining the data from the *petA* overexpressor lines and the $\Delta psbB$ mutant, it is tempting to speculate that the expression of *petA* and *petB* is closely co-regulated at the translational level by a positive feedback mechanism that was not described before in *Chlamydomonas*.

4.1.4 Search for potential CES regulation in the ATP synthase complex

The CES interaction between AtpB and AtpA described in *Chlamydomonas* (Drapier et al., 2007) was previously shown not to occur in maize and tobacco (Trösch et al., 2018; Zoschke et al., 2013). Results presented in this work verified whether this regulation may be found at an early developmental stage and whether the second CES interaction observed in *Chlamydomonas* between AtpC and AtpB is conserved in tobacco (Figure 1.3). My data showed that even at the cotyledon stage where the chloroplast biogenesis is going on at the highest activity, the synthesis rate of AtpA remains unchanged in the background of the *atpB* knock-down mutant. This substantiates the previous notion that the AtpB-dependent CES regulation of AtpA is not conserved in embryophytes. Conversely, the absence of the nucleus-encoded AtpC subunit in an *AtpC* antisense knockdown mutant caused a decrease in the translation output of *atpA*, *atpH*, and *atpF* together with subunits from other photosynthetic complexes (Figure 3.18). This effect could be due to an off-targeting effect of the antisense RNA that is supposed to be targeted to the *AtpC* mRNA. This hypothesis was later partially disproved based on the Ribo-seq analysis of the $\Delta atpB$ knockout mutant. In this mutant, the translation of all photosynthesis genes (except *ndhB*) decreased whereas that of the genetic system genes increased, suggesting that a severe defect in the assembly of the ATP synthase complex leads to a general defect in chloroplast gene expression, potentially due to energy depletion. That said, inference about CES regulation in the ATP synthase complex from data of constitutive mutants is rather problematic due to possible secondary effects resulting from the defect in the accumulation of this complex. In future experiments, an inducible mutant could be used to minimize such secondary effects by enabling a time-resolved analysis of translation in plants with reduced levels of AtpC or AtpB.

4.1.5 Preliminary analyses do not provide evidence for CES regulation in the NDH complex

The NDH complex was discovered based on the homology of its subunits with those of the mitochondrial respiratory complex I. The NDH complex was first identified in tobacco (Shinozaki et al., 1986) and *Marchantia polymorpha* (liverwort) (Ohyama et al., 1986). Given the fact that the CES model organism *Chlamydomonas* lacks a chloroplast NDH complex, no CES interactions that connect the synthesis of its subunits have been previously identified. Furthermore, the hierarchical assembly of the NDH complex is only partially understood (Shikanai, 2016) and knowledge of early assembly intermediates is missing. Consequently, it is unknown which core subunits are required for the efficient synthesis of later assembling subunits. Furthermore, the NDH complex is not essential for autotrophic growth in standard growth conditions, which complicates the designation of core subunits and peripheral/accessory subunits. In this work, two triple mutants were analyzed: $\Delta ndhA/H/I$ and $\Delta ndhC/K/J$. The first mutant was picked because of the position of NdhA as a connector between the membrane-associated subunits and hydrophilic membrane-integral subunits of the complex, suggesting that the NdhA subunit may assemble early. Additionally, based on the structure of the ubiquinone-binding channel in the mammalian complex I (Fedor et al., 2017), NdhA is part of the quinone binding site. The knockout of *ndhA/H/I* caused a downregulation in the translation output of *ndhG* and *ndhE*, both of which are located in the very same polycistronic transcription unit as *ndhA* in close proximity to the *aadA* insertion site (Figure 3.19B). Although it is tempting to speculate that this effect could be explained with a translation feedback regulation, a processing defect or another gene expression defect caused by the nearby *aadA* insertion cannot be excluded. RNA hybridization experiments are planned to reveal a potential processing defect and thereby rule out or validate a translational feedback regulation of *ndhG* and *ndhE*. The second mutant, $\Delta ndhC/K/J$, was analyzed to assess three additional subunits of the complex. Based on the results obtained, no CES regulation has been observed (Figure 3.19A), however, this does not exclude the potential presence of CES regulation between other subunits in the NDH complex. A comprehensive investigation of the assembly of the NDH complex seems to be valuable in order to be able to target the early-assembled subunits and conduct a more systematic search for CES interactions in the NDH complex.

4.1.6 Towards the confident identification of CES regulation and the examination of its interplay with protein degradation

Altogether, the results discussed above point to potentially diverged CES regulation between *Chlamydomonas* and embryophytes, which is not surprising considering the different gene expression mechanisms and involved *trans*-factors and *cis*-elements (Nickelsen et al., 2014; Sun and Zerges, 2015; Zoschke and Bock, 2018). Nevertheless, my data provide good evidence that assembly-dependent regulation of translation does also exist in embryophytes.

As an alternative experimental setup to validate the potential translation feedback regulation detected with ribosome profiling, *in vivo* protein pulse labeling with ^{35}S -methionine and ^{35}S -cysteine was established for tobacco to assess the protein synthesis rate of chloroplast proteins. Despite the technical difficulties of applying this technique for tobacco (Wittenberg et al., 2017), the approach was successfully established and used to confirm the CES regulation of RbcL and PsbA (Figure 3.2 and Figure 3.7). In a subsequent step, I combined *in vivo* pulse labeling with immunoprecipitation, which enabled the confirmation of the positive feedback regulation of *petB* translation (Figure 3.16), the whose synthesized product is not visible in standard pulse labeling approaches. This methodological setup will allow in future the pulse labeling analysis of protein synthesis for any potential CES subunit for which an antibody is available. Thereby, I solved the problem that many photosynthetic proteins are not visible or cannot be resolved in tobacco pulse labeling analyses.

Further investigation of how the CES process controls the biogenesis of photosynthetic complexes in embryophytes could be realized with the use of chimeric reporter genes expressed under the translational control of the 5' UTR from the mRNA of interest. This approach was used in *Chlamydomonas* to confirm the CES regulation in the Cyt *b₆f* complex. Thereby, it was shown that the *petA* 5' UTR is sufficient to reduce the synthesis of reporter proteins when they were expressed in the absence of the assembly partner PetB (Choquet et al., 2003). Another way would be to exchange the 5' UTR of the CES subunit and then to examine whether its rate of synthesis responds to the presence or absence of its assembly partner. The introduction of a reporter gene or exchange of the 5' UTR would require laborious chloroplast transformation. Therefore, only candidates whose CES regulation was validated with independent experimental approaches, i.e., by ribosome profiling and pulse labeling, will be considered as promising candidates for such follow-up experiments.

Besides the feedback regulation of translation, the assembly-dependent homeostasis of photosynthetic subunits depends on proteases that degrade the unassembled proteins. These two modes of regulation are complementary and not exclusive. Therefore, CES regulation might be more relevant during the biogenesis of photosynthetic complexes if the other mode of regulation, proteolysis, is limited. To test this hypothesis, some of the above-described mutants with defects in the assembly of specific photosynthetic complexes were crossed with ethanol inducible RNAi-based knockdown mutants of the major thylakoid and stroma FtsH and Clp proteases (Moreno et al., 2018). Future immunoblotting analyses shall reveal the accumulation of the subunits of the disrupted photosynthetic complex in a time-course manner and also reveal the substrate specificity of the protease. This will deepen the understanding of photosynthetic complexes assembly and the coordinated degradation of unassembled subunits.

4.1.7 General lessons on chloroplast translation gained from the analysis of numerous chloroplast mutants

In addition to the search for potential CES mechanism, the question whether overlapping or neighboring ORFs in the chloroplast are translationally coupled was addressed for several reading frames in my work. Translational coupling is known as the interdependence of translation efficiency between genes located in polycistronic mRNAs where alteration of the translation of the upstream gene would affect the translation output of the downstream gene. Translational coupling was observed in several prokaryotes and bacteriophages (Berkhout and van Duin, 1985; Schümperli et al., 1982; van de Guchte et al., 1991). Out of the four overlapping gene pairs identified in the chloroplast genome of tobacco (Adachi et al., 2012; Yukawa et al., 2005), I examined translation in *atpB-atpE* and *psbD-psbC*. For the pair of overlapping genes *atpB-atpE*, it was previously shown that they are not translationally coupled (Trösch et al., 2018; Zoschke et al., 2013). The data presented in this work confirmed the lack of translational coupling in this gene pair as no effect was observed on the translation output of *atpE* in the *atpB-GTG* mutant, which exhibits a strong *atpB* translation defect (Supplemental Figure 4). Similarly, despite the pronounced defect in *psbD* translation caused by the point mutation in its start codon in the *psbD-TTG* mutant, only a little defect was observed for *psbC*, arguing also against a strict translational coupling for this pair of genes (Figure 3.4). My results, however, suggest that the translation of *psbD* and *psbC* is partially coupled *in vivo*, which supports what was previously demonstrated *in vitro* (Adachi et al., 2012). It is important to mention that *psbC* has a separate promoter that produces a monocistronic *psbC* transcript, which can be translated without coupling (Yao et al., 1989). Taking this into consideration, it is possible that the translation defect of *psbC* located in the dicistronic *psbD/psbC* transcript might be attenuated. Furthermore, the presence of a translational coupling between *psaA* and *psaB* was clarified: The mutation of the SD sequence upstream of *psaA* affected its translation, however, the translation output of *psaB* and *rps14* remained unchanged suggesting that translational coupling in this transcription unit does not occur (Figure 3.13C). Overall, this work and previous work has shown only one case of only partial translational coupling between *psbD* and *psbC*.

In this work, two developmental stages were used for tobacco mutants: seedlings with four true leaves, and seedlings in a cotyledon stage just before the emergence of true leaves. In the cotyledon stage, the cotyledons make the largest fraction of the material used for ribosome profiling as the true leaves and hypocotyl of the seedlings are small. In contrast, the older seedlings consist mainly of mature photosynthetic leaves and young developing leaves. Taken this into account, the portion of the photosynthetically active tissue at the older developmental stage may be smaller than in the cotyledon stage. Therefore, the cotyledon tissue possesses a higher degree of maturity in comparison to seedlings with four true leaves. This was highlighted by the higher similarity in the translation output of KD-*psbD* mutant at the cotyledon stage to the *Chlamydomonas psbD* mutant, considering the mature nature of the chloroplast in vegetative *Chlamydomonas* cells cultures (Trösch et al., 2018).

4.1.8 Are there no true “neutral” insertion sites in the chloroplast genome?

All of the tobacco chloroplast mutants used in this work possess an *aadA* cassette used as a selection marker and in some mutants also for interrupting or deleting target genes. Usually, the *aadA* cassette is driven by a strong promoter derived from a highly expressed gene, e.g., 16S rRNA or *psbA*, to ensure the high expression needed for selection (Zhang et al., 2012). Transcription termination is inefficient in the chloroplast (Stern and Gruissem, 1987), which was also highlighted in the data presented in this work. The overexpression driven by the *aadA* overactive promoter coupled with inefficient transcription termination might cause a simultaneous overexpression of downstream transcripts or produce a high amount of antisense transcripts against these transcripts (e.g., Figure 3.15). These read-through transcripts may interfere with the expression, processing, or stability of mRNAs of the downstream genes. In most of the mutants analyzed in this work, the *aadA* insertion caused overexpression of the downstream genes located on the same strand for example in *psbD*-GTG, *psbD*-TTG, Δ *psaI*, Δ *ycf10*, and Δ *ycf4*. In addition, a defect in the expression of the downstream genes located on the opposite strand was observed for example *rps15* in *ndhA/H/I* mutant and *clpP* in Δ *psbB* operon mutant. These defects are potentially caused by the overexpression of the antisense transcript driven by the upstream *aadA* cassette. To control for the general effects that *aadA* insertion may impose on chloroplast gene expression (i.e., the burden of strong expression of an additional gene), plants containing an *aadA* cassette in the chloroplast genome were used as controls. Whenever possible, a control with an *aadA* cassette inserted at the same position was used, such as the control for the KD-*psaA* mutant (section 3.1.13). For the Δ *psbN* mutant, the RB70 line was used as a control. In this line, the *aadA* cassette was introduced in the intergenic region between the *trnG* and *trnfM* genes, a position considered to be neutral (Ruf et al., 2001). Interestingly, no effects were observed on the downstream genes in this line. For most of the other mutants, pRB8 plants were used as controls. In the pRB8 control, the *aadA* cassette is inserted downstream of *psbE* on the same strand. By virtue of the insertion site, read-through transcription may interfere with the expression of *petA* located downstream of the *aadA* cassette in the antisense direction. The read-through transcripts might form a double stranded RNA thus sequestering the *petA* mRNA and blocking its translation. Alternatively, they might interfere with RNA-binding of the factors involved in the stabilization or translation of the *petA* transcript. Another possibility could be that the strong *aadA* promoter reduces the expression of the weaker *petA* promoter in what is known as “transcriptional interference” (Shearwin et al., 2005). Consequently, the insertion of the *aadA* cassette may cause a mild decrease in the expression of *petA* in pRB8 control. Indeed, it was recently shown that pRB8 lines contain reduced Cyt *b₆f* levels (Loiacono et al., 2019). Hence, by calculating the expression ratios between mutant and pRB8 control, the expression of *petA* might seem slightly increased in the mutants versus the pRB8 control. In the mutants where *petA* is overexpressed, it is important to take into account that the overexpression effect might be slightly amplified by normalization to the pRB8 control. Taken together, finding a neutral insertion site for chloroplast genome transformation remains challenging as long as transcription termination cannot be tightly

controlled. Although more analysis is needed to draw a firm conclusion regarding the RB70 line, it is tempting to speculate that it could be an advantage to place the *aadA* cassette upstream of a tRNA whose secondary structure may serve as a terminator.

Furthermore, some effects were observed for genes that were not in close proximity to the *aadA* insertion sites. Lowly expressed genes, such as *ycf10*, *matK*, and the *rpo* genes exhibited altered expression in several mutants, and also a slight defect was observed for *rbcL* in some of the analyzed mutants. These alterations might be a result of secondary effects of disrupted photosynthesis.

4.2 Towards unraveling of the RNA-binding proteomes and localizations of *psbA* and *rbcL* mRNAs using an aptamer-based affinity purification

In order to identify the translation factors involved in CES regulation, an aptamer-based affinity purification was adapted to chloroplasts. Two independent aptamers were used to specifically enrich for *psbA* and *rbcL* transcripts, confirming the reliability of the approach. Furthermore, the results presented in this work showed that the aptamers can be employed to examine the suborganellar localization of chloroplast-encoded transcripts.

4.2.1 Chloroplast transformation and homoplasmy of transplastomic plants

Aptamer tagging is an established technique that has been used for the affinity purification of RNPs in *E.coli*, yeast, and human cells (section 1.7.2), however, it was not previously used in chloroplasts. In this work, aptamer-based affinity purification was adapted for chloroplast transcripts. For that, two aptamers that were identified by SELEX (Sephadex-binding and streptavidin-binding) and the MS2 aptamer derived from the bacteriophage MS2, one of the most widely used aptamers, were inserted into the 3' UTR of *psbA* and *rbcL*. Plants tagged with the Sephadex- and the streptavidin-binding aptamers were previously created (Dr. Reimo Zoschke, MPIMP). *psbA*-MS2 and *rbcL*-MS2 tags were designed and plants were created in this work by chloroplast transformation. The homoplasmy of all these lines was tested by PCR and Southern blot analysis (Figure 3.27A and Figure 3.28A). Several homoplastomic lines were obtained for each aptamer, demonstrating that there is no major negative selection against the tag. Furthermore, all tagged lines developed like the wild type (Figure 3.27C and Figure 3.28C). However, *psbA*-seph and *rbcL*-ctrl lines were male sterile, drawback phenotype that is regularly observed after chloroplast transformation, and subsequent *in vitro* propagation in tissue culture (personal communication with Dr. Stephanie Ruf, MPIMP). Given the absence of a similar phenotype in the other tagged and control lines, it is unlikely that the integration of an aptamer approximately 100 bp downstream of the translation stop codon of *psbA* and *rbcL* does substantially affects the expression of those genes or the physiology of the plants.

4.2.2 Establishment of an aptamer-based affinity purification for chloroplast transcripts

In this work, MS2-based affinity purification has been successfully adapted for chloroplasts and allowed the purification of *psbA* and *rbcl* transcripts. The specificity of the purification has been verified by microarray hybridization (Figure 3.30E). In addition to the targeted transcripts, rRNAs were co-purified in both pulldowns (Figure 3.30E). These rRNAs most likely derive from the ribosomes that remain attached to the mRNAs due to the applied elongation inhibitor (chloramphenicol). Notably, more rRNA signals and more ribosomal proteins were detected in the *rbcl*-MS2 pulldown in comparison to the *psbA*-MS2 pulldown. One possible reason could be explained by the fact that the *psbA* mRNA is distributed into two fractions: a polysome-associated fraction and an untranslated mRNA pool, hence a mixture of translated and untranslated *psbA* mRNAs are purified, whereas *rbcl* mRNAs are predominantly found in polysomes. Furthermore, streptavidin-mediated purification has been successfully optimized to enrich *psbA* mRNA, however, attempts to purify *rbcl* using the streptavidin-binding aptamer failed. One possible explanation could be that the *rbcl* transcript is folded in a way that the tag is occluded in a secondary structure and inaccessible to the streptavidin beads. On the other hand, *rbcl* was specifically purified using the Sephadex-binding aptamer but it was partially degraded (Figure 3.33).

Comparison of the purification of *psbA* using the different aptamers showed that the MS2 and the streptavidin-binding aptamers are equally efficient to purify *psbA*, and that both performed better than the Sephadex-binding aptamer (Figure 3.34B, C). In contrast, for *rbcl*, the pulldown with the Sephadex-binding aptamer yielded the highest enrichment, however, the transcript remained intact only when purified with the MS2 aptamer (Figure 3.34A, C). Based on the enrichment, the integrity of the purified RNA, and the specificity of the purification I could show that the MS2 aptamer is the most suitable for affinity purification of chloroplast transcripts. Altogether, the data shown in this work highlight the possibility to specifically purify chloroplast transcripts using an aptamer-based approach. Despite the rigorous studies to identify the factors that bind to the *psbA* 5' UTR and trigger its translation, so far HCF173 is the only known translational activator that directly binds the *psbA* transcripts (McDermott et al., 2019; Watkins et al., 2019). Therefore, the detection of HCF173 in the pulldown can be used as a positive control to validate the enrichment in experiments intended to identify the *psbA*-binding proteome. In addition, other proteins that were shown to bind to the *psbA* mRNA could also be used as positive controls (e.g., CP33B (Teubner et al., 2020; Watkins et al., 2019), CP33C, and SRRP1 (McDermott et al., 2019; Watkins et al., 2019)). On the other hand, in *rbcl*-MS2 pulldowns, MRL1 PPR protein can be used as a positive control. In the first mass spectrometry trials, HCF173 was not detected in the *psbA*-MS2 pulldown. This suggests that either the *psbA*-binding proteome is not sufficiently enriched to be detected by mass spectrometry or that HCF173 is lost during the purification procedure. Furthermore, apart from the ribosomal proteins, no chloroplast-localized RNA-binding proteins were detected (section 3.2.5). Similarly, in the *rbcl*-MS2 pulldown, no chloroplast-localized RNA-binding proteins apart from ribosomal proteins were detected, again reflecting a lack of proteome

enrichment. Overall, the results indicate that further optimizations need to be implemented to enrich the protein fraction of the purified RNPs and to optimize their detection by mass spectrometry. To do so, different modifications are planned to be tested. This includes:

- 1) Addition of a cross-linking step using 1 % formaldehyde, which may, however, results in the increase of false-positive proteins by stabilizing transient interactions.
- 2) Enrichment of the translated transcripts by including a polysome pre-purification step and using the polysomes for pulldown.
- 3) Enrichment of chloroplast transcripts by using the chloroplast lysate for pulldown instead of the whole cell lysate. On the other hand, this would require a chloroplast isolation step. However, it should be considered that the longer the isolation process the less likely to capture the innate *trans*-acting factors therefore including a chloroplast isolation step might introduce artificial effects.

