RESEARCH ARTICLE

Dynamics of Chloroplast Translation during Chloroplast Differentiation in Maize

Prakitchai Chotewutmontri, Alice Barkan*

Institute of Molecular Biology, University of Oregon, Eugene, Oregon, United States of America

* abarkan@uoregon.edu

Abstract

Chloroplast genomes in land plants contain approximately 100 genes, the majority of which reside in polycistronic transcription units derived from cyanobacterial operons. The expression of chloroplast genes is integrated into developmental programs underlying the differentiation of photosynthetic cells from non-photosynthetic progenitors. In C4 plants, the partitioning of photosynthesis between two cell types, bundle sheath and mesophyll, adds an additional layer of complexity. We used ribosome profiling and RNA-seq to generate a comprehensive description of chloroplast gene expression at four stages of chloroplast differentiation, as displayed along the maize seedling leaf blade. The rate of protein output of most genes increases early in development and declines once the photosynthetic apparatus is mature. The developmental dynamics of protein output fall into several patterns. Programmed changes in mRNA abundance make a strong contribution to the developmental shifts in protein output, but output is further adjusted by changes in translational efficiency. RNAs with prioritized translation early in development are largely involved in chloroplast gene expression, whereas those with prioritized translation in photosynthetic tissues are generally involved in photosynthesis. Differential gene expression in bundle sheath and mesophyll chloroplasts results primarily from differences in mRNA abundance, but differences in translational efficiency amplify mRNA-level effects in some instances. In most cases, rates of protein output approximate steady-state protein stoichiometries, implying a limited role for proteolysis in eliminating unassembled or damaged proteins under non-stress conditions. Tuned protein output results from gene-specific trade-offs between translational efficiency and mRNA abundance, both of which span a large dynamic range. Analysis of ribosome footprints at sites of RNA editing showed that the chloroplast translation machinery does not generally discriminate between edited and unedited RNAs. However, editing of ACG to AUG at the *rpl2* start codon is essential for translation initiation, demonstrating that ACG does not serve as a start codon in maize chloroplasts.

Author Summary

Chloroplasts are subcellular organelles in plants and algae that carry out the core reactions of photosynthesis. Chloroplasts originated as cyanobacterial endosymbionts. Subsequent

coevolution with their eukaryotic host resulted in a massive transfer of genes to the nuclear genome, the acquisition of new gene expression mechanisms, and the integration of chloroplast functions into host programs. Chloroplasts in multicellular plants develop from non-photosynthetic proplastids, a process that involves a prodigious increase in the expression of chloroplast genes encoding components of the photosynthetic apparatus. We used RNA sequencing and ribosome profiling to generate a comprehensive description of the dynamics of chloroplast gene expression during the transformation of proplastids into the distinct chloroplast types found in bundle sheath and mesophyll cells in maize. Genes encoding proteins that make up the chloroplast gene expression machinery peak in protein output earlier in development than do those encoding proteins that function in photosynthesis. Programmed changes in translational efficiencies superimpose on changes in mRNA abundance to shift the balance of protein output as chloroplast development proceeds. We also mined the data to gain insight into general features of chloroplast gene expression, such as relative translational efficiencies, the impact of RNA editing on translation, and the identification of rate limiting steps in gene expression. The findings clarify the parameters that dictate the abundance of chloroplast gene products and revealed unanticipated phenomena to be addressed in future studies.

Introduction

The evolution of chloroplasts from a cyanobacterial endosymbiont was accompanied by a massive transfer of bacterial genes to the nuclear genome, and by the integration of chloroplast processes into the host's developmental and physiological programs [1]. In multicellular plants, chloroplasts differentiate from non-photosynthetic proplastids in concert with the differentiation of meristematic cells into photosynthetic leaf cells. This transformation is accompanied by a prodigious increase in the abundance of the proteins that make up the photosynthetic apparatus, which contribute more than half of the protein mass in photosynthetic leaf tissue [2]. Both nuclear and chloroplast genes contribute subunits to the multisubunit complexes that participate in photosynthesis. The expression of these two physically separated gene sets is coordinated by nucleus-encoded proteins that control chloroplast gene expression, and by signals emanating from chloroplasts that influence nuclear gene expression [1, 3]. Beyond these general concepts, however, little is known about the mechanisms that coordinate chloroplast and nuclear gene expression in the context of the proplastid to chloroplast transition. Furthermore, a thorough description of the dynamics of chloroplast gene expression during this process is currently lacking.

Despite roughly one billion years of evolution, the bacterial ancestry of the chloroplast genome is readily apparent in its gene organization and gene expression mechanisms. Most chloroplast genes in land plants are grouped into polycistronic transcription units [4] that are transcribed by a bacterial-type RNA polymerase [5] and translated by 70S ribosomes that strongly resemble bacterial ribosomes [6]. As in bacteria, chloroplast ribosomes bind mRNA at ribosome binding sites near start codons, sometimes with the assistance of a Shine-Dalgarno element [6]. Superimposed on this ancient scaffold are numerous features that arose post-endosymbiosis [7]. For example, a phage-type RNA polymerase collaborates with an RNA polymerase of cyanobacterial origin [5], and chloroplast RNAs are modified by RNA editing, RNA splicing, and other events that are either unusual or absent in bacteria [8].

Ribosome profiling data from *E. coli* revealed that the rate of protein output from genes encoding subunits of multisubunit complexes is proportional to subunit stoichiometry, and

that proportional synthesis is typically achieved by differences in the translational efficiency of genes residing in the same operon [9, 10]. As the majority of chloroplast gene products are components of multisubunit complexes, it is of interest to know whether similar themes apply. Furthermore, the gene content of polycistronic transcription units in chloroplasts has diverged from that in the cyanobacterial ancestor. Has "tuned" protein output been maintained in chloroplasts despite this disrupted operon organization? If so, what mechanisms achieve this tuning in light of the new gene arrangements and the new features of mRNA metabolism?

In this work, we used ribosome profiling to address these and other questions of chloroplast gene regulation in the context of the proplastid to chloroplast transition. For this purpose, we took advantage of the natural developmental gradient of the maize seedling leaf blade, where cells and plastids at increasing stages of photosynthetic differentiation form a developmental gradient from base to tip [11]. By using the normalized abundance of ribosome footprints as a proxy for rates of protein synthesis, we show that the rate of protein output from many chloroplast genes is tuned to protein stoichiometry, and that tuned protein output is achieved through gene-specific balancing of mRNA abundance with translational efficiency. This comprehensive analysis revealed developmentally programmed changes in translational efficiencies, which superimpose on programmed changes in mRNA abundance to shift the balance of protein output as chloroplast development proceeds.

Results

Experimental design

We analyzed tissues from the same genetic background and developmental stage as used in previous proteome [2] and nuclear transcriptome [12, 13] studies of photosynthetic differentiation in maize. Four leaf sections were harvested from the third leaf to emerge in 9-day old seedlings (Fig 1A): the leaf base (segment 1), which harbors non-photosynthetic proplastids; 3–4 cm above the base (segment 4), representing the sink-source transition and a region of active chloroplast biogenesis; 8–9 cm above the base (segment 9), representing young chloroplasts; and a section near the tip (segment 14) harboring mature bundle sheath and mesophyll chloroplasts [2, 12]. The developmental transitions represented by these fractions are illustrated in the immunoblot assays shown in Fig 1B. The mitochondrial protein Atp6 is most abundant in the two basal sections, subunits of photosynthetic complexes (AtpB, PetD, PsaD, PsbA, NdhH, RbcL) are most abundant in the two apical sections, and a chloroplast ribosomal protein (Rpl2) exhibits peak abundance in the two middle sections. These developmental profiles are consistent with prior proteome data [2].

To explore the contribution of differential chloroplast gene expression to the distinct proteomes in bundle sheath and mesophyll cells, we also analyzed bundle sheath and mesophyll-enriched fractions from the apical region of seedling leaves. Standard protocols for the separation of bundle sheath and mesophyll cells involve lengthy incubations that are likely to cause changes in ribosome position. We used a rapid mechanical fractionation method that minimizes the time between tissue disruption and the generation of ribosome footprints (see Materials and Methods). Markers for each cell type were enriched 5- to 10-fold in the corresponding fraction (Fig 1C). This degree of enrichment is comparable to that of the fractions used to define mesophyll and bundle sheath-enriched proteomes in maize [14].

We modified our previous method for preparing ribosome footprints from maize leaf tissue [15] to reduce the amount of time and tissue required, and to reduce contamination by non-ribosomal ribonucleoprotein particles (RNPs). In brief, leaf tissue was flash frozen and ground
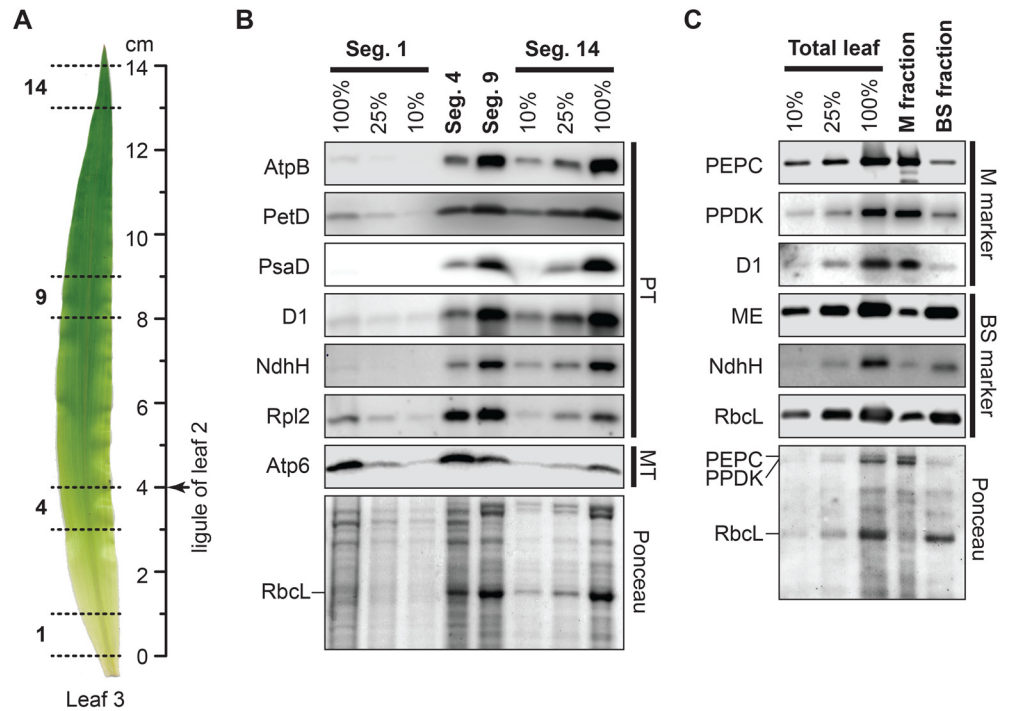
Fig 1. Tissue samples used for ribosome profiling. (A) Leaf segments analyzed in this study. The indicated segments were excised from leaf 3 of nine-day-old seedlings. The segments are numbered according to the nomenclature of [13]. **(B)** Immunoblots showing abundance of marker proteins in the leaf gradient tissue used for ribosome profiling. Replicate immunoblots were probed with antibodies to subunits of the chloroplast ATP synthase (AtpB), cytochrome b_6f complex (PetD), Photosystem I (PsaD), Photosystem II (D1), the NDH-like complex NdhH), chloroplast ribosomes (Rpl2), and mitochondrial ATP synthase (Atp6). Samples were loaded on the basis of equal protein. An image of one blot stained with Ponceau S (below) serves as a loading control and illustrates the abundance of the large subunit of Rubisco (RbcL). PT, plastid proteins; MT, mitochondrial protein. **(C)** Replicate immunoblots showing abundance of marker proteins in mesophyll (M) and bundle sheath (BS) fractions used for ribosome profiling. Samples were loaded on the basis of equal protein. An image of one blot stained with Ponceau S (below) illustrates the abundance of several marker proteins. PEPC, phosphoenolpyruvate carboxylase; PPDK, pyruvate orthophosphate dikinase; ME, NADP-dependent malic enzyme; other abbreviations as in panel (B).

doi:10.1371/journal.pgen.1006106.g001

in liquid N_2, thawed in a standard polysome extraction buffer, and treated with Ribonuclease I to liberate monosomes. Ribosomes were purified by pelleting through a sucrose cushion under conditions that leave chloroplast group II intron RNPs (~600 kDa) [16] in the supernatant (S1A Fig). RNAs between approximately 20 and 35 nucleotides (nt) were gel purified and converted to a sequencing library with a commercial small RNA library kit that has minimal ligation bias [17]. rRNA contaminants were depleted after first strand cDNA synthesis by hybridization to biotinylated oligonucleotides designed to match abundant contaminants detected in pilot experiments (S1 Table). Approximately 35 million reads were obtained for each "Ribo-seq" replicate, roughly 50% of which aligned to mRNA (S2 Table). RNA-seq data was generated from RNA extracted from aliquots of each lysate taken prior to addition of RNAse I. Replicate RNA-seq and Ribo-seq assays showed high reproducibility (Pearson correlation of >0.98, S2 Fig). Almost all plastid genes were represented by at least 100 reads per replicate in all datasets (S3 Fig). Several clusters of low abundance reads mapped to small unannotated ORFs, but further investigation is required to evaluate which, if any, of these are the footprints of translating ribosomes.

