A role for chloroplast RNA binding protein CP29A in *rbcL* expression during cold acclimation

Benjamin Lenzen[1$], Florian Rösch[1$], Hannes Ruwe[1], Julia Legen[1], Ian Small[2], Christian Schmitz-Linneweber[1]

[1] Molecular Genetics, Humboldt Universität zu Berlin, Philippstr.13, 10115 Berlin, Germany.

[2] Australian Research Council Centre of Excellence in Plant Energy Biology, School of Molecular Sciences, The University of Western Australia, Crawley 6009, Australia.

Corresponding author: christian.schmitz-linneweber@rz.hu-berlin.de

[$] equal contribution

Abstract

The chloroplast genome encodes key components of the photosynthetic light reaction machinery and the large subunit of the enzyme central for carbon fixation, RuBisCo. Plants constantly face the challenge of balancing light and dark reactions under varying environmental conditions. Nuclear RNA binding proteins (RBPs) play a crucial role in plant acclimation to these changes through post-transcriptional processes. Mutants of chloroplast gene expression factors often exhibit impaired chloroplast biogenesis, especially in cold conditions. Cold temperatures pose a challenge for plants as they slow down Calvin Cycle enzymes, potentially leading to a shortage of electron acceptors and oxidative damage from excess electrons in the thylakoid membrane. A well-known response of plants to this problem is to increase the production of RuBisCo and other Calvin Cycle enzymes in the cold. The chloroplast RNA binding protein CP29A targets *rbcL* mRNA and is essential for cold resistance in *Arabidopsis thaliana*. This effect is confined to the youngest leaf tissue and is linked to its role in enhancing the splicing of various chloroplast RNAs in cold conditions.

n this study, we utilized enhanced cross-linking and immunoprecipitation (eCLIP) and RNA-Bind-N-Seq (RBNS) to investigate the RNA targets of CP29A, achieving nucleotide-resolution insights into protein-RNA interaction sites. We discovered that CP29A preferentially binds to mRNAs encoding subunits of photosystem II. Notably, one of the most confidently identified targets of CP29A is the 5'-UTR of *rbcL*, where it interacts with a site downstream of the pentatricopeptide repeat protein MRL1, a crucial player in *rbcL* accumulation. Arabidopsis mutants lacking CP29A showed no significant *rbcL* changes, possibly due to CP29A's restricted role in a limited number of cells at the base of leaves. In contrast, CRISPR mutants

of tobacco *NtCP29A* exhibit photosynthetic deficiencies throughout the entire leaf blade, correlating with a substantial decrease in both *rbcL* mRNA and RbcL protein levels. Conclusively, our study establishes CP29A as a pioneer regulator in sustaining optimal RuBisCo expression during cold acclimation, highlighting its integral role in plant cold response mechanisms.

Introduction

The chloroplast is crucial in how plants adapt to cold stress (Crosatti et al., 2013). Cold tolerance in many plants relies on light and photosynthesis during low-temperature growth. When exposed to light in cold conditions, two key changes occur: Calvin cycle enzyme activity declines and thus the availability of electron acceptors becomes limited, while at the same time the plastoquinone pool is fully reduced. Eventually, this leads to a buildup of reactive oxygen species. These metabolic shifts send signals that control the expression of nuclear genes. The degree of the chloroplast's response to cold often directly relates to the plant's overall ability to withstand chilling and freezing (Crosatti et al., 2013). This suggests that how well the chloroplast adjusts to cold might be a critical factor in a plant's overall capacity to adapt to low temperatures.

Acclimation of chloroplasts to cold includes changes in chloroplast gene expression given that the chloroplast is encoding key components of the light and dark reactions of photosynthesis. Among these genes is *rbcL*, encoding the large subunit of RuBisCo. An important factor required for the expression of RuBisCo is the pentatricopeptide repeat (PPR) protein MRL1. MRL1 has been shown to target the 5'-UTR of the *rbcL* mRNA and is required for normal levels of *rbcL* mRNA accumulation, likely by blocking exonuclease-mediated RNA degradation (Johnson et al., 2010). Another factor associated with *rbcL* mRNA is the chloroplast ribonucleoprotein (cpRNP) CP29A (Kupsch et al., 2012). The cpRNP family, composed of ten nucleus-encoded, chloroplast-targeted RNA-binding proteins (RBPs), is conserved across angiosperms (Ruwe et al., 2011). The RNA-binding ability of cpRNPs is contingent upon two RNA recognition motif (RRM) binding domains, connected by a linker domain whose size and composition differ between members. cpRNPs were initially discovered in chloroplast lysates through affinity pulldown experiments using synthetic nucleic acid homopolymers (Li and Sugiura, 1990). These proteins demonstrated a high affinity for RNA homopolymers composed of uracil (U) and guanine (G), but showed lower affinity for cytosine (C) and no affinity for adenine (A) homopolymers (Li and Sugiura, 1991). The interaction of cpRNPs with chloroplast RNAs was examined through a series of *in vitro* and *in vivo* tests. Intron-lacking tRNAs and rRNAs were found to have little or no affinity for cpRNPs (Nakamura et al., 1999). In contrast, cpRNPs were found to associate with various mRNAs and some intron-containing tRNAs (Kupsch et al., 2012; Teubner et al., 2020, 2017; Nakamura et al., 1999). A top-

enriched transcript in IPs of CP29A was *rbcL* (Kupsch et al., 2012). Beyond the identification of co-precipitating transcripts, no details on binding sites *in vivo* nor whether binding of RNA is direct *in vivo* are known.

Crosslinking followed by immunoprecipitation (CLIP) approaches employ crosslinking by UV light, which affects only direct protein-RNA interactions (Licatalosi et al., 2008; Meyer et al., 2017; Konig et al., 2011). The resulting covalent bonds preserve *in vivo* binding events, are non-reversible, and resist harsh washing conditions. Thereby, the number of potential false positive interactions is reduced and the identification of specific targets of an RBP is facilitated. Another advantage of CLIP-based protocols is the higher resolution and thus the ability to identify target sites rather than target RNAs. As the molecular function of RBPs is highly dependent on their binding position within target transcripts, knowledge of binding sites can provide helpful information with respect to an RBP's biological role. Therefore, a CLIP-based approach was established for the investigation of the molecular role of CP29A in organellar gene expression.

Results

eCLIP demonstrates that CP29A binds RNA directly *in vivo* and has a preference for U-rich sequence elements *in vivo* and *in vitro*

Only a limited number of high-throughput UV-crosslinking experiments have been conducted in plants so far (Zhang et al., 2015, 2016; Reichel et al., 2016; Meyer et al., 2017), and none on chloroplasts. We adapted the eCLIP protocol (Van Nostrand et al., 2016, 2017) to the chloroplast using specific antibodies for CP29A and CP33B, another member of the cpRNP family. The binding pattern of CP33B is, based on previous RIP-Chip results (Kupsch et al., 2012; Watkins et al., 2019; Teubner et al., 2020), substantially different from CP29A. The parallel analysis of two RBPs with different target specificities can therefore uncover potential bias and false-positives in the procedure.

We used 10 billion chloroplasts per CLIP sample and crosslinked them with UV light, followed by lysis and immunoprecipitation using antibodies against CP29A and CP33B. Successful precipitation of protein-RNA complexes was verified by western blot and radio-labeling of bound RNA (Figure S1). RNA-protein complexes vary in size based on the length of the target RNA, and they are larger than the ribonucleoprotein (RBP) by itself. This size variation is evident on the autoradiogram, where it appears as a smear of radioactive signal above the respective cpRNP size (figure S1B). The area of the blot corresponding to this signal was excised and then treated with protease. This treatment facilitates the release of RNAs from the nitrocellulose. Subsequently, the released RNA was processed for library preparation compatible with Illumina sequencing, enabling high-throughput sequencing analysis.

The analysis of the CP29A CLIP dataset revealed 69 peaks with significant enrichment, distributed across 15 chloroplast mRNAs and two tRNAs (Figure 1A). The CP33B CLIP dataset analysis identified 52 peaks among 5 mRNAs and one tRNA (Figure 1B). Notably, none of the significantly enriched peaks in the CP29A and CP33B datasets overlapped at the sequence level, indicating distinct specificities. The *psbA* mRNA was confirmed as a target for CP33B, with no corresponding peak in the CP29A dataset (Figure 1C). In contrast, *rbcL* mRNA was a target of CP29A but not CP33B, aligning with previous RIP-Chip and dot blot analyses (Figure 1D; Kupsch et al., 2012; Watkins et al., 2019; Teubner et al., 2020).

CP29A peaks were found in all transcriptional units encoding proteins of the photosystem II complex (p*sbB-psbT-psbH-petB-petD*; *psbD-psbC-psbZ*; *psbE-psbF-psbL-psbJ*; *psbK-psbI*; *psbM*), except in the monocistronic *psbN* and *psbA* mRNAs. Targets included nine out of fifteen *psb* transcripts, along with significant peaks in *petB* and *petD* of the *psbB* operon, encoding cytochrome b_6f complex subunits. The *atpF* mRNA, related to the ATP synthase complex, was also a target of CP29A. Additionally, monocistronic mRNAs *psaI*, *petN*, and the already mentioned *rbcL* were identified as targets. Most of the significant CP29A peaks were detected in coding sequences. Next to these, there were peaks in the tRNA genes *trnH* and *trnT* and also some peaks in 3' and 5' untranslated regions (UTRs) (Figure 1A and Figure S2). UTR regions are known to be of prime importance for the regulation of transcript stability and translation in chloroplasts (Small et al., 2023) and therefore of special interest. Target UTRs of CP29A include the 3'-UTRs of *atpF* (partly overlapping with the CDS), *petN*, *psbM* and the 5'-UTR of *rbcL* and *petB* (Figure 1D and S2).

For CP33B, significant peaks were also predominantly located in the CDS of target genes, except for *atpH*, where peaks extended partially into the 3' UTR (Figure S2). Besides the top peak in *psbA*, peaks were identified in the second exon of *petB*, in *psbB* and *psbD* mRNA, and in *trnH*.

It is important to note that the CLIP analysis for CP29A and CP33B might be biased towards well-expressed genes, as all target genes in this study are in the top two tertiles when ranked by read counts in the wildtype Col-0 RNA-seq dataset (Legen et al., 2023). Hence, the potential for false negatives due to low expression levels cannot be ruled out, which may include many intronic regions. The identified transcripts in the CP29A and CP33B datasets through CLIP are consistent with previously identified target mRNAs (Kupsch et al., 2012; Watkins et al., 2019; Teubner et al., 2020) and demonstrate the utility of CLIP-based assays in chloroplasts.