As reported previously, *psbA* translation is light-induced (Chotewutmontri and Barkan, 2018; Schuster et al., 2020). Therefore, the detection of novel or known regulators of *psbA* might be influenced by the growth conditions. Moreover, based on the data presented here, the CES regulation of *psbA* was seen at the cotyledon stage (section 3.1.7.2), thus the developmental stage might additionally influence translational regulation. Accordingly, affinity purification experiments under different growth conditions (e.g., light quantity and quality) and different developmental stages would enable the identification of factors that control the transcript expression at specific stages or growth conditions.

My data demonstrate that an aptamer-based transcript affinity purification is in general applicable for chloroplasts. Nonetheless, the protocol still needs further refinements to improve the mass spectrometric detection of RNA-binding proteins including translation factors. From these proteins, candidates with known RNA-binding domains (e.g., PPR or RRM) and domains of unknown function will be selected for further analysis. Since chloroplast RNP complexes contain many RNA-binding proteins with no direct relevance for translational regulation (e.g., RNA processing/splicing/editing factors, etc.), the functions of the identified factors need to be further elucidated by genetic and biochemical approaches. Most importantly, this method can be applied to identify *trans*-factors involved in CES regulation of *psbA* and *rbcL* and consequently help to unravel the molecular mechanism of translational feedback regulation in embryophytes. Overall, this approach can be basically applied to purify any chloroplast RNA. However, purifying transcripts of a lowly expressed gene might be less straightforward, given that the lower the targeted RNA is expressed the more input is needed, and hence, the higher the background of other abundant RNAs. Also if the transcript resides in a polycistronic transcription unit verifying the co-purified factors might be challenging.

4.2.3 Pros and cons of the aptamer-based affinity purification of chloroplast RNPs

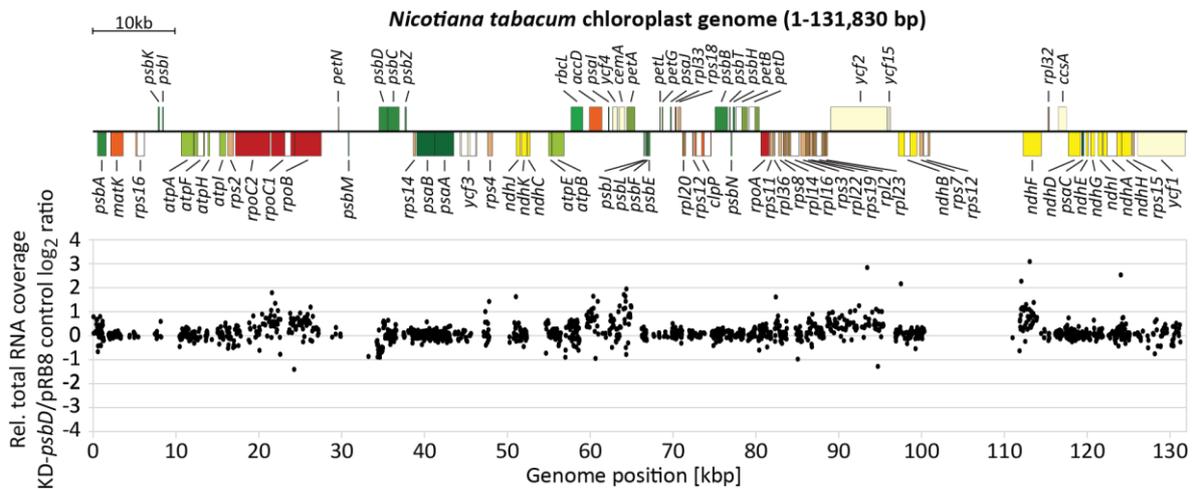
Two different approaches were recently described to specifically purify the RNA-binding proteome of chloroplast transcripts. A designer PPR protein that binds specifically to *psbA* was used to pulldown the transcript together with RNP. The sequence specificity of the PPR protein relies on many features like the RNA structure and the protein concentration. It was shown that the binding of the designed PPR protein is inhibited by RNA secondary structure (McDermott et al., 2018) rendering the binding of the designer PPR protein more favorable for less structured RNAs. This was indeed the case in McDermott et al. (2019) where the off-targeting binding was observed. Additionally, biotinylated antisense oligonucleotides previously used to unravel the proteins binding to non-coding RNAs in mammalian cells was applied in chloroplasts (Chu et al., 2015; McHugh et al., 2015; Watkins et al., 2019). This RNA-based approach enabled the purification of *psbA* and its bound proteins validating the data from McDermott et al. (2019). However, the major limitation of this technique is that the antisense oligonucleotides might sequester the mRNA and block the binding of RBPs, which subsequently affect the trafficking and formation of the RNP complex. The aptamer-based approach presented in this work permit to overcome these limitations. High-specific enrichment of *psbA* and *rbcL* was achieved using the MS2 aptamer and only *psbA* using the streptavidin-binding aptamer. However, introducing a tagging aptamer is only possible by chloroplast transformation, which very laborious. Furthermore, it should be noted, although this was not the case in this work, that the insertion of an RNA element can alter the RNA structure and function of the target transcript.

4.2.4 Using an aptamer to study the suborganellar localization of chloroplast transcripts

Aptamers were also used to unravel the localization of transcripts (Forrest and Gavis, 2003; Morisaki et al., 2016; Sheth and Parker, 2003; Wang et al., 2012; Xiao et al., 2005). Hence, I sought to examine the suborganellar localization of *psbA* and *rbcL* mRNAs using the MS2 aptamer as a bait. Most of the *rbcL* transcripts were localized in the stroma as expected due to the stromal localization of *rbcL* translation and Rubisco assembly (Figure 3.31B) (Hauser et al., 2015; Vitlin Gruber and Feiz, 2018; Zoschke and Barkan, 2015). Conversely, *psbA* mRNA was detected in the stroma as well as near the thylakoid membrane (Figure 3.31A). This dual distribution supports the presence of two pools of *psbA* transcripts in the chloroplast: actively translated *psbA* tethered to the thylakoid membrane and ribosome-free stromal *psbA* (Legen and Schmitz-Linneweber, 2017; Zoschke and Barkan, 2015). Though preliminary, these results are very promising and show the potential of this technique to be applied for other chloroplast RNAs. In order to obtain a better resolution, further optimizations are needed including the adjustment of the incubation times with the primary and secondary antibody and the amount of MS2-MBP fusion protein added. As a next step, co-localization of the transcript and the translating ribosomes would enable the visualization of the translation sites within the chloroplast. In mammalian cells, fluorescence-based co-localization of the transcript and the translating ribosomes has been established and enabled the quantification of single RNA translation dynamics (Katz et al., 2016;

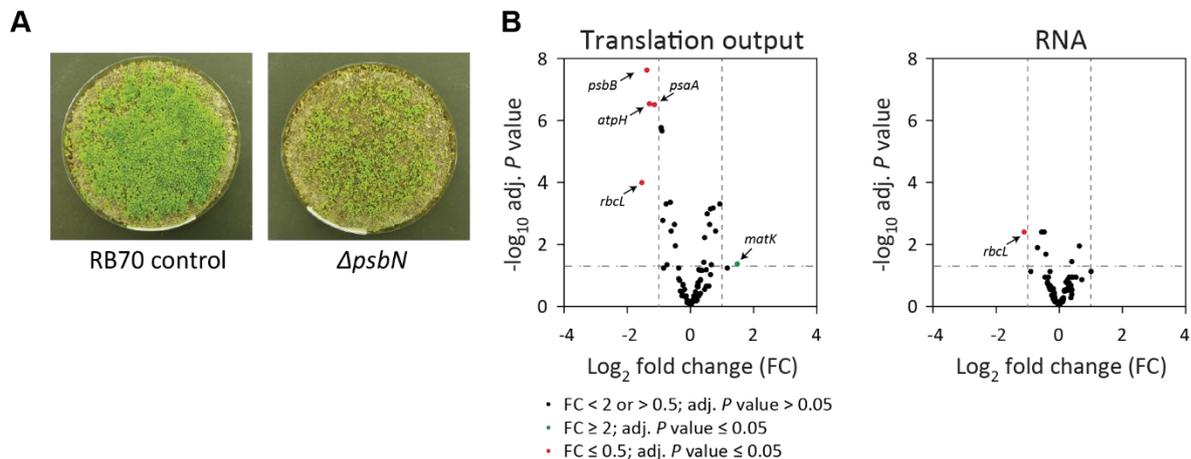
Morisaki et al., 2016; Wu et al., 2016). Lines expressing GFP-tagged nucleus-encoded chloroplast ribosomal proteins are available (unpublished lines provided by Prof. Dr. Ralph Bock, MPIMP). On the other hand, the MS2 coat protein can be fused to a fluorescent protein and thus used to mark the MS2-tagged transcripts. Subsequently, the MS2 lines can be crossed with the GFP-tagged ribosomal proteins lines and a fluorescently-based co-localization analysis can be performed. Eventually, this analysis shall mark the transcripts undergoing translation and the translation ‘hot-spots’ in the chloroplast. This could be used to identify *psbA* and *rbcL* translation sites in land plant chloroplasts and compare these to the ones demonstrated in the *Chlamydomonas* chloroplast (Uniacke and Zerges, 2009). However, it should be noted that fluorescence-based suborganellar co-localization is technically still challenging due to the autofluorescence emitted by the chlorophyll.

5 Supplemental information



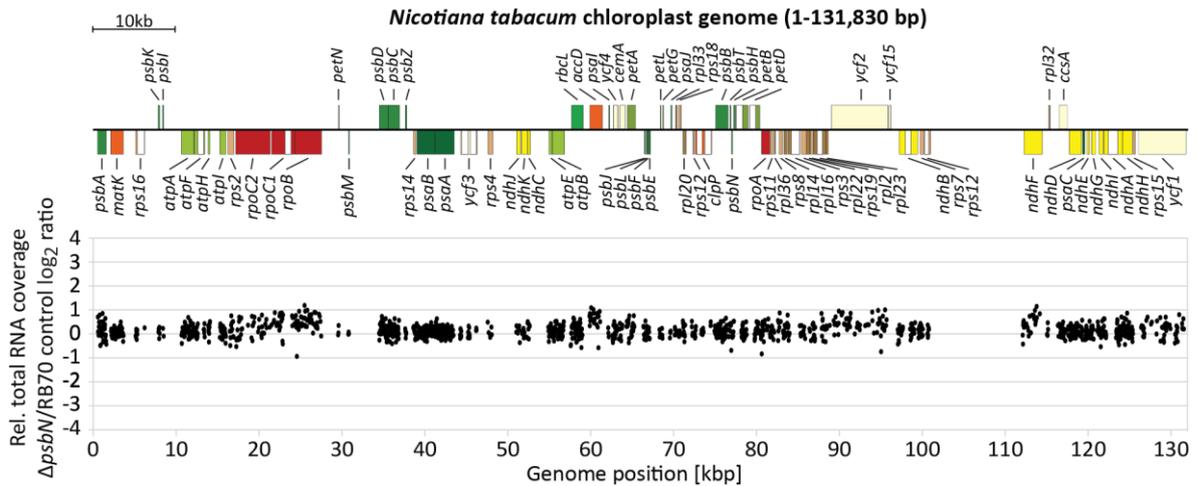
Supplemental Figure 1: Transcriptome-wide analysis of the total RNA coverage in KD-*psbD* mutant reveals no significant alterations.

Chloroplast genome-wide comparison of the relative local total RNA coverage. For each probe located in an ORF, the total RNA signal was normalized to the sum of the signals of all the probes located in the same ORF. Ratios of the total RNA coverage in KD-*psbD* in comparison to pRB8 control were calculated and plotted according to the position in the chloroplast genome. Results were obtained from three biological replicates. Statistical test was performed as described in section 3.1.8 and no significant changes were observed. Refer to Figure 3.8A for labeling details.



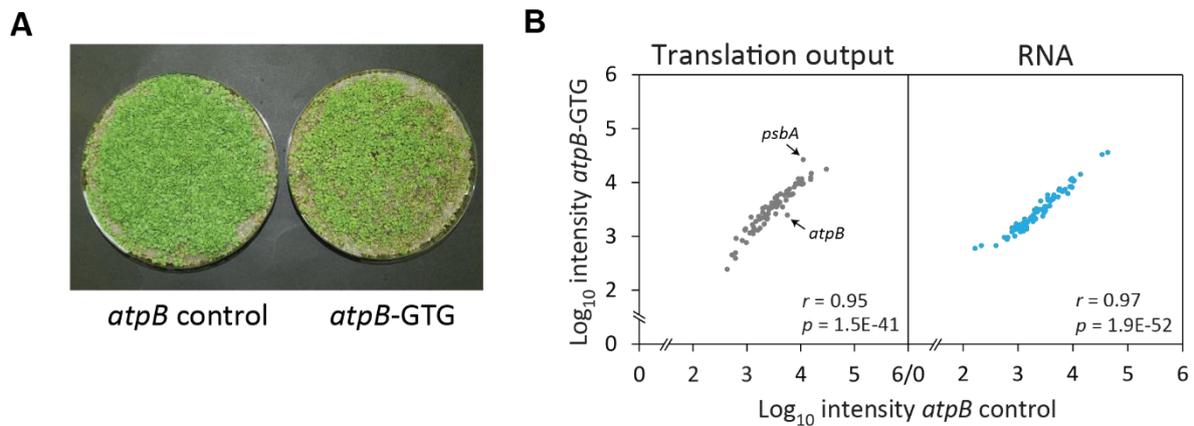
Supplemental Figure 2: Assembly defect of PSII in $\Delta psbN$ causes *psbB* downregulation at the cotyledon stage.

A. Mutant and control were grown as described before (see Figure 3.6A). **B.** Translation output and transcript abundance from three biological replicates were calculated and represented as volcano plots as described in Figure 3.4B.



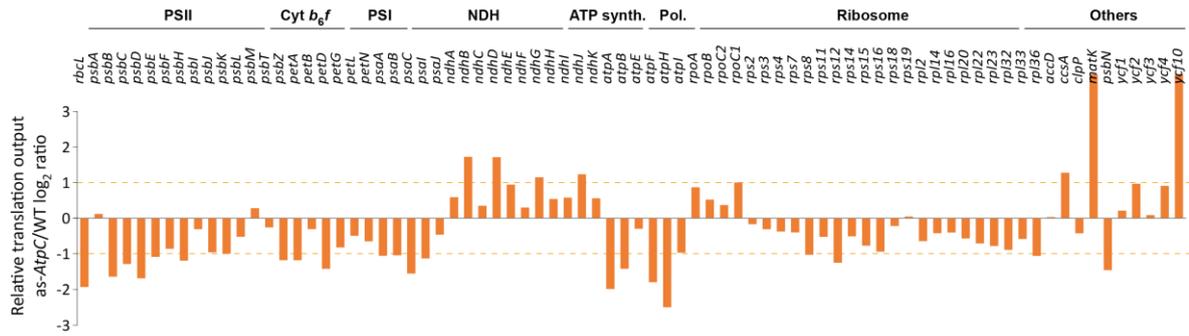
Supplemental Figure 3: Transcriptome-wide analysis of the total RNA coverage in $\Delta psbN$ mutant at the cotyledon stage.

Ratios of local total RNA coverage in $\Delta psbN$ compared to RB70 control at the cotyledon stage are plotted according to the position in the tobacco chloroplast genome. The results were obtained from three biological replicates and the average values are plotted. Labeling details are given in Figure 3.8A.

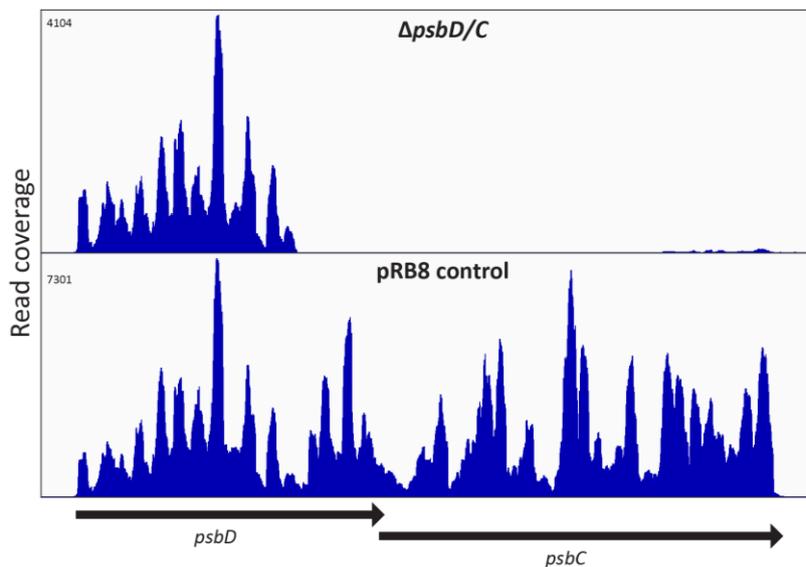


Supplemental Figure 4: Ribosome profiling of *atpB*-GTG at the cotyledon stage.

A. *atpB*-GTG mutant and *atpB* control were grown and harvested as described in Figure 3.6A. **B.** Transcript abundance and translation output from one biological replicate were calculated and plotted as described in Figure 3.4E.

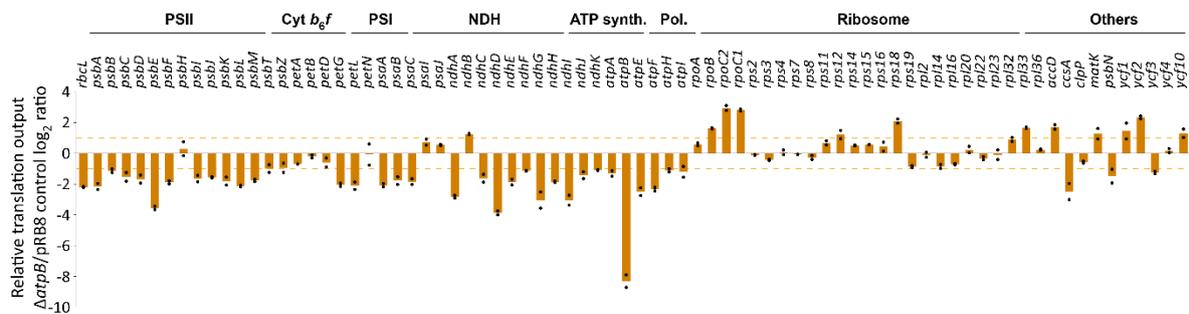


Supplemental Figure 5: Ratio of the translation output in *as-AtpC* in comparison to WT.
 Ratio of the translation output in the mutant in comparison to the WT. Data was collected from two biological replicates.