Characteristics of ribosome footprints in the chloroplast, mitochondrion, and cytosol

Ribosomes in the cytosol, mitochondria, and chloroplasts have distinct genetic origins. Accordingly, the ribosome footprints from each compartment displayed different size distributions (Fig 2A). The cytosolic ribosome footprints showed a minor peak at 23 nucleotides and a major peak at 31 nucleotides, similar to observations in yeast [18]. The mitochondrial data showed a major peak at 28–29 nucleotides and a minor peak at 36 nucleotides, similar to the 27 and 33-nt peaks reported for human mitochondria [19]. The plastid ribosome footprints had a broad size distribution suggestive of two populations, with peaks at approximately 30 and 35 nucleotides. A similar distribution was observed in pilot experiments involving the gel purification of RNAs up to 40-nt (S1B Fig) indicating that the peak at 35-nt was not an artifact of our gel purification strategy. A broad and bimodal size distribution was also observed for chloroplast ribosome footprints from the single-celled alga *Chlamydomonas reinhardtii*, albeit with peaks at slightly different positions [20]. The two prior reports of ribosome footprint size distributions in plants [21, 22] did not parse the data from the three compartments, but the 31-nucleotide modal size reported in those studies is consistent with our data. Our data show the 3-nucleotide periodicity expected for ribosome footprints (Fig 2B and 2C). Interestingly, the degree of periodicity varies with footprint size (S4 Fig). The reads are largely restricted to open reading frames in the cytosol (Fig 2C) and chloroplast (Fig 2D). Taken together, these results provide strong evidence that the vast majority of the Ribo-seq reads come from bonafide ribosome footprints.

The placement of ribosome P and A sites with respect to ribosome footprint termini has not been reported for any organellar ribosomes or for cytosolic ribosomes in maize. A meta analysis of our data showed that the position of the 3' end of ribosome footprints from initiating and terminating ribosomes in chloroplasts and mitochondria is constant with respect to start and stop codons, respectively, regardless of footprint size; however, the position of the 5' ends varies with footprint size (Fig 2E, S4C Fig). Therefore, the positions of the A and P sites in organellar ribosomes can be inferred based on the 3'-ends of their footprints, as is also true for bacterial ribosomes [23, 24]. The modal distance between the start of the P site in chloroplast ribosomes and the 3'-ends of chloroplast ribosome footprints is 7 nucleotides. By contrast, cytosolic ribosome footprints are approximately centered on the P site regardless of footprint size (S4B Fig).

The partitioning of ribosome footprints among the three genetic compartments shifts dramatically during the course of leaf development (Fig 2G). The contribution of cytosolic translation drops from 99% at the leaf base to 57% in the apical leaf sections due to the increasing contribution of ribosome footprints from chloroplasts. This shift of cellular resources towards chloroplast translation corresponds with the massive increase in the content of photosynthetic complexes harboring plastid-encoded subunits (Rubisco, PSII, PSI, cytochrome b_6f, ATP synthase, NDH) (Fig 1). Ribosome footprints from mitochondria accounted for a very small fraction of the total at all stages. However, our protocol was not optimized for the quantitative recovery of mitochondrial ribosomes so these data may not reflect the total mitochondrial ribosome population.

The translational output of most chloroplast genes is tuned to the stoichiometry of their products

In the discussion below we define the "translational output" of a gene as the abundance of ribosome footprints per kb per million reads mapped to nuclear coding sequences (RPKM), and we use this value to compare rates of protein synthesis among genes on a molar basis. This is a typical interpretation of Ribo-seq data, and it is based on evidence that the bulk rate of
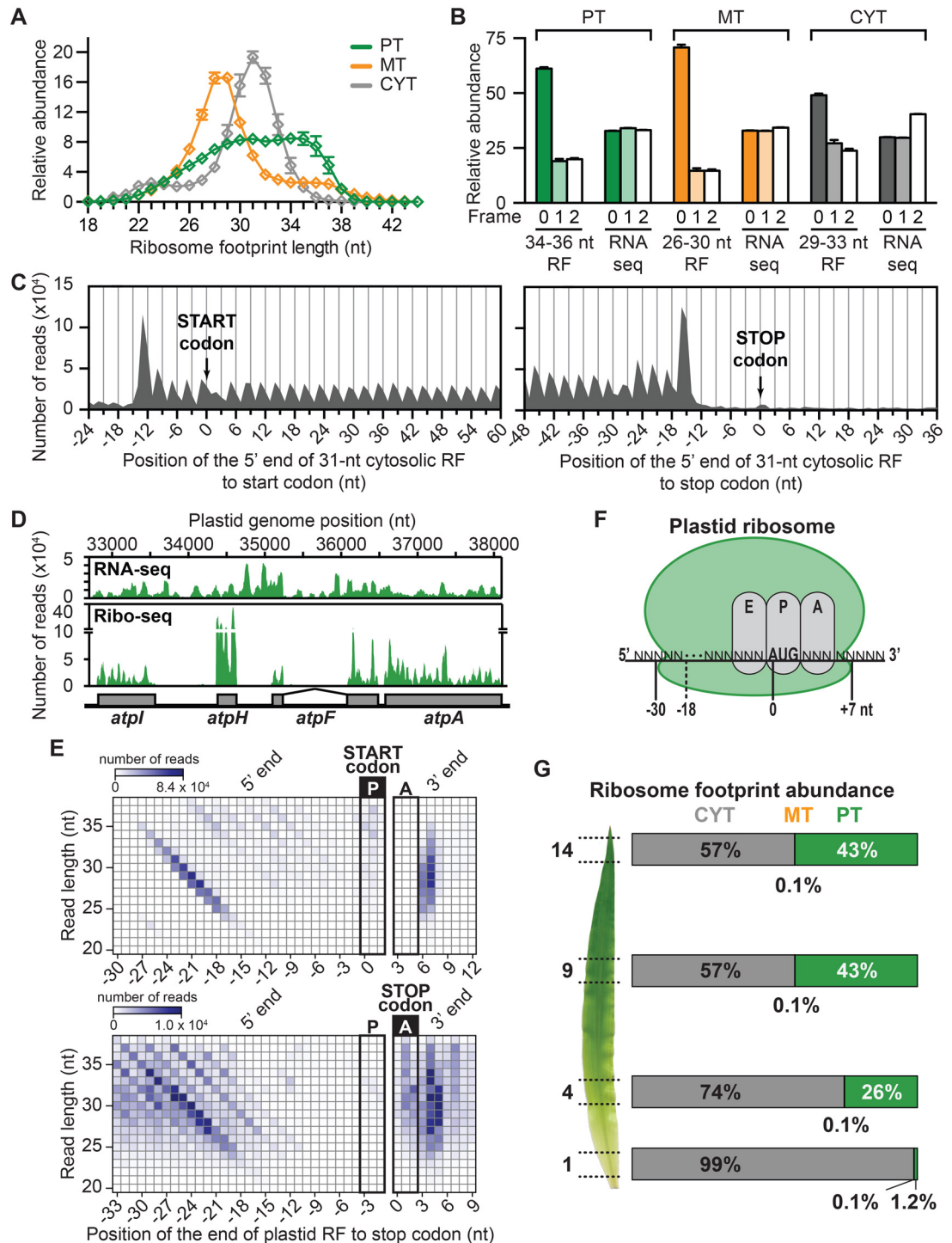
Fig 2. Characteristics of ribosome footprints from the cytosol (CYT), plastids (PT), and mitochondria (MT). (A)
Size distribution of ribosome footprints. Values are the mean ± Standard Error of the Mean (SEM) from the twelve leaf segment datasets. **(B)** Three-nucleotide periodicity of Ribo-seq data. Histograms show the abundance of reads representing P-site placement in each of the three reading frames. P-site placements were inferred based on the footprint size and the size-dependent P-site positions inferred from our data (see panel **E** and S4 Fig). Values are the mean ± SEM from the 12 leaf segment datasets. **(C)** Meta-analysis of cytosolic ribosome footprints mapping near the start and stop codons of cytosolic ORFs in the leaf gradient datasets. Values show the number of 31-nt footprints with 5' ends at each position. **(D)** Example of Ribo-seq and RNA-seq coverage in a chloroplast polycistronic transcription unit. Reads are combined from the twelve leaf segment datasets. The group II intron in the *atpF* gene is marked. **(E)**

Meta-analysis of plastid ribosome footprints that map to start and stop codons. The data from the twelve leaf gradient datasets are parsed by footprint size. **(F)** Placement of plastid ribosome footprints with respect to the E, P, and A sites of the ribosome. The 5'-end placement varies with footprint size, while the 3'-end is constant at 7-nt downstream from the start of the P site (see data in panel **E**). **(G)** Partitioning of translational output among the three genetic compartments during photosynthetic differentiation. Values represent the average from three biological replicates. Note that the footprints from mitochondrial ribosomes may be under-represented due to the protocol used here.

translation elongation on all ORFs is similar under any particular condition, despite the fact that ribosome pausing can lead to the over-representation of ribosomes at specific positions [9, 25]. Although this may be an over simplification in some instances, this interpretation of our data produced results that are generally coherent with current understanding of chloroplast biogenesis (see below). Group II introns interrupt eight protein-coding genes in maize chloroplasts. These present a challenge for data analysis because the unspliced transcripts make up a substantial fraction of the RNA pool [16] and translation can initiate on unspliced RNAs and terminate within introns [15]. We therefore calculated translational output based solely on the last exon (normalized to exon length). Data summaries presented below include RNA-seq data only for that subset of intron-containing genes for which multiple methods of analysis provided consistent values for the abundance of spliced RNA isoforms (see Materials and Methods).

Fig 3 summarizes the abundance of Ribo-seq and RNA-seq reads from protein-coding chloroplast genes in each of the four leaf segments. To display the low values from Segment 1, they are replotted with a smaller Y-axis scale in S5 Fig. The abundance of mRNA from genes in the same transcription unit (Fig 3A and S5A Fig, bracketed arrows) is typically similar, but the protein output of co-transcribed genes varies considerably. Translational efficiency (translational output /mRNA abundance) varies widely among genes (Fig 3A and S5A Fig, bottom). The *atpH* mRNA is the most efficiently translated of any chloroplast mRNA at all four developmental stages, surpassing even *psbA*, whose product is the most rapidly synthesized protein in photosynthetic tissues [26]. Prodigious *psbA* expression results from very high mRNA abundance in combination with a translational efficiency that is comparable to that of other photosystem genes.