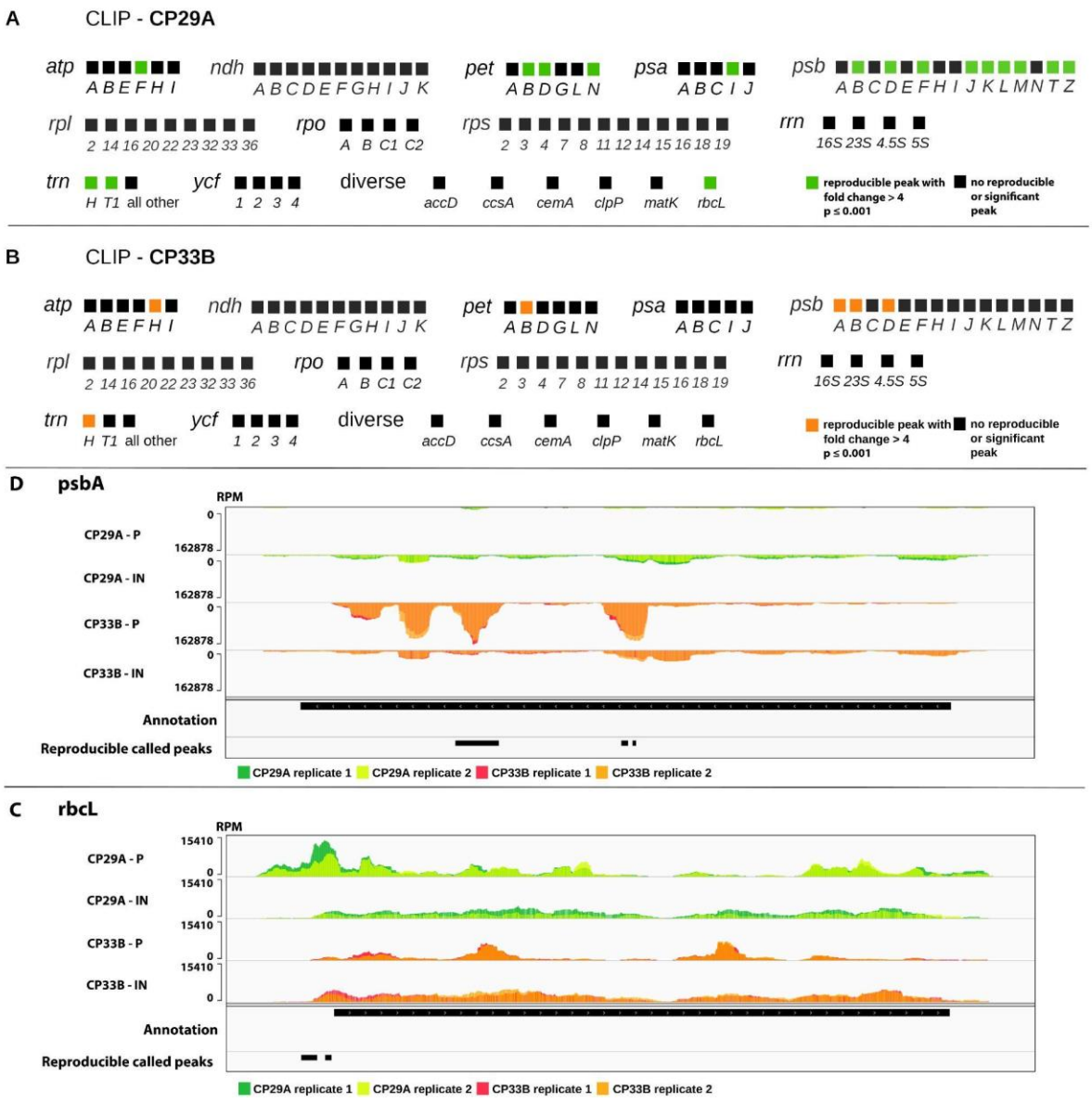

Figure 1: Binding analysis of CP29A and CP33B using an eCLIP-derived approach. CLIP libraries (P) and corresponding size-matched input (IN) libraries were prepared from UV-crosslinked chloroplasts. The library preparation workflow was adapted to chloroplasts based on the eCLIP protocol (Van Nostrand et al., 2016). Data analysis was performed using the eCLIP pipeline (v0.3.99) and the chloroplast genome (TAIR10; customized AtRTD2 annotation with designated 100 nt of 5'- and 3'-UTR). Identified peaks were considered significant and reproducible if fold-enrichment was > 4 and $p < 0.001$ in CLIP versus size-matched input, in both replicates. (A) Identified reproducible peaks for CP29A. (B) Identified reproducible peaks for CP33B. In detail CLIP data for CP29A and CP33B at the (C) *rbcL* and (D) *psbA* locus. Read coverage (in reads per million = RPM) plots are shown for both biological replicates. Identified reproducible peaks are annotated as black rectangles.

In vitro binding data reveal the sequence specificity of CP29A and CP33B

The binding of RNA-binding proteins (RBPs) to their target RNAs depends on specific structural and sequence characteristics. RRM domains, as found in cpRNPs, exhibit a variety of binding patterns, rendering the prediction of target sequences currently unfeasible (Maris et al., 2005). Here, we utilized an *in vitro* method, RNA Bind-n-Seq (RBNS; Lambert et al.,

2014), to delineate the binding preferences of cpRNPs and to supplement *in vivo* CLIP data. RBNS employs recombinant RBPs to isolate target RNAs from a synthesized pool of random RNAs, followed by analysis through high-throughput sequencing.

We engineered CP29A and CP33B with an N-terminal glutathione-S-transferase (GST) tag and a streptavidin-binding protein (SBP) tag. These were heterologously expressed in *E. coli* and subsequently purified using the GST tag, which was removed during purification. The purified recombinant RBPs were incubated with a mixture of random 40-mer RNA oligonucleotides at various concentrations (0, 10, 100, and 1000 mM). The RBPs were then isolated using their SBP tag, and the associated RNAs were extracted. We prepared sequencing libraries from both the original RNA pool and the immunoprecipitated (IP) samples, which were sequenced on an Illumina NextSeq platform. The resulting libraries, generating approximately 20 million reads each, were analyzed using the RBNS pipeline (Dominguez et al., 2018). Briefly, we counted all possible k-mers of a specified length in both the IP samples and the input RNA pool. We calculated the enrichment of each k-mer as its frequency in the IP samples relative to its frequency in the input pool, denoted as the R-value. Significantly enriched k-mers (z-score > 3) from the library with the highest R-value were selected for further downstream analysis.

RRM domains typically bind sequences ranging from 2 to 4 nucleotides in length. Given that cpRNPs contain two RRM domains, we focused our RBNS library analysis on 6-mers. The highest R-values for both CP29A and CP33B were found in libraries derived from IPs conducted with a protein concentration of 100 nM. The control samples, which contained no protein, did not show significant enrichment for any 6-mers. For CP29A, the most highly enriched 6-mers were UUUUUU (R-value=2.26), UUUUUA (R-value=1.98), and UAUUUU (R-value=1.97). In the case of CP33B, the 6-mers with the highest enrichment were GUUACU (R-value= 3.47), GGUUAC (R-value=3.39), and GCUACU (R-value= 2.67). These enriched 6-mers for each protein exhibited a notable sequence similarity and maintained a consistent ranking among all significantly enriched 6-mers, regardless of the protein concentration used in the experiments. All significantly enriched 6-mers were then utilized to construct consensus sequence motifs and to develop respective position weight matrices for each motif, as shown in Figure 2.

The binding preference of CP29A was most accurately characterized by a degenerate U-rich motif, with A, U, or G being almost equally probable at the third position. The position weight matrix (PWM)-based consensus motif for CP33B corresponded closely to the consensus sequence GHUAUY. To examine the occurrence of these RBNS-derived PWM-based consensus motifs within CLIP peaks compared to control peaks, we employed the Analysis of Motif Enrichment (AME) tool from the MEME suite (McLeay and Bailey, 2010). For the

sequence analysis of CLIP-derived peaks, we first merged all peaks located within 10 base pairs of each other. These merged peaks were then extended to match the size of the largest observed peak (53 nucleotides for CP29A and 82 nucleotides for CP33B). The control peak set was generated in accordance with the guidelines in the AME manual. Briefly, all chloroplast sequences annotated in AtRTD2, including 100 nucleotides of 5'- and 3'- pseudo-UTRs as used in the CLIP analysis, were shuffled ten times. This process produced ten randomly shuffled versions of each gene. Using these sequences, 1000 peaks, each matching the size of the largest merged CLIP peak, were randomly selected to create a control set for each protein. Additionally, the chloroplast gene sequences were utilized to construct a 0-order background model for the AME tool.

The PWM-based RBNS-derived motif of CP29A was identified in 13 out of 23 CLIP-derived merged peaks, as determined by AME. This enrichment was statistically significant, with a multiple testing corrected p-value of less than 0.05 (Figure 2A). Notably, this motif appeared twice in the *rbcL* 5'-UTR CLIP peak, which exhibited the lowest p-value in the peak enrichment analysis (Figure 2B). In contrast, the AME analysis for CP33B did not reveal any statistically significant enrichment of the RBNS motif in the merged CLIP peaks. However, this motif was found in 5 of the 8 merged peaks, including the *psbA* CLIP peak, which had the lowest p-value in the CP33B CLIP analysis (Figures 2C,D).

The correlation between the sequence preferences of CP29A and CP33B identified by RBNS and their respective *in vivo* binding sites revealed through CLIP provides a foundation for establishing a high-confidence set of binding sites.