Supplemental Figure 6: Truncated PsbD is produced by the *aadA* insertion.

Screenshot from the Integrated Genome Viewer (IGV) showing the ribosome footprints distribution along *psbD* and *psbC* ORFs in $\Delta psbD/C$ and pRB8 control. The maximum y-axis values are shown in the upper left corner.



Supplemental Figure 7: Ratio of the translation output in $\Delta atpB$ in comparison to pRB8.
 Data was collected from two biological replicates. Values from each biological replicate are shown.

Supplemental Table 1: Reproducibility of ORF average of transcripts between the biological replicates

Correlation analysis of the ORF average of transcript abundance between the biological replicates was performed. The Pearson's correlation R values are given in the table. Rep: Replicate.

<i>as-RBCS</i>	SR1	Rep. 1	Rep. 2	Rep. 3	<i>as-RBCS</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.984	1		Rep. 2	0.882	1	
	Rep. 3	0.979	0.992	1	Rep. 3	0.946	0.983	1
<i>psbD-TTG</i>	WT	Rep. 1	Rep. 2	Rep. 3	<i>psbD-TTG</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.986	1		Rep. 2	0.991	1	
	Rep. 3	0.993	0.997	1	Rep. 3	0.992	0.997	1
<i>psbD-GTG</i>	WT	Rep. 1	Rep. 2	Rep. 3	<i>psbD-GTG</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.991	1		Rep. 2	0.994	1	
	Rep. 3	0.991	0.985	1	Rep. 3	0.992	0.980	1
KD- <i>psbD</i> 21 days stage	PRB8 control	Rep. 1	Rep. 2	Rep. 3	KD- <i>psbD</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.991	1		Rep. 2	0.996	1	
	Rep. 3	0.986	0.995	1	Rep. 3	0.993	0.994	1
KD- <i>psbD</i> cotyled on stage	PRB8 control	Rep. 1	Rep. 2	Rep. 3	KD- <i>psbD</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.946	1		Rep. 2	0.942	1	
	Rep. 3	0.952	0.994	1	Rep. 3	0.964	0.995	1
<i>hcf111-1</i>	WT Col-0	Rep. 1	Rep. 2	Rep. 3	<i>hcf111-1</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.991	1		Rep. 2	0.989	1	
	Rep. 3	0.991	0.999	1	Rep. 3	0.990	0.999	1
Δ <i>psbN</i> 21 days stage	RB70 control	Rep. 1	Rep. 2	Rep. 3	Δ <i>psbN</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.985	1		Rep. 2	0.991	1	
	Rep. 3	0.961	0.97	1	Rep. 3	0.978	0.966	1

<i>ΔpsbN</i> cotyled on stage	RB70 control	Rep. 1	Rep. 2	Rep. 3		<i>ΔpsbN</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.971	1			Rep. 2	0.982	1	
	Rep. 3	0.985	0.989	1		Rep. 3	0.982	0.988	1
KD- <i>psaA</i>	<i>psaA</i> - control	Rep. 1	Rep. 2			KD- <i>psaA</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.976	1			Rep. 2	0.894	1	
<i>psad1-1</i>	Col-0	Rep. 1	Rep. 2			<i>psad1-1</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.901	1			Rep. 2	0.937	1	
<i>Δycf4</i>	PRB8 control	Rep. 1	Rep. 2	Rep. 3		<i>Δycf4</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.998	1			Rep. 2	0.997	1	
	Rep. 3	0.993	0.991	1		Rep. 3	0.988	0.991	1
<i>ΔpsaI</i>	PRB8 control	Rep. 1	Rep. 2	Rep. 3		<i>ΔpsaI</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.990	1			Rep. 2	0.987	1	
	Rep. 3	0.969	0.984	1		Rep. 3	0.968	0.988	1
<i>Δycf10</i>	PRB8 control	Rep. 1	Rep. 2	Rep. 3		<i>Δycf10</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.996	1			Rep. 2	0.995	1	
	Rep. 3	0.996	0.997	1		Rep. 3	0.993	0.995	1
<i>ΔpetL</i>	PRB8 control	Rep. 1	Rep. 2			<i>ΔpetL</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.992	1			Rep. 2	0.996	1	
as- <i>AtpC</i>	WT	Rep. 1	Rep. 2			as- <i>AtpC</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.949	1			Rep. 2	0.975	1	

Supplemental Table 2: Reproducibility of ORF average of ribosome footprint between the biological replicates

Correlation analysis of the ORF average of ribosome footprint between the biological replicates was performed. The Pearson's correlation R values are given in the table. Rep: Replicate.

<i>as-RBCS</i>	SR1	Rep. 1	Rep. 2	Rep. 3	<i>as-RBCS</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.94	1		Rep. 2	0.904	1	
	Rep. 3	0.906	0.971	1	Rep. 3	0.921	0.982	1
<i>psbD-TTG</i>	WT	Rep. 1	Rep. 2	Rep. 3	<i>psbD-TTG</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.996	1		Rep. 2	0.992	1	
	Rep. 3	0.997	0.996	1	Rep. 3	0.991	0.986	1
<i>psbD-GTG</i>	WT	Rep. 1	Rep. 2	Rep. 3	<i>psbD-GTG</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.993	1		Rep. 2	0.990	1	
	Rep. 3	0.986	0.975	1	Rep. 3	0.993	0.985	1
KD- <i>psbD</i> 21 days stage	PRB8 control	Rep. 1	Rep. 2	Rep. 3	KD- <i>psbD</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.965	1		Rep. 2	0.960	1	
	Rep. 3	0.970	0.982	1	Rep. 3	0.976	0.968	1
KD- <i>psbD</i> Cotelyd ons stage	PRB8 control	Rep. 1	Rep. 2	Rep. 3	KD- <i>psbD</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.934	1		Rep. 2	0.921	1	
	Rep. 3	0.930	0.980	1	Rep. 3	0.873	0.971	1
<i>hcf111-1</i>	WT Col-0	Rep. 1	Rep. 2	Rep. 3	<i>hcf111-1</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.990	1		Rep. 2	0.981	1	
	Rep. 3	0.983	0.993	1	Rep. 3	0.931	0.976	1
Δ <i>psbN</i> 21 days stage	RB70 control	Rep. 1	Rep. 2	Rep. 3	Δ <i>psbN</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.993	1		Rep. 2	0.990	1	
	Rep. 3	0.957	0.969	1	Rep. 3	0.913	0.945	1

<i>ΔpsbN</i> cotyled on stage	RB70 control	Rep. 1	Rep. 2	Rep. 3		<i>ΔpsbN</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.976	1			Rep. 2	0.900	1	
	Rep. 3	0.988	0.991	1		Rep. 3	0.951	0.981	1
KD- <i>psaA</i>	<i>psaA</i> - control	Rep. 1	Rep. 2			KD- <i>psaA</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.946	1			Rep. 2	0.923	1	
<i>psad1-1</i>	Col-0	Rep. 1	Rep. 2			<i>psad1-1</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.933	1			Rep. 2	0.905	1	
<i>Δycf4</i>	PRB8 control	Rep. 1	Rep. 2	Rep. 3		<i>Δycf4</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.981	1			Rep. 2	0.983	1	
	Rep. 3	0.985	0.976	1		Rep. 3	0.971	0.966	1
<i>ΔpsaI</i>	PRB8 control	Rep. 1	Rep. 2	Rep. 3		<i>ΔpsaI</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.958	1			Rep. 2	0.957	1	
	Rep. 3	0.959	0.989	1		Rep. 3	0.966	0.992	1
<i>Δycf10</i>	PRB8 control	Rep. 1	Rep. 2	Rep. 3		<i>Δycf10</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.983	1			Rep. 2	0.990	1	
	Rep. 3	0.982	0.982	1		Rep. 3	0.987	0.988	1
<i>ΔpetL</i>	PRB8 control	Rep. 1	Rep. 2			<i>ΔpetL</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.997	1			Rep. 2	0.989	1	
as- <i>AtpC</i>	WT	Rep. 1	Rep. 2			as- <i>AtpC</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.932	1			Rep. 2	0.977	1	

Supplemental Table 3: Reproducibility of the probe signals of the transcript between the biological replicates

Correlation analysis of the transcript probe signals located in the protein-coding regions from the biological replicates was undertaken. The Pearson's correlation R values are given in the table. Rep: Replicate.

<i>as-RBCS</i>	SR1	Rep. 1	Rep. 2	Rep. 3	<i>as-RBCS</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.887	1		Rep. 2	0.793	1	
	Rep. 3	0.906	0.976	1	Rep. 3	0.874	0.978	1
<i>psbD-TTG</i>	WT	Rep. 1	Rep. 2	Rep. 3	<i>psbD-TTG</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.979	1		Rep. 2	0.984	1	
	Rep. 3	0.987	0.992	1	Rep. 3	0.988	0.995	1
<i>psbD-GTG</i>	WT	Rep. 1	Rep. 2	Rep. 3	<i>psbD-GTG</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.986	1		Rep. 2	0.984	1	
	Rep. 3	0.966	0.954	1	Rep. 3	0.971	0.943	1
KD- <i>nsbD</i> 21 days stage	PRB8 control	Rep. 1	Rep. 2	Rep. 3	KD- <i>psbD</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.962	1		Rep. 2	0.970	1	
	Rep. 3	0.963	0.989	1	Rep. 3	0.969	0.983	1
KD- <i>psbD</i> Cotelyd ons stage	PRB8 control	Rep. 1	Rep. 2	Rep. 3	KD- <i>psbD</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.876	1		Rep. 2	0.920	1	
	Rep. 3	0.870	0.980	1	Rep. 3	0.930	0.986	1
<i>hcf111-1</i>	WT Col-0	Rep. 1	Rep. 2	Rep. 3	<i>hcf111-1</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.980	1		Rep. 2	0.961	1	
	Rep. 3	0.984	0.993	1	Rep. 3	0.965	0.995	1
Δ <i>psbN</i> 21 days stage	RB70 control	Rep. 1	Rep. 2	Rep. 3	Δ <i>psbN</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.938	1		Rep. 2	0.957	1	

	Rep. 3	0.918	0.898	1		Rep. 3	0.926	0.854	1
<i>ΔpsbN</i> cotyled on stage	RB70 control	Rep. 1	Rep. 2	Rep. 3		<i>ΔpsbN</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.954	1			Rep. 2	0.968	1	
	Rep. 3	0.958	0.977	1		Rep. 3	0.963	0.984	1
KD- <i>psaA</i>	<i>psaA</i> - control	Rep. 1	Rep. 2			KD- <i>psaA</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.952	1			Rep. 2	0.889	1	
<i>psad1-1</i>	Col-0	Rep. 1	Rep. 2			<i>psad1-1</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.826	1			Rep. 2	0.906	1	
<i>Δycf4</i>	PRB8 control	Rep. 1	Rep. 2	Rep. 3		<i>Δycf4</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.991	1			Rep. 2	0.993	1	
	Rep. 3	0.966	0.967	1		Rep. 3	0.965	0.968	1
<i>ΔpsaI</i>	PRB8 control	Rep. 1	Rep. 2	Rep. 3		<i>ΔpsaI</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.932	1			Rep. 2	0.937	1	
	Rep. 3	0.85	0.937	1		Rep. 3	0.834	0.954	1
<i>Δycf10</i>	PRB8 control	Rep. 1	Rep. 2	Rep. 3		<i>Δycf10</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.991	1			Rep. 2	0.984	1	
	Rep. 3	0.991	0.987	1		Rep. 3	0.978	0.992	1
<i>ΔpetL</i>	PRB8 control	Rep. 1	Rep. 2			<i>ΔpetL</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.981	1			Rep. 2	0.986	1	
as- <i>AtpC</i>	WT	Rep. 1	Rep. 2			as- <i>AtpC</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.854	1			Rep. 2	0.930	1	

Supplemental Table 4: Reproducibility of the probe signals of ribosome footprints between the biological replicates

Correlation analysis of the footprints probe signals located in the protein-coding regions from the biological replicates was undertaken. The Pearson's correlation R values are given in the table. Rep: Replicate.

<i>as-RBCS</i>	SR1	Rep. 1	Rep. 2	Rep. 3	<i>as-RBCS</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.854	1		Rep. 2	0.845	1	
	Rep. 3	0.808	0.968	1	Rep. 3	0.838	0.974	1
<i>psbD-TTG</i>	WT	Rep. 1	Rep. 2	Rep. 3	<i>psbD-TTG</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.976	1		Rep. 2	0.975	1	
	Rep. 3	0.991	0.985	1	Rep. 3	0.982	0.973	1
<i>psbD-GTG</i>	WT	Rep. 1	Rep. 2	Rep. 3	<i>psbD-GTG</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.984	1		Rep. 2	0.972	1	
	Rep. 3	0.969	0.946	1	Rep. 3	0.979	0.966	1
KD- <i>psbD</i> 21 days stage	PRB8 control	Rep. 1	Rep. 2	Rep. 3	KD- <i>psbD</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.942	1		Rep. 2	0.938	1	
	Rep. 3	0.938	0.953	1	Rep. 3	0.928	0.936	1
KD- <i>psbD</i> Cotelyd ons stage	PRB8 control	Rep. 1	Rep. 2	Rep. 3	KD- <i>psbD</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.796	1		Rep. 2	0.813	1	
	Rep. 3	0.814	0.965	1	Rep. 3	0.804	0.950	1
<i>hcf111-1</i>	WT col-0	Rep. 1	Rep. 2	Rep. 3	<i>hcf111-1</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.981	1		Rep. 2	0.967	1	
	Rep. 3	0.957	0.977	1	Rep. 3	0.898	0.95	1
Δ <i>psbN</i> 21 days stage	RB70 control	Rep. 1	Rep. 2	Rep. 3	Δ <i>psbN</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1			Rep. 1	1		
	Rep. 2	0.967	1		Rep. 2	0.972	1	
	Rep. 3	0.837	0.850	1	Rep. 3	0.870	0.894	1

<i>ΔpsbN</i> cotyled on stage	RB70 control	Rep. 1	Rep. 2	Rep. 3		<i>ΔpsbN</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.947	1			Rep. 2	0.899	1	
	Rep. 3	0.954	0.983	1		Rep. 3	0.933	0.982	1
KD- <i>psaA</i>	<i>psaA</i> - control	Rep. 1	Rep. 2			KD- <i>psaA</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.869	1			Rep. 2	0.887	1	
<i>psad1-1</i>	Col-0	Rep. 1	Rep. 2			<i>psad1-1</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.798	1			Rep. 2	0.759	1	
<i>Δycf4</i>	PRB8 control	Rep. 1	Rep. 2	Rep. 3		<i>Δycf4</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.941	1			Rep. 2	0.942	1	
	Rep. 3	0.940	0.968	1		Rep. 3	0.924	0.943	1
<i>ΔpsaI</i>	PRB8 control	Rep. 1	Rep. 2	Rep. 3		<i>ΔpsaI</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.857	1			Rep. 2	0.891	1	
	Rep. 3	0.872	0.976	1		Rep. 3	0.903	0.988	1
<i>Δycf10</i>	PRB8 control	Rep. 1	Rep. 2	Rep. 3		<i>Δycf10</i>	Rep. 1	Rep. 2	Rep. 3
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.971	1			Rep. 2	0.981	1	
	Rep. 3	0.971	0.976	1		Rep. 3	0.980	0.976	1
<i>ΔpetL</i>	PRB8 control	Rep. 1	Rep. 2			<i>ΔpetL</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.990	1			Rep. 2	0.981	1	
as- <i>AtpC</i>	WT	Rep. 1	Rep. 2			as- <i>AtpC</i>	Rep. 1	Rep. 2	
	Rep. 1	1				Rep. 1	1		
	Rep. 2	0.844	1			Rep. 2	0.932	1	

Supplemental Table 5: Mapping statistics of Ribo-seq data

For each Ribo-seq library, the reads mapped to rRNAs and each compartment (chloroplast, mitochondria and nucleus) are shown. Number of unique and multimapped reads and their percentages are shown for each compartment.

Sample	$\Delta aptB$ R1	$\Delta atpB$ R2	$\Delta psbD/C$ R1	$\Delta psbD/C$ R2	pRB8 R1	pRB8 R2	$\Delta psbB$ R1	$\Delta psbB$ R2
Library	52,181,34	52,778,12	56,503,16	43,250,91	27,679,65	38,801,53	32,296,96	25,517,64
rRNA alignments								
Unique reads	33,681,90 4	37,142,35 5	28,122,26 9	22,485,92 2	9,663,345	17,327,85 1	13,837,33 7	10,000,02 1
Multimapped reads	7,549,653	5,899,897	16,388,20 2	12,695,02 1	8,464,068	11,305,20 4	11,000,83 8	10,109,15 3
Total reads	41,231,55 7	43,042,25 2	44,510,47 1	35,180,94 3	18,127,41 3	28,633,05 5	24,838,17 5	20,109,17 4
% unique reads	64.55	70.37	49.77	51.99	34.91	44.66	42.84	39.19
% multimapped reads	14.47	11.18	29.00	29.35	30.58	29.14	34.06	39.62
% total reads	79.02	81.55	78.78	81.34	65.49	73.79	76.91	78.80
Chloroplast alignments								
Input	10,949,78	9,735,873	11,992,69	8,069,974	9,552,245	10,168,48	7,458,790	5,408,467
Unique reads	241827	208481	809541	644898	2622266	2285856	776315	540204
Multimapped reads	26416	21860	34119	24896	27330	22418	24450	39100
Total reads	268,243	230,341	843,660	669,794	2,649,596	2,308,274	800,765	579,304
% unique reads	0.46	0.40	1.43	1.49	9.47	5.89	2.40	2.12
% multimapped reads	0.05	0.04	0.06	0.06	0.10	0.06	0.08	0.15
% total reads	0.51	0.44	1.49	1.55	9.57	5.95	2.48	2.27
Mitochondria alignments								

Supplemental information

Input	10,681,54	9,505,532	11,149,03	7,400,180	6,902,649	7,860,207	6,658,025	4,829,163
Unique reads	87444	69301	87121	43687	59076	48274	42030	28208
Multimapped reads	9884	9117	9535	4586	7589	6234	6393	4074
Total reads	97,328	78,418	96,656	48,273	66,665	54,508	48,423	32,282
% unique reads	0.17	0.13	0.15	0.10	0.21	0.12	0.13	0.11
% multimapped reads	0.02	0.02	0.02	0.01	0.03	0.02	0.02	0.02
% total reads	0.19	0.15	0.17	0.11	0.24	0.14	0.15	0.13
Nuclear alignments								
Input	10,584,21	9,427,114	11,052,37	7,351,907	6,835,984	7,805,699	6,609,602	4,796,881
Unique reads	5,481,899	4,667,970	5,667,627	3,447,730	3,763,339	4,029,669	3,539,113	2,313,084
Multimapped reads	2,776,839	2,307,523	2,711,837	1,841,680	1,723,802	1,774,146	1,715,778	1,370,299
Total reads	8,258,738	6,975,493	8,379,464	5,289,410	5,487,141	5,803,815	5,254,891	3,683,383
% unique reads	10.51	8.84	10.03	7.97	13.60	10.39	10.96	9.06
% multimapped reads	5.32	4.37	4.80	4.26	6.23	4.57	5.31	5.37
% total reads	15.83	13.22	14.83	12.23	19.82	14.96	16.27	14.43

Supplemental Table 6: Proteins enriched in the *psbA*-MS2 affinity purification

Based on the intensity data obtained from MaxQuant. A twofold threshold was set to filter the most enriched proteins. Arabidopsis gene identification numbers (AGIs) are shown or the *Nicotiana tomentosiformis* transcript identifier. The description includes known functions and/or domains of the protein. The uncharacterized protein may represent either an interesting candidate or a contaminant. Extended descriptions can be searched in TAIR (www.arabidopsis.org).