When the data are grouped according to gene function, correlations between function and translational output become apparent (Fig 3B). For example, the translational output of genes encoding subunits of ribosomes and the NDH complex are consistently very low, whereas the translational output of genes encoding subunits of PSI, PSII, the ATP synthase, and the cytochrome b_6f complex are consistently much higher. These trends mirror the abundance of these complexes as inferred from proteome data [27]. The data for complexes whose subunits are not found in a 1:1 ratio show further that translational output is tuned to subunit stoichiometry. For example, the chloroplast-encoded subunits of the ATP synthase (AtpA, AtpB, AtpE, AtpF, AtpH, AtpI) are found in a 3: 3: 1: 1: 14: 1 molar ratio in the complex [28, 29]. The translational output of their genes mirrors this stoichiometry quite well, whereas mRNA abundance does not (Fig 4A). These genes are distributed between two transcription units (Fig 4A). A single mRNA encodes AtpB and AtpE, whose rates of synthesis are tuned via differences in translational efficiency. The *atpI-atpH-atpF-atpA* primary transcript is processed to yield various smaller isoforms [30] but the abundance of RNA from each gene is nonetheless quite similar (Fig 4A). The translational output of the *atpH* gene is boosted relative to that of its neighbors primarily through exceptionally high translational efficiency (Fig 4A bottom). In a second example, the unequal stoichiometry of subunits of the plastid-encoded RNA polymerase (PEP) (2 RpoA:1 RpoB:1 RpoC1:1 RpoC2) [5] is mirrored by the relative translational output of the
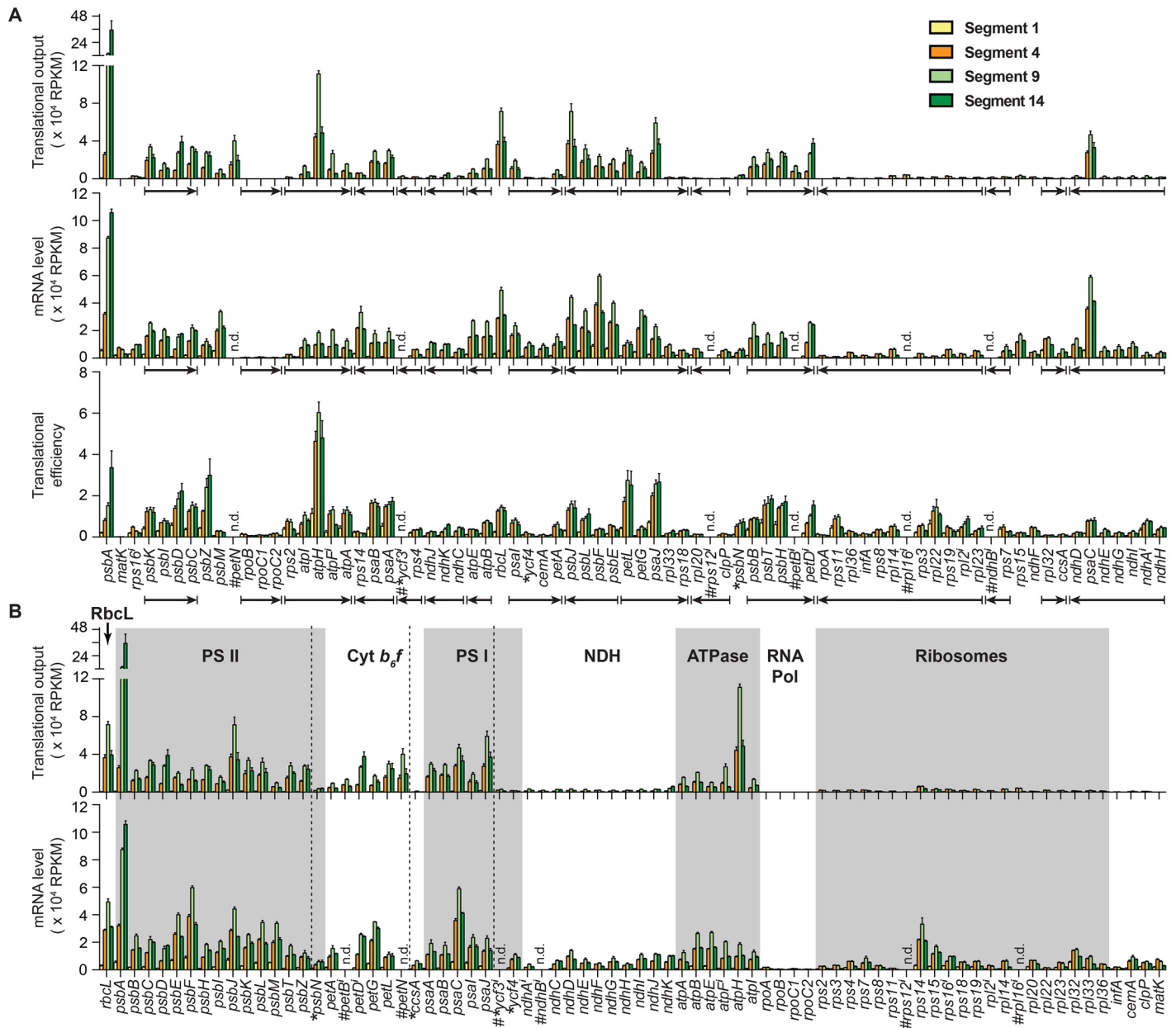
Fig 3. Genome-wide views of the plastid translatome and transcriptome during the proplastid to chloroplast transition. The data are expressed as reads per kilobase per million reads mapping to nuclear genome coding sequences (RPKM). Translational output is defined as Ribo-seq RPKM. mRNA level is defined as RNA-seq RPKM. Translational efficiency is calculated as the ratio of translational output to mRNA level. Values are the mean ± SEM from three replicates. Intron-containing genes are marked with a superscript *i*. Genes for which RNA levels and translational efficiency were not determined (n.d.) are marked (#). These include intron-containing genes for which the fraction of reads derived from spliced transcripts is uncertain, and *petN*, whose short mRNA is not represented quantitatively in the RNA-seq data. Genes encoding assembly factors are marked with asterisks. Other genes encode structural components of the complexes indicated in panel B. **(A)** Translational output, RNA abundance and translational efficiency displayed according to native gene order. Co-transcribed genes are marked with arrows that indicate the direction of transcription. **(B)** Translational output and RNA abundance displayed according to gene function. Genes encoding assembly factors are demarcated from the structural genes with dashed lines. The data for each functional group are plotted separately in Fig 4 and S6 Fig using Y-axis scales suited for the relevant values. The data for Segment 1 are displayed with a different Y-axis scale in S5 Fig.

doi:10.1371/journal.pgen.1006106.g003
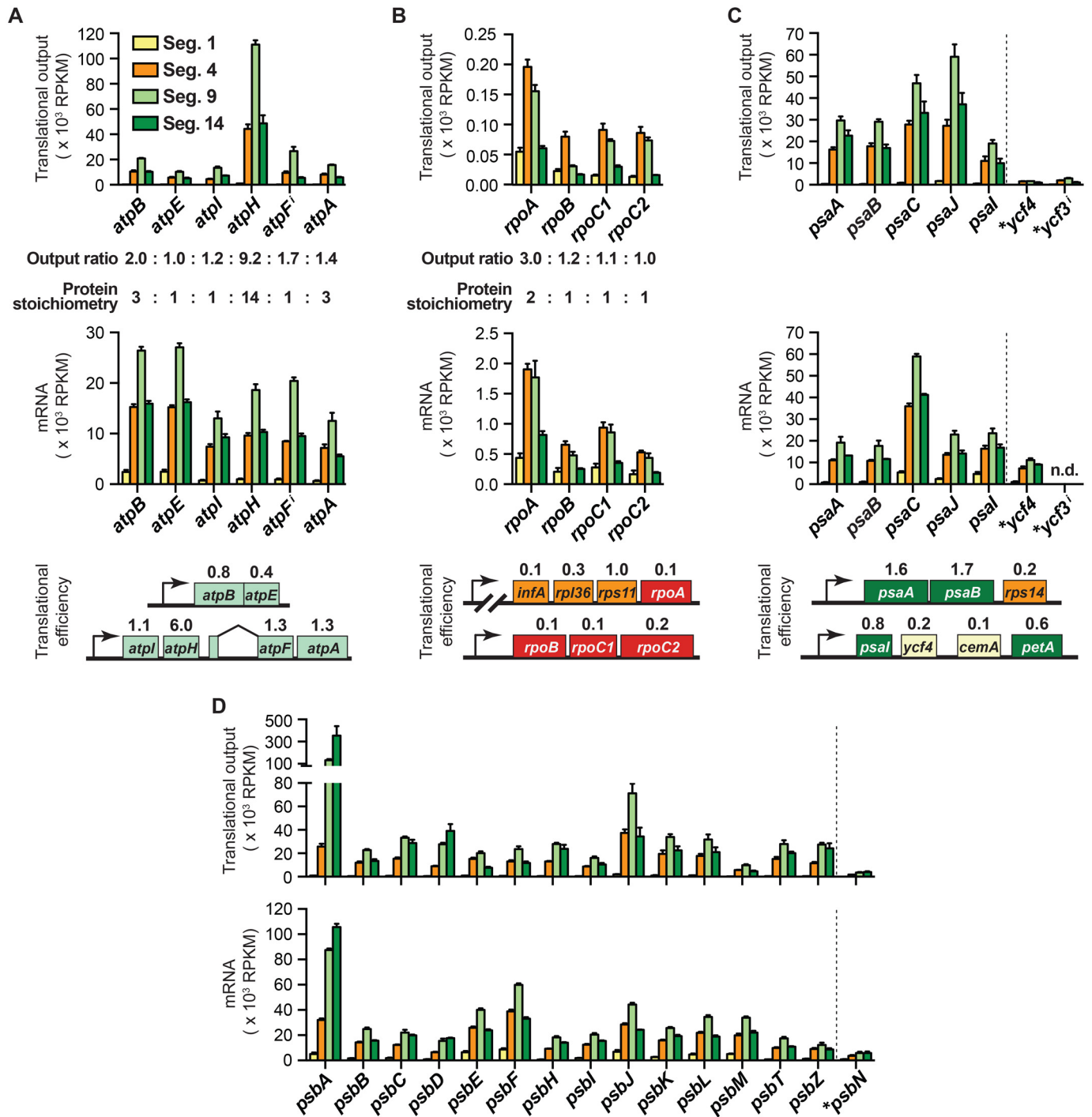
Fig 4. Comparison of protein stoichiometry, translational output and mRNA abundance for several complexes. Values represent the mean ± SEM from three replicates. RNA-seq data are not provided for *ycf3* due to uncertainty about the proportion of its transcripts that are fully spliced. Genes with introns are marked with superscript *i*. Genes encoding assembly factors are marked with asterisks. n.d. not determined. **(A)** Plastid genes encoding ATP synthase subunits. The Output Ratio values are averages of those in leaf segments 4, 9 and 14. The chloroplast transcription units encoding ATP synthase subunits are shown below, and are annotated with translational efficiencies from leaf segment 9. **(B)** Plastid genes encoding RNA polymerase subunits. The Output Ratio values are averages of those in leaf segments 1 and 4. The chloroplast transcription units encoding RNA polymerase subunits are shown below, and are annotated with translational efficiencies from leaf segment 9. Genes upstream of *rpoA* on the same polycistronic transcript are included to provide context. **(C)** Plastid genes encoding PSI subunits and assembly factors. Two of the chloroplast transcription units encoding PSI-related proteins are diagrammed below and annotated with translational efficiencies from leaf segment 9. **(D)** Plastid genes encoding PSII subunits and assembly factors.

doi:10.1371/journal.pgen.1006106.g004

corresponding genes (Fig 4B). In this case, however, tuning occurs primarily at the level of mRNA accumulation.

The plastid-encoded subunits of PSI, PSII, the cytochrome b_6f complex, the NDH complex, and chloroplast ribosomes are found in equal numbers in their respective complex. Genes encoding subunits of each of these complexes are distributed across multiple transcription units, many of which also encode subunits of other complexes. This gene organization sometimes results in considerable disparity in mRNA level among subunits of the same complex (Fig 3B bottom). In general, such differences are buffered by opposing changes in translational efficiency, such that translational outputs more closely reflect protein stoichiometry than does mRNA abundance (see, for example, the NDH complex in S6B Fig). In the case of PSI (Fig 4C), the structural genes (*psaA*, *psaB*, *psaC*, *psaJ*, *psaI*) exhibit an approximately three-fold range of translational output, but all of these genes vastly out produce two genes encoding PSI assembly factors (*ycf3* and *ycf4*) [31–33]. The *psaI* and *ycf4* genes are adjacent in the same polycistronic transcription unit (Fig 4C bottom), and their difference in translational output is programmed primarily by a difference in translational efficiency. The translational output of *psbN*, which encodes a PSII assembly factor [34], is likewise much less than that of structural genes for PSII (Fig 4D). Taken together, this body of data shows that the tuning of translational output to protein stoichiometries is accomplished via trade-offs between mRNA level and translational efficiency, with this balance differing from one gene to the next. Where mRNA abundance closely matches protein stoichiometry, differences in translational efficiency make only a small contribution (as observed for *rpoA*, *rpoB*, *rpoC1* and *rpoC2*). Where mRNAs are severely out of balance with protein stoichiometry, differences in translational efficiency compensate.