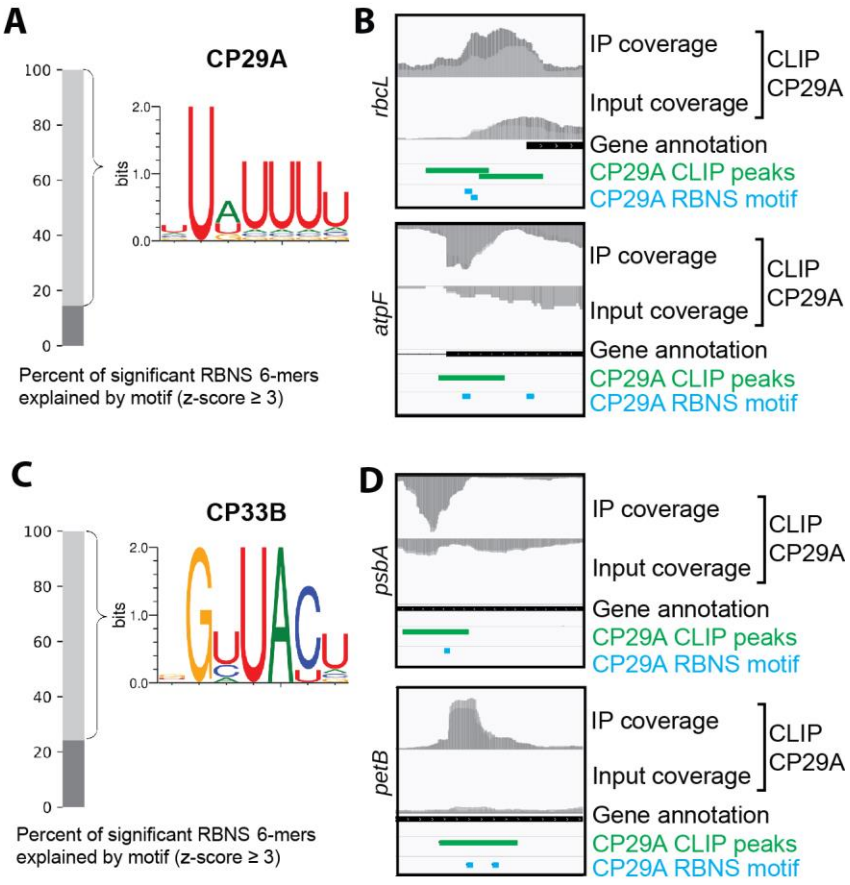

Figure 2: In vitro binding data of CP29A and CP33B complement *in vivo* CLIP data. Sequence preferences of CP29A and CP33B were analyzed using the RNA Bind-N-Seq protocol (RBNS; Dominguez et al., 2018). The *in vitro* approach relies on the enrichment of specific sequence elements from a random RNA input pool by purified, recombinant RNA-binding proteins. Significantly enriched 6-mers in RBNS data sets were used to construct a sequence logo, which represents most of the observed binding of (A) CP29A and (B) CP33B. Examples of overlapping RBNS motifs and CLIP peaks are shown for CP29A in *rbcL* and *atpF* (C) and for CP33B in *psbA* and *petB* (D). Coverage graphs (RPM) of normalized CLIP and size-matched input libraries are shown in addition.

CP29A binds adjacent to the predicted binding sites of the PPR protein MRL1 in the *rbcL* 5'-UTR

In our study, the primary binding site for CP29A was identified in the *rbcL* 5'UTR (Figure 1C, Table 1). A crucial factor known to influence the expression of *rbcL* is the pentatricopeptide repeat (PPR) protein MRL1 (Johnson et al., 2010). PPR proteins are known to produce short RNA fragments that are protected from nuclease degradation due to their tight binding (Ruwe and Schmitz-Linneweber, 2012; Ruwe et al., 2016; Pfalz et al., 2009). We previously discovered a short RNA fragment, designated as C28, which precisely matches the end of the mature, 5′-processed *rbcL* mRNA (Ruwe and Schmitz-Linneweber, 2012; Ruwe et al., 2016). The formation of the processed 5' end of the *rbcL* RNA is dependent on the PPR protein MRL1 (Johnson et al., 2010).

To investigate whether MRL1 is essential for the accumulation of the *rbcL* 5' sRNA/C28 sRNA, we conducted small RNA sequencing on the Arabidopsis *mrl1* mutant and compared it with wild-type (wt) plants. Our analysis revealed that C28 was the sole small RNA (sRNA) displaying differential accumulation across the plastid transcriptome (Figure 3A,B). This finding was further validated through RNA gel blot hybridization (Figure 3C). As controls, we also analyzed mutants deficient in the PPR protein HCF152 and the H-TRP protein HCF107. These mutants were observed to accumulate the *rbcL* 5' footprint, with a slight reduction in C28 signal (Figure 3C). We attribute this reduction to the photosynthetic defects and the resultant paleness in these two control mutants, which likely led to secondary effects on chloroplast RNA. Importantly, the accumulation of C28 in the control mutants demonstrates the specificity of the loss of this sRNA in *mrl1* mutants.

The binding site of MRL1 within the *rbcL* mRNA was hypothesized to be near its 5'-end (Johnson et al., 2010). Utilizing the PPR code (Barkan et al., 2012), we predicted the binding sequence of MRL1 based on the specific amino acids in each repeat of its PPR tract. This prediction yielded an ambiguous sequence that partially aligns with the 5'-*rbcL* sRNA, corroborating the region proposed by Johnson et al. (2010). However, the sRNA extends considerably beyond the length of the predicted MRL1 binding site. Interestingly, when analyzing the CP29A CLIP-data, we noticed that one peak is close to the MRL1-dependent C28 sRNA. eCLIP reads often terminate at the site of protein-RNA cross-link, thus giving base-specific resolution. We therefore mapped reads ends of the CLIP data and found a peak located downstream of the predicted MRL1 binding site, but still within the C28 sRNA (Figure 4). This alignment is particularly notable as the U-rich target sequence identified in the *rbcL* peak aligns with the consensus sequence determined by our RBNS analysis (as shown in Figure 2A, B and Figure 3). This suggests a potential functional interplay or overlap in the binding regions of MRL1 and CP29A within the *rbcL* mRNA, with CP29A binding downstream of the MRL1 site.

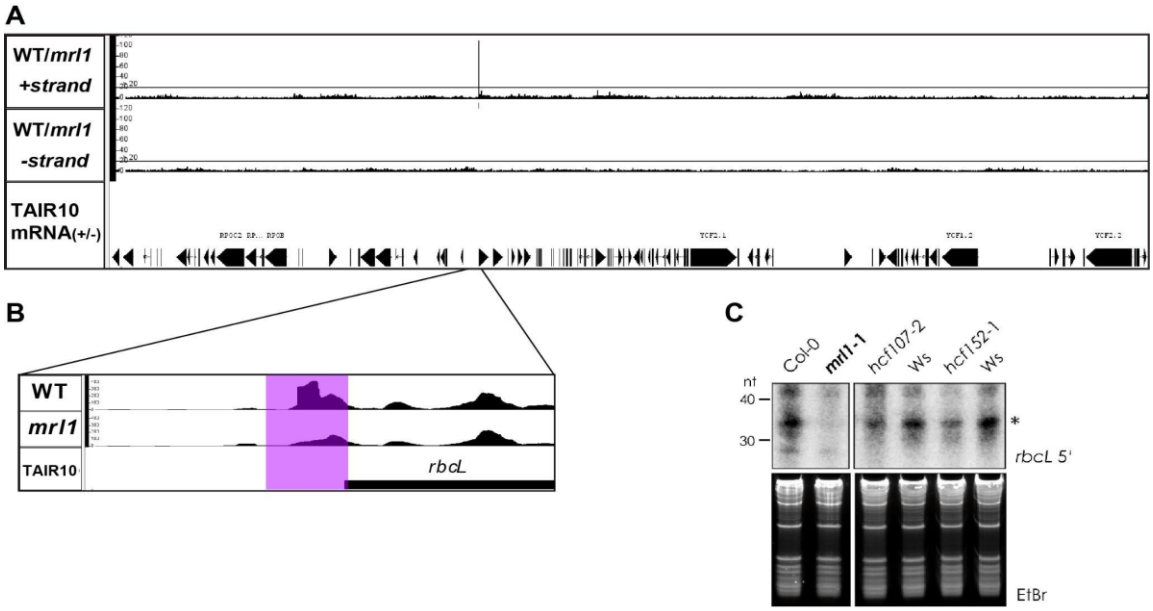

Figure 3: Identification of differential small RNA accumulation in *mrl1* mutants. (A) Small RNA libraries were prepared from three week old plants. 5' position of small RNA alignments were extracted from small RNA mappings using BEDTools (Quinlan and Hall 2010). Counts for every genome position were normalized to reads per million reads mapped to the chloroplast genome. Ratios of WT and mutant samples were calculated for positive and negative strand separately and visualized using the Integrated Genome Browser (Nicol et al. 2009). Ratios for the positive and negative strand are shown. (B) For only the sRNA C28 in the *rbcL* 5'-UTR, a high ratio of reads in wt versus *mrl1* was identified. An sRNAcoverage graph of the corresponding genomic region is shown here. The area of differential RNA accumulation is shaded in violet. Please note the sharp edges of the coverage in the center of the shading, which correspond to the 5' and 3'-end of sRNA C28. (C) 3µg total RNA from the genotypes indicated were separated on denaturing polyacrylamide gels, and transferred to nylon membranes. Small RNAs were detected with ^{32}P end-labeled oligo-nucleotides antisense to the small RNA sequence. The approximate sizes as compared to DNA oligonucleotides are indicated in nucleotides (nt). The *mrl1* mutant and the corresponding WT in the Col-0 background were grown on soil for three weeks. Mutants with a high chlorophyll fluores-cence (hcf) phenotype were grown for three weeks on MS-plates containing 3% sucrose and plants showing a pale phenotype were selected. Plants with WT phenotype were used as control (Ws). Both *hcf107*-2 and *hcf152*-1 are in the Wassilewskija (Ws) background. The ethidium bromide stain controls for equal loading. The asterisk marks the band overlapping with the size expected from small RNA sequencing.

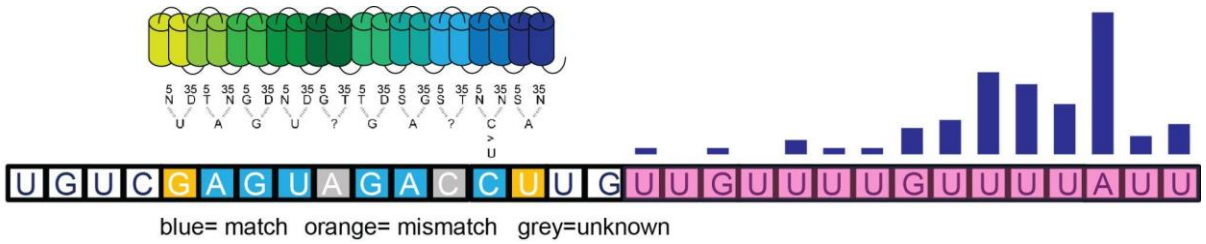

Figure 4: Mapping of the predicted MRL binding site, CP29A CLIP data and CP29A RBNS data onto the 5'-*rbcL* sRNA sequence. The sequence of the 5'-rbcL sRNA is shown with framed capital letters. MRL1 has 10 PPRs, which are schematically shown as an array of two joined helices. The key amino acids are shown below together with their predicted base target. Some amino acid combinations have not yet been assigned to a base (marked by question mark). Bases matching the prediction are marked in blue, those not matching in orange. CP29A CLIP-coverage data of read termini are shown as bar graph on top of each base in the sRNA sequence. The sequences matching the RBNS consensus shown in Figure 2 are highlighted in magenta.