POT number	AGI number	Fold change (tag/control)	Localization	Description
POT001608	AT5G20610	4973900	Nucleus	Plastid movement impaired-related 1, PMIR1
POT010659	AT5G53220	4419200	Nucleus	Hypothetical protein
POT010560	AT4G35987	3756300	Plastid	S-adenosyl-L-methionine-dependent methyltransferases superfamily
POT001784	XP_009630213.1	2547200	-	Uncharacterized, LOC104120879, <i>Nicotiana tomentosiformis</i> , ncRNA
POT000115	AT1G31950	1921800	Mitochondrion, endoplasmic reticulum, plastid, plasma membrane	Terpenoid cyclases/Protein prenyltransferases superfamily protein
POT010454	AT5G06610	336690	Cytosol	Protein of unknown function (DUF620)
POT005153	AT4G37550	297140	Cytosol, peroxisome	Acetamidase/Formamidase family protein
POT005317	AT1G06720	160760	Nucleus	P-loop containing nucleoside triphosphate hydrolases superfamily protein
POT004960	AT3G17900	158300	Nucleus	Unknown protein
POT001119	AT3G12060	118640	Plasma membrane, extracellular	Plant protein of unknown function (DUF828)
POT011234	AT4G11440	79178	Nucleus	Mitochondrial substrate carrier family protein
POT000421	AT1G61670	78316	Plasma membrane	Lung seven transmembrane receptor family protein
POT000340	AT2G28620	72870	Plasma membrane, nucleus	P-loop containing nucleoside triphosphate hydrolases superfamily protein
POT002739	AT1G02660	69681	Plasma membrane	Alpha/beta-Hydrolases superfamily protein
POT001576	AT3G17690	66068	Mitochondrion, plasma membrane	Cyclic nucleotide gated channel 19
POT000737	AT1G47980	46571	Plasma membrane, vacuole	Dessication-like protein

POT010711	AT3G48480	37614	Nucleus	Cysteine proteinases superfamily protein
POT007387	AT1G07310	31949	Nucleus	Calcium-dependent lipid-binding (CaLB domain) family protein
POT001553	AT1G28480	24126	Cytosol	Thioredoxin superfamily protein
POT000184	AT1G33720	23635	Cytosol, endoplasmic reticulum	Cytochrome P450, family 76, subfamily C, polypeptide 6
POT008879	AT4G15890	22518	Nucleus	Condensin, cap-d8
POT013646	AT2G40260	21908	Nucleus	Homeodomain-like superfamily protein
POT006916	AT2G29640	19953	Nucleus	JOSEPHIN-like protein
POT019605	XP_009608487.1	14445	-	Uncharacterized, LOC104102476, <i>Nicotiana tomentosiformis</i>
POT007986	AT4G28260	14155	Nucleus	Acyl-UDP-N-acetylglucosamine O-acyltransferase
POT008474	AT3G22680	12928	Nucleus	RNA-DIRECTED DNA METHYLATION 1
POT004532	AT2G34780	12642	Nucleus	Maternal effect embryo arrest 22
POT005089	AT5G05480	11643	Plasma membrane	Peptide-N4-(N-acetyl-beta-glucosaminy)asparagine amidase A protein
POT000910	AT2G26330	7210.3	Plasma membrane	Leucine-rich receptor-like protein kinase family protein
POT000058	AT1G17060	6390.9	Endoplasmic reticulum, plasma	Cytochrome p450 72c1
POT000042	AT1G04160	4415.6	Nucleus, cytosol, vacuole,	Myosin XI B
POT000698	AT2G33210	3699.8	Mitochondrion	heat shock protein 60-2
POT000993	XP_009608178.1	3469.2	-	Uncharacterized, LOC104105360, <i>Nicotiana tomentosiformis</i>
POT005976	AT1G44575	3163.9	Plastid	Chlorophyll A-B binding family protein
POT009728	AT4G27120	2856.8	Endoplasmic reticulum, cytosol	DDR GK domain protein
POT000721	AT3G01300	2252.7	Cytosol	Protein kinase superfamily protein
POT006330	AT5G26570	1588.9	Plastid	Catalytics; carbohydrate kinases; phosphoglucan, water dikinases
POT010325	AT2G33680	1242.6	Plastid	Tetratricopeptide repeat (TPR)-like superfamily protein

POT014546	XP_009785113.1	869.82	-	UTP--glucose-1-phosphate uridylyltransferase-like, LOC104094835, <i>Nicotiana tomentosiformis</i> , transcript variant X1
POT000207	XP_009781584.1	847.77	-	Uncharacterized, LOC104105277, <i>Nicotiana tomentosiformis</i>
POT002506	AT3G51780	628.33	Cytosol	BCL-2-associated athanogene 4
POT016701	XP_009620515.1	514.67	-	F-box protein CPR30-like, LOC104112333, <i>Nicotiana tomentosiformis</i>
POT004765	XP_009627933.1	415.91	-	Cytochrome P450 94A1-like, LOC104103499, <i>Nicotiana tomentosiformis</i> partial mRNA
POT000207	XP_009596268.1	367.9	-	Uncharacterized, LOC104105277, <i>Nicotiana tomentosiformis</i>
POT001905	AT2G19680	6.186037287	Mitochondrion	Mitochondrial ATP synthase subunit G protein
PCP000032	NP_054506.2	4.015440161	Plastid	ATP synthase CF1 beta subunit
POT001772	AT2G36170	3.467099034	Cytosol, nucleus	Ubiquitin supergroup; Ribosomal protein L40e
POT007575	AT3G01500	2.165838678	Plastid	Carbonic anhydrase 1

Supplemental Table 7: Proteins enriched in the *psbA* affinity purification experiments (with MS2 aptamer, with streptavidin-binding aptamer without elution and with 30-min elution)

Based on the intensity data obtained from MaxQuant. A 2-fold threshold was set to filter the most enriched proteins in all three experiments.

POT number	AGI number	Fold change (tag/control)	Localization	Description
POT000042	AT1G04160	6.76	Nucleus, cytosol, vacuole,	Myosin XI B
POT000615	AT3G17360	5.76	Nucleus	Phragmoplast orienting kinesin 1, POK1
POT002564	XP_009780432.1	4.83	-	Uncharacterized protein LOC104229477, predicted, partial <i>Nicotiana sylvestris</i> , 174 aa

POT000115	AT1G31950	3.88	Mitochondrion, endoplasmic	Terpenoid cyclases/Protein prenyltransferases superfamily protein
POT000967a	AT1G12900	3.40	Plastid	Glyceraldehyde 3-phosphate dehydrogenase A subunit 2
PCP000033	NP_054507.1	3.04	Chloroplast	RbcL
POT000207	XP_009607033.1	3.00	-	Uncharacterized protein LOC104101292, <i>Nicotiana tomentosiformis</i> Similarity to hypothetical protein in Arabidopsis
POT007575	AT3G01500	2.71	Plastid	Carbonic anhydrase 1, CA1
POT007387	AT1G07310	2.64	Nucleus	Calcium-dependent lipid-binding (CaLB domain) family protein
POT004779	AT4G12800	2.41	Plastid	Photosystem I subunit 1, PsaL
POT000744	AT2G28720	2.19	Nucleus	Histone superfamily protein
POT008712	AT5G13120	2.10	Plastid	Cyclophilin 20-2
POT005089	AT5G05480	2.01	Plasma membrane	Peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase A protein

Supplemental Table 8: Summary of the proteins enriched in the *rbcL*-MS2 purification

Based on the intensity data obtained from MaxQuant. A threshold of twofold was applied.

POT number	AGI number	Fold change (tag/control)	Localization	Description
POT001608	AT5G20610	5821800	Nucleus	Plastid movement impaired-related 1, PMIR1
POT000699	AT2G22450	4263400	Peroxisome, plastid	Riboflavin biosynthesis protein, putative
POT005550	AT2G38300	244220	Nucleus	Myb-like HTH transcriptional regulator family protein
POT001905	AT2G19680	197830	Mitochondrion	Mitochondrial ATP synthase subunit G protein
POT005177	AT4G24900	170920	Nucleus	Titan-like, transthyretin-like protein, TTL

Supplemental information

POT010763	XP_009782999.1	163130	Nucleus	Transcriptional activator Myb-like, predicted, LOC104085082, <i>Nicotiana sylvestris</i>
POT000988	AT5G42650	60050	Plastid	Allene oxide synthase, AOS
POT004623	AT5G11590	44667	Nucleus	Integrase-type DNA-binding superfamily protein, TINY2
POT012281	AT4G24500	11176	Nucleus	Hydroxyproline-rich glycoprotein family protein
POR120138	XP_009617324.1	11053	-	Uncharacterized, LOC104109648, predicted, <i>Nicotiana tomentosiformis</i>
POT004563	AT3G27830	10982	Plastid	Ribosomal protein L12-A, RPL12A
POT006197	AT2G42260	6954.2	Nucleus, cytosol	UV-B-insensitive 4, UVI4
POT009684	AT4G01310	4912	plastid	Ribosomal L5P family protein, RPL5
POT010326	AT2G33800	4807.1	Plastid	Ribosomal S5 family protein, RPS5
POT006330	AT5G26570	2770.1	Plastid	Catalytics, carbohydrate kinases, phosphoglucan, water dikinases
POT058939	XP_009610762.1	2329.1	-	Anthocyanidin 3-O-glucosyltransferase 2-like, predicted, <i>Nicotiana tomentosiformis</i>
POT005673	AT5G30510	1784.5	Plastid	Ribosomal protein S1, RPS1
POT000828	AT2G42740	1725.9	Cytosol	Ribosomal protein large subunit 16A, RPL16A
POT007048	AT3G63490	1637.2	Plastid	Ribosomal protein L1p/L10e family, RPL1
POT011398	AT3G52150	930.88	Plastid	RNA-binding (RRM/RBD/RNP motifs) family protein, plastid-specific ribosomal protein 2, PSRP2
POT000277	AT1G07920	2.216216	Cytosol	GTP binding Elongation factor Tu family protein

6 References

- Adachi, Y., Kuroda, H., Yukawa, Y., and Sugiura, M. (2012). Translation of partially overlapping *psbD-psbC* mRNAs in chloroplasts: the role of 5'-processing and translational coupling. *Nucleic Acids Res* 40:3152-3158.
- Adam, Z. (1996). Protein stability and degradation in chloroplasts. *Plant molecular biology* 32:773-783.
- Adam, Z. (2000). Chloroplast proteases: possible regulators of gene expression? *Biochimie* 82:647-654.
- Adam, Z. (2007). Protein stability and degradation in plastids. In: *Cell and Molecular Biology of Plastids*--Bock, R., ed.: Springer Berlin Heidelberg. 315-338.
- Adam, Z., Rudella, A., and van Wijk, K.J. (2006). Recent advances in the study of Clp, FtsH and other proteases located in chloroplasts. *Curr Opin Plant Biol* 9:234-240.
- Ahmed, T., Yin, Z., and Bhushan, S. (2016). Cryo-EM structure of the large subunit of the spinach chloroplast ribosome. *Sci Rep* 6:35793.
- Albus, C.A., Ruf, S., Schottler, M.A., Lein, W., Kehr, J., and Bock, R. (2010). Y3IP1, a nucleus-encoded thylakoid protein, cooperates with the plastid-encoded Ycf3 protein in photosystem I assembly of tobacco and Arabidopsis. *The Plant cell* 22:2838-2855.
- Alkatib, S., Scharff, L.B., Rogalski, M., Fleischmann, T.T., Matthes, A., Seeger, S., Schöttler, M.A., Ruf, S., and Bock, R. (2012). The contributions of wobbling and superwobbling to the reading of the genetic code. *PLoS Genet* 8:e1003076.
- Allen, J.F. (2015). Why chloroplasts and mitochondria retain their own genomes and genetic systems: Colocation for redox regulation of gene expression. *Proceedings of the National Academy of Sciences of the United States of America* 112:10231-10238.
- Allison, L.A. (2000). The role of sigma factors in plastid transcription. *Biochimie* 82:537-548.
- Allison, L.A., Simon, L.D., and Maliga, P. (1996). Deletion of *rpoB* reveals a second distinct transcription system in plastids of higher plants. *Embo J* 15:2802-2809.
- Amunts, A., Drory, O., and Nelson, N. (2007). The structure of a plant photosystem I supercomplex at 3.4 Å resolution. *Nature* 447:58-63.
- Amunts, A., and Nelson, N. (2009). Plant photosystem I design in the light of evolution. *Structure* 17:637-650.
- Andersson, I., and Backlund, A. (2008). Structure and function of Rubisco. *Plant Physiol Biochem* 46:275-291.
- Antonkine, M.L., Jordan, P., Fromme, P., Krauss, N., Golbeck, J.H., and Stehlik, D. (2003). Assembly of protein subunits within the stromal ridge of photosystem I. Structural changes between unbound and sequentially PS I-bound polypeptides and correlated changes of the magnetic properties of the terminal iron sulfur clusters. *J Mol Biol* 327:671-697.
- Baniulis, D., Yamashita, E., Whitelegge, J.P., Zatsman, A.I., Hendrich, M.P., Hasan, S.S., Ryan, C.M., and Cramer, W.A. (2009). Structure-function, stability, and chemical modification of the cyanobacterial cytochrome *b₆f* complex from *Nostoc* sp. PCC 7120. *J Biol Chem* 284:9861-9869.
- Bardwell, V.J., and Wickens, M. (1990). Purification of RNA and RNA-protein complexes by an R17 coat protein affinity method. *Nucleic Acids Res* 18:6587-6594.
- Barkan, A. (1998). Approaches to investigating nuclear genes that function in chloroplast biogenesis in land plants. *Methods Enzymol* 297:38-57.
- Barkan, A. (2009). Genome-wide analysis of RNA-protein interactions in plants. *Methods in molecular biology* 553:13-37.

- Barkan, A. (2011). Expression of plastid genes: organelle-specific elaborations on a prokaryotic scaffold. *Plant physiology* 155:1520-1532.
- Barkan, A., Rojas, M., Fujii, S., Yap, A., Chong, Y.S., Bond, C.S., and Small, I. (2012). A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet* 8:e1002910.
- Barkan, A., and Small, I. (2014). Pentatricopeptide repeat proteins in plants. *Annual review of plant biology* 65:415-442.
- Barrientos, A., Zambrano, A., and Tzagoloff, A. (2004). Mss51p and Cox14p jointly regulate mitochondrial Cox1p expression in *Saccharomyces cerevisiae*. *EMBO J* 23:3472-3482.
- Barta, I., and Iggo, R. (1995). Autoregulation of expression of the yeast Dbp2p 'DEAD-box' protein is mediated by sequences in the conserved *DBP2* intron. *EMBO J* 14:3800-3808.
- Benjamini, Y., and 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.
- Berkhout, B., and van Duin, J. (1985). Mechanism of translational coupling between coat protein and replicase genes of RNA bacteriophage MS2. *Nucleic Acids Res* 13:6955-6967.
- Bertani, G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol* 62:293-300.
- Bhuiyan, N.H., Friso, G., Poliakov, A., Ponnala, L., and van Wijk, K.J. (2015). MET1 is a thylakoid-associated TPR protein involved in photosystem II supercomplex formation and repair in *Arabidopsis*. *The Plant cell* 27:262-285.
- Bieri, P., Leibundgut, M., Saurer, M., Boehringer, D., and Ban, N. (2017). The complete structure of the chloroplast 70S ribosome in complex with translation factor pY. *EMBO J* 36:475-486.
- Blencowe, B.J., Sproat, B.S., Ryder, U., Barabino, S., and Lamond, A.I. (1989). Antisense probing of the human U4/U6 snRNP with biotinylated 2'-OMe RNA oligonucleotides. *Cell* 59:531-539.
- Bock, R. (2015). Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology. *Annual review of plant biology* 66:211-241.
- Bock, R., Kössel, H., and Maliga, P. (1994). Introduction of a heterologous editing site into the tobacco plastid genome: the lack of RNA editing leads to a mutant phenotype. *EMBO J* 13:4623-4628.
- Bock, R., and Timmis, J.N. (2008). Reconstructing evolution: gene transfer from plastids to the nucleus. *Bioessays* 30:556-566.
- Boeck, R., and Kolakofsky, D. (1994). Positions +5 and +6 can be major determinants of the efficiency of non-AUG initiation codons for protein synthesis. *EMBO J* 13:3608-3617.
- Bohne, A.V., Schwenkert, S., Grimm, B., and Nickelsen, J. (2016). Roles of tetratricopeptide repeat proteins in biogenesis of the photosynthetic apparatus. *Int Rev Cell Mol Biol* 324:187-227.
- Börner, T., Aleynikova, A.Y., Zubo, Y.O., and Kusnetsov, V.V. (2015). Chloroplast RNA polymerases: Role in chloroplast biogenesis. *Biochimica et biophysica acta* 1847:761-769.
- Boudreau, E., Nickelsen, J., Lemaire, S.D., Ossenbühl, F., and Rochaix, J.D. (2000). The *Nac2* gene of *Chlamydomonas* encodes a chloroplast TPR-like protein involved in *psbD* mRNA stability. *EMBO J* 19:3366-3376.
- Boudreau, E., Takahashi, Y., Lemieux, C., Turmel, M., and Rochaix, J.D. (1997). The chloroplast *ycf3* and *ycf4* open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem I complex. *EMBO J* 16:6095-6104.
- Boulouis, A., Raynaud, C., Bujaldon, S., Aznar, A., Wollman, F.A., and Choquet, Y. (2011). The nucleus-encoded trans-acting factor MCA1 plays a critical role in the regulation of cytochrome *f* synthesis in *Chlamydomonas* chloroplasts. *The Plant cell* 23:333-349.