The translational output of PSII structural genes is well matched, with the notable exception of *psbA* (Fig 4D), whose output vastly exceeds that of other genes in photosynthetic leaf segments (segments 9 and 14). This behavior is consistent with the known properties of the *psbA* gene product, whose damage and rapid turnover during active photosynthesis is compensated by a high rate of synthesis to support PSII repair [26]. Setting *psbA* aside, the relative translational outputs of other genes only approximate the stoichiometries of their products: several-fold differences between relative output and stoichiometry are common among subunits of a particular complex, suggesting that proteolysis of unassembled subunits serves to fine-tune protein stoichiometries. It is also possible that the calculated translational outputs do not perfectly reflect rates of protein synthesis due to differences in translation elongation rates among mRNAs. That said, instances in which translational outputs are particularly discordant among subunits of the same complex are worthy of note, as this may reflect physiologically relevant behaviors. For example, the translational output of *ndhK* is balanced with other *ndh* genes in non-photosynthetic leaf segments but *ndhK* substantially out produces the other *ndh* genes in mature chloroplasts (S6B Fig). This behavior is reminiscent of *psbA*, and suggests that NdhK may be damaged and replaced during active photosynthesis.

Developmental dynamics of the chloroplast transcriptome and translatome

To explore the dynamics of chloroplast gene expression during the proplastid to chloroplast transition, we calculated standardized values for translational output, mRNA abundance and translational efficiency such that developmental shifts can be compared despite large differences in signal magnitude. This analysis shows that the developmental dynamics of translational output varies widely among genes (Fig 5A top). The standardized values were used as the input for hierarchical clustering, which produced four clusters from the translational
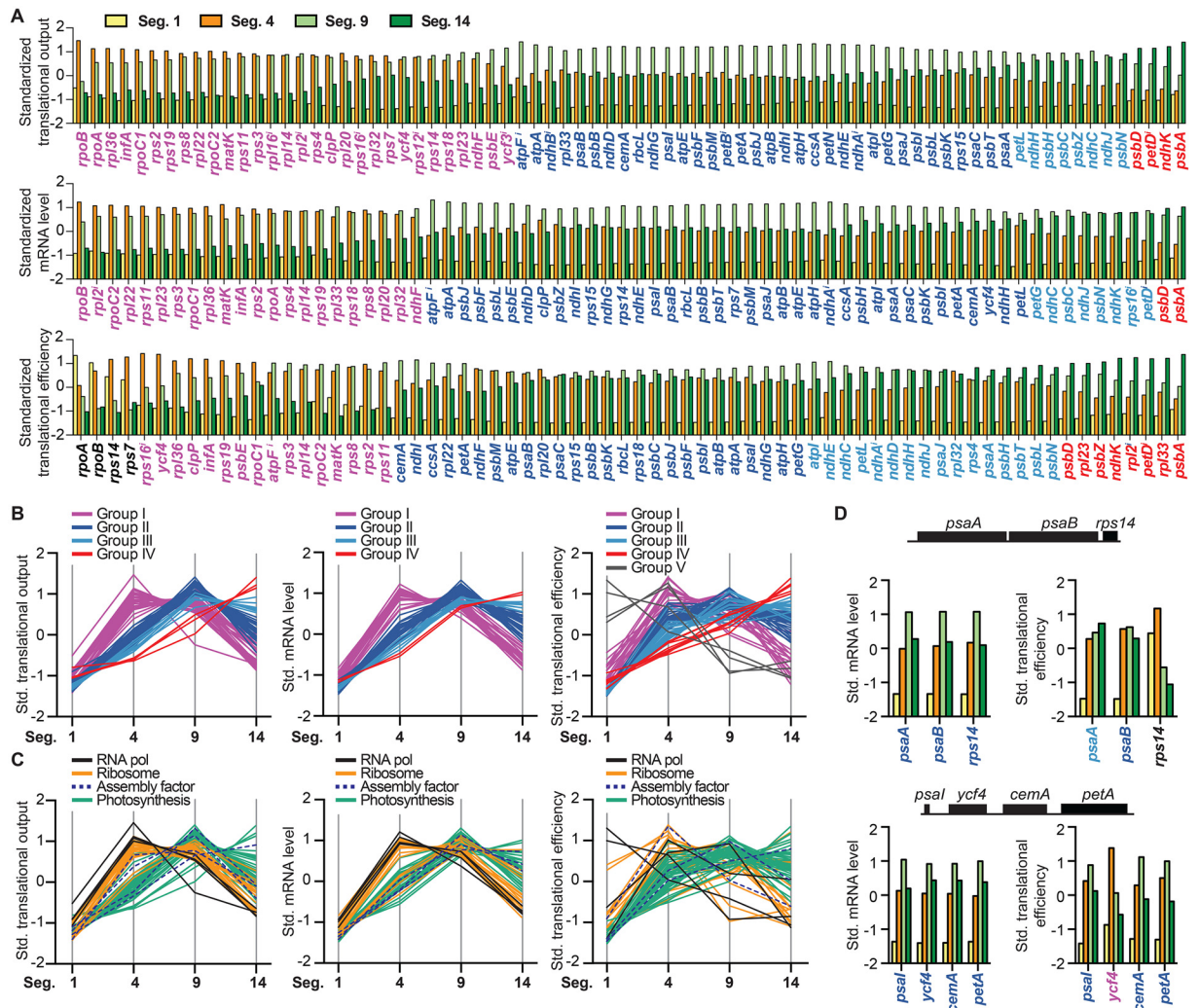
Fig 5. Dynamics of chloroplast gene expression during the proplastid to chloroplast transition. Values for each gene in the four leaf segments were standardized to have a mean of 0 and a standard deviation of 1. Standardized values for translational output, mRNA abundance, and translational efficiency were used for hierarchical clustering. Genes are color-coded according to the cluster they reside in, as defined by the clustograms shown in S7 Fig. Values represent the mean from three replicates. (**A**) Genes are ranked based on the clustograms shown in S7 Fig. Genes for which mRNA levels are uncertain (as explained above) are not included in the plots of mRNA level and translational efficiency. Intron-containing genes are marked with superscript *i*. (**B**) Developmental dynamics of each cluster derived from the data for translational output, mRNA abundance, or translational efficiency. Each line represents data from one gene. (**C**) Developmental dynamics of translational output, mRNA level, and translational efficiency color-coded according to gene function. Each line represents data from one gene. (**D**) Excerpts of the data for two transcription units encoding proteins that function in both photosynthesis (*psaI, petA, psaA, psaB*) and biogenesis of the photosynthetic apparatus (*ycf4, rps14*).

doi:10.1371/journal.pgen.1006106.g005

output data, four from the mRNA data, and five from the translational efficiency data (Fig 5B, S7 Fig). The genes in each cluster are identified by color in Fig 5A. Although the transitions between clusters are not marked by obvious distinctions, the distinct trends defining each cluster are clear in the plots in Fig 5B. Genes whose translational output and mRNA abundance peak early in development (segment 4) generally encode components of the chloroplast gene expression machinery (*rpl, rps, rpo, matK*) (Fig 5A and 5C). Most genes encoding components of the photosynthetic apparatus (*psb, psa, atp, pet* genes) have peak mRNA and translational output in young chloroplasts (segment 9). A handful of photosynthesis genes either

maintain or increase translational output and mRNA in mature chloroplasts (segment 14) (Fig 5A and 5C).

There is considerable similarity among the clusters produced from the translational output and mRNA data (Fig 5A and 5B), implying that programmed changes in mRNA abundance underlie the majority of developmental shifts in translational output. However, changes in translational efficiency also influence the developmental shifts in translational output (Fig 5A bottom). In general, ORFs encoding proteins involved in photosynthesis are more efficiently translated later in development and those encoding gene expression factors are more efficiently translated early in development, albeit with numerous exceptions (Fig 5A bottom, 5C right). Transcription units that encode both photosynthesis and gene expression factors provide revealing examples of distinct translational dynamics. In the *psaA-psaB-rps14* transcription unit, for example, *rps14* is found in a translational output cluster with other genes involved in gene expression, whereas *psaA* and *psaB* reside in a translational output cluster with other photosynthesis genes (Fig 5A top). This results from distinct developmental shifts in translational efficiency: the *rps14* ORF is translated more efficiently early in development whereas *psaA* and *psaB* are more efficiently translated later in development (Fig 5D). The *psaI-ycf4-cemA-petA* transcription unit provides a second example. The translational output of *psaI*, *cemA*, and *petA* show similar developmental dynamics, but *ycf4* clusters with different genes due to more efficient translation earlier in development (Fig 5D). Again, these distinct patterns correlate with function, as *psaI* and *petA* encode components of the photosynthetic apparatus, whereas *ycf4* encodes an assembly factor for PSI [31, 32].

Many polycistronic RNAs in chloroplasts are processed to smaller isoforms. Although the impact of processing on translational efficiencies remains unclear [35, 36], it is plausible that programmed changes in the accumulation of processed isoforms could uncouple the expression of cotranscribed genes during development. To address this possibility, we used RNA gel blot hybridization to analyze transcripts from two transcription units that include genes whose translational efficiencies exhibit distinct developmental dynamics: *psaI-ycf4-cemA-petA* and *psaA-psaB-rps14* transcription units (S8 Fig). Processed *rps14*-specific transcripts accumulate preferentially in immature chloroplasts (segment 4), correlating with the stage at which *rps14* is most efficiently translated. Analogously, a monocistronic *psaI* isoform accumulates preferentially in segments 4 and 9 where *psaI* is most efficiently translated. Various cause and effect relationships may underlie these correlations, as is discussed below.

Differential gene expression in bundle sheath and mesophyll chloroplasts

In maize and other C4 plants, photosynthesis is partitioned between mesophyll (M) and bundle sheath (BS) cells. Three protein complexes that include plastid-encoded subunits accumulate differentially in the two cell types: Rubisco and the NDH complex are enriched in BS cells whereas PSII is enriched in M cells [2, 14]. Differential accumulation of several chloroplast mRNAs in the two cell types has been reported [37–41], but a comprehensive comparison of chloroplast gene expression in BS and M cells has been lacking. To address this issue we performed RNA-seq and Ribo-seq analyses of BS- and M- enriched leaf fractions. The translational output of genes encoding subunits of Rubisco, PSII, and the NDH complex (Fig 6A) correlated well with the relative abundance of subunits of these complexes in the same sample preparations (Fig 1C), and with quantitative proteome data [2]. Cell-type specific differences in mRNA accumulation (Fig 6B) can account for many of the differences in translational output (Fig 6A), indicating that differences in transcription and/or RNA stability make a strong contribution to preferential gene expression in one cell type or the other. However, the data
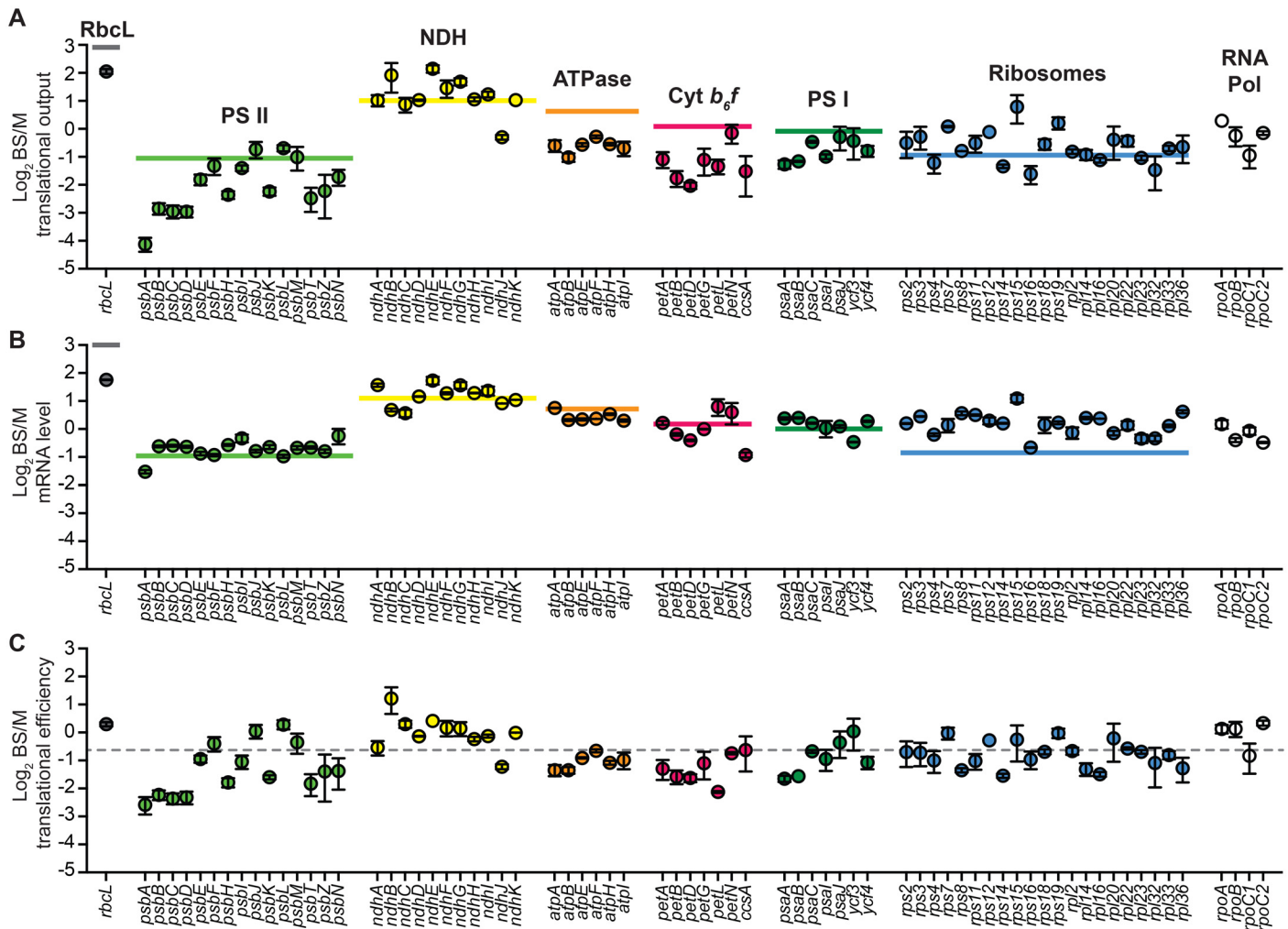
Fig 6. Differential gene expression in bundle sheath (BS) and mesophyll (M) chloroplasts. Values are the mean ± SEM from two replicates. The horizontal lines show the average BS/M ratios for subunits of each complex determined from quantitative proteome data [27]. Immunoblots illustrating the abundance of marker proteins in these samples are shown in Fig 1C. **(A)** Relative translational output in BS versus M fractions. **(B)** Relative mRNA abundance in BS versus M fractions. **(C)** Relative translational efficiency in BS versus M fractions. The dashed line marks the average ratio of translational efficiencies in BS versus M fractions.