Table 1: Top eCLIP peaks for AtCP29A.

Gene	*Start*[1]	*Stop*[1]	*Length*	*log_2 (IP/input)*[2]	*log_{10} p value*[3]	*Strand*[1]
rbcL 5'UTR-1	54945	54951	7	2,66	58,20	+
rbcL 5'UTR-2	54880	54951	72	3,51	57,61	+
trnH	19	46	28	3,19	23,06	-
psbL	63831	63881	51	3,85	17,13	-
psbB-1	73108	73144	37	2,70	14,97	+
psbT	74117	74161	45	4,20	12,68	+
atpF	11525	11561	37	4,11	11,74	-
psbB-2	73607	73616	10	3,29	11,56	+
psbD	33181	33188	8	2,66	11,51	+
petB	76214	76242	29	2,40	11,26	+

[1] genbank acc. no. NC_000932
[3] The number of eCLIP reads overlapping CLIPper-identified peaks and the number overlapping the identical genomic region in the paired size-matched Input sample were counted and used to calculate fold enrichment (normalized by total usable read counts in each data set)
[2] enrichment p-value of reads in the CLIP precipitated versus reads in the size-matched input calculated by Yates' Chi-Square test

Arabidopsis CP29A mutants do not show a change in *rbcL* mRNA levels nor is translation of the *rbcL* mRNA reduced

Considering the role of MRL1 in the stabilization and translation of *rbcL* (Johnson et al., 2010), we investigated the expression of *rbcL* in *cp29a* null mutants. Given CP29A's role for the cold acclimation process of chloroplast biogenesis, our analysis encompassed plants under both standard growth conditions and cold acclimation conditions. We treated the plants for three days with 12°C, conditions that do not yet lead to a photosynthetic or macroscopic phenotype, thus avoiding secondary effects due to failed chloroplast biogenesis (Legen et al. 2023). We observed no significant changes in the accumulation of *rbcL* mRNA in the *cp29a* null mutants under either condition, nor of the *psbA* control mRNA (Figure S3A). We next assayed accumulation of the RbcL protein by immunoblot analysis, but again saw no difference between wt and *cp29a* mutants (Figure S3B). RuBisCo is a particularly abundant and stable enzyme with a half-live of 7 days (Piques et al., 2009; Simpson, 1981). Thus smaller changes on the level of translation activity may not become visible after three days on the backdrop of the large pool of stable RubisCo protein. We therefore next tested the synthesis of chloroplast proteins by labeling with ^{35}S methionine. After three days of cold exposure, no significant alterations were detected in protein synthesis of either RbcL or D1 as a control protein in the chloroplasts of the *cp29a* null mutants (Figure S3C). These results suggest that the absence of CP29A does not significantly impact the accumulation or translation of *rbcL* mRNA in short-term cold treated *Arabidopsis* tissue.

CRIPRS-Cas9-induced null mutants of tobacco CP29A show photosynthetic deficiency in the cold

Arabidopsis thaliana, a cold-adapted species, thrives in temperate zones and can accumulate biomass even at temperatures as low as 4°C. We asked whether disrupting a relative of CP29A in a tropical plant, which does not often need to acclimate to cold, might yield similar outcomes. We focused on tobacco NtCP29A for this study. This choice was informed by evidence showing that NtCP29A localizes to the chloroplast, can bind *rbcL* and other RNAs *in vitro* and *in vivo* (Ye et al., 1991; Nakamura et al., 1999), and stabilize RNAs *in vitro* (Nakamura et al., 2001), mirroring some of the functions of AtCP29A in *A. thaliana*.

Tobacco is an allotetraploid formed from the hybridization of *N. sylvestris* (maternal) and *N. tomentosiformis* (paternal). We refer to the alleles from these species as *NtsCP29A* and *NttCP29A*, respectively. To generate null alleles, we utilized CRISPR/Cas9 mutagenesis with guide RNAs (gRNAs) designed to target all four alleles (Figure S4A). In the T1 generation, we successfully isolated plants deficient for either the *NtsCP29A* alleles or the *NttCP29A allele*, but not plants with deletions in both alleles of tobacco's parents. We self-fertilized these T1 lines and crossed the resulting homozygous *Ntsp29a* and *Nttcp29a* plants, followed by self-fertilization of their progeny. This process yielded a homozygous *Ntscp29a/Nttcp29a* line. PCR analysis confirmed the desired deletion in all alleles in this line (Figure S4B). Sequencing of the deletion alleles verified that the deletions commenced within the gRNA sequences (Figure S4C). We designate these plant lines as *Ntcp29a* mutants.

Upon standard growth conditions, *Ntcp29a* plants showed no color deficiencies or developmental defects, appearing phenotypically similar to wild type (Figure S4D) and exhibiting normal fertility. After exposure to low temperatures for 14 days, the *Ntcp29a* mutants still displayed no macroscopic defects. Given that *Atcp29a* mutants experience photosynthetic impairments under prolonged cold exposure (Legen et al., 2023), we evaluated the quantum yield efficiency of photosystem II in *Ntcp29a* mutants in both normal and low-temperature conditions. A slight but significant decrease in Fv/Fm values was observed even at standard growth temperatures (as shown in Figure 5A,B), worsening in the cold with a pronounced reduction in Fv/Fm in the mutant compared to wt (Figure 5A,B). This reduction mirrors observations in *Arabidopsis (Legen et al., 2023)*. In *Atcp29a* mutants, the decrease in Fv/Fm was however limited to the youngest leaf tissue (Legen et al., 2023, and Figure 5C) , whereas in *Ntcp29a* mutants, the entire leaf was affected (Figure 5A).

To verify that the reduction in photosynthetic efficiency was exclusively caused by the lack of *NtCP29A* expression, we performed a complementation test on the null mutant. Following the *Agrobacterium*-mediated introduction of the full-length genomic versions of both the *N.*

sylvestris and *N. tomentosiformis* alleles on a single T-DNA, we observed a recovery of photosynthetic activity (Figure 5A, B). The degree of recovery is varying in the independent complementation lines, which could be caused by different expression levels of the transgenes depending on the specific T-DNA integration sites. Overall, these analyses confirm that the tobacco *NtCP29A* is essential for maintaining full photosynthetic performance during cold acclimation.

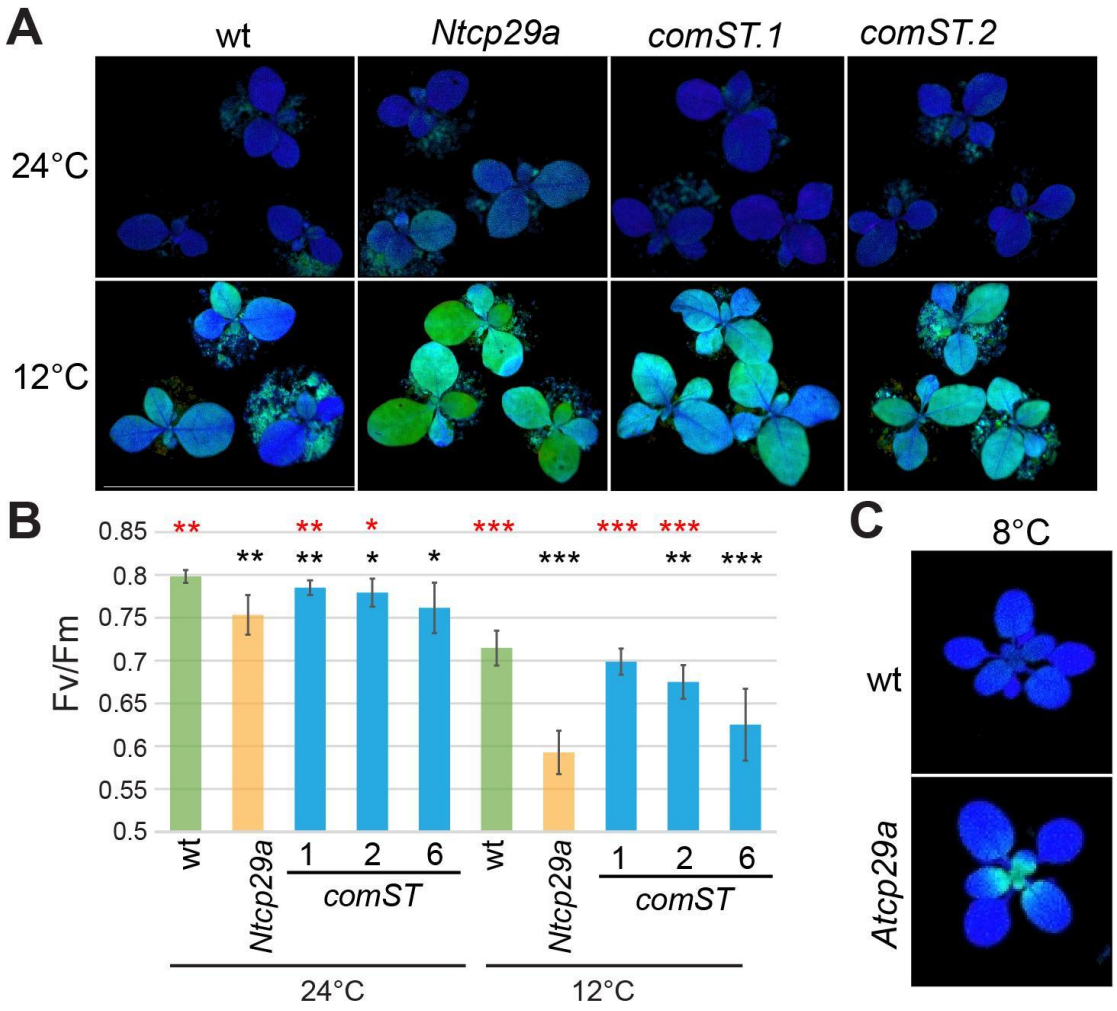

Figure 5: Chlorophyll a fluorescence analysis of *Ntcp29a* mutants. (A) Visualization of the maximum quantum yield of photosystem II (FV/FM) with an imaging PAM under normal growth conditions and after cold treatment. Wt, *Ntcp29a* mutants and three independent complementation lines (*comST*) were grown for 21 days at 24°C and then transferred to 12°C for ten days. (B) Fv/Fm values based on the measurements of eight individual plants for each plant line. Bars indicate the standard deviation and the asterisks represent the statistical significance (*p-value < 0,05; **p-value < 0,01; ***p-value < 0,001), as evaluated by Student's t-test. red asterisks = comparison to *Ntcp29a* mutants; black asterisks = comparison to wt. © Phenotype of Arabidopsis *Atcp29a* mutants in comparison to wt after 14 days at 21°C and 10 days at 8°C.