- Boynton, J.E., Gillham, N.W., Harris, E.H., Hosler, J.P., Johnson, A.M., Jones, A.R., Randolph-Anderson, B.L., Robertson, D., Klein, T.M., Shark, K.B., et al. (1988). Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. *Science* 240:1534-1538.
- Burrows, P.A., Sazanov, L.A., Svab, Z., Maliga, P., and Nixon, P.J. (1998). Identification of a functional respiratory complex in chloroplasts through analysis of tobacco mutants containing disrupted plastid *ndh* genes. *EMBO J* 17:868-876.
- Butler, J.S., Springer, M., Dondon, J., Graffe, M., and Grunberg-Manago, M. (1986). *Escherichia coli* protein synthesis initiation factor IF3 controls its own gene expression at the translational level *in vivo*. *J Mol Biol* 192:767-780.
- Cabral, F., and Schatz, G. (1978). Identification of cytochrome *c* oxidase subunits in nuclear yeast mutants lacking the functional enzyme. *J Biol Chem* 253:4396-4401.
- Cahoon, E.B., Shanklin, J., and Ohlrogge, J.B. (1992). Expression of a coriander desaturase results in petroselinic acid production in transgenic tobacco. *Proceedings of the National Academy of Sciences of the United States of America* 89:11184-11188.
- Calder, K.M., and McEwen, J.E. (1991). Deletion of the *COX7* gene in *Saccharomyces cerevisiae* reveals a role for cytochrome *c* oxidase subunit VII in assembly of remaining subunits. *Mol Microbiol* 5:1769-1777.
- Cavaiuolo, M., Kuras, R., Wollman, F.A., Choquet, Y., and Vallon, O. (2017). Small RNA profiling in *Chlamydomonas*: insights into chloroplast RNA metabolism. *Nucleic Acids Res* 45:10783-10799.
- Cavanagh, A.P., and Kubien, D.S. (2014). Can phenotypic plasticity in Rubisco performance contribute to photosynthetic acclimation? *Photosynth Res* 119:203-214.
- Celedon, J.M., and Cline, K. (2013). Intra-plastid protein trafficking: how plant cells adapted prokaryotic mechanisms to the eukaryotic condition. *Biochimica et biophysica acta* 1833:341-351.
- Chateigner-Boutin, A.L., Colas des Francs-Small, C., Fujii, S., Okuda, K., Tanz, S.K., and Small, I. (2013). The E domains of pentatricopeptide repeat proteins from different organelles are not functionally equivalent for RNA editing. *The Plant journal : for cell and molecular biology* 74:935-945.
- Chen, J., Nikolaitchik, O., Singh, J., Wright, A., Bencsics, C.E., Coffin, J.M., Ni, N., Lockett, S., Pathak, V.K., and Hu, W.S. (2009). High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. *Proceedings of the National Academy of Sciences of the United States of America* 106:13535-13540.
- Chevalier, F., Ghulam, M.M., Rondet, D., Pfannschmidt, T., Merendino, L., and Lerbs-Mache, S. (2015). Characterization of the *psbH* precursor RNAs reveals a precise endoribonuclease cleavage site in the *psbT/psbH* intergenic region that is dependent on *psbN* gene expression. *Plant molecular biology* 88:357-367.
- Choquet, Y., Stern, D.B., Wostrikoff, K., Kuras, R., Girard-Bascou, J., and Wollman, F.A. (1998). Translation of cytochrome *f* is autoregulated through the 5' untranslated region of *petA* mRNA in *Chlamydomonas* chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America* 95:4380-4385.
- Choquet, Y., and Wollman, F.A. (2009). The CES process. In: *Chlamydomonas Source Book*--Harris, E., Stern, D.B., and Whitman, G., eds. New York: Academic Press/Elsevier. 1027–1064.
- Choquet, Y., Zito, F., Wostrikoff, K., and Wollman, F.A. (2003). Cytochrome *f* translation in *Chlamydomonas* chloroplast is autoregulated by its carboxyl-terminal domain. *The Plant cell* 15:1443-1454.
- Chotewutmontri, P., and Barkan, A. (2016). Dynamics of chloroplast translation during chloroplast differentiation in maize. *PLoS Genet* 12:e1006106.

- Chotewutmontri, P., and Barkan, A. (2018). Multilevel effects of light on ribosome dynamics in chloroplasts program genome-wide and *psbA*-specific changes in translation. *PLoS Genet* 14:e1007555.
- Chotewutmontri, P., and Barkan, A. (2020). Light-induced *psbA* translation in plants is triggered by photosystem II damage via an assembly-linked autoregulatory circuit. [bioRxiv:2020.2004.2027.061879](https://doi.org/10.1101/2020.2004.2027.061879).
- Chotewutmontri, P., Williams-Carrier, R., and Barkan, A. (2020). Exploring the link between photosystem II assembly and translation of the chloroplast *psbA* mRNA. *Plants* 9:152.
- Chu, C., Zhang, Q.C., da Rocha, S.T., Flynn, R.A., Bharadwaj, M., Calabrese, J.M., Magnuson, T., Heard, E., and Chang, H.Y. (2015). Systematic discovery of Xist RNA binding proteins. *Cell* 161:404-416.
- Chung, B.Y., Hardcastle, T.J., Jones, J.D., Irigoyen, N., Firth, A.E., Baulcombe, D.C., and Brierley, I. (2015). The use of duplex-specific nuclease in ribosome profiling and a user-friendly software package for Ribo-seq data analysis. *RNA* 21:1731-1745.
- Chung, S., and Takizawa, P.A. (2010). Multiple Myo4 motors enhance *ASH1* mRNA transport in *Saccharomyces cerevisiae*. *The Journal of cell biology* 189:755-767.
- Cox, J., and 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.
- Cramer, W.A., and Zhang, H. (2006). Consequences of the structure of the cytochrome *b₆f* complex for its charge transfer pathways. *Biochimica et biophysica acta* 1757:339-345.
- De Cosa, B., Moar, W., Lee, S.B., Miller, M., and Daniell, H. (2001). Overexpression of the *Bt cry2Aa2* operon in chloroplasts leads to formation of insecticidal crystals. *Nat Biotechnol* 19:71-74.
- De Marchis, F., Wang, Y., Stevanato, P., Arcioni, S., and Bellucci, M. (2009). Genetic transformation of the sugar beet plastome. *Transgenic Res* 18:17-30.
- Dephoure, N., Hwang, S., O'Sullivan, C., Dodgson, S.E., Gygi, S.P., Amon, A., and Torres, E.M. (2014). Quantitative proteomic analysis reveals posttranslational responses to aneuploidy in yeast. *Elife* 3:e03023.
- Ding, S., Zhang, Y., Hu, Z., Huang, X., Zhang, B., Lu, Q., Wen, X., Wang, Y., and Lu, C. (2019). mTERF5 acts as a transcriptional pausing factor to positively regulate transcription of chloroplast *psbEFLJ*. *Mol Plant* 12:1259-1277.
- Dingwall, C., Lomonosoff, G.P., and Laskey, R.A. (1981). High sequence specificity of micrococcal nuclease. *Nucleic Acids Res* 9:2659-2673.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15-21.
- Drapier, D., Rimbault, B., Vallon, O., Wollman, F.A., and Choquet, Y. (2007). Intertwined translational regulations set uneven stoichiometry of chloroplast ATP synthase subunits. *EMBO J* 26:3581-3591.
- Drechsel, O., and Bock, R. (2011). Selection of Shine-Dalgarno sequences in plastids. *Nucleic Acids Res* 39:1427-1438.
- Drescher, A., Ruf, S., Calsa, T., Jr., Carrer, H., and Bock, R. (2000). The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. *The Plant journal : for cell and molecular biology* 22:97-104.
- Du, T.G., Schmid, M., and Jansen, R.P. (2007). Why cells move messages: the biological functions of mRNA localization. *Semin Cell Dev Biol* 18:171-177.
- Edwards, K.D., Fernandez-Pozo, N., Drake-Stowe, K., Humphry, M., Evans, A.D., Bombarely, A., Allen, F., Hurst, R., White, B., Kernodle, S.P., et al. (2017). A reference genome for *Nicotiana*

- tabacum* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. *BMC Genomics* 18:448.
- Egea, I., Barsan, C., Bian, W., Purgatto, E., Latche, A., Chervin, C., Bouzayen, M., and Pech, J.C. (2010). Chromoplast differentiation: current status and perspectives. *Plant & cell physiology* 51:1601-1611.
- Ellington, A.D., and Szostak, J.W. (1990). *In vitro* selection of RNA molecules that bind specific ligands. *Nature* 346:818-822.
- Ellington, A.D., and Szostak, J.W. (1992). Selection *in vitro* of single-stranded DNA molecules that fold into specific ligand-binding structures. *Nature* 355:850-852.
- Ellis, R.J. (1979). The most abundant protein in the world. *Trends in Biochemical Sciences* 4:241-244.
- Elson, S.L., Noble, S.M., Solis, N.V., Filler, S.G., and Johnson, A.D. (2009). An RNA transport system in *Candida albicans* regulates hyphal morphology and invasive growth. *PLoS Genet* 5:e1000664.
- Endo, T., Ishida, S., Ishikawa, N., and Sato, F. (2008). Chloroplastic NAD(P)H dehydrogenase complex and cyclic electron transport around photosystem I. *Mol Cells* 25:158-162.
- Fedor, J.G., Jones, A.J.Y., Di Luca, A., Kaila, V.R.I., and Hirst, J. (2017). Correlating kinetic and structural data on ubiquinone binding and reduction by respiratory complex I. *Proceedings of the National Academy of Sciences of the United States of America* 114:12737-12742.
- Felder, S., Meierhoff, K., Sane, A.P., Meurer, J., Driemel, C., Plucken, H., Klaff, P., Stein, B., Bechtold, N., and Westhoff, P. (2001). The nucleus-encoded *HCF107* gene of *Arabidopsis* provides a link between intergenic RNA processing and the accumulation of translation-competent *psbH* transcripts in chloroplasts. *The Plant cell* 13:2127-2141.
- Fica, S.M., Oubridge, C., Wilkinson, M.E., Newman, A.J., and Nagai, K. (2019). A human postcatalytic spliceosome structure reveals essential roles of metazoan factors for exon ligation. *Science* 363:710-714.
- Forrest, K.M., and Gavis, E.R. (2003). Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for *nanos* mRNA localization in *Drosophila*. *Curr Biol* 13:1159-1168.
- Freedman, L.P., Zengel, J.M., Archer, R.H., and Lindahl, L. (1987). Autogenous control of the S10 ribosomal protein operon of *Escherichia coli*: genetic dissection of transcriptional and posttranscriptional regulation. *Proceedings of the National Academy of Sciences of the United States of America* 84:6516-6520.
- Freyer, R., Kiefer-Meyer, M.C., and Kössel, H. (1997). Occurrence of plastid RNA editing in all major lineages of land plants. *Proceedings of the National Academy of Sciences of the United States of America* 94:6285-6290.
- Fu, H.-Y. (2012). Reduction of PSII accumulation through manipulating *psbD* translational initiation. Universität Potsdam.
- Gagliardi, M., and Matarazzo, M.R. (2016). RIP: RNA Immunoprecipitation. *Methods in molecular biology* 1480:73-86.
- Gawroński, P., Pałac, A., and Scharff, L.B. (2020). Secondary structure of chloroplast mRNAs *in vivo* and *in vitro*. *Plants (Basel)* 9.
- Goldschmidt-Clermont, M. (1991). Transgenic expression of aminoglycoside adenine transferase in the chloroplast: a selectable marker of site-directed transformation of chlamydomonas. *Nucleic Acids Res* 19:4083-4089.
- Graf, M., Arenz, S., Huter, P., Dönhöfer, A., Nováček, J., and Wilson, D.N. (2016). Cryo-EM structure of the spinach chloroplast ribosome reveals the location of plastid-specific ribosomal proteins and extensions. *Nucleic Acids Res* 45:2887-2896.

- Greiner, S., Lehwark, P., and Bock, R. (2019). OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res* 47:W59-W64.
- Groth, G., and Pohl, E. (2001). The structure of the chloroplast F1-ATPase at 3.2 Å resolution. *J Biol Chem* 276:1345-1352.
- Gulati, A., and Mahadevan, S. (2001). The *Escherichia coli* antiterminator protein BglG stabilizes the 5' region of the *bgl* mRNA. *J Biosci* 26:193-203.
- Hager, M. (2002). Anwendung reverser Genetik zur Funktionsanalyse plastidenkodierter Gene. Albert-Ludwigs-Universität Freiburg.
- Hager, M., Biehler, K., Illerhaus, J., Ruf, S., and Bock, R. (1999). Targeted inactivation of the smallest plastid genome-encoded open reading frame reveals a novel and essential subunit of the cytochrome *b6f* complex. *EMBO J* 18:5834-5842.
- Haldrup, A., Lunde, C., and Scheller, H.V. (2003). *Arabidopsis thaliana* plants lacking the PSI-D subunit of photosystem I suffer severe photoinhibition, have unstable photosystem I complexes, and altered redox homeostasis in the chloroplast stroma. *J Biol Chem* 278:33276-33283.
- Halperin, T., and Adam, Z. (1996). Degradation of mistargeted OEE33 in the chloroplast stroma. *Plant molecular biology* 30:925-933.
- Hammani, K., Bonnard, G., Bouchoucha, A., Gobert, A., Pinker, F., Salinas, T., and Giegé, P. (2014). Helical repeats modular proteins are major players for organelle gene expression. *Biochimie* 100:141-150.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* 109:5651-5656.
- Hammel, A., Zimmer, D., Sommer, F., Mühlhaus, T., and Schroda, M. (2018). Absolute quantification of major photosynthetic protein complexes in *Chlamydomonas reinhardtii* using quantification concatamers (QconCATs). *Frontiers in plant science* 9:1265.
- Hasan, S.S., Yamashita, E., Baniulis, D., and Cramer, W.A. (2013). Quinone-dependent proton transfer pathways in the photosynthetic cytochrome *b6f* complex. *Proceedings of the National Academy of Sciences of the United States of America* 110:4297-4302.
- Hauser, T., Popilka, L., Hartl, F.U., and Hayer-Hartl, M. (2015). Role of auxiliary proteins in Rubisco biogenesis and function. *Nat Plants* 1:15065.
- Hayes, M.L., and Hanson, M.R. (2008). High conservation of a 5' element required for RNA editing of a C target in chloroplast *psbE* transcripts. *J Mol Evol* 67:233-245.
- Hedtke, B., Börner, T., and Weihe, A. (1997). Mitochondrial and chloroplast phage-type RNA polymerases in *Arabidopsis*. *Science* 277:809-811.
- Hedtke, B., Börner, T., and Weihe, A. (2000). One RNA polymerase serving two genomes. *EMBO Rep* 1:435-440.
- Hendrick, J.L., Wilson, P.G., Edelman, II, Sandbaken, M.G., Ursic, D., and Culbertson, M.R. (2001). Yeast frameshift suppressor mutations in the genes coding for transcription factor Mbf1p and ribosomal protein S3: evidence for autoregulation of S3 synthesis. *Genetics* 157:1141-1158.
- Hershey, J.W., Sonenberg, N., and Mathews, M.B. (2012). Principles of translational control: an overview. *Cold Spring Harb Perspect Biol* 4:a011528.
- Hess, W.R., and Börner, T. (1999). Organellar RNA polymerases of higher plants. *Int Rev Cytol* 190:1-59.
- Hirose, T., Ideue, T., Wakasugi, T., and Sugiura, M. (1999). The chloroplast *infA* gene with a functional UUG initiation codon. *FEBS Lett* 445:169-172.

- Hirose, T., Kusumegi, T., and Sugiura, M. (1998). Translation of tobacco chloroplast *rps14* mRNA depends on a Shine-Dalgarno-like sequence in the 5'-untranslated region but not on internal RNA editing in the coding region. *FEBS Lett* 430:257-260.
- Hirose, T., and Sugiura, M. (2004a). Functional Shine-Dalgarno-like sequences for translational initiation of chloroplast mRNAs. *Plant & cell physiology* 45:114-117.
- Hirose, T., and Sugiura, M. (2004b). Multiple elements required for translation of plastid *atpB* mRNA lacking the Shine-Dalgarno sequence. *Nucleic Acids Res* 32:3503-3510.
- Hoch, B., Maier, R.M., Appel, K., Igloi, G.L., and Kössel, H. (1991). Editing of a chloroplast mRNA by creation of an initiation codon. *Nature* 353:178-180.
- Hristou, A., Gerlach, I., Stolle, D.S., Neumann, J., Bischoff, A., Dünschede, B., Nowaczyk, M.M., Zoschke, R., and Schünemann, D. (2019). Ribosome-associated chloroplast SRP54 enables efficient co-translational membrane insertion of key photosynthetic proteins. *The Plant cell:tpc.00169.02019*.
- Hsu, P.Y., Calviello, L., Wu, H.L., Li, F.W., Rothfels, C.J., Ohler, U., and Benfey, P.N. (2016). Super-resolution ribosome profiling reveals unannotated translation events in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* 113:E7126-E7135.
- Ihnatowicz, A., Pesaresi, P., Varotto, C., Richly, E., Schneider, A., Jahns, P., Salamini, F., and Leister, D. (2004). Mutants for photosystem I subunit D of *Arabidopsis thaliana*: effects on photosynthesis, photosystem I stability and expression of nuclear genes for chloroplast functions. *The Plant journal : for cell and molecular biology* 37:839-852.
- Ingolia, N.T. (2014). Ribosome profiling: new views of translation, from single codons to genome scale. *Nat Rev Genet* 15:205-213.
- Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* 324:218-223.
- Ingolia, N.T., Hussmann, J.A., and Weissman, J.S. (2018). Ribosome profiling: global views of translation. *Cold Spring Harb Perspect Biol*.
- Ishikawa, K., Makanae, K., Iwasaki, S., Ingolia, N.T., and Moriya, H. (2017). Post-translational dosage compensation buffers genetic perturbations to stoichiometry of protein complexes. *PLoS Genet* 13:e1006554.
- Ishizaki, Y., Tsunoyama, Y., Hatano, K., Ando, K., Kato, K., Shinmyo, A., Kobori, M., Takeba, G., Nakahira, Y., and Shiina, T. (2005). A nuclear-encoded sigma factor, *Arabidopsis* SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. *The Plant journal : for cell and molecular biology* 42:133-144.
- Itzhaki, H., Naveh, L., Lindahl, M., Cook, M., and Adam, Z. (1998). Identification and characterization of DegP, a serine protease associated with the luminal side of the thylakoid membrane. *J Biol Chem* 273:7094-7098.
- Izumi, M., Tsunoda, H., Suzuki, Y., Makino, A., and Ishida, H. (2012). *RBCS1A* and *RBCS3B*, two major members within the *Arabidopsis* *RBCS* multigene family, function to yield sufficient Rubisco content for leaf photosynthetic capacity. *J Exp Bot* 63:2159-2170.
- Jagannathan, B., Dekat, S., Golbeck, J.H., and Lakshmi, K.V. (2010). The assembly of a multisubunit photosynthetic membrane protein complex: a site-specific spin labeling EPR spectroscopic study of the PsaC subunit in photosystem I. *Biochemistry* 49:2398-2408.
- Jagendorf, A.T., and Michaels, A. (1990). Rough thylakoids: translation on photosynthetic membranes. *Plant Science* 71:137-145.
- Jarvis, P., and Lopez-Juez, E. (2013). Biogenesis and homeostasis of chloroplasts and other plastids. *Nature reviews. Molecular cell biology* 14:787-802.