suggest that differences in translational efficiency contribute in certain instances (Fig 6C). Four genes encoding PSII core subunits (*psbA*, *psbB*, *psbC*, *psbD*) provide the most compelling examples, as their translational output is considerably more biased toward M cells than are their mRNA levels.

Translation of edited and unedited RNAs

Organellar RNAs in land plants are often modified by an editing process that converts specific cytidine residues to uridine [42, 43]. Some sites are inefficiently edited, which raises the question of whether the translation machinery discriminates between edited and unedited RNAs. The protein products of several unedited mitochondrial RNAs have been detected in plants [44, 45]. We used our Ribo-seq and RNA-seq data to examine this issue for chloroplast RNAs. Fig 7 summarizes the data for those sites of editing that are represented by at least 100 reads in both the Ribo-seq and RNA-seq data in at least two replicates (17 of the 28 edited sites in the
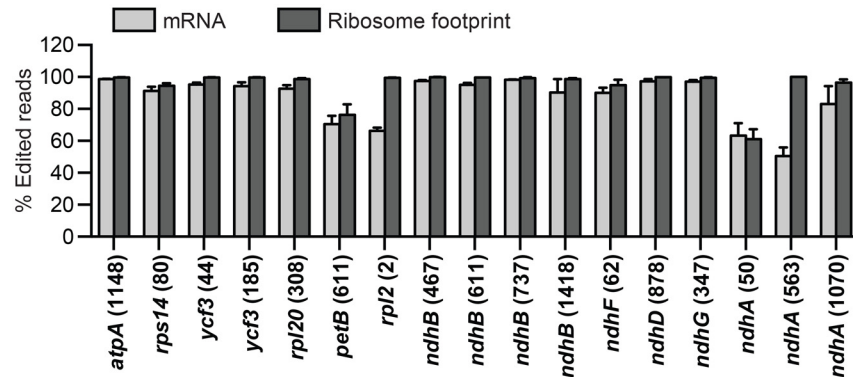
Fig 7. Abundance of ribosome footprints representing edited and unedited mRNA positions. Data are shown for the subset of editing sites that were represented by at least 100 reads in both the RNA-seq and Ribo-seq data in at least two different samples from the leaf gradient analysis. Values are the mean ± SEM.

doi:10.1371/journal.pgen.1006106.g007

maize chloroplast transcriptome). In general, the percent editing was similar in the RNA-seq and Ribo-seq data, implying little discrimination between edited and unedited RNAs by the translation machinery. There were, however, two major exceptions: *rpl2* (nt 2) and *ndhA* (nt 563). In these cases a large fraction of the RNA-seq reads came from unedited RNA, whereas virtually all of the Ribo-seq reads came from edited sites. These two sites have unusual features that can account for the preferential translation of the edited RNAs. Editing at the *ndhA* site is linked to the splicing of the group II intron in the *ndhA* pre-mRNA: the site is not edited in unspliced transcripts and it is fully edited in spliced transcripts [46–48]. Failure to edit unspliced RNA is presumably due to the position of the intron between the edited site and the cis-element that specifies it. Translation that initiates on unspliced *ndhA* RNA would terminate at an in-frame stop codon within the intron. Thus, exon 2 is translated only from spliced RNAs, and these are 100% edited. In the case of *rpl2*, the editing event creates an AUG start codon from an ACG precursor; this is the only editing event in maize chloroplasts that creates a canonical start codon. Although it has been reported that ACG can function as a start codon in chloroplasts [49, 50], our data show that this particular ACG is strongly discriminated against by initiating ribosomes.

The fact that the Ribo-seq data show the expected strong bias toward edited *rpl2* and *ndhA* (563) instills confidence that valid conclusions can be made from our data for other edited sites. Approximately 40% of the *petB* and *ndhA*(nt 50) sequences are unedited in both the RNA-seq and Ribo-seq data, indicating that these unedited sequences give rise to a considerable fraction of the translational output of the corresponding genes. Editing of the *petB* site is essential for the function of its gene product (cytochrome b_6) [51]. It seems likely that the product of this unedited RNA is either unstable or selected against during complex assembly, as has also been suggested for the products of two unedited transcripts in mitochondria [52, 53]. The remaining sites show almost complete editing in the RNA-seq data and, as expected, in the Ribo-seq data as well. That said, there is an overall trend toward less representation of unedited sequences in the Ribo-seq data than in the RNA-seq data. This may simply be a kinetic effect as would be expected if ribosome binding is slow in comparison to editing, such that ribosomes generally translate older (and therefore more highly edited) mRNAs.

Discussion

Ribosome profiling has provided a wealth of new insights into translation and associated processes in a wide variety of organisms [54], but its application to questions in organellar biology

is just beginning. The method has been used to analyze the effects of a disease-associated mutation in mitochondria [19], to define targets of nucleus-encoded translational activators in chloroplasts [15] and to characterize the cotranslational targeting of chloroplast-encoded proteins to the thylakoid membrane [36]. The results reported here provide the first comprehensive description of an organellar transcriptome and translatome in a developmental context. The data revealed dynamic changes in RNA abundance and translational efficiency during the differentiation of proplastids into chloroplasts, elucidated mechanisms that dictate the abundance of chloroplast-encoded proteins, clarified the relationship between RNA editing and translation, and provided new insights that suggest hypotheses to be explored in future studies.

Tuning of protein synthesis to protein stoichiometry: Close but not quite

Ribosome profiling data from bacteria revealed a striking correspondence between the stoichiometry of subunits of multisubunit complexes and their relative rates of synthesis [9, 10]. Our results show that the relative translational outputs of chloroplast genes likewise approximate the relative abundance of the gene products. This tuning is apparent when comparing sets of genes encoding different complexes (e.g. compare genes encoding the low abundance NDH complex to genes encoding the highly abundant PSI and PSII complexes) (Fig 3B), and when comparing genes encoding subunits of the same complex (e.g. the PEP RNA polymerase and the ATP synthase) (Fig 4A and 4B).

Our calculations of translational output rest on the assumption that the rate of translation elongation on all mRNAs is similar under any particular condition. This same assumption produced remarkable concordance between protein stoichiometry and inferred translational output in bacteria [9, 10]. Although our results show a clear trend toward "proportional synthesis", they also suggest that the tuning of protein output to stoichiometry is less precise in chloroplasts than it is in bacteria. Subunits of photosynthetic complexes are subject to proteolysis when their assembly is disrupted [55], and a similar (albeit wasteful) mechanism could contribute to balancing stoichiometries when proteins are synthesized in excess under normal conditions. That said, instances in which inferred translational outputs are particularly incongruent with protein stoichiometries may reflect physiologically informative behaviors. The most prominent examples of "over-produced" proteins in our data are PsbA and PsbJ in PSII, PsaC and PsaJ in PSI, NdhK in the NDH complex, Rps14 in ribosomes, and PetD, PetL and PetN in the cytochrome b_6f complex (Fig 4 and S6 Fig). Disproportionate synthesis of PsbA is well known, and compensates for its damage and proteolysis during photosynthesis [26]. The other proteins suggested by our data to be produced in excess may likewise be subject to more rapid turnover than their partners in the assembled complex. A proteomic study in barley demonstrated that subunits of each photosynthetic complex generally turn over at similar rates [56], but data for these particular proteins were not reported.

Interestingly, the inferred rates of synthesis of PsbA, PetD, and NdhK are well matched to those of their partner subunits early in development, but outpace those of their partners in mature chloroplasts (Fig 4D, S6 Fig). This feature of *psbA* expression coincides with the need to replace its gene product, D1, following photo-induced damage and proteolysis [26]. By extension, the developmental dynamics of *petD* and *ndhK* expression suggest that their gene products may turn over more rapidly than their partners as a consequence of photosynthetic activity.

Varying contributions of mRNA abundance and translational efficiency to the tuning of protein output

In bacteria, proportional synthesis of subunits within a complex is achieved largely through the tuning of translational efficiencies among ORFs on the same mRNA [9, 10]. In chloroplasts,

genes encoding subunits of the same complex are generally distributed among multiple transcription units [4] and RNA segments within a transcription unit often accumulate to different levels [8]. It is interesting to consider how this shift in the gene expression landscape is reflected in the mechanisms that balance protein output among genes. In the case of the four genes encoding the PEP RNA polymerase, relative translational outputs closely match the 2:1:1:1 protein stoichiometry, and this is programmed primarily at the level of mRNA abundance (Fig 4B). By contrast, widely varying translational efficiencies are superimposed on small variations in mRNA abundance to tune translational output to protein stoichiometry in the ATP synthase complex (Fig 4A). Genes for ribosomal proteins are distributed among ten transcription units, several of which also encode proteins involved in photosynthesis (see Fig 3A). For example, *rps14* is cotranscribed with genes encoding the reaction center proteins of PSI (*psaA/psaB*), and translational outputs within this transcription unit are balanced by large differences in translational efficiency (Fig 4C). Similarly, the *psaI* transcription unit encodes subunits of the abundant PSI and cytochrome b_6f complexes, a low abundance PSI assembly factor (Ycf4) and a protein of unknown function (CemA); large differences in translational efficiency adjust the translational outputs to meet these different needs (Fig 4C bottom). For complexes harboring plastid-encoded subunits in equal stoichiometries (ribosomes, NDH, PSI, PSII, cytochrome b_6f), compensating differences in translational efficiency generally buffer differences in mRNA level. Taken together, these results imply that mRNA abundance and translational efficiencies have coevolved in chloroplasts to produce proteins in close to the optimal amounts. In some instances, mRNA levels are sharply out of balance with protein stoichiometries, in which case differential translational efficiencies compensate. In other instances, mRNA levels approximate protein stoichiometries, and translational efficiencies are similar. These observations further suggest that for most genes in maize chloroplasts, mRNA levels and translational efficiencies are poised such that they limit the rate of protein synthesis to a similar extent. This view is further supported by the developmental dynamics discussed below.