Tobacco CP29A mutants have reduced levels of *rbcL* mRNA and RbcL protein in the cold

To identify the molecular cause of the diminished photosynthetic performance in *Ntcp29a* mutants, we examined chloroplast RNA processing and accumulation. This analysis was conducted under normal growth conditions (21 days at 24°C) and after exposure to cold (21 days at 24°C followed by ten days at 12°C). Total RNA was extracted from one half of the fourth primary leaf, split along the middle vein, while the other half was reserved for protein analyses (details below). We performed Illumina-based RNA-seq on the extracted RNA after rRNA depletion, adhering to standard procedures.

Considering that *Atcp29a* mutants exhibited splicing defects after short-term cold exposure before visible bleaching (Legen et al., 2023), we investigated RNA splicing in the tobacco mutant. The chloroplast genome of tobacco contains 21 introns. To evaluate the splicing efficiency of these introns, we used the Chloro-Seq pipeline (Castandet et al., 2016). This method compares the number of reads spanning exon-exon junctions with those crossing intron-exon junctions. Our analysis enabled us to determine the splicing efficiencies for 14 mRNAs. Assessing introns in tRNAs was not feasible due to their short exons, which hindered the detection of reads spanning intron-exon boundaries. Our results showed no significant differences in splicing efficiency between the two genotypes under either normal or cold conditions (Figure S5A,B). Additionally, we quantified chloroplast RNA editing in wt and mutants under both standard and cold conditions using Chloro-Seq. This involved counting edited versus unedited reads. Here too, no significant changes were observed (Figure S5C,D).

Next, we investigated RNA accumulation differences between the two genotypes under normal and cold temperatures. To ensure accurate comparisons, we normalized the data for library size before conducting a Principal Component Analysis (PCA), representing all cellular transcripts. This analysis involved creating a scatterplot of the first two principal components. The resulting data showed good reproducibility, as evidenced by the close clustering of replicates (Figure 6A). Notably, the PCA revealed distinct patterns: the first principal component primarily distinguished between the two growth conditions, while the second principal component representing much lesser variance differentiated between the mutant and wild type (wt) samples (Figure 6B). This observation aligns with the expectation that cold significantly alters RNA pools in plant cells, leading to marked differences in transcript accumulation under the tested conditions. Furthermore, our analysis highlighted consistent differences in transcript abundance between the two genotypes.

We then examined chloroplast gene expression differences between *Ntcp29a* mutants and wild type (wt), focusing on the fold change in normalized RNA coverage. 94 chloroplast genes

met our detection criteria. Among these, only *rbcL* exhibited a significant decrease, reaching only 49% of wt levels, and this reduction occurred exclusively under cold conditions. While there were minor expression changes in several other genes, including some for eCLIP targets of CP29A, none reached our threshold for fold change. In summary, *Ntcp29a* mutants demonstrated a significant and specific reduction in *rbcL* mRNA at low temperatures.

To investigate whether this reduction in *rbcL* mRNA affects RbcL protein accumulation, we performed immunoblot analyses. In addition to using an antiserum against RbcL, we also examined the accumulation of subunits from Photosystem I (PsaB), Photosystem II (PsbO), the ATPase (AtpB), and the cytochrome b_6f complex (PetA). We normalized the signals by detecting actin protein on the same blots. The results revealed no significant changes in the accumulation of any proteins, except for RbcL, which showed a reduction to less than 50% of the levels observed in wt.

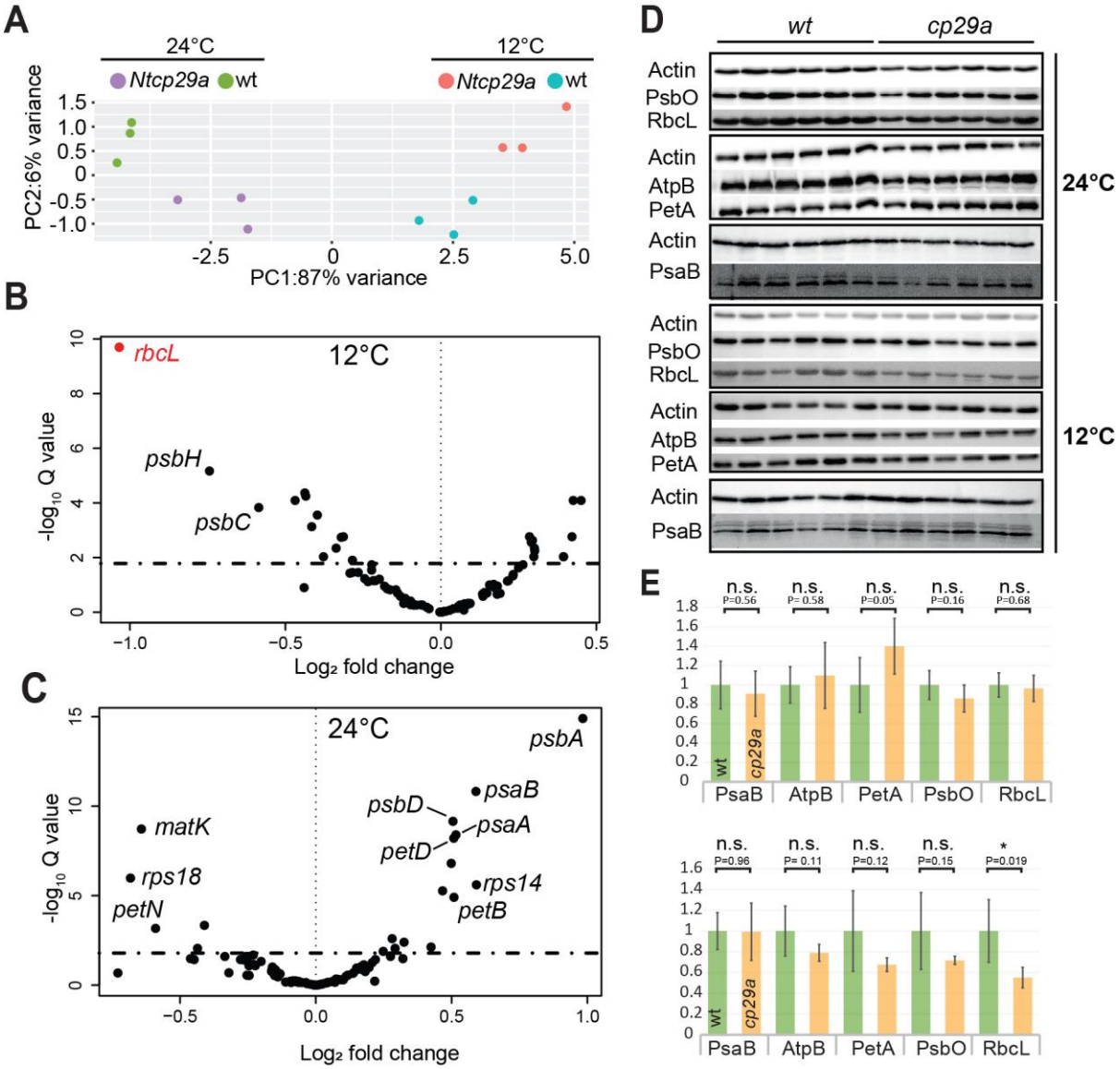

Figure 6: Analysis of *rbcL* expression in *Ntcp29a* mutants. (A) Principal component analysis of *Ntcp29a* mutant and wt leaf RNA-Seq samples (n=3). Plants were grown for 21 days at 24°C and then directly analyzed or transferred for 10 days to 12°C and then analyzed. (B) Volcano plot of RNA-Seq results for chloroplast RNAs from cold-treated mutant versus wt plants after cold treatment. Differential expression was analyzed using DESeq2 (Love et al., 2014). Only the red dot represents a significantly downregulated RNA, the *rbcL* mRNA ($p<0.01$; significance test: wald test adjusted with Benjamini Hochberg). (C) As in (B), but for plants grown at standard growth temperature, 24°C. No significant expression changes were detected. (D) Immunoblot analysis of leaf extracts of tobacco seedlings. Samples from plants analyzed for RNA accumulation in A-C were subject to immunological analyses. A total of six wt and six *Ntcp29a* mutant plants were analyzed. (E) Quantification of immunoblot analysis from (D). * = $p<0.05$; student's t-test.

Discussion

eCLIP reveals a post-transcriptional operon rich in PSII mRNAs for AtCP29A

We here adopt the eCLIP method to a chloroplast RBP which allows to demonstrate direct RNA interactions of a plant organellar RBP. CLIP-Seq has proven its value for RBPs in yeast and human mitochondria, but has not been applied to plant organelles (e.g. Hillen et al., 2021; Jones et al., 2019; Hornig-Do et al., 2014; Siira et al., 2017; Van Nostrand et al., 2020). We demonstrate that despite high pigment concentrations in chloroplasts UV-cross-linking is efficient enough to allow precipitation of RBP:RNA complexes. We further demonstrate that the identified peaks are highly specific, with no overlaps between the two related cpRNPs analyzed. This allowed us to assess the specific direct ligands of AtCP29A, while previous attempts using RIP could not distinguish between such direct RNA interactions and precipitations of larger RNP complexes possibly containing multiple proteins (Kupsch et al., 2012; Nakamura et al., 1999). The CLIP-based analysis corroborated previous findings, identifying CP29A as a chloroplast RNA-binding protein (RBP) with a multitude of target transcripts. This includes the reconfirmation of previously identified targets like *rbcL*, *psbD*, and *psbB*. Other previously identified targets such as *atpH*, *psaA*, and *atpB* were only moderately enriched in CP29A quantitative dot blots (Kupsch et al., 2012) and many of these genes were not identified in the current CLIP analysis. This discrepancy could be attributed to the absence of fixed cut-off criteria in the earlier RIP-Chip experiments, suggesting that these genes might represent minor targets of CP29A. They possibly did not meet the more stringent threshold criteria applied in the present CLIP experiment.