- Jean-Francois, M.J.B., Lukins, H.B., and Marzuki, S. (1986). Post-transcriptional defects in the synthesis of the mitochondrial H⁺-ATPase subunit 6 in yeast mutants with lesions in the subunit 9 structural gene. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 868:178-182.
- Jin, H., Fu, M., Duan, Z., Duan, S., Li, M., Dong, X., Liu, B., Feng, D., Wang, J., Peng, L., et al. (2018). LOW PHOTOSYNTHETIC EFFICIENCY 1 is required for light-regulated photosystem II biogenesis in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*.
- Johnson, X., Wostrikoff, K., Finazzi, G., Kuras, R., Schwarz, C., Bujaldon, S., Nickelsen, J., Stern, D.B., Wollman, F.A., and Vallon, O. (2010). MRL1, a conserved Pentatricopeptide repeat protein, is required for stabilization of *rbcL* mRNA in *Chlamydomonas* and *Arabidopsis*. *The Plant cell* 22:234-248.
- Jurica, M.S., Licklider, L.J., Gygi, S.R., Grigorieff, N., and Moore, M.J. (2002). Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA* 8:426-439.
- Kato, Y., and Sakamoto, W. (2018). FtsH protease in the thylakoid membrane: physiological functions and the regulation of protease activity. *Frontiers in plant science* 9:855.
- Katz, Z.B., English, B.P., Lionnet, T., Yoon, Y.J., Monnier, N., Ovrzyn, B., Bathe, M., and Singer, R.H. (2016). Mapping translation 'hot-spots' in live cells by tracking single molecules of mRNA and ribosomes. *Elife* 5.
- Kauss, D., Bischof, S., Steiner, S., Apel, K., and Meskauskiene, R. (2012). FLU, a negative feedback regulator of tetrapyrrole biosynthesis, is physically linked to the final steps of the Mg⁺⁺-branch of this pathway. *FEBS Lett* 586:211-216.
- Khrebtukova, I., and Spreitzer, R.J. (1996). Elimination of the *Chlamydomonas* gene family that encodes the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. *Proceedings of the National Academy of Sciences of the United States of America* 93:13689-13693.
- Kikuchi, S., Bedard, J., Hirano, M., Hirabayashi, Y., Oishi, M., Imai, M., Takase, M., Ide, T., and Nakai, M. (2013). Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* 339:571-574.
- Klaff, P., and Gruissem, W. (1991). Changes in chloroplast mRNA stability during leaf development. *The Plant cell* 3:517-529.
- Klug, G. (1993). The role of mRNA degradation in the regulated expression of bacterial photosynthesis genes. *Mol Microbiol* 9:1-7.
- Kode, V., Mudd, E.A., Iamtham, S., and Day, A. (2005). The tobacco plastid *accD* gene is essential and is required for leaf development. *The Plant journal : for cell and molecular biology* 44:237-244.
- Kofer, W., Koop, H.U., Wanner, G., and Steinmüller, K. (1998). Mutagenesis of the genes encoding subunits A, C, H, I, J and K of the plastid NAD(P)H-plastoquinone-oxidoreductase in tobacco by polyethylene glycol-mediated plastome transformation. *Mol Gen Genet* 258:166-173.
- Komenda, J., Nickelsen, J., Tichý, M., Prásil, O., Eichacker, L.A., and Nixon, P.J. (2008). The cyanobacterial homologue of HCF136/YCF48 is a component of an early photosystem II assembly complex and is important for both the efficient assembly and repair of photosystem II in *Synechocystis* sp. PCC 6803. *J Biol Chem* 283:22390-22399.
- Komenda, J., Reisinger, V., Müller, B.C., Dobáková, M., Granvogl, B., and Eichacker, L.A. (2004). Accumulation of the D2 protein is a key regulatory step for assembly of the photosystem II reaction center complex in *Synechocystis* PCC 6803. *J Biol Chem* 279:48620-48629.
- Krech, K., Fu, H.Y., Thiele, W., Ruf, S., Schöttler, M.A., and Bock, R. (2013). Reverse genetics in complex multigene operons by co-transformation of the plastid genome and its application to the open reading frame previously designated *psbN*. *The Plant journal : for cell and molecular biology* 75:1062-1074.

- Krech, K., Ruf, S., Masduki, F.F., Thiele, W., Bednarczyk, D., Albus, C.A., Tiller, N., Hasse, C., Schöttler, M.A., and Bock, R. (2012). The plastid genome-encoded Ycf4 protein functions as a nonessential assembly factor for photosystem I in higher plants. *Plant physiology* 159:579-591.
- Krichevsky, A., Meyers, B., Vainstein, A., Maliga, P., and Citovsky, V. (2010). Autoluminescent plants. *PLoS One* 5:e15461.
- Kühn, K., Richter, U., Meyer, E.H., Delannoy, E., de Longevialle, A.F., O'Toole, N., Borner, T., Millar, A.H., Small, I.D., and Whelan, J. (2009). Phage-type RNA polymerase RPOTmp performs gene-specific transcription in mitochondria of *Arabidopsis thaliana*. *The Plant cell* 21:2762-2779.
- Kuras, R., and Wollman, F.A. (1994). The assembly of cytochrome *b₆f* complexes: an approach using genetic transformation of the green alga *Chlamydomonas reinhardtii*. *EMBO J.* 13:1019-1027.
- Kwon, K.C., Chan, H.T., Leon, I.R., Williams-Carrier, R., Barkan, A., and Daniell, H. (2016). Codon optimization to enhance expression yields insights into chloroplast translation. *Plant physiology* 172:62-77.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.
- Lapointe, C.P., Wilinski, D., Saunders, H.A., and Wickens, M. (2015). Protein-RNA networks revealed through covalent RNA marks. *Nat Methods* 12:1163-1170.
- Lareau, L.F., Hite, D.H., Hogan, G.J., and Brown, P.O. (2014). Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. *Elife* 3:e01257.
- Larson, D.R., Zenklusen, D., Wu, B., Chao, J.A., and Singer, R.H. (2011). Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science* 332:475-478.
- LeCuyer, K.A., Behlen, L.S., and Uhlenbeck, O.C. (1995). Mutants of the Bacteriophage MS2 Coat Protein That Alter Its Cooperative Binding to RNA. *Biochemistry* 34:10600-10606.
- Lee, S.M., Kang, K., Chung, H., Yoo, S.H., Xu, X.M., Lee, S.B., Cheong, J.J., Daniell, H., and Kim, M. (2006). Plastid transformation in the monocotyledonous cereal crop, rice (*Oryza sativa*) and transmission of transgenes to their progeny. *Mol Cells* 21:401-410.
- Legen, J., and Schmitz-Linneweber, C. (2017). Stable membrane-association of mRNAs in etiolated, greening and mature plastids. *Int J Mol Sci* 18.
- Lein, W., Usadel, B., Stitt, M., Reindl, A., Ehrhardt, T., Sonniewald, U., and Börnke, F. (2008). Large-scale phenotyping of transgenic tobacco plants (*Nicotiana tabacum*) to identify essential leaf functions. *Plant Biotechnol J* 6:246-263.
- Lemaire, C., Girardbascou, J., Wollman, F.A., and Bennoun, P. (1986). Studies on the cytochrome *b₆f* complex. I. Characterization of the complex subunits in *Chlamydomonas reinhardtii*. *Biochim. Biophys. Acta* 851:229-238.
- Lennartz, K., Plücken, H., Seidler, A., Westhoff, P., Bechtold, N., and Meierhoff, K. (2001). HCF164 encodes a thioredoxin-like protein involved in the biogenesis of the cytochrome *b₆f* complex in *Arabidopsis*. *The Plant cell* 13:2539-2551.
- Leonov, A.A., Sergiev, P.V., Bogdanov, A.A., Brimacombe, R., and Dontsova, O.A. (2003). Affinity purification of ribosomes with a lethal G2655C mutation in 23 S rRNA that affects the translocation. *J Biol Chem* 278:25664-25670.
- Leppek, K., and Stoecklin, G. (2014). An optimized streptavidin-binding RNA aptamer for purification of ribonucleoprotein complexes identifies novel ARE-binding proteins. *Nucleic Acids Res* 42:e13.
- Levey, T., Westhoff, P., and Meierhoff, K. (2014). Expression of a nuclear-encoded *psbH* gene complements the plastidic RNA processing defect in the PSII mutant *hcf107* in *Arabidopsis thaliana*. *The Plant journal : for cell and molecular biology* 80:292-304.

- Li, G.W., Burkhardt, D., Gross, C., and Weissman, J.S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* 157:624-635.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078-2079.
- Li, X., Lindahl, L., and Zengel, J.M. (1996). Ribosomal protein L4 from *Escherichia coli* utilizes nonidentical determinants for its structural and regulatory functions. *RNA* 2:24-37.
- Li, Y., and Altman, S. (2002). Partial reconstitution of human RNase P in HeLa cells between its RNA subunit with an affinity tag and the intact protein components. *Nucleic Acids Res* 30:3706-37011.
- Li, Y., Liu, B., Zhang, J., Kong, F., Zhang, L., Meng, H., Li, W., Rochaix, J.D., Li, D., and Peng, L. (2019). OHP1, OHP2, and HCF244 form a transient functional complex with the photosystem II reaction center. *Plant physiology* 179:195-208.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30:923-930.
- Lichtenthaler, H.K. (1999). The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Annu Rev Plant Physiol Plant Mol Biol* 50:47-65.
- Lim, F., and Peabody, D.S. (1994). Mutations that increase the affinity of a translational repressor for RNA. *Nucleic Acids Res* 22:3748-3752.
- Lindahl, M., Tabak, S., Cseke, L., Pichersky, E., Andersson, B., and Adam, Z. (1996). Identification, characterization, and molecular cloning of a homologue of the bacterial FtsH protease in chloroplasts of higher plants. *J Biol Chem* 271:29329-29334.
- Link, S., Engelmann, K., Meierhoff, K., and Westhoff, P. (2012). The atypical short-chain dehydrogenases HCF173 and HCF244 are jointly involved in translational initiation of the *psbA* mRNA of *Arabidopsis*. *Plant physiology* 160:2202-2218.
- Liu, J., Valencia-Sanchez, M.A., Hannon, G.J., and Parker, R. (2005). MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* 7:719-723.
- Liu, M.J., Wu, S.H., Wu, J.F., Lin, W.D., Wu, Y.C., Tsai, T.Y., Tsai, H.L., and Wu, S.H. (2013). Translational landscape of photomorphogenic *Arabidopsis*. *The Plant cell* 25:3699-7310.
- Loiacono, F.V., Thiele, W., Schöttler, M.A., Tillich, M., and Bock, R. (2019). Establishment of a heterologous RNA editing event in chloroplasts. *Plant physiology* 181:891-900.
- Loiselay, C., Gumpel, N.J., Girard-Bascou, J., Watson, A.T., Purton, S., Wollman, F.A., and Choquet, Y. (2008). Molecular identification and function of *cis*- and *trans*-acting determinants for *petA* transcript stability in *Chlamydomonas reinhardtii* chloroplasts. *Molecular and cellular biology* 28:5529-5542.
- Lopez-Juez, E., and Pyke, K.A. (2005). Plastids unleashed: their development and their integration in plant development. *Int J Dev Biol* 49:557-577.
- Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B., et al. (2004). Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *The Plant cell* 16:2089-2103.
- Lysenko, E.A. (2007). Plant sigma factors and their role in plastid transcription. *Plant Cell Rep* 26:845-859.
- Lyska, D., Meierhoff, K., and Westhoff, P. (2013). How to build functional thylakoid membranes: from plastid transcription to protein complex assembly. *Planta* 237:413-428.
- Macias, S., Bragulat, M., Tardiff, D.F., and Vilardell, J. (2008). L30 binds the nascent RPL30 transcript to repress U2 snRNP recruitment. *Mol Cell* 30:732-742.

- Maliga, P., Sz-Breznovits, A., and Marton, L. (1973). Streptomycin-resistant plants from callus culture of haploid tobacco. *Nat New Biol* 244:29-30.
- Marsh, J.A., and Teichmann, S.A. (2015). Structure, dynamics, assembly, and evolution of protein complexes. *Annu Rev Biochem* 84:551-575.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17:10-12.
- Martín, M., Noarbe, D.M., Serrot, P.H., and Sabater, B. (2015). The rise of the photosynthetic rate when light intensity increases is delayed in *ndh* gene-defective tobacco at high but not at low CO₂ concentrations. *Frontiers in plant science* 6:34.
- Martin, W. (2003). Gene transfer from organelles to the nucleus: frequent and in big chunks. *Proceedings of the National Academy of Sciences of the United States of America* 100:8612-8614.
- Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M., and Penny, D. (2002). Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proceedings of the National Academy of Sciences of the United States of America* 99:12246-12251.
- Mattheakis, L.C., and Nomura, M. (1988). Feedback regulation of the *spc* operon in *Escherichia coli*: translational coupling and mRNA processing. *J Bacteriol* 170:4484-4492.
- McDermott, J.J., Civic, B., and Barkan, A. (2018). Effects of RNA structure and salt concentration on the affinity and kinetics of interactions between pentatricopeptide repeat proteins and their RNA ligands. *PLoS One* 13:e0209713.
- McDermott, J.J., Watkins, K.P., Williams-Carrier, R., and Barkan, A. (2019). Ribonucleoprotein capture by *in vivo* expression of a designer pentatricopeptide repeat protein in *Arabidopsis*. *The Plant cell* 31:1723-1733.
- McFarlane, H.E., Young, R.E., Wasteneys, G.O., and Samuels, A.L. (2008). Cortical microtubules mark the mucilage secretion domain of the plasma membrane in *Arabidopsis* seed coat cells. *Planta* 227:1363-1375.
- McHugh, C.A., Chen, C.K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., et al. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* 521:232-236.
- Mereschkowski, C. (1905). Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol Centralbl* 25:593-604.
- Méteignier, L.V., Ghandour, R., Meierhoff, K., Zimmerman, A., Chicher, J., Baumberger, N., Alioua, A., Meurer, J., Zoschke, R., and Hammani, K. (2020). The *Arabidopsis* mTERF-repeat MDA1 protein plays a dual function in transcription and stabilization of specific chloroplast transcripts within the *psbE* and *ndhH* operons. *New Phytol*.
- Minai, L., Wostrikoff, K., Wollman, F.A., and Choquet, Y. (2006). Chloroplast biogenesis of photosystem II cores involves a series of assembly-controlled steps that regulate translation. *The Plant cell* 18:159-175.
- Miyazawa, Y., Sakai, A., Miyagishima, S., Takano, H., Kawano, S., and Kuroiwa, T. (1999). Auxin and cytokinin have opposite effects on amyloplast development and the expression of starch synthesis genes in cultured bright yellow-2 tobacco cells. *Plant physiology* 121:461-469.
- Mohammad, F., Woolstenhulme, C.J., Green, R., and Buskirk, A.R. (2016). Clarifying the Translational Pausing Landscape in Bacteria by Ribosome Profiling. *Cell Rep* 14:686-694.
- Monde, R.A., Zito, F., Olive, J., Wollman, F.A., and Stern, D.B. (2000). Post-transcriptional defects in tobacco chloroplast mutants lacking the cytochrome *b6/f* complex. *The Plant journal : for cell and molecular biology* 21:61-72.

- Morais, F., Barber, J., and Nixon, P.J. (1998). The chloroplast-encoded α subunit of cytochrome *b*₅₅₉ is required for assembly of the photosystem two complex in both the light and the dark in *Chlamydomonas reinhardtii*. *J Biol Chem* 273:29315-29320.
- Moreno, J.C., Martinez-Jaime, S., Schwartzmann, J., Karcher, D., Tillich, M., Graf, A., and Bock, R. (2018). Temporal proteomics of inducible RNAi lines of Clp protease subunits identifies putative protease substrates. *Plant physiology* 176:1485-1508.
- Morisaki, T., Lyon, K., DeLuca, K.F., DeLuca, J.G., English, B.P., Zhang, Z., Lavis, L.D., Grimm, J.B., Viswanathan, S., Looger, L.L., et al. (2016). Real-time quantification of single RNA translation dynamics in living cells. *Science* 352:1425-1429.
- Mullet, J.E. (1993). Dynamic regulation of chloroplast transcription. *Plant physiology* 103:309-313.
- Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassay with tobacco tissue culture. *Physiol Plant* 15:473-497.
- Nagashima, A., Hanaoka, M., Shikanai, T., Fujiwara, M., Kanamaru, K., Takahashi, H., and Tanaka, K. (2004). The multiple-stress responsive plastid sigma factor, SIG5, directs activation of the *psbD* blue light-responsive promoter (BLRP) in *Arabidopsis thaliana*. *Plant & cell physiology* 45:357-368.
- Nakamura, T., Ohta, M., Sugiura, M., and Sugita, M. (2001). Chloroplast ribonucleoproteins function as a stabilizing factor of ribosome-free mRNAs in the stroma. *J Biol Chem* 276:147-152.
- Neuhaus, H.E., and Emes, M.J. (2000). Nonphotosynthetic Metabolism in Plastids. *Annu Rev Plant Physiol Plant Mol Biol* 51:111-140.
- Nickelsen, J., Bohne, A.-V., and Westhoff, P. (2014). Chloroplast gene expression - translation. In: *Plastid Biology--Theg, S.M., and Wollman, F.-A., eds.: Springer New York.* 49-78.
- Nickelsen, J., and Rengstl, B. (2013). Photosystem II assembly: from cyanobacteria to plants. *Annual review of plant biology* 64:609-635.
- 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. *The Plant journal : for cell and molecular biology* 63:766-777.
- Nishimura, K., Kato, Y., and Sakamoto, W. (2016). Chloroplast proteases: updates on proteolysis within and across suborganellar compartments. *Plant physiology* 171:2280-2293.
- Nishimura, K., and van Wijk, K.J. (2015). Organization, function and substrates of the essential Clp protease system in plastids. *Biochimica et biophysica acta* 1847:915-930.
- O'Toole, N., Hattori, M., Andres, C., Lida, K., Lurin, C., Schmitz-Linneweber, C., Sugita, M., and Small, I. (2008). On the expansion of the pentatricopeptide repeat gene family in plants. *Mol Biol Evol* 25:1120-1128.
- Oey, M., Lohse, M., Kreikemeyer, B., and Bock, R. (2009). Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. *The Plant journal : for cell and molecular biology* 57:436-445.
- Oh, E., Becker, A.H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., Nichols, R.J., Typas, A., Gross, C.A., Kramer, G., et al. (2012). Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor *in vivo*. *Cell* 147:1295-1308.
- Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesone, K., Shiki, Y., Takeuchi, M., Chang, Z., et al. (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* 322:572-574.
- Okita, T.W., and Choi, S.B. (2002). mRNA localization in plants: targeting to the cell's cortical region and beyond. *Curr Opin Plant Biol* 5:553-559.
- Okuda, K., Chateigner-Boutin, A.L., Nakamura, T., Delannoy, E., Sugita, M., Myouga, F., Motohashi, R., Shinozaki, K., Small, I., and Shikanai, T. (2009). Pentatricopeptide repeat proteins with the