In Chlamydomonas chloroplasts, synthesis of subunits within the same photosynthetic complex is coordinated through assembly-dependent auto-regulatory mechanisms [57]. By contrast, current data for angiosperm chloroplasts suggest that translational efficiencies are generally independent of the assembly status of the gene products [15, 58]. It seems likely that translational efficiencies are dictated by the interplay between the sequence and structure of RNA proximal to start codons and the proteins that bind this region. Translation initiation in chloroplasts sometimes involves a Shine-Dalgarno interaction and is facilitated by an unstructured translation initiation region [6, 59]. Additionally, the translation of some chloroplast ORFs requires the participation of gene-specific translation activators [15, 60–75]. Such proteins provide a means for tuning protein synthesis within and between transcription units. The *atpH* ORF and its nucleus-encoded translational activator PPR10 exemplify this mechanism. The exceptionally high translational efficiency of *atpH* (Fig 3A) boosts its translational output to match the high stoichiometry of AtpH in the ATP synthase complex (Fig 4A); this high translational efficiency requires the binding of PPR10 adjacent to the *atpH* ribosome binding site, an interaction that prevents the formation of inhibitory RNA structures involving the translation initiation region [15, 30, 62].

Developmental dynamics of chloroplast gene expression

Our results provide a comprehensive view of the dynamics of chloroplast mRNA abundance and translation during the proplastid to chloroplast transition. The majority of genes involved in chloroplast gene expression exhibit peak mRNA abundance and translational output in developing chloroplasts (segment 4) whereas the majority of genes encoding subunits of the

photosynthetic apparatus exhibit peak mRNA abundance and translational output in young chloroplasts (segment 9) (Fig 5C). That said, even in proplastids (segment 1), genes involved in photosynthesis are generally represented by more mRNA and a higher translational output than are those involved in chloroplast gene expression (S5 Fig). Our data show that programmed changes in translational efficiency combine with changes in mRNA abundance to produce developmental shifts in translational output (Fig 5A). In general, translational efficiency is lowest at the leaf base, reflecting the low ribosome content in proplastids. The translational efficiency of most ORFs peaks in young chloroplasts (segment 9). In this context, it is intriguing that one subset of genes exhibit peak translational efficiency in the basal leaf segments (Fig 5A, bottom; Fig 5C, right), whereas another subset increases in translational efficiency right out to the leaf tip (Fig 5A and 5C). The former group is strongly enriched for "biogenesis" genes (RNA polymerase, ribosomes, assembly factors), and the latter for photosynthesis genes. Possible mechanisms underlying these distinct "translational regulons" are discussed below.

A study in Chlamydomonas showed that changes in chloroplast mRNA abundance are not reflected by corresponding changes in rates of protein synthesis, leading to the conclusion that translation is the primary rate-limiting step [76]. The data presented here suggest that this is not the case in maize chloroplasts. The developmental shifts in mRNA abundance were largely mirrored by shifts in translational output (Fig 5A and 5B), implying that mRNA abundance has considerable impact on the output of most chloroplast genes in maize. Likewise, chloroplast DNA copy number limits gene expression in developing maize chloroplasts [77] but does not limit gene expression in Chlamydomonas chloroplasts [76]. It is perhaps unsurprising that mechanisms of gene regulation have diverged in the chloroplasts of vascular plants and single-celled algae, given their very different developmental and ecological contexts.

Mechanisms underlying developmental shifts in the chloroplast transcriptome and translatome

Our data revealed a strong correlation between gene function and the developmental dynamics of mRNA abundance (Fig 5C middle): mRNAs encoding proteins involved in gene expression generally peak in abundance earlier in development than do those encoding components of the photosynthetic apparatus. This finding was foreshadowed by analyses of several chloroplast mRNAs during leaf development in barley and Arabidopsis [78–80]. Land plant chloroplasts harbor two types of RNA polymerase, a single-subunit nucleus-encoded polymerase (NEP) and a bacterial-type plastid-encoded polymerase (PEP) [5]. The ratio of NEP to PEP drops precipitously during chloroplast development, and this likely makes a large contribution to the changes in chloroplast mRNA pools [5, 78, 80, 81]. There is evidence that NEP plays an especially important role in the transcription of "house keeping" genes, and PEP in the transcription of photosynthesis genes [5]; however, most chloroplast genes can be transcribed by both NEP and PEP [82], and the degree to which each polymerase contributes to the transcription of each gene during the course of chloroplast development remains unknown. Chloroplasts harbor several nucleus-encoded sigma factors that target PEP to distinct promoters [83], and these provide an additional means to tune transcription rates in a developmental context.

Changes in RNA stability combine with changes in transcription to modulate mRNA pools during chloroplast development [78, 80, 81, 84, 85]. Determinants of chloroplast mRNA stability include various ribonucleases, RNA structure, ribosome occupancy, and proteins that protect RNAs from nuclease attack [8]. Most mRNA termini in chloroplasts are protected by helical repeat RNA binding proteins that provide a steric blockade to exoribonucleases [8, 61, 62, 86]. The majority of such proteins belong to the pentatricopeptide repeat (PPR) family, a

large family of sequence-specific RNA binding proteins that influence virtually every post-transcriptional step in gene expression in mitochondria and chloroplasts [87]. In addition, chloroplasts harbor abundant hnRNP-like proteins, and these have been shown to impact the stability of several chloroplast mRNAs [88, 89]. Programmed changes in the abundance and/or activities of PPR and hnRNP-like proteins might contribute to the shifting mRNA pools during the proplastid to chloroplast transition.

Changes in translational efficiency superimpose on changes in mRNA abundance to modulate the output of plastid genes during the transformation of proplastids into chloroplasts. ORFs encoding proteins involved in photosynthesis generally exhibit maximal translational efficiency in young or mature chloroplasts (segments 9 and 14), whereas those that function in gene expression generally peak in translational efficiency earlier in development (Fig 5C right). Furthermore, our data suggest that mRNAs encoding PSII reaction center proteins are translated with higher efficiency in mesophyll chloroplasts than in bundle sheath chloroplasts (Fig 6C). It will be interesting to explore the mechanisms that underlie these differential effects on translational efficiency. Some possibilities include shifts in stromal pH, Mg^{++}, or the polymerase generating the mRNA (NEP versus PEP), which might impact the formation of RNA structures at specific ribosome binding sites. Programmed changes in the activities of nucleus-encoded gene-specific translational activators could modulate translational efficiencies in a developmental context. Most such proteins in land plant chloroplasts are PPR (or PPR-like) proteins, and several of these also stabilize processed mRNAs with a 5' end at the 5' boundary of their binding site [15, 30, 60–62, 64–67, 90–93]. Indeed, many polycistronic transcripts in chloroplasts are processed to smaller isoforms whose ends are defined and stabilized by PPR-like proteins [7, 8, 86]. The impact of this type of RNA processing on translational efficiencies *in vivo* remains unclear. The removal of upstream ORFs is not required for the translation of several ORFs that are found on processed RNAs with a proximal 5'-terminus [35, 36]. Some proteins have dual translation activation and RNA processing/stabilization functions, implying that the two activities are coupled [15, 30, 60–62, 65, 66, 91–93]; however, the translation activation and RNA processing/ stabilization effects of such proteins could be independent consequences of their binding upstream of an ORF [62, 86]. We showed here that there is a correlation between the accumulation of processed RNA isoforms and changes in relative translational efficiencies in two polycistronic transcription units (S8 Fig). Deciphering the cause and effect relationships underling these correlations presents a challenge for the future.

The data presented here lead to numerous new questions for future exploration. Is the synthesis of nucleus-encoded subunits of photosynthetic complexes tuned to that of their chloroplast-encoded partners? What is the mechanistic basis for the preferential translation of some mRNAs in developing chloroplasts and others in photosynthetic chloroplasts? To what extent do environmental inputs such as light and temperature modify the developmental dynamics of chloroplast mRNA abundance and translation? The use of ribosome profiling can be anticipated to accelerate progress in addressing these and many other long-standing questions relating to the biology of organelles.

Materials and Methods

Plant material

For the developmental analysis, *Zea mays* (inbred line B73) was grown under diurnal cycles for 9 days and harvested as described [12]. Leaf sections from twelve plants were pooled for each of three replicates; each pool contained between 0.15 g and 0.3 g tissue. Plants used to prepare mesophyll and bundle sheath fractions were grown similarly, except the light was set at 300 µmol·m^{-2}·s^{-1} and the tissues were harvested 13 days after planting, 2 hours into the light

cycle. The apical one-third of leaf two and three were pooled from fifteen seedlings for each replicate, and the bundle sheath and mesophyll-enriched fractions were obtained with a rapid mechanical procedure. The tissue was cut into ~ 1 cm-sections, placed in a pre-chilled mortar and pestle, and lightly ground for 2 min in 5 ml of ice-cold modified polysome extraction buffer lacking detergents (0.2 M sucrose, 0.2 M KCl, 50 mM Tris-acetate, pH 8.0, 15 mM MgCl$_2$, 20 mM 2-mercaptoethanol, 2 µg/ml pepstatin A, 2 µg/ml leupeptin, 2 mM phenyl-methanesulfonyl fluoride, 100 µg/ml chloramphenicol, 100 µg/ml cycloheximide). The material that was released into solution constituted the mesophyll cell-enriched fraction. A portion of this was frozen in liquid N$_2$ for RNA isolation, and the remainder was stored on ice while bundle sheath strands were purified from the tissue remaining in the mortar. The tissue was subject to four additional rounds of light grinding (2 min), each time in a fresh aliquot of 5-ml modified polysome extraction buffer. The light green fibers remaining constituted the bundle sheath enriched fraction; these cells were broken by hard grinding in 5 ml of modified polysome extraction buffer. A portion of this material was flash frozen for future RNA isolation and the remainder was used immediately for ribosome footprint isolation. Polyoxyethylene (10) tridecyl ether and Triton X-100 were added to the mesophyll and bundle sheath fractions retained for ribosome profiling (final concentrations of 2% and 1%, respectively), and the material was filtered through glass wool. The isolation of ribosome footprints and total RNA were performed as described below.

Preparation of ribosome footprints

Ribosome footprints were prepared using a protocol similar to that described in [15], but with two key modifications: (i) RNAse I rather than micrococcal nuclease was used to generate monosomes, and (ii) the centrifugation time used to pellet ribosomes through the sucrose cushion was shortened to reduce contamination by other RNPs. Tissues were pulverized in liquid N$_2$ with a mortar and pestle, and thawed in 5 ml of polysome extraction buffer (0.2 M sucrose, 0.2 M KCl, 50 mM Tris-acetate, pH 8.0, 15 mM MgCl$_2$, 20 mM 2-mercaptoethanol, 2% polyoxyethylene (10) tridecyl ether, 1% Triton X-100, 100 µg/ml chloramphenicol, 100 µg/ml cycloheximide). A 2.4-ml aliquot was removed and frozen in liquid N$_2$ for total RNA isolation. The remaining suspension was filtered through glass wool and centrifuged at 15,000xg for 10 min. The supernatant was digested with 3,500 units of RNAse I (Ambion) at 23°C for 30 min. 2.5 ml lysate was layered on a 2 ml sucrose cushion (1 M sucrose, 0.1 M KCl, 40 mM Tris-acetate, pH 8.0, 15 mM MgCl$_2$, 10 mM 2-mercaptoethanol, 100 µg/ml chloramphenicol, and 100 µg/ml cycloheximide) in a 16 x 76 mm tube and centrifuged in a Type 80 Ti rotor for 1.5 h at 55,000 rpm. The pellet was dissolved in 0.7 mL of ribosome dissociation buffer (10 mM Tris-Cl, pH 8.0, 10 mM EDTA, 5 mM EGTA, 100 mM NaCl, 1% SDS). RNA was isolated with Tri reagent (Molecular Research Center). RNAs between ~20 and ~35 nt were purified on a denaturing polyacrylamide gel, eluted, extracted with phenol/chloroform, precipitated with ethanol, and suspended in water. We have subsequently modified our protocol to purify RNAs between 20 and 40 nt; this results in a small shift in the size distribution of the reads (S1B Fig).