Many of these binding sites are in mRNAs coding for photosystem II subunits as well as for other photosynthetic electron transport components. For example, within the *psbB* operon, genes like *psbB*, *psbT*, *petB*, and *petD* were identified as targets. Similarly, in the *psbE* operon, *psbF*, *psbL*, and *psbJ* emerged as targets. This finding is particularly noteworthy because the input RNA in the CLIP-based approach undergoes fragmentation during cell lysis, and the resulting library is size-selected. This process allows for the differentiation of binding sites even within the same gene, as exemplified by the identification of distinct binding sites in the

5' UTR and exon 2 of *petB* (Figure S2). Additionally to polycistronic photosystem II transcripts, the functionally related monocistronic transcripts of *psbM*, *psbK*, and *petN* were also identified as targets. The preferential association of AtCP29A with these functionally related transcripts suggests the potential existence of a post-transcriptional operon, a concept previously described in various species (Keene, 2007). Another cpRNP, AtCP31A has been demonstrated to associate preferentially with mRNAs encoding subunits of the NDH complex (Lenzen et al., 2020), which are not targets of AtCP29A. AtCP33B with its preference for the *psbA* mRNA that is neither targeted by AtCP29A nor by AtCP31A has yet another non-overlapping target range. Therefore, the data presented here suggest that cpRNPs might serve as operon-defining chloroplast RNA-binding proteins, indicating a sophisticated level of post-transcriptional regulation within the chloroplast.

While our study identified an intriguing target profile for AtCP29A, most of these mRNAs did not exhibit significant accumulation changes in previous RNA-Seq analyses conducted on *Atcp29a* mutants (Legen et al., 2023), nor in the RNA-Seq analysis of *Ntcp29a* mutants. Despite this, the possibility of other impacts, such as effects on translation or more nuanced RNA processing alterations like end maturation in *cp29a* mutants, cannot be ruled out and warrants further investigation. Notably, most of the AtCP29A binding sites identified in this study are located within coding regions. This placement hints at a potential role in the translation process. Supporting this idea, a Ribo-Seq analysis indicated a reduction in translation under cold conditions in *Atcp29a* mutants (Legen et al., 2023). Thus, while the direct impact of CP29A binding on mRNA accumulation appears limited, its potential role in regulating translation or in other subtle aspects of RNA processing in *cp29a* mutants presents a compelling avenue for further research, particularly in understanding the response to cold stress in plants.

Sequence Preferences of AtCP29A

The RBNS method was instrumental in revealing the sequence preferences of cpRNPs, which were previously considered to be relatively nonspecific RBPs (Li and Sugiura, 1991). This approach demonstrated that the most enriched k-mers were closely related across all libraries associated with a specific cpRNP but not in the zero-protein control. This finding validates the specificity of the RBNS approach when using recombinant cpRNP proteins. The identification of distinct target sequence motifs for CP29A and CP33B indicates that, although cpRNPs bind to many chloroplast RNAs, their binding is specific and not merely due to a general affinity for RNA. The discovery of a polyU motif for Arabidopsis CP29A aligns with previous research on tobacco homologs CP29A and CP29B, which showed the ability to bind U homopolymers even in high ionic strength buffers (Ye and Sugiura, 1992). In addition to the enrichment of pure U k-mers in ACP29A RBNS data, U-rich k-mers with interspersed A/G/C nucleotides were

similarly enriched. The extended glycine-rich linker domain in CP29A might contribute to its structural flexibility, enabling the binding of non-consecutive motifs. Such a binding pattern has been observed in several RBPs with RRM domains (Loughlin et al., 2019; Dorn et al., 2017).

Extensive profiling of numerous human RNA-binding proteins (RBPs) has revealed a pronounced preference for target motifs with low compositional complexity (Dominguez et al., 2018). Given that chloroplast genomes typically have low GC-content, the predominance of simple U or A binding motifs in chloroplast RBPs might represent an evolutionary adaptation. Such a motif could be functionally significant to co-regulate multiple transcripts post-transcriptionally. In contrast to CP29A, CP33B binds to a more complex motif, GHUAUY. This complexity in the binding motif might partly explain why CP33B, unlike CP29A and CP31A, binds to fewer target RNAs *in vivo*. This distinction highlights the diversity in binding preferences among chloroplast RBPs and suggests that the complexity of binding motifs may influence the range of target RNAs an RBP can regulate *in vivo*.

NtCP29A supports *rbcL* mRNA accumulation in the cold in tobacco

RNAs often interact with multiple RBPs that work together to regulate the expression or repression of their target transcripts. Many RBPs with RRM domains, such as twin RRM-proteins in the heterogeneous nuclear RNP (hnRNP) class, function in coordination with other RBPs (Han et al., 2010). In this context, it was observed that a high-confidence CLIP binding site on the *rbcL* mRNA in chloroplasts is located near a region proposed to be targeted by the MRL1 protein. Previous genetic analyses of *mrl1* mutants in *Chlamydomonas reinhardtii* and *Arabidopsis (Johnson et al., 2010)*, along with predictions of MRL1-binding sites and identification of the MRL1-dependent accumulation of the *rbcL* sRNA, suggest that MRL1 binds to the 5'-end of the *rbcL* mRNA, adjacent to the AtCP29A binding site. In tobacco, it was noted that the *rbcL* mRNA level decreases by about 50% in cold conditions, a change paralleled by a similar reduction in the RbcL protein. This suggests that mRNA levels directly limit protein levels under such conditions, positioning CP29A as a potentially critical factor for RuBisCo expression in the cold. Given CP29A's location downstream of the MRL1 binding site, it is hypothesized that CP29A might assist MRL1 in stabilizing the *rbcL* mRNA against exonucleolytic degradation (Johnson et al., 2010), possibly by aiding MRL1 to bind its target sequence. PPR proteins like MRL1 typically prefer single-stranded RNA targets, and secondary structures in 5'-UTRs are known to adversely affect the translation of chloroplast mRNAs (Prikryl et al., 2011; Hammani et al., 2012). Low temperatures tend to stabilize RNA secondary structures, leading to speculation that CP29A may facilitate the association of MRL1 with the *rbcL* 5'-UTR in the presence of potentially detrimental RNA structures. This

hypothesis can be further investigated through structural analysis of the *rbcL* mRNA under various temperature conditions and in *cp29a* mutants.

The anticipated outcome of a 50% reduction in RbcL, and consequently in RuBisCo, is a decrease in carbon fixation. Mutants with alterations in the catalytic center of RuBisCo (center is encoded by *rbcL*), as well as *rbcL* deletion mutants, exhibit pronounced photosynthetic deficiencies and associated growth defects (Whitney et al., 1999; Allahverdiyeva et al., 2005). Hypomorphic RuBisCo mutants caused by lowered RbcS levels, show reduced Fv/Fm values similar to the *Ntcp29a* plants described here, though this occurs even at normal temperatures (Donovan et al., 2020). When comparing plants on the basis of leaf area, those grown in low temperatures typically exhibit higher concentrations and activities of Rubisco and other enzymes involved in photosynthetic carbon metabolism than plants grown at higher temperatures (Badger et al., 1982; Holaday et al., 1992; Hurry et al., 1995; Strand et al., 1999). This is supported in Arabidopsis by an increase of *rbcL* mRNA in the cold and a switch of nuclear-encoded RbcS gene isoforms (Donovan et al., 2020; Cavanagh et al., 2023). Since *rbcL* is the rate-limiting factor for RuBisCo assembly (Wostrikoff and Stern, 2007), *rbcL* expression needs to be sustained or possibly increased in cold environments. The basis for cold-driven expression of RuBisCo were so far unclear, but our data suggest that NtCP29A plays a role here at least in tobacco.

Materials and Methods

Plant Material and Growth

Arabidopsis *cp29a-6* mutants were described previously (Figure 1D; Kupsch et al., 2012). Arabidopsis seeds were stratified at 4°C in the dark for 3 days and then grown at 21°C with a light intensity of 120 µmol $m^{-2}s^{-1}$. For cold treatment, plants were grown for 14 days at 21°C and then transferred to 8°C for 10 days. Tobacco plants were grown at 24°C with a light intensity of 300 µmol $m^{-2}s^{-1}$. For cold treatment, plants were grown for 21 days at 24°C and then transferred to 12°C for 7 days

Chlorophyll fluorescence analysis

Chlorophyll a fluorescence *in vivo* was measured using the Imaging PAM chlorophyll fluorimeter (M-Series; Walz, Effeltrich, Germany). Prior to measurement, plants were dark-adapted for 15 minutes, following standard protocols (Klughammer and Schreiber, 2008).

Vector construction

CRISPR/Cas vector

Vector constructs were assembled as described (Yan et al., 2016).

Complementation constructs

The genomic DNA regions of NtsCP29A and NttCP29A including the UTRs and the promoter regions were amplified with gene-specific primers. The PCR products were cloned together with a spacer region into the pB7WG vector (Lukan et al. 2018), which was then used in the tobacco transformation.

Tobacco transformation

Wild-type tobacco plants (Petit havana) were transformed using the leaf disc transformation method (Horsch et al. 1989). For this purpose, the tobacco plants were first cultivated for 6-8 weeks on MS medium under sterile conditions. True leaves were then cut into discs of about 1 cm² and incubated with *Agrobacterium tumefaciens* in an MS solution containing glucose and acetosyringone for 30 minutes in the dark. The discs were then placed upside down on a solid MS medium and incubated at 28°C for 2 days. For callus and shoot induction, the discs were transferred to a solid MS medium containing glucose, 6-benzylaminopurine, 1-naphthalene acetic acid, cefotaxime and the corresponding antibiotic. As soon as the sprouts had formed roots, the plantlets were placed in the soil.

RNA gel blot hybridization

Small RNA gel blot analysis was performed as described in Loizeau et al. (2012) using an oligonucleotide probe with the sequence GCAATAAAACAAAACAACAAGGTCTACTCGACA. Standard RNA gel blot hybridization were carried out as described in Kupsch et al., 2012. RNA probes were prepared by *in vitro* transcription in the presence of 5-azido-C3-UTP nucleotides (Jena Biosciences) with final linking to sulpho-cyanine cy5.5 or cy7.5 dye (Lumiprobe) according to the manufacturers protocol. Primer used for amplifying the probes are as following: CK_rbcL_for (GCAGCATTCCGAGTAACTCC) and CK_rbcL_T7 (GTAATACGACTCACTATAGGGCCACGTAGACATTCATAAACTGC) resulting in a fragment of 473bp covering the N terminal part of *rbcL* mRNA. The *psbA* probe was generated with the use of psbA.T7 primer (GTAATCGACTCACTATAGGGATTCCTAGAGGCATACCATCAG) and psbAfw primer (GAAAGCGAAAGCCTATGGGG) resulting in a PCR fragment of 503bp.

Small RNA sequencing

Small RNA sequencing and bioinformatic analysis were done as described earlier (Ruwe et al. 2019).