- DYW motif have distinct molecular functions in RNA editing and RNA cleavage in *Arabidopsis* chloroplasts. *The Plant cell* 21:146-156.
- Ooi, B.G., Lukins, H.B., Linnane, A.W., and Nagley, P. (1987). Biogenesis of mitochondria: a mutation in the 5'-untranslated region of yeast mitochondrial *oli1* mRNA leading to impairment in translation of subunit 9 of the mitochondrial ATPase complex. *Nucleic Acids Res* 15:1965-1977.
- Orsat, B., Monfort, A., Chatellard, P., and Stutz, E. (1992). Mapping and sequencing of an actively transcribed *Euglena gracilis* chloroplast gene (*ccsA*) homologous to the *Arabidopsis thaliana* nuclear gene *cs(ch-42)*. *FEBS Lett* 303:181-184.
- Ostersetzer, O., Kato, Y., Adam, Z., and Sakamoto, W. (2007). Multiple intracellular locations of Lon protease in *Arabidopsis*: evidence for the localization of AtLon4 to chloroplasts. *Plant & cell physiology* 48:881-885.
- Ouyang, M., Li, X., Zhang, J., Feng, P., Pu, H., Kong, L., Bai, Z., Rong, L., Xu, X., Chi, W., et al. (2020). Liquid-liquid phase transition drives intra-chloroplast cargo sorting. *Cell* 180:1144-1159 e1120.
- Pakrasi, H.B., Diner, B.A., Williams, J., and Arntzen, C.J. (1989). Deletion mutagenesis of the cytochrome *b₅₅₉* protein inactivates the reaction center of photosystem II. *The Plant cell* 1:591-597.
- Palmer, J.D. (2003). The symbiotic birth and spread of plastids: How many times and whodunit? *Journal of Phycology* 39:4-12.
- Peabody, D.S. (1993). The RNA binding site of bacteriophage MS2 coat protein. *EMBO J* 12:595-600.
- Peacock, S., Brot, N., and Weissbach, H. (1983). Translational control of the expression of the β subunit gene of *E. coli* RNA polymerase. *Biochemical and biophysical research communications* 113:1018-1025.
- Perez-Martinez, X., Broadley, S.A., and Fox, T.D. (2003). Mss51p promotes mitochondrial Cox1p synthesis and interacts with newly synthesized Cox1p. *EMBO J* 22:5951-5961.
- Plücken, H., Müller, B., Grohmann, D., Westhoff, P., and Eichacker, L.A. (2002). The HCF136 protein is essential for assembly of the photosystem II reaction center in *Arabidopsis thaliana*. *FEBS Lett* 532:85-90.
- Portier, C., Dondon, L., and Grunberg-Manago, M. (1990). Translational autocontrol of the *Escherichia coli* ribosomal protein S15. *J Mol Biol* 211:407-414.
- Poutre, C.G., and Fox, T.D. (1987). *PET111*, a *Saccharomyces cerevisiae* nuclear gene required for translation of the mitochondrial mRNA encoding cytochrome *c* oxidase subunit II. *Genetics* 115:637-647.
- Pribil, M., Labs, M., and Leister, D. (2014). Structure and dynamics of thylakoids in land plants. *J Exp Bot* 65:1955-1972.
- Prikryl, J., Rojas, M., Schuster, G., and Barkan, A. (2011). Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. *Proceedings of the National Academy of Sciences of the United States of America* 108:415-420.
- Privat, I., Hakimi, M.A., Buhot, L., Favory, J.J., and Mache-Lerbs, S. (2003). Characterization of *Arabidopsis* plastid sigma-like transcription factors SIG1, SIG2 and SIG3. *Plant molecular biology* 51:385-399.
- Qbadou, S., Becker, T., Bionda, T., Reger, K., Ruprecht, M., Soll, J., and Schleiff, E. (2007). Toc64-a preprotein-receptor at the outer membrane with bipartite function. *J Mol Biol* 367:1330-1346.
- Quick, W.P., Fichtner, K., Schulze, E.D., Wendler, R., Leegood, R.C., Mooney, H., Rodermeil, S.R., Bogorad, L., and Stitt, M. (1992). Decreased ribulose-1,5-bisphosphate carboxylase-oxygenase in transgenic tobacco transformed with "antisense" *rbcS*. *Planta* 188:522-531.

- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841-842.
- Rex, G., Surin, B., Besse, G., Schneppe, B., and McCarthy, J.E. (1994). The mechanism of translational coupling in *Escherichia coli*. Higher order structure in the *atpHA* mRNA acts as a conformational switch regulating the access of *de novo* initiating ribosomes. *J Biol Chem* 269:18118-18127.
- Robles, P., Micol, J.L., and Quesada, V. (2012). Arabidopsis MDA1, a nuclear-encoded protein, functions in chloroplast development and abiotic stress responses. *PLoS One* 7:e42924.
- Rodermel, S., Haley, J., Jiang, C.Z., Tsai, C.H., and Bogorad, L. (1996). A mechanism for intergenomic integration: abundance of ribulose biphosphate carboxylase small-subunit protein influences the translation of the large-subunit mRNA. *Proceedings of the National Academy of Sciences of the United States of America* 93:3881-3885.
- Rodermel, S.R., Abbott, M.S., and Bogorad, L. (1988). Nuclear-organelle interactions: nuclear antisense gene inhibits ribulose biphosphate carboxylase enzyme levels in transformed tobacco plants. *Cell* 55:673-681.
- Rogalski, M., Karcher, D., and Bock, R. (2008). Superwobbling facilitates translation with reduced tRNA sets. *Nat Struct Mol Biol* 15:192-198.
- Röhl, T., and van Wijk, K.J. (2001). *In vitro* reconstitution of insertion and processing of cytochrome *f* in a homologous chloroplast translation system. *J Biol Chem* 276:35465-35472.
- Rojas, M., Ruwe, H., Miranda, R.G., Zoschke, R., Hase, N., Schmitz-Linneweber, C., and Barkan, A. (2018). Unexpected functional versatility of the pentatricopeptide repeat proteins PGR3, PPR5 and PPR10. *Nucleic Acids Res* 46:10448-10459.
- Rojas, M., Yu, Q., Williams-Carrier, R., Maliga, P., and Barkan, A. (2019). Engineered PPR proteins as inducible switches to activate the expression of chloroplast transgenes. *Nat Plants* 5:505-511.
- Rolland, N., Dorne, A.J., Amoroso, G., Sultemeyer, D.F., Joyard, J., and Rochaix, J.D. (1997). Disruption of the plastid *ycf10* open reading frame affects uptake of inorganic carbon in the chloroplast of *Chlamydomonas*. *EMBO J* 16:6713-6726.
- Rott, M., Martins, N.F., Thiele, W., Lein, W., Bock, R., Kramer, D.M., and Schöttler, M.A. (2011). ATP synthase repression in tobacco restricts photosynthetic electron transport, CO₂ assimilation, and plant growth by overacidification of the thylakoid lumen. *The Plant cell* 23:304-321.
- Ruf, S., Biehler, K., and Bock, R. (2000). A small chloroplast-encoded protein as a novel architectural component of the light-harvesting antenna. *The Journal of cell biology* 149:369-378.
- Ruf, S., and Bock, R. (2011). *In vivo* analysis of RNA editing in plastids. *Methods in molecular biology* 718:137-150.
- Ruf, S., Forner, J., Hasse, C., Kroop, X., Seeger, S., Schollbach, L., Schadach, A., and Bock, R. (2019). High-efficiency generation of fertile transplastomic *Arabidopsis* plants. *Nat Plants* 5:282-289.
- Ruf, S., Hermann, M., Berger, I.J., Carrer, H., and Bock, R. (2001). Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit. *Nat Biotechnol* 19:870-875.
- Ruf, S., Karcher, D., and Bock, R. (2007). Determining the transgene containment level provided by chloroplast transformation. *Proceedings of the National Academy of Sciences of the United States of America* 104:6998-7002.
- Sagan, L. (1967). On the origin of mitosing cells. *J Theor Biol* 14:255-274.
- Said, N., Rieder, R., Hurwitz, R., Deckert, J., Urlaub, H., and Vogel, J. (2009). *In vivo* expression and purification of aptamer-tagged small RNA regulators. *Nucleic Acids Res* 37:e133.
- Sakamoto, W. (2006). Protein degradation machineries in plastids. *Annual review of plant biology* 57:599-621.

- Sasaki, Y., Hakamada, K., Suama, Y., Nagano, Y., Furusawa, I., and Matsuno, R. (1993a). Chloroplast-encoded protein as a subunit of acetyl-CoA carboxylase in pea plant. *J Biol Chem* 268:25118-25123.
- Sasaki, Y., Konishi, T., and Nagano, Y. (1995). The Compartmentation of Acetyl-Coenzyme A Carboxylase in Plants. *Plant physiology* 108:445-449.
- Sasaki, Y., Sekiguchi, K., Nagano, Y., and Matsuno, R. (1993b). Chloroplast envelope protein encoded by chloroplast genome. *FEBS Lett* 316:93-98.
- Schägger, H. (2006). Tricine-SDS-PAGE. *Nat Protoc* 1:16-22.
- Scharff, L.B., Childs, L., Walther, D., and Bock, R. (2011). Local absence of secondary structure permits translation of mRNAs that lack ribosome-binding sites. *PLoS Genet* 7:e1002155.
- Scharff, L.B., Ehrnthaler, M., Janowski, M., Childs, L.H., Hasse, C., Gremmels, J., Ruf, S., Zoschke, R., and Bock, R. (2017). Shine-Dalgarno sequences play an essential role in the translation of plastid mRNAs in tobacco. *The Plant cell* 29:3085-3101.
- Schmidt, M.G., and Oliver, D.B. (1989). SecA protein autogenously represses its own translation during normal protein secretion in *Escherichia coli*. *J Bacteriol* 171:643-649.
- Schmitz-Linneweber, C., Lampe, M.K., Sultan, L.D., and Ostersetzer-Biran, O. (2015). Organellar maturases: A window into the evolution of the spliceosome. *Biochimica et biophysica acta* 1847:798-808.
- Schmitz-Linneweber, C., and Small, I. (2008). Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends in Plant Science* 13:663-670.
- 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. *The Plant cell* 17:2791-2804.
- Schöttler, M.A., Albus, C.A., and Bock, R. (2011). Photosystem I: its biogenesis and function in higher plants. *J Plant Physiol* 168:1452-1461.
- Schöttler, M.A., Flügel, C., Thiele, W., and Bock, R. (2007). Knock-out of the plastid-encoded PetL subunit results in reduced stability and accelerated leaf age-dependent loss of the cytochrome *b₆f* complex. *J Biol Chem* 282:976-985.
- Schöttler, M.A., Thiele, W., Belkuis, K., Bergner, S.V., Flügel, C., Wittenberg, G., Agrawal, S., Stegemann, S., Ruf, S., and Bock, R. (2017). The plastid-encoded PsaI subunit stabilizes photosystem I during leaf senescence in tobacco. *J Exp Bot* 68:1137-1155.
- Schöttler, M.A., Tóth, S.Z., Boulouis, A., and Kahlau, S. (2015). Photosynthetic complex stoichiometry dynamics in higher plants: biogenesis, function, and turnover of ATP synthase and the cytochrome *b₆f* complex. *J Exp Bot* 66:2373-2400.
- Schult, K., Meierhoff, K., Paradies, S., Töller, T., Wolff, P., and Westhoff, P. (2007). The nuclear-encoded factor HCF173 is involved in the initiation of translation of the *psbA* mRNA in *Arabidopsis thaliana*. *The Plant cell* 19:1329-1346.
- Schümperli, D., McKenney, K., Sobieski, D.A., and Rosenberg, M. (1982). Translational coupling at an intercistronic boundary of the *Escherichia coli* galactose operon. *Cell* 30:865-871.
- Schuster, M., Gao, Y., Schöttler, M.A., Bock, R., and Zoschke, R. (2020). Limited responsiveness of chloroplast gene expression during acclimation to high light in tobacco. *Plant physiology* 182:424-435.
- Schweer, J., Türkeri, H., Kolpack, A., and 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.
- Schwenkert, S., Legen, J., Takami, T., Shikanai, T., Herrmann, R.G., and Meurer, J. (2007). Role of the low-molecular-weight subunits PetL, PetG, and PetN in assembly, stability, and dimerization of the cytochrome *b₆f* complex in tobacco. *Plant physiology* 144:1924-1935.

- Seelert, H., Dencher, N.A., and Müller, D.J. (2003). Fourteen protomers compose the oligomer III of the proton-rotor in spinach chloroplast ATP synthase. *J Mol Biol* 333:337-344.
- Shamimuzzaman, M., and Vodkin, L. (2018). Ribosome profiling reveals changes in translational status of soybean transcripts during immature cotyledon development. *PLoS One* 13:e0194596.
- Shanklin, J., DeWitt, N.D., and Flanagan, J.M. (1995). The stroma of higher plant plastids contain ClpP and ClpC, functional homologs of *Escherichia coli* ClpP and ClpA: an archetypal two-component ATP-dependent protease. *The Plant cell* 7:1713-1722.
- Sharma, M.R., Dönhöfer, A., Barat, C., Marquez, V., Datta, P.P., Fucini, P., Wilson, D.N., and Agrawal, R.K. (2010). PSRP1 is not a ribosomal protein, but a ribosome-binding factor that is recycled by the ribosome-recycling factor (RRF) and elongation factor G (EF-G). *J Biol Chem* 285:4006-4014.
- Sharma, M.R., Wilson, D.N., Datta, P.P., Barat, C., Schluenzen, F., Fucini, P., and Agrawal, R.K. (2007). Cryo-EM study of the spinach chloroplast ribosome reveals the structural and functional roles of plastid-specific ribosomal proteins. *Proceedings of the National Academy of Sciences of the United States of America* 104:19315-19320.
- Shearwin, K.E., Callen, B.P., and Egan, J.B. (2005). Transcriptional interference - a crash course. *Trends in genetics : TIG* 21:339-345.
- Shen-Orr, S.S., Milo, R., Mangan, S., and Alon, U. (2002). Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat Genet* 31:64-68.
- Sheth, U., and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science (New York, N.Y.)* 300:805-808.
- Shi, L.X., and Schröder, W.P. (2004). The low molecular mass subunits of the photosynthetic supracomplex, photosystem II. *Biochimica et biophysica acta* 1608:75-96.
- Shikanai, T. (2016). Chloroplast NDH: A different enzyme with a structure similar to that of respiratory NADH dehydrogenase. *Biochimica et biophysica acta* 1857:1015-1022.
- Shikanai, T., Endo, T., Hashimoto, T., Yamada, Y., Asada, K., and Yokota, A. (1998). Directed disruption of the tobacco *ndhB* gene impairs cyclic electron flow around photosystem I. *Proceedings of the National Academy of Sciences of the United States of America* 95:9705-9709.
- Shikanai, T., and Fujii, S. (2013). Function of PPR proteins in plastid gene expression. *RNA Biol* 10:1446-1456.
- Shikanai, T., Shimizu, K., Ueda, K., Nishimura, Y., Kuroiwa, T., and Hashimoto, T. (2001). The chloroplast *clpP* gene, encoding a proteolytic subunit of ATP-dependent protease, is indispensable for chloroplast development in tobacco. *Plant & cell physiology* 42:264-273.
- Shine, J., and Dalgarno, L. (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America* 71:1342-1346.
- Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, N., Zaita, N., Cunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., et al. (1986). The complete nucleotide sequence of tobacco chloroplast genome: its gene organization and expression. *Embo J* 5:2043-2049.
- Sidorov, V.A., Kasten, D., Pang, S.Z., Hajdukiewicz, P.T., Staub, J.M., and Nehra, N.S. (1999). Stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker. *The Plant journal : for cell and molecular biology* 19:209-216.
- Slobodin, B., and Gerst, J.E. (2010). A novel mRNA affinity purification technique for the identification of interacting proteins and transcripts in ribonucleoprotein complexes. *RNA* 16:2277-2290.
- Small, I.D., and Peeters, N. (2000). The PPR motif - a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem Sci* 25:46-47.

- Small, I.D., Schallenberg-Rüdinger, M., Takenaka, M., Mireau, H., and Ostersetzer-Biran, O. (2020). Plant organellar RNA editing: what 30 years of research has revealed. *The Plant journal : for cell and molecular biology* 101:1040-1056.
- Smith, P.M., Fox, J.L., and Winge, D.R. (2012). Biogenesis of the cytochrome *bc₁* complex and role of assembly factors. *Biochimica et biophysica acta* 1817:276-286.
- Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res* 27:491-499.
- Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3:Article3.
- Springer, M., Weissman, J.S., and Kirschner, M.W. (2010). A general lack of compensation for gene dosage in yeast. *Mol Syst Biol* 6:368.
- Srisawat, C., and Engelke, D.R. (2001). Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins. *RNA* 7:632-641.
- Srisawat, C., Goldstein, I.J., and Engelke, D.R. (2001). Sephadex-binding RNA ligands: rapid affinity purification of RNA from complex RNA mixtures. *Nucleic Acids Res* 29:E4.
- Stampacchia, O., Girard-Bascou, J., Zanasco, J.L., Zerges, W., Bennoun, P., and Rochaix, J.D. (1997). A nuclear-encoded function essential for translation of the chloroplast *psaB* mRNA in *Chlamydomonas*. *The Plant cell* 9:773-782.
- Steitz, J.A. (1969). Polypeptide chain initiation: nucleotide sequences of the three ribosomal binding sites in bacteriophage R17 RNA. *Nature* 224:957-964.
- Stern, D.B., Goldschmidt-Clermont, M., and Hanson, M.R. (2010). Chloroplast RNA metabolism. *Annual review of plant biology* 61:125-155.
- Stern, D.B., and Grussem, W. (1987). Control of plastid gene expression: 3' inverted repeats act as mRNA processing and stabilizing elements, but do not terminate transcription. *Cell* 51:1145-1157.
- Stoppel, R., and Meurer, J. (2013). Complex RNA metabolism in the chloroplast: an update on the psbB operon. *Planta* 237:441-449.
- Sugimoto, Y., König, J., Hussain, S., Zupan, B., Curk, T., Frye, M., and Ule, J. (2012). Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions. *Genome Biol* 13:R67.
- Sun, Y., and Zerges, W. (2015). Translational regulation in chloroplasts for development and homeostasis. *Biochimica et biophysica acta* 1847:809-820.
- Svab, Z., Hajdukiewicz, P., and Maliga, P. (1990). Stable transformation of plastids in higher plants. *Proceedings of the National Academy of Sciences of the United States of America* 87:8526-8530.
- Svab, Z., and Maliga, P. (1993). High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. *Proceedings of the National Academy of Sciences of the United States of America* 90:913-917.
- Swiatecka-Hagenbruch, M., Emanuel, C., Hedtke, B., Liere, K., and Börner, T. (2008). Impaired function of the phage-type RNA polymerase RpoTp in transcription of chloroplast genes is compensated by a second phage-type RNA polymerase. *Nucleic Acids Res* 36:785-792.
- Swiatek, M., Regel, R.E., Meurer, J., Wanner, G., Pakrasi, H.B., Ohad, I., and Herrmann, R.G. (2003). Effects of selective inactivation of individual genes for low-molecular-mass subunits on the assembly of photosystem II, as revealed by chloroplast transformation: the *psbEFLJ* operon in *Nicotiana tabacum*. *Mol Genet Genomics* 268:699-710.
- Szabo, E.X., Reichert, P., Lehniger, M.K., Ohmer, M., de Francisco Amorim, M., Gowik, U., Schmitz-Linneweber, C., and Laubinger, S. (2020). Metabolic labeling of RNAs uncovers hidden features and dynamics of the *Arabidopsis* transcriptome. *The Plant cell* 32:871-887.