Preparation of sequencing libraries

The ribosome footprint preparation was treated with T4 polynucleotide kinase. Twenty ng of the kinased RNA was converted to a sequencing library using the NEXTflex Small RNA Sequencing Kit v2 (Bioo Scientific), which minimizes ligation bias by introducing four randomized bases at the 3' ends of the adapters [17]. rRNA fragments were depleted by subtractive hybridization after first-strand cDNA synthesis, using 54 biotinylated DNA oligonucleotides corresponding to the most abundant rRNA fragments detected in pilot experiments (see S1 Table). 10 µl of the

oligonucleotide mixture (concentrations as in S1 Table) was added to 40-μl of the first-strand synthesis reaction and heated to 95°C for 2 min. A 50-μl aliquot of pre-warmed 2X hybridization buffer (10 mM Tris-Cl pH 7.5, 1 mM EDTA, 2 M NaCl) was added and incubated at 55°C for 30 min. The solution was transferred to a new tube containing 1 mg of prewashed Dynabeads M-270 Streptavidin (Invitrogen) and incubated at room temperature for 15 min with frequent agitation. The tube was placed on a magnet for 5 min and the supernatant was collected and desalted using Sephadex G-25 Fine (GE Healthcare). The sample was concentrated to 18 μl and used as input for the PCR amplification step in the library construction protocol. After 14 cycles, PCR products were separated by electrophoresis through a 5% polyacrylamide gel and a gel slice corresponding to DNA fragments between markers at 147 and 180 bp (representing insert sizes of 20–53 bp) was excised. The DNA was eluted overnight, phenol/chloroform extracted, precipitated with ethanol, suspended in water, and stored at -20°C.

For RNA-seq, rRNA was depleted from the RNA samples using the Ribo-Zero rRNA Removal Kit (Plant Leaf) (Epicentre). One hundred ng of the rRNA-depleted RNA was used for library construction using the NEXTflex Rapid Directional qRNA-Seq Kit (Bioo Scientific) according the manufacturer's instructions. The adapters provided with the kit include 8-nt molecular labels that were used during data processing to remove PCR bias. The libraries were combined and sequenced using a HiSeq 2500 or NextSeq 500 instrument (Illumina). The read lengths were 50 or 75 nt for Ribo-seq and 75 nt for mRNA-seq.

Sequence read processing, alignment, and analysis

Adapter sequences were trimmed using cutadapt [94]. Ribo-seq reads between 18 and 40 nt were used as input for alignments. Alignments were performed using Bowtie 2 with default parameters [95], which permits up to 2 mismatches, thereby allowing edited sequences to align. Reads were aligned to the following gene sets, with unaligned reads from each step used as input for the next round of alignment: (i) chloroplast tRNA and rRNA; (ii) chloroplast genome; (iii) mitochondrial tRNA and rRNA; (iv) mitochondrial genome (B73 AGP v3); (v) nuclear tRNA and rRNA; nuclear genome (B73 AGP v3).

Maize genome annotation 6a (phytozome.jgi.doe.gov) was reduced to the gene set annotated in 5b+ (60,211 transcripts) (gramene.org). For metagene analysis, all coding sequence (CDS) coordinates from all transcript variants were combined to make a union CDS coordinate. Custom Perl scripts extracted mapping information using SAMtools [96] and analyzed mapped reads as follows. The distribution of ribosome footprint lengths and the RPKM for both the Ribo-seq and RNA-seq data were calculated based only on reads mapping to CDS regions. For RPKM calculations, we defined the total number of mapped reads as the number of reads mapping to nuclear CDSs. Translation efficiency was calculated from the division of ribosome footprint RPKM by RNA-seq RPKM.

Because unspliced RNAs constitute a substantial fraction of the RNA pool from intron-containing genes in chloroplasts, these genes require special treatment to infer the abundance of spliced (functional) mRNA. The fraction spliced at each intron was calculated in several ways. (i) RNA-seq reads were aligned to the chloroplast genome with splicing-aware software TopHat 2.0.11 [97]. The number of reads spanning each exon-exon junction (spliced) was divided by the sum of spliced (exon-exon) and unspliced (exon 1-intron or intron-exon 2) reads; (ii) RNA-seq reads were aligned with Bowtie2 to a reference gene set that included both spliced and unspliced forms (100-nt on each side of each junction). The fraction of spliced RNA was calculated as for method (i); (iii) RNA-seq reads were aligned with TopHat to the genome and the spliced fraction was calculated from (exon RPKM—intron RPKM)/exon RPKM. Values calculated by each method are provided in S3 Table. Summary plots report

mRNA abundance and translational efficiencies only for genes for which all of these methods gave similar results. We cannot confidently infer the amount of fully spliced RNAs from genes with two introns (*ycf3* and *rps12)*, so these are also excluded.

Hierarchical clustering was performed using the Bioinformatic Toolbox of MATLAB software (Mathworks) using standardized values as input: values from the four leaf segments for each gene were standardized to have a mean of 0 and a standard deviation of 1 such that developmental shifts can be compared among genes despite differences in signal magnitude. Hierarchical clustering was performed using Pearson correlation coefficient values and unweighted average distance.

Antibodies

Antibodies to AtpB, D1 and PetD were raised by our group and have been described previously [30]. Antibodies to Atp6, NdhH, PPDK, and Rpl2 were generously provided by Christine Chase (University of Florida), Tsuyoshi Endo (Kyoto University), Kazuko Aoyagi (UC Berkeley), and Alap Subramanian (University of Arizona), respectively. Antibodies to PEPC, malic enzyme, and RbcL were generous gifts of William Taylor (University of California, Berkeley).

Accession numbers

Illumina read sequences were deposited at the NCBI Sequence Read Archive with accession number SRP070787. Alignments of reads to the maize chloroplast genome used Genbank accession X86563.

Supporting Information

S1 Fig. Optimized ribosome purification reduces contamination by non-ribosomal ribonucleoprotein particles (RNPs). (A) Immunoblots showing that a marker for chloroplast ribosomes (Rpl2) was highly enriched in the pellet after sedimentation of nuclease-treated extract through a sucrose cushion, whereas a subunit of a ~600 kDa group II intron RNP (CFM2) [16] remained in the supernatant. An equal proportion of the starting material, the supernatant above the sucrose cushion, and the pellet fraction was analyzed. **(B)** Comparison of size distribution of chloroplast ribosome footprints resulting from two different size selection strategies. The experiments in this study used gel purified RNA fragments between approximately 20 and 35-nt (green). A pilot experiment used gel-purified RNA fragments between approximately 20 and 40-nt (blue). The size distribution of the sequence reads was nonetheless similar.
(TIF)

S2 Fig. Correlation of data among biological replicates. Pearson correlation coefficients for each sample pair combination were calculated using $\log_{10}$ of RPKM values for each protein-coding gene in the chloroplast genome. The correlation coefficients were used as the input for hierarchical clustering. The replicate number of each leaf segment sample is indicated after the hyphen. **(A)** Leaf segment data. **(B)** Bundle sheath (BS) and mesophyll (M) data.
(TIF)

S3 Fig. Summary of read counts per chloroplast gene. The values displayed are the mean from replicate assays. The identities of genes with low read counts are indicated. The *petN* mRNA is under represented in the RNA-seq data due to its small size, which is below the cut-off used for library preparation. **(A)** Read counts/gene for leaf gradient samples. **(B)** Read counts/gene for bundle sheath (BS) and mesophyll (M) samples.
(TIF)

S4 Fig. Characteristics of ribosome footprints in each compartment. The data plotted come from the twelve leaf gradient samples. **(A)** Three-nucleotide periodicity of P site position as a function of ribosome footprint length in the plastid, cytosol, and mitochondrion. Values shown are the mean ± SEM. The position of the P site in each footprint was inferred from the footprint size distributions at start/stop codons (Fig 2E and S4B-C). **(B)** Placement of cytosolic ribosome footprints with respect to the A and P sites of the ribosome based on reads aligning to start and stop codons. A diagram of the placement of 31-nucleotide cytosolic ribosome footprints is shown below. **(C)** Placement of mitochondrial ribosome footprints with respect to the A and P sites of the ribosome based on reads aligning to start and stop codons. A diagram of the placement of 28-nucleotide mitochondrial ribosome footprints is shown below.
(TIF)

S5 Fig. Genome-wide views of the plastid translatome and transcriptome in Segments 1 and 14 displayed using different Y-axis scales. The data are expressed as reads per kilobase per million reads mapping to nuclear genome coding sequences (RPKM). Values are the mean ± SEM from three replicates. Intron-containing genes are marked with a superscript *i*. Genes for which RNA levels and translational efficiency were not determined (n.d.) are marked (#). These include intron-containing genes for which the fraction of reads derived from spliced transcripts is uncertain, and *petN*, whose short mRNA is not represented quantitatively in the RNA-seq data. Genes encoding assembly factors are marked with asterisks. **(A)** Translational output, RNA abundance and translational efficiency displayed according to native gene order. Co-transcribed genes are marked with arrows that indicate the direction of transcription. **(B)** Translational output and RNA abundance displayed according to gene function. Genes encoding assembly factors are demarcated from the structural genes with dashed lines.
(TIF)

S6 Fig. Data for translational output, mRNA abundance and translational efficiency parsed according to gene function. The data are expressed as reads per kilobase per million reads mapping to nuclear coding sequences (RPKM). Values represent the mean ± SEM from three replicates. Intron-containing genes are marked with a superscript *i*. Genes for which RNA levels and translational efficiency were not calculated are marked with a hashtag (#). These include intron-containing genes for which the fraction of reads derived from spliced transcripts is uncertain, and *petN* whose very short mRNA is not represented quantitatively in the RNA-seq data due to the library protocol. n.d., not determined. **(A)** Genes related to cytochrome b_6f function. The *pet* genes encode cytochrome b_6f subunits and the *ccsA* gene encodes a protein involved in heme attachment [98]. **(B)** Genes encoding subunits of the NDH complex. **(C)** Genes encoding ribosomal proteins.
(TIF)

S7 Fig. Heat map representation of the results of hierarchical clustering of plastid genes according to their developmental dynamics. Clusters were generated independently from the data for translational output, mRNA level, and translational efficiency. The genes in each cluster are shown in Fig 5A.
(TIF)

S8 Fig. Developmental dynamics of processed mRNA isoforms from two polycistronic transcription units whose genes exhibit distinct developmental shifts in translational efficiency. The transcription units and probes used for RNA gel blot hybridizations are shown at top. Lanes contain an equal mass of total RNA, as illustrated by the methylene blue-stained blots shown below. RNAs were extracted from aliquots of the leaf lysates used for the ribosome footprint preparation prior to RNAse I addition, and therefore suffered slight degradation. The

rps14 gene is represented by several small transcripts that are over-represented with respect to the precursor at early developmental stages. The *psaI* gene is represented on a monocistronic mRNA that accumulates preferentially in segments 4 and 9 and on the polycistronic primary transcript.
(TIF)

S1 Table. List of DNA oligonucleotides used in the rRNA depletion step of Ribo-seq library preparation.
(XLSX)

S2 Table. Mapping statistics of leaf segment data.
(XLSX)

S3 Table. RNA-seq data at splice junctions of plastid genes.
(XLSX)

S4 Table. Values for translational output, mRNA level and translational efficiency from the leaf segment data.
(XLSX)

S5 Table. Hierarchical clustering of the leaf segment data.
(XLSX)

S6 Table. Values for translational output, mRNA level and translational efficiency from the mesophyll and bundle sheath data.
(XLSX)

S7 Table. Ribo-seq and RNA-seq data at sites of plastid RNA editing.
(XLSX)

Acknowledgments

Author Contributions

Conceived and designed the experiments: PC AB. Performed the experiments: PC. Analyzed the data: PC AB. Contributed reagents/materials/analysis tools: PC. Wrote the paper: PC AB.

References

1. Jarvis P, Lopez-Juez E (2013) Biogenesis and homeostasis of chloroplasts and other plastids. Nat Rev Mol Cell Biol 14:787–802. doi: 10.1038/nrm3702 PMID: 24263360

2. Majeran W, Friso G, Ponnala L, Connolly B, Huang M, Reidel E, et al. (2010) Structural and metabolic transitions of C4 leaf development and differentiation defined by microscopy and quantitative proteomics in maize. Plant Cell 22:3509–42. doi: 10.1105/tpc.110.079764 PMID: 21081695

3. Lyska D, Meierhoff K, Westhoff P (2013) How to build functional thylakoid membranes: from plastid transcription to protein complex assembly. Planta 237:413–28. doi: 10.1007/s00425-012-1752-5 PMID: 22976450

4. Bock R. Structure, function, and inheritance of plastid genomes. In: Bock R, editor. Cell and molecular biology of plastids. Heidelberg: Springer-Verlag; 2007. p. 29–63.