RNA Seq analysis

RNA Seq data were processed using the nf-core/rnaseq (https://github.com/nf-core/rnaseq/) pipeline for quality control, trimming and mapping. Reads were only mapped to the tobacco chloroplast (Nicotiana tabacum plastid, NC_001879.2). Reads were counted using featureCounts (Rsubread, v2.0.6) with the parameter Min overlaping bases: 50. Transcripts with <50 counts in at least one sample as well as rRNAs and tRNAs were removed from further

analysis. Differentially expression analysis was done with DESeq2. Splicing and editing efficiencies were calculated using the chloroseq pipeline (Castandet et al., 2016).

eCLIP analysis

In this experiment, approximately 4×10^9 isolated chloroplasts were resuspended in RB-buffer and placed onto a glass petri dish. The chloroplasts were then exposed to 500 mJ/cm^2 of UV light at a wavelength of 254 nm while being kept on ice. After UV exposure, the chloroplasts were centrifuged at 500 × g for 5 minutes at 4 °C, forming a pellet. The supernatant was discarded, and the resulting pellet was flash-frozen in liquid nitrogen and stored at -80°C. For further processing, the chloroplast pellets were thawed on ice and resuspended in 2 ml of CLIP lysis-buffer each. The resulting lysates were then centrifuged at 20,000 × g at 4°C for 30 minutes. Specific antibodies targeting the RNA-binding protein (RBP) of interest (10 µl for both anti-CP29A and anti-CP33B) were attached to 50 µl of Dynabeads Protein G (Invitrogen) and then resuspended in 500 µl of CO-IP buffer. The supernatant from the chloroplast lysate pellet was transferred to a new tube, combined with 500 µl of the prepared bead suspension, and 4 units of Turbo DNase (Invitrogen). This mixture was incubated for 1 hour at 4°C with rotation. After incubation, 5% of the bead-lysate mixture was set aside for preparing size-matched input libraries and for western blot analysis of the immunoprecipitation. Finally, the supernatant was removed from the post-immunoprecipitation mixture and retained for western blot analysis.

The library preparation protocol for this study closely followed the established eCLIP protocol (Van Nostrand et al., 2016). Initially, the beads were washed twice with high-salt CLIP washing buffer, once with CLIP washing buffer, and once with FastAP buffer. On-bead RNA dephosphorylation was performed by incubating with FastAP at 37°C for 15 minutes, followed by a 20-minute incubation with T4 polynucleotide kinase at the same temperature. Turbo DNase was added to both dephosphorylation steps. Subsequently, the beads underwent two washes in CLIP washing buffer and one in T4 RNA ligase 1 buffer (omitting DTT). A color-balanced pair of two 3'-CLIP-RNA adapters was ligated on-bead to the RNAs in each sample using T4 RNA ligase 1 (New England Biolabs), with the ligation reaction incubated for 75 minutes at 21°C. Post-ligation, the beads were washed three times in CLIP washing buffer. Both the crosslinked, adapter-ligated protein-RNA complexes and the previously set aside input samples were size-separated via polyacrylamide gel electrophoresis and transferred to nitrocellulose membranes. The RNA, positioned 75 kDa above the RBP of interest, was released from the nitrocellulose membrane using Proteinase K (New England Biolabs) digestion. The released RNA of the size-matched input samples was dephosphorylated and adapter-ligated, similar to the IP samples, but using a single 3'-input-RNA adapter. Both the IP and size-matched input samples' adapter-ligated RNA was reverse transcribed using SuperScript II Reverse Transcriptase (Invitrogen) at 42°C for 45 minutes. Following ExoSAP-

IT (Applied Biosystems) treatment and chemical hydrolysis of residual RNA, cDNA was recovered using MyONE Silane beads (ThermoFisher Scientific). A 5'-DNA adapter, including unique molecular identifiers, was ligated to all cDNA samples using T4 RNA ligase 1 (New England Biolabs). This ligated cDNA was purified again using MyONE Silane beads (ThermoFisher Scientific). Quantification of the cDNA samples was done via qPCR, followed by PCR amplification using Q5 High-Fidelity DNA Polymerase (New England Biolabs) and indexed primers for Illumina sequencing. The resulting libraries were purified using the GeneJET PCR Purification Kit (ThermoFisher Scientific) and separated on 6% polyacrylamide gels. Library fragments ranging from 170 bp to 350 bp were extracted and subjected to Illumina sequencing.

RB-buffer: 0.3 M sorbitol, 20 mM tricine-KOH (pH 8.4), 2.5 mM EDTA, 5 mM MgCl2
CLIP lysis-buffer: 50 mM Tris-HCl (pH 7.4), 100 mM NaCl, 0.5% Nonidet P-40, 0.5% sodium deoxycholate, 1xcOmpleteTM EDTA-free Protease Inhibitor Cocktail (Roche)
CO-IP-buffer: 150 mM NaCl, 20 mM Tris-HCl pH (7.5), 2 mM MgCl2, 5 µg/mL aprotinin, 0.5% Nonidet P-40
High salt CLIP washing-buffer: 1000 mM NaCl, 20 mM Tris-HCl (pH 7.4), 1 mM EDTA, 0.5% Nonidet P-40
CLIP washing buffer: 20 mM Tris-HCl (pH 7.4), 10 mM MgCl2, 0.5% Nonidet P-40
FastAP buffer: 10 mM Tris-HCl (pH 7.5), 5 mM MgCl2, 100 KCl, 0.02% Triton X-100
T4 RNA ligase 1 buffer (without DTT): 50 mM Tris-HCl (pH 7.5), 10 mM MgCl2

Computational analysis of CLIP data

The analysis of the CLIP libraries and corresponding size-matched input controls was conducted using the CWL-based and dockerized eCLIP pipeline version 0.3.99 (available at github.com/YeoLab/eclip/releases/tag/v0.3.99). The general methodology followed the approach described previously (Van Nostrand et al., 2016), with some minor modifications. A custom docker container was created, incorporating an annotation of the *Arabidopsis thaliana* chloroplast chromosome for use with the CLIPPER tool. This custom annotation was derived from the AtRTD2 annotation, enhanced by extending all chloroplast coding sequences by 100 base pairs at both the 5'- and 3'-UTR ends. The criteria for identifying significant peaks in the analysis were stringent: a minimum of 4-fold enrichment of IP over the input control was required, along with an adjusted p-value of ≤ 0.001. Additionally, peaks were only considered significant if they were identified as such in both biological replicates.

RBNS analysis

We slightly modified the RNA Bind-n-Seq protocol (Lambert et al., 2014). AtCP29A and AtCP33B were tagged with N-terminal streptavidin-binding protein (SBP) and produced using

the pGEX system. Their quality and quantity were verified by SDS-PAGE and a protein standard dilution series. The *in vitro* transcribed RNA pool (0.5 µM) was then mixed with various concentrations of these RBPs (0, 10, 100 and 1000 µM) and incubated for 3-hour at 21°C. Magnetic Dynabeads MyOne Streptavidin C1 were then added, followed by another hour of incubation. The protein-RNA complexes were magnetically separated and washed twice. RNA was eluted in SDS-containing buffer at 70°C for 10 minutes and purified using the RNA Clean & Concentrator-5 kit. The RNA was reverse transcribed using ProtoScript II Reverse Transcriptase, and 0.5 pm of the RNA pool was also reverse transcribed. cDNA libraries were amplified with NEXTflex® Unique Dual Index Barcodes and Q5® High-Fidelity DNA Polymerase, then size-selected and purified via gel electrophoresis for sequencing on the Illumina NextSeq500 platform. Binding-buffer composition was 25 mM Tris-HCl (pH 7.5), 150 mM KCl, 3 mM MgCl2, 0.01% Tween, and 1 mM DTT, 1 mg/mL BSA; Washing-buffer had 25 mM Tris-HCl (pH 7.5), 150 mM KCl, 0.5 mM EDTA, 0.01% Tween, 60 µg/mL BSA; Elution-buffer contained 10 mM Tris-HCl (pH 7.0), 1 mM EDTA, 1% SDS.

Data analysis RNA Bind-N-Seq libraries was performed using a published computational workflow (Dominguez et al., 2018, bitbucket.org/pfreese/rbns_pipeline). The kmer enrichment analysis was performed for k=6. We also checked 6, 7, 8, and 9 mers. All resulted in very similar variations of the 6mer motifs, i.e. the polyU one for CP29A and the more complex one for CP33B. We also checked for gapped motifs, which resulted in the same pattern.

Immunoblot analysis

Total protein preparations were loaded based on equal fresh weight. The proteins were separated by SDS-PAGE and transferred to a PVDF membrane. Protein integrity and loading were tested by Ponceau S staining. Hybridization with the primary antibody was carried out overnight at 4°C and with the second antibody for 1 hour at room temperature. Stripping of antibodies was done with 1.5% glycine, 0.1% SDS, 1 & Tween 20, pH = 2.2.

***In Vivo* Pulse-Chase Labeling of Chloroplast-Encoded Proteins**

Young leaves from 21 day old Arabidopsis plants were used for 35S methionine labeling of proteins as previously described (Meurer et al., 1998).

Data availability

Next generation sequencing data associated with this manuscript can be found at SRA, BioProject PRJNA1043672

Acknowledgements

The vectors for CRISPR/Cas9 mutagenesis were kindly provided by Kerstin Kaufmann (Humboldt Universität zu Berlin). We thank DFG TR175 projects A02 to CSL and A07 to HR for funding.

References

Allahverdiyeva, Y., Mamedov, F., Mäenpää, P., Vass, I., and Aro, E.-M. (2005). Modulation of photosynthetic electron transport in the absence of terminal electron acceptors: characterization of the rbcL deletion mutant of tobacco. Biochim. Biophys. Acta **1709**: 69–83.

Badger, M.R., Björkman, O., and Armond, P.A. (1982). An analysis of photosynthetic response and adaptation to temperature in higher plants: temperature acclimation in the desert evergreen Nerium oleander L*. Plant Cell Environ. **5**: 85–99.

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.

Castandet, B., Hotto, A.M., Strickler, S.R., and Stern, D.B. (2016). ChloroSeq, an Optimized Chloroplast RNA-Seq Bioinformatic Pipeline, Reveals Remodeling of the Organellar Transcriptome Under Heat Stress. G3 **6**: 2817–2827.

Cavanagh, A.P., Slattery, R., and Kubien, D.S. (2023). Temperature-induced changes in Arabidopsis Rubisco activity and isoform expression. J. Exp. Bot. **74**: 651–663.

Crosatti, C., Rizza, F., Badeck, F.W., Mazzucotelli, E., and Cattivelli, L. (2013). Harden the chloroplast to protect the plant. Physiol. Plant. **147**: 55–63.

Dominguez, D. et al. (2018). Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. Mol. Cell **70**: 854–867.e9.