- Taggart, J.C., and Li, G.W. (2018). Production of protein-complex components is stoichiometric and lacks general feedback regulation in eukaryotes. *Cell Syst* 7:580-589 e584.
- Takahashi, Y., Goldschmidt-Clermont, M., Soen, S.Y., Franzén, L.G., and Rochaix, J.D. (1991). Directed chloroplast transformation in *Chlamydomonas reinhardtii*: insertional inactivation of the *psaC* gene encoding the iron sulfur protein destabilizes photosystem I. *EMBO J* 10:2033-2040.
- 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 molecular biology* 24:779-788.
- Teubner, M., Lenzen, B., Espenberger, L.B., Fuss, J., Nickelsen, J., Krause, K., Ruwe, H., and Schmitz-Linneweber, C. (2020). The chloroplast ribonucleoprotein CP33B quantitatively binds the *psbA* mRNA. *Plants (Basel)* 9.
- Theil, K., Imami, K., and Rajewsky, N. (2019). Identification of proteins and miRNAs that specifically bind an mRNA *in vivo*. *Nature communications* 10:4205.
- Tiller, N., and Bock, R. (2014). The translational apparatus of plastids and its role in plant development. *Mol Plant* 7:1105-1120.
- Tillich, M., Beick, S., and Schmitz-Linneweber, C. (2010). Chloroplast RNA-binding proteins: repair and regulation of chloroplast transcripts. *RNA Biol* 7:172-178.
- Timmis, J.N., Ayliffe, M.A., Huang, C.Y., and Martin, W. (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* 5:123-135.
- Trösch, R., Barahimipour, R., Gao, Y., Badillo-Corona, J.A., Gotsmann, V.L., Zimmer, D., Mühlhaus, T., Zoschke, R., and Willmund, F. (2018). Commonalities and differences of chloroplast translation in a green alga and land plants. *Nat Plants* 4:564-575.
- Tuerk, C., and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249:505-510.
- Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 13:731-740.
- Uniacke, J., and Zerges, W. (2009). Chloroplast protein targeting involves localized translation in *Chlamydomonas*. *Proceedings of the National Academy of Sciences of the United States of America* 106:1439-1444.
- van Bezouwen, L.S., Caffarri, S., Kale, R.S., Kouřil, R., Thunnissen, A.W.H., Oostergetel, G.T., and Boekema, E.J. (2017). Subunit and chlorophyll organization of the plant photosystem II supercomplex. *Nat Plants* 3:17080.
- van de Guchte, M., Kok, J., and Venema, G. (1991). Distance-dependent translational coupling and interference in *Lactococcus lactis*. *Mol Gen Genet* 227:65-71.
- van Wijk, K.J. (2015). Protein maturation and proteolysis in plant plastids, mitochondria, and peroxisomes. *Annual review of plant biology* 66:75-111.
- van Wijk, K.J., Bingsmark, S., Aro, E.M., and Andersson, B. (1995). *In vitro* synthesis and assembly of photosystem II core proteins. The D1 protein can be incorporated into photosystem II in isolated chloroplasts and thylakoids. *J Biol Chem* 270:25685-256895.
- Verma, D., and Daniell, H. (2007). Chloroplast vector systems for biotechnology applications. *Plant physiology* 145:1129-1143.
- Vitlin Gruber, A., and Feiz, L. (2018). Rubisco assembly in the chloroplast. *Front Mol Biosci* 5:24.
- Vollmar, M., Schlieper, D., Winn, M., Buchner, C., and Groth, G. (2009). Structure of the c14 rotor ring of the proton translocating chloroplast ATP synthase. *J Biol Chem* 284:18228-18235.

- Walker, S.C., Scott, F.H., Srisawat, C., and Engelke, D.R. (2008). RNA affinity tags for the rapid purification and investigation of RNAs and RNA-protein complexes. *Methods in molecular biology* 488:23-40.
- Walter, B., Hristou, A., Nowaczyk, M.M., and Schünemann, D. (2015). *In vitro* reconstitution of co-translational D1 insertion reveals a role of the cpSec-Alb3 translocase and Vipp1 in photosystem II biogenesis. *Biochem J* 468:315-324.
- Wang, F., Flanagan, J., Su, N., Wang, L.C., Bui, S., Nielson, A., Wu, X., Vo, H.T., Ma, X.J., and Luo, Y. (2012). RNAscope: a novel *in situ* RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J Mol Diagn* 14:22-29.
- Watkins, K.P., Williams-Carrier, R., Chotewutmontri, P., Friso, G., Teubner, M., Belcher, S., Ruwe, H., Schmitz-Linneweber, C., van Wijk, K.J., and Barkan, A. (2019). Exploring the proteome associated with the mRNA encoding the D1 reaction center protein of Photosystem II in plant chloroplasts. *The Plant journal : for cell and molecular biology*.
- Wei, X., Su, X., Cao, P., Liu, X., Chang, W., Li, M., Zhang, X., and Liu, Z. (2016). Structure of spinach photosystem II-LHCII supercomplex at 3.2 Å resolution. *Nature* 534:69-74.
- Weis, B.L., Schleiff, E., and Zerges, W. (2013). Protein targeting to subcellular organelles via mRNA localization. *Biochimica et biophysica acta* 1833:260-273.
- Whitfeld, P.R., Leaver, C.J., Bottomley, W., and Atchison, B.A. (1978). Low-molecular-weight (4.5S) ribonucleic acid in higher-plant chloroplast ribosomes. *Biochemical Journal* 175:1103-1112.
- Williams-Carrier, R., Brewster, C., Belcher, S.E., Rojas, M., Chotewutmontri, P., Ljungdahl, S., and Barkan, A. (2019). The *Arabidopsis* pentatricopeptide repeat protein LPE1 and its maize ortholog are required for translation of the chloroplast *psbJ* RNA. *The Plant journal : for cell and molecular biology* 99:56-66.
- Wittenberg, G., Järvi, S., Hojka, M., Tóth, S.Z., Meyer, E.H., Aro, E.M., Schöttler, M.A., and Bock, R. (2017). Identification and characterization of a stable intermediate in photosystem I assembly in tobacco. *The Plant journal : for cell and molecular biology* 90:478-490.
- Wolin, S.L., and Walter, P. (1988). Ribosome pausing and stacking during translation of a eukaryotic mRNA. *EMBO J* 7:3559-3569.
- Wollman, F.A., Minai, L., and Nechushtai, R. (1999). The biogenesis and assembly of photosynthetic proteins in thylakoid membranes I. *Biochimica et biophysica acta* 1411:21-85.
- Wostrikoff, K., Choquet, Y., Wollman, F.A., and Girard-Bascou, J. (2001). TCA1, a single nuclear-encoded translational activator specific for *petA* mRNA in *Chlamydomonas reinhardtii* chloroplast. *Genetics* 159:119-132.
- Wostrikoff, K., Girard-Bascou, J., Wollman, F.A., and Choquet, Y. (2004). Biogenesis of PSI involves a cascade of translational autoregulation in the chloroplast of *Chlamydomonas*. *EMBO J* 23:2696-2705.
- Wostrikoff, K., and Stern, D. (2007). Rubisco large-subunit translation is autoregulated in response to its assembly state in tobacco chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America* 104:6466-6471.
- Wu, B., Eliscovich, C., Yoon, Y.J., and Singer, R.H. (2016). Translation dynamics of single mRNAs in live cells and neurons. *Science* 352:1430-1435.
- Wu, C.C., Zinshteyn, B., Wehner, K.A., and Green, R. (2019). High-resolution ribosome profiling defines discrete ribosome elongation states and translational regulation during cellular stress. *Mol Cell* 73:959-970 e955.
- Xiao, J., Li, J., Ouyang, M., Yun, T., He, B., Ji, D., Ma, J., Chi, W., Lu, C., and Zhang, L. (2012). DAC is involved in the accumulation of the cytochrome *b₆f* complex in *Arabidopsis*. *Plant physiology* 160:1911-1922.

- Xiao, S., Day-Storms, J.J., Srisawat, C., Fierke, C.A., and Engelke, D.R. (2005). Characterization of conserved sequence elements in eukaryotic RNase P RNA reveals roles in holoenzyme assembly and tRNA processing. *RNA* 11:885-896.
- Xie, Z., and Merchant, S. (1996). The plastid-encoded *ccsA* gene is required for heme attachment to chloroplast *c*-type cytochromes. *J Biol Chem* 271:4632-4639.
- Xiong, S., Zhang, L., and He, Q.Y. (2008). Fractionation of proteins by heparin chromatography. *Methods in molecular biology* 424:213-221.
- 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. *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.
- Yao, W.B., Meng, B.Y., Tanaka, M., and Sugiura, M. (1989). An additional promoter within the protein-coding region of the *psbD-psbC* gene cluster in tobacco chloroplast DNA. *Nucleic Acids Res* 17:9583-9591.
- Yates, J.L., and Nomura, M. (1980). *E. coli* ribosomal protein L4 is a feedback regulatory protein. *Cell* 21:517-522.
- Yoon, B.C., Zivraj, K.H., and Holt, C.E. (2009). Local translation and mRNA trafficking in axon pathfinding. *Results Probl Cell Differ* 48:269-288.
- Yosef, I., Irihimovitch, V., Knopf, J.A., Cohen, I., Orr-Dahan, I., Nahum, E., Keasar, C., and 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.
- Yu, Q., Lutz, K.A., and Maliga, P. (2017). Efficient Plastid Transformation in *Arabidopsis*. *Plant physiology* 175:186-193.
- Yukawa, M., Tsudzuki, T., and Sugiura, M. (2005). The 2005 version of the chloroplast DNA sequence from tobacco (*Nicotiana tabacum*). *Plant Molecular Biology Reporter* 23:359-365.
- Zambrano, A., Fontanesi, F., Solans, A., de Oliveira, R.L., Fox, T.D., Tzagoloff, A., and Barrientos, A. (2007). Aberrant translation of cytochrome *c* oxidase subunit 1 mRNA species in the absence of Mss51p in the yeast *Saccharomyces cerevisiae*. *Mol Biol Cell* 18:523-535.
- Zara, V., Conte, L., and Trumpower, B.L. (2009). Biogenesis of the yeast cytochrome *bc1* complex. *Biochimica et biophysica acta* 1793:89-96.
- Zghidi-Abouzid, O., Merendino, L., Buhr, F., Malik Ghulam, M., and Lerbs-Mache, S. (2011). Characterization of plastid psbT sense and antisense RNAs. *Nucleic Acids Res* 39:5379-5387.
- Zhang, J., Ruf, S., Hasse, C., Childs, L., Scharff, L.B., and Bock, R. (2012). Identification of *cis*-elements conferring high levels of gene expression in non-green plastids. *The Plant journal : for cell and molecular biology* 72:115-128.
- Zhang, L., Paakkanen, V., Suorsa, M., and Aro, E.M. (2001). A SecY homologue is involved in chloroplast-encoded D1 protein biogenesis. *J Biol Chem* 276:37809-37814.
- 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. *Photosynthesis Research* 38:369-377.
- Zhelyazkova, P., Sharma, C.M., Forstner, K.U., Liere, K., Vogel, J., and Börner, T. (2012). The primary transcriptome of barley chloroplasts: numerous noncoding RNAs and the dominating role of the plastid-encoded RNA polymerase. *The Plant cell* 24:123-136.

- Zhou, Z., Sim, J., Griffith, J., and Reed, R. (2002). Purification and electron microscopic visualization of functional human spliceosomes. *Proceedings of the National Academy of Sciences of the United States of America* 99:12203-12207.
- Zoschke, R., and Barkan, A. (2015). Genome-wide analysis of thylakoid-bound ribosomes in maize reveals principles of cotranslational targeting to the thylakoid membrane. *Proceedings of the National Academy of Sciences of the United States of America* 112:E1678-E1687.
- Zoschke, R., and Bock, R. (2018). Chloroplast translation: structural and functional organization, operational control, and regulation. *The Plant cell* 30:745-770.
- Zoschke, R., Chotewutmontri, P., and Barkan, A. (2017). Translation and co-translational membrane engagement of plastid-encoded chlorophyll-binding proteins are not influenced by chlorophyll availability in maize. *Frontiers in plant science* 8:385.
- Zoschke, R., Nakamura, M., Liere, K., Sugiura, M., Börner, T., and Schmitz-Linneweber, C. (2010). An organellar maturase associates with multiple group II introns. *Proceedings of the National Academy of Sciences of the United States of America* 107:3245-3250.
- Zoschke, R., Watkins, K.P., and Barkan, A. (2013). A rapid ribosome profiling method elucidates chloroplast ribosome behavior *in vivo*. *The Plant cell* 25:2265-2275.
- Zoschke, R., Watkins, K.P., Miranda, R.G., and Barkan, A. (2016). The PPR-SMR protein PPR53 enhances the stability and translation of specific chloroplast RNAs in maize. *The Plant journal : for cell and molecular biology* 85:594-606.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31:3406-3415.

7 Publications and conference attendance

7.1 Publications

Published :

Méteignier, L.V., **Ghandour, R.**, Meierhoff, K., Zimmerman, A., Chicher, J., Baumberger, N., Alioua, A., Meurer, J., Zoschke, R., and Hammani, K. (2020). The *Arabidopsis* mTERF-repeat MDA1 protein plays a dual function in transcription and stabilization of specific chloroplast transcripts within the *psbE* and *ndhH* operons. *New Phytol.*

Joret, C., Capeyrou, R., Belhabich-Baumas, K., Plisson-Chastang, C., **Ghandour, R.**, Humbert, O., Fribourg, S., Leulliot, N., Lebaron, S., Henras, A.K., et al. (2018). The Npa1p complex chaperones the assembly of the earliest eukaryotic large ribosomal subunit precursor. *PLoS Genet.*

In revision:

Méteignier, L.V., **Ghandour, R.**, Zimmerman, A., Kuhn, L., Meurer, J., Zoschke, R., and Hammani, K. (2020). *Arabidopsis* mTERF-repeat mTERF9 protein promotes chloroplast ribosomal assembly and translation via ribonucleoprotein *in vivo* interactions. *The Plant cell.*

Fu, H.-Y., **Ghandour, R.**, Ruf, S., Zoschke, R., Bock, R., and Schöttler, M.A. (2020). PSII biogenesis in tobacco is neither limited by transcript accumulation nor translation of D2 and CP43.

7.2 Conference and workshops attendance

- 09/2019 The 14th International Colloquium in Endocytobiology and Symbiosis, Lille, France.
- 01/2019 Gordon Research Conference for Chloroplast Biotechnology: Redesigning Plastids for Novel Functions, Ventura, CA, US.
- 01/2019 Gordon Research Seminar for Chloroplast Biotechnology: Redesigning Plastids for Novel Functions, Ventura, CA, US.
- 08/2018 Summer school: “Plant Genome editing”, Berlin, Germany (SFB TRR175).
- 07/2018 Summer school: “Data analysis in R”, Vrije University, Amsterdam, Netherlands.
- 04/2018 Havel Spree Colloquium, Humboldt University in Berlin, Germany.
- 08/2017 Summer school: “Omics and Quantitative Biology”, in Bad Münster am Stein-Ebernburg.

8 Eidesstattliche Erklärung

Hiermit versichere ich, die vorliegende Arbeit selbstständig angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben. Ich versichere ebenfalls, dass die Arbeit an keiner anderen Hochschule als der Universität Potsdam eingereicht wurde.

Potsdam-Golm, Juni 2020

Rabea Ghandour

9 Acknowledgements

First, I would like to acknowledge my supervisor, Reimo, for his guidance, advices, and constructive support throughout my thesis. Also I would like to thank Ralph for his valuable discussions and suggestions.

Special thanks to Kerstin, Jacqueline, Birgitt, and Stefanie for all their help with all the administrative stuff and bureaucracy. Thanks to Valerie for being always so nice and helpful.

Thanks to the infrastructure groups in the MPIMP. Big thanks to the greenteam especially Britta, Florian, and Sven who always took very good care of my plants. Special thanks for IT from helping me installing lots of programs and solving all my computer problems. Thanks for haustechnik and the media kitchen for their support.

I am thankful for the service groups in the MPIMP, especially the trafo team (thanks to Steph and Margit) and the microscopy service group (thanks to Arun and Anja) for their help and patience with all my questions. I am grateful for Nooshin for her help with the statistical analysis.

Thanks to my PAC (Christian, Dirk, and Zoran) for their good advice and helping to narrow the focus of my work and make a plan for publishing (and always remind me to keep an eye on the time).

Big thanks to the whole AG Zoschke. First, a huge thank for Yang who teached me how to do ribosome profiling and for every day discussion scientifically and personal. Thanks for Ines for her kindness and help in the lab especially with approving all my orders! Thanks for Mehrdad and Maja for being nice, fun, and active. To Mike, thanks for all the great activities outside the lab especially the robo! Thanks for handling my Ribo-seq data analysis. Thanks for Raphael, Reimo, and Yang for proof reading my thesis. Thanks to all the members of the group for the great working atmosphere (Chen, Lorenz, Kenny, Julia, Marcin, Jesus, Dascha, Bai, Robin, Maik, Maruf, Josi and her super delicious cheesecake). A very special thanks to my first student ever, Carla, who was not only a great help in the lab but also a great person and friend. Thank you.

More thanks to all the friends I made at the MPIMP: Sandrine, Irina, Vanessa, Gian, Fabio, and to all the game night gang and the IMPRS students. Thanks for Ina for her great help with sorting all the things with the university, for organizing the progress seminar, the PAC meetings, and also helping me to submit my thesis. Thank you Ina for being the organization master.

I am thankful to the people in AG Bock who were always so helpful with small questions and scientific discussions. Thanks to Guo Zhang for answering my questions on Arabidopsis and for Etienne for introducing me to mass spectrometry. A very special thanks for Mark Aurel for his help with organizing the phytotron space, for doing some photosynthetic measurements for me and for providing many of the unpublished mutants I studied in my thesis.

Thanks to my family for their support and encouragement. Biggest thanks of all goes to my mother whom I admire the most. Thank you for all your sacrifices and for always being there for me.