Donovan, S., Mao, Y., Orr, D.J., Carmo-Silva, E., and McCormick, A.J. (2020). CRISPR-Cas9-Mediated Mutagenesis of the Rubisco Small Subunit Family in Nicotiana tabacum. Front Genome Ed **2**: 605614.

Dorn, G., Leitner, A., Boudet, J., Campagne, S., von Schroetter, C., Moursy, A., Aebersold, R., and Allain, F.H.-T. (2017). Structural modeling of protein-RNA complexes using crosslinking of segmentally isotope-labeled RNA and MS/MS. Nat. Methods **14**: 487–490.

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

Han, S.P., Tang, Y.H., and Smith, R. (2010). Functional diversity of the hnRNPs: past, present and perspectives. Biochem. J **430**: 379–392.

Hillen, H.S., Markov, D.A., Wojtas, I.D., Hofmann, K.B., Lidschreiber, M., Cowan, A.T., Jones, J.L., Temiakov, D., Cramer, P., and Anikin, M. (2021). The pentatricopeptide repeat protein Rmd9 recognizes the dodecameric element in the 3'-UTRs of yeast mitochondrial mRNAs. Proc. Natl. Acad. Sci. U. S. A. **118**.

Holaday, A.S., Martindale, W., Alred, R., Brooks, A.L., and Leegood, R.C. (1992). Changes in Activities of Enzymes of Carbon Metabolism in Leaves during Exposure of Plants to Low Temperature. Plant Physiol. **98**: 1105–1114.

Hornig-Do, H.T., Montanari, A., Rozanska, A., Tuppen, H.A., Almalki, A.A., Abg-Kamaludin, D.P., Frontali, L., Francisci, S., Lightowlers, R.N., and Chrzanowska-Lightowlers, Z.M. (2014). Human mitochondrial leucyl tRNA synthetase can suppress non cognate pathogenic mt-tRNA mutations. EMBO Mol. Med. **6**: 183–193.

Hurry, V.M., Keerberg, O., Pärnik, T., Gardeström, P., and Öquist, G. (1995). Cold-hardening results in increased activity of enzymes involved in carbon metabolism in leaves of winter rye (Secale cereale L.). Planta **195**: 554–562.

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. Plant Cell **22**: 234–248.

Jones, J.L., Hofmann, K.B., Cowan, A.T., Temiakov, D., Cramer, P., and Anikin, M. (2019). Yeast mitochondrial protein Pet111p binds directly to two distinct targets in COX2 mRNA, suggesting a mechanism of translational activation. J. Biol. Chem. **294**: 7528–7536.

Keene, J.D. (2007). RNA regulons: coordination of post-transcriptional events. Nat. Rev. Genet. **8**: 533–543.

Klughammer, C. and Schreiber, U. (2008). Complementary PS II quantum yields calculated from simple fluorescence parameters measured by PAM fluorometry and the Saturation Pulse method. PAM application notes **1**: 201–247.

Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2011). iCLIP--transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution. J. Vis. Exp.

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

Lambert, N., Robertson, A., Jangi, M., McGeary, S., Sharp, P.A., and Burge, C.B. (2014). RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. Mol. Cell **54**: 887–900.

Legen, J., Lenzen, B., Kachariya, N., Feltgen, S., Gao, Y., Mergenthal, S., Weber, W., Klotzsch, E., Zoschke, R., Sattler, M., and Schmitz-Linneweber, C. (2023). The prion-like domain of the chloroplast RNA binding protein CP29A is required for cold-induced phase separation next to nucleoids and supports RNA splicing and translation during cold acclimation. bioRxiv.

Lenzen, B., Rühle, T., Lehniger, M.-K., Okuzaki, A., Labs, M., Muino, J.M., Ohler, U., Leister, D., and Schmitz-Linneweber, C. (2020). The Chloroplast RNA Binding Protein CP31A Has a Preference for mRNAs Encoding the Subunits of the Chloroplast NAD(P)H Dehydrogenase Complex and Is Required for Their Accumulation. Int. J. Mol. Sci. **21**.

Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., Darnell, J.C., and Darnell, R.B. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature **456**: 464–469.

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

Li, Y.Q. and Sugiura, M. (1990). Three distinct ribonucleoproteins from tobacco chloroplasts: each contains a unique amino terminal acidic domain and two ribonucleoprotein consensus motifs. EMBO J. **9**: 3059–3066.

Loughlin, F.E. et al. (2019). The Solution Structure of FUS Bound to RNA Reveals a Bipartite Mode of RNA Recognition with Both Sequence and Shape Specificity. Mol. Cell **73**: 490–504.e6.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. **15**: 550.

Maris, C., Dominguez, C., and Allain, F.H.-T. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. FEBS J. **272**: 2118–2131.

McLeay, R.C. and Bailey, T.L. (2010). Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. BMC Bioinformatics **11**: 165.

Meurer, J., Plücken, H., Kowallik, K.V., and Westhoff, P. (1998). A nuclear-encoded protein of prokaryotic origin is essential for the stability of photosystem II in Arabidopsis thaliana. EMBO J. **17**: 5286–5297.

Meyer, K., Köster, T., Nolte, C., Weinholdt, C., Lewinski, M., Grosse, I., and Staiger, D. (2017). Adaptation of iCLIP to plants determines the binding landscape of the clock-regulated RNA-binding protein AtGRP7. Genome Biol. **18**: 204.

Nakamura, T., Ohta, M., Sugiura, M., and Sugita, M. (1999). Chloroplast ribonucleoproteins are associated with both mRNAs and intron-containing precursor tRNAs. FEBS Lett. **460**: 437–441.

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.

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

Piques, M., Schulze, W.X., Höhne, M., Usadel, B., Gibon, Y., Rohwer, J., and Stitt, M. (2009). Ribosome and transcript copy numbers, polysome occupancy and enzyme dynamics in

Arabidopsis. Mol. Syst. Biol. **5**: 314.

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

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

Ruwe, H., Gutmann, B., Schmitz-Linneweber, C., Small, I., and Kindgren, P. (2019). The E domain of CRR2 participates in sequence-specific recognition of RNA in plastids. New Phytol. **222**: 218–229.

Ruwe, H., Kupsch, C., Teubner, M., and Schmitz-Linneweber, C. (2011). The RNA-recognition motif in chloroplasts. J. Plant Physiol. **168**: 1361–1371.

Ruwe, H. and Schmitz-Linneweber, C. (2012). Short non-coding RNA fragments accumulating in chloroplasts: footprints of RNA binding proteins? Nucleic Acids Res. **40**: 3106–3116.

Ruwe, H., Wang, G., Gusewski, S., and Schmitz-Linneweber, C. (2016). Systematic analysis of plant mitochondrial and chloroplast small RNAs suggests organelle-specific mRNA stabilization mechanisms. Nucleic Acids Res. **44**: 7406–7417.

Sasaki, T., Yukawa, Y., Miyamoto, T., Obokata, J., and Sugiura, M. (2003). Identification of RNA editing sites in chloroplast transcripts from the maternal and paternal progenitors of tobacco (Nicotiana tabacum): comparative analysis shows the involvement of distinct trans-factors for ndhB editing. Mol. Biol. Evol. **20**: 1028–1035.

Siira, S.J., Spåhr, H., Shearwood, A.-M.J., Ruzzenente, B., Larsson, N.-G., Rackham, O., and Filipovska, A. (2017). LRPPRC-mediated folding of the mitochondrial transcriptome. Nat. Commun. **8**: 1532.

Simpson, E. (1981). Measurement of Protein Degradation in Leaves of Zea mays Using [H]Acetic Anhydride and Tritiated Water. Plant Physiol. **67**: 1214–1219.

Small, I., Melonek, J., Bohne, A.-V., Nickelsen, J., and Schmitz-Linneweber, C. (2023). Plant organellar RNA maturation. Plant Cell **35**: 1727–1751.

Strand, A., Hurry, V., Henkes, S., Huner, N., Gustafsson, P., Gardeström, P., and Stitt, M. (1999). Acclimation of Arabidopsis leaves developing at low temperatures. Increasing cytoplasmic volume accompanies increased activities of enzymes in the Calvin cycle and in the sucrose-biosynthesis pathway. Plant Physiol. **119**: 1387–1398.

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

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 **9**.

Van Nostrand, E.L. et al. (2020). A large-scale binding and functional map of human RNA-binding proteins. Nature **583**: 711–719.

Van Nostrand, E.L. et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat. Methods **13**: 508–514.

Van Nostrand, E.L., Nguyen, T.B., Gelboin-Burkhart, C., Wang, R., Blue, S.M., Pratt, G.A., Louie, A.L., and Yeo, G.W. (2017). Robust, Cost-Effective Profiling of RNA Binding Protein Targets with Single-end Enhanced Crosslinking and Immunoprecipitation (seCLIP). Methods Mol. Biol. **1648**: 177–200.

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. Plant J.

Whitney, S.M., von Caemmerer, S., Hudson, G.S., and Andrews, T.J. (1999). Directed mutation of the Rubisco large subunit of tobacco influences photorespiration and growth. Plant Physiol. **121**: 579–588.

Wostrikoff, K. and Stern, D. (2007). Rubisco large-subunit translation is autoregulated in response to its assembly state in tobacco chloroplasts. Proc. Natl. Acad. Sci. U. S. A. **104**: 6466–6471.

Yan, W., Chen, D., and Kaufmann, K. (2016). Efficient multiplex mutagenesis by RNA-guided Cas9

and its use in the characterization of regulatory elements in the AGAMOUS gene. Plant Methods **12**: 23.

Ye, L.H., Li, Y.Q., Fukami-Kobayashi, K., Go, M., Konishi, T., Watanabe, A., and Sugiura, M. (1991). Diversity of a ribonucleoprotein family in tobacco chloroplasts: two new chloroplast ribonucleoproteins and a phylogenetic tree of ten chloroplast RNA-binding domains. Nucleic Acids Res. **19**: 6485–6490.

Ye, L. and Sugiura, M. (1992). Domains required for nucleic acid binding activities in chloroplast ribonucleoproteins. Nucleic Acids Res. **20**: 6275–6279.

Zhang, Y., Gu, L., Hou, Y., Wang, L., Deng, X., Hang, R., Chen, D., Zhang, X., Zhang, Y., Liu, C., and Cao, X. (2015). Integrative genome-wide analysis reveals HLP1, a novel RNA-binding protein, regulates plant flowering by targeting alternative polyadenylation. Cell Res. **25**: 864–876.

Zhang, Z., Boonen, K., Ferrari, P., Schoofs, L., Janssens, E., van Noort, V., Rolland, F., and Geuten, K. (2016). UV crosslinked mRNA-binding proteins captured from leaf mesophyll protoplasts. Plant Methods **12**: 42.