5. Borner T, Aleynikova AY, Zubo YO, Kusnetsov VV (2015) Chloroplast RNA polymerases: Role in chloroplast biogenesis. Biochim Biophys Acta 1847:761–9. doi: 10.1016/j.bbabio.2015.02.004 PMID: 25680513

6. Sugiura M (2014) Plastid mRNA translation. Methods Mol Biol 1132:73–91. doi: 10.1007/978-1-62703-995-6_4 PMID: 24599847

7. Barkan A (2011) Expression of plastid genes: organelle-specific elaborations on a prokaryotic scaffold. Plant Physiol 155:1520–32. pp.110.171231 [pii] doi: 10.1104/pp.110.171231 PMID: 21346173

8. Germain A, Hotto AM, Barkan A, Stern DB (2013) RNA processing and decay in plastids. Wiley Interdiscip Rev RNA 4:295–316. doi: 10.1002/wrna.1161 PMID: 23536311

9. Li GW, Burkhardt D, Gross C, Weissman JS (2014) Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell 157:624–35. doi: 10.1016/j.cell.2014.02.033 PMID: 24766808

10. Quax TE, Wolf YI, Koehorst JJ, Wurtzel O, van der Oost R, Ran W, et al. (2013) Differential translation tunes uneven production of operon-encoded proteins. Cell Rep 4:938–44. doi: 10.1016/j.celrep.2013.07.049 PMID: 24012761

11. Leech R, Rumsby M, Thompson W (1973) Plastid differentiation, acyl lipid and fatty acid changes in developing green maize leaves. Plant Physiol 52:240–5. PMID: 16658539

12. Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, et al. (2010) The developmental dynamics of the maize leaf transcriptome. Nat Genet 42:1060–7. ng.703 [pii] doi: 10.1038/ng.703 PMID: 21037569

13. Wang L, Czedik-Eysenberg A, Mertz RA, Si Y, Tohge T, Nunes-Nesi A, et al. (2014) Comparative analyses of C(4) and C(3) photosynthesis in developing leaves of maize and rice. Nat Biotechnol 32:1158–65. doi: 10.1038/nbt.3019 PMID: 25306245

14. Friso G, Majeran W, Huang M, Sun Q, van Wijk KJ (2010) Reconstruction of metabolic pathways, protein expression, and homeostasis machineries across maize bundle sheath and mesophyll chloroplasts: large-scale quantitative proteomics using the first maize genome assembly. Plant Physiol 152:1219–50. doi: 10.1104/pp.109.152694 PMID: 20089766

15. Zoschke R, Watkins K, Barkan A (2013) A rapid microarray-based ribosome profiling method elucidates chloroplast ribosome behavior in vivo Plant Cell 25:2265–75. doi: 10.1105/tpc.113.111567 PMID: 23735295

16. Asakura Y, Barkan A (2007) A CRM domain protein functions dually in group I and group II intron splicing in land plant chloroplasts. Plant Cell 19:3864–75. PMID: 18065687

17. Jayaprakash AD, Jabado O, Brown BD, Sachidanandam R (2011) Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. Nucleic Acids Res 39:e141. doi: 10.1093/nar/gkr693 PMID: 21890899

18. Lareau LF, Hite DH, Hogan GJ, Brown PO (2014) Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. Elife 3:e01257. doi: 10.7554/eLife.01257 PMID: 24842990

19. Rooijers K, Loayza-Puch F, Nijtmans LG, Agami R (2013) Ribosome profiling reveals features of normal and disease-associated mitochondrial translation. Nat Commun 4:2886. doi: 10.1038/ncomms3886 PMID: 24301020

20. Chung BY, Hardcastle TJ, Jones JD, Irigoyen N, Firth AE, Baulcombe DC, et al. (2015) The use of duplex-specific nuclease in ribosome profiling and a user-friendly software package for Ribo-seq data analysis. RNA. doi: 10.1261/rna.052548.115

21. Juntawong P, Girke T, Bazin J, Bailey-Serres J (2014) Translational dynamics revealed by genome-wide profiling of ribosome footprints in Arabidopsis. Proc Natl Acad Sci U S A 111:E203–12. doi: 10.1073/pnas.1317811111 PMID: 24367078

22. Lei L, Shi J, Chen J, Zhang M, Sun S, Xie S, et al. (2015) Ribosome profiling reveals dynamic translational landscape in maize seedlings under drought stress. Plant J. doi: 10.1111/tpj.13073

23. Woolstenhulme CJ, Guydosh NR, Green R, Buskirk AR (2015) High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. Cell Rep 11:13–21. doi: 10.1016/j.celrep.2015.03.014 PMID: 25843707

24. Martens AT, Taylor J, Hilser VJ (2015) Ribosome A and P sites revealed by length analysis of ribosome profiling data. Nucleic Acids Res 43:3680–7. doi: 10.1093/nar/gkv200 PMID: 25805170

25. Ingolia NT, Lareau LF, Weissman JS (2011) Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell 147:789–802. doi: 10.1016/j.cell.2011.10.002 PMID: 22056041

26. Jarvi S, Suorsa M, Aro EM (2015) Photosystem II repair in plant chloroplasts—Regulation, assisting proteins and shared components with photosystem II biogenesis. Biochim Biophys Acta 1847:900–9. doi: 10.1016/j.bbabio.2015.01.006 PMID: 25615587

27. Ponnala L, Wang Y, Sun Q, van Wijk KJ (2014) Correlation of mRNA and protein abundance in the developing maize leaf. Plant J 78:424–40. doi: 10.1111/tpj.12482 PMID: 24547885

28. Vollmar M, Schlieper D, Winn M, Buchner C, Groth G (2009) Structure of the c14 rotor ring of the proton translocating chloroplast ATP synthase. J Biol Chem 284:18228–35. doi: 10.1074/jbc.M109.006916 PMID: 19423706

29. Groth G, Pohl E (2001) The structure of the chloroplast F1-ATPase at 3.2 A resolution. J Biol Chem 276:1345–52. doi: 10.1074/jbc.M008015200 PMID: 11032839

30. Pfalz J, Bayraktar O, Prikryl J, Barkan A (2009) Site-specific binding of a PPR protein defines and stabi-lizes 5' and 3' mRNA termini in chloroplasts. EMBO J 28:2042–52. doi: 10.1038/emboj.2009.121 PMID: 19424177

31. Boudreau E, Takahashi Y, Lemieux C, Turmel M, Rochaix JD (1997) The chloroplast ycf3 and ycf4 open reading frames of Chlamydomonas reinhardtii are required for the accumulation of the photosys-tem I complex. EMBO J 16:6095–104. doi: 10.1093/emboj/16.20.6095 PMID: 9321389

32. Krech K, Ruf S, Masduki FF, Thiele W, Bednarczyk D, Albus CA, et al. (2012) The plastid genome-encoded Ycf4 protein functions as a nonessential assembly factor for photosystem I in higher plants. Plant Physiol 159:579–91. doi: 10.1104/pp.112.196642 PMID: 22517411

33. Ruf S, Kossel H, Bock R (1997) Targeted inactivation of a tobacco intron-containing open reading frame reveals a novel chloroplast-encoded photosystem I-related gene. J Cell Biol 139:95–102. PMID: 9314531

34. Torabi S, Umate P, Manavski N, Plochinger M, Kleinknecht L, Bogireddi H, et al. (2014) PsbN is required for assembly of the photosystem II reaction center in Nicotiana tabacum. Plant Cell 26:1183–99. doi: 10.1105/tpc.113.120444 PMID: 24619613

35. Barkan A (1988) Proteins encoded by a complex chloroplast transcription unit are each translated from both monocistronic and polycistronic mRNAs. EMBO J 7:2637–44. PMID: 2460341

36. Zoschke R, Barkan A (2015) Genome-wide analysis of thylakoid-bound ribosomes in maize reveals principles of cotranslational targeting to the thylakoid membrane. Proc Natl Acad Sci U S A 112: E1678–87. doi: 10.1073/pnas.1424655112 PMID: 25775549

37. Barkan A (1989) Tissue-dependent plastid RNA splicing in maize: Transcripts from four plastid genes are predominantly unspliced in leaf meristems and roots. Plant Cell 1:437–45. PMID: 2562564

38. Sheen J-Y, Bogorad L (1988) Differential expression in bundle sheath and mesophyll cells of maize genes for photosystem II components encoded by the plastid genome. Plant Physiol 86:1020–6. PMID: 16666025

39. Kubicki A, Steinmuller K, Westhoff P (1994) Differential transcription of plastome-encoded genes in the mesophyll and bundle-sheath chloroplasts of the monocotyledonous NADP-malic enzyme-type C4 plants maize and Sorghum. Plant Mol Biol 25:669–79. PMID: 8061319

40. Westhoff P, Offermann-Steinhard K, Hofer M, Eskins K, Oswald A, Streubel M (1991) Differential accu-mulation of plastid transcripts encoding photosystem II components in the mesophyll and bundle-sheath cells of monocotyledonous NADP-malic enzyme-type C4 plants. Planta 184:377–88. doi: 10.1007/BF00195340 PMID: 24194156

41. Sharpe RM, Mahajan A, Takacs EM, Stern DB, Cahoon AB (2011) Developmental and cell type charac-terization of bundle sheath and mesophyll chloroplast transcript abundance in maize. Curr Genet 57:89–102. doi: 10.1007/s00294-010-0329-8 PMID: 21152918

42. Shikanai T (2015) RNA editing in plants: Machinery and flexibility of site recognition. Biochim Biophys Acta 1847:779–85. doi: 10.1016/j.bbabio.2014.12.010 PMID: 25585161

43. Takenaka M, Zehrmann A, Verbitskiy D, Hartel B, Brennicke A (2013) RNA editing in plants and its evo-lution. Annu Rev Genet 47:335–52. doi: 10.1146/annurev-genet-111212-133519 PMID: 24274753

44. Phreaner CG, Williams MA, Mulligan RM (1996) Incomplete editing of rps12 transcripts results in the synthesis of polymorphic polypeptides in plant mitochondria. Plant Cell 8:107–17. PMID: 8597655

45. Lu B, Wilson RK, Phreaner CG, Mulligan RM, Hanson MR (1996) Protein polymorphism generated by differential RNA editing of a plant mitochondrial rps12 gene. Mol Cell Biol 16:1543–9. PMID: 8657128

46. del Campo EM, Sabater B, Martin M (2000) Transcripts of the ndhH-D operon of barley plastids: possi-ble role of unedited site III in splicing of the ndhA intron. Nucleic Acids Res 28:1092–8. PMID: 10666448

47. Schmitz-Linneweber C, Tillich M, Herrmann RG, Maier RM (2001) Heterologous, splicing-dependent RNA editing in chloroplasts: allotetraploidy provides trans-factors. EMBO J 20:4874–83. doi: 10.1093/emboj/20.17.4874 PMID: 11532951

48. Peeters N, Hanson M (2002) Transcript abundance supercedes editing efficiency as a factor in developmental variation of chloroplast gene expression. RNA 8:497–511. PMID: 11991643

49. Zandueta-Criado A, Bock R (2004) Surprising features of plastid ndhD transcripts: addition of non-encoded nucleotides and polysome association of mRNAs with an unedited start codon. Nucleic Acids Res 32:542–50. doi: 10.1093/nar/gkh217 PMID: 14744979

50. Chen X, Kindle KL, Stern DB (1995) The initiation codon determines the efficiency but not the site of translation initiation in Chlamydomonas chloroplasts. Plant Cell 7:1295–305. doi: 10.1105/tpc.7.8.1295 PMID: 7549485

51. Zito F, Kuras R, Choquet Y, Kossel H, Wolman F (1997) Mutations of cytochrome b6 in Chlamydomonas reinhardtii disclose the funcitonal significance for proline to leucine conversion by petB editing in maize and tobacco. Plant Mol Biol 33:79–86. PMID: 9037161

52. Lu B, Hanson MR (1994) A single homogeneous form of ATP6 protein accumulates in petunia mitochondria despite the presence of differentially edited atp6 transcripts. Plant Cell 6:1955–68. doi: 10.1105/tpc.6.12.1955 PMID: 7866035

53. Williams MA, Tallakson WA, Phreaner CG, Mulligan RM (1998) Editing and translation of ribosomal protein S13 transcripts: unedited translation products are not detectable in maize mitochondria. Curr Genet 34:221–6. PMID: 9745026

54. Brar GA, Weissman JS (2015) Ribosome profiling reveals the what, when, where and how of protein synthesis. Nat Rev Mol Cell Biol 16:651–64. doi: 10.1038/nrm4069 PMID: 26465719

55. Adam Z. Protein stability and degradation in plastids. In: Bock R, editor. Cell and molecular biology of plastids. Heidelberg: Springer-Verlag; 2007. p. 315–38.

56. Nelson CJ, Alexova R, Jacoby RP, Millar AH (2014) Proteins with high turnover rate in barley leaves estimated by proteome analysis combined with in planta isotope labeling. Plant Physiol 166:91–108. doi: 10.1104/pp.114.243014 PMID: 25082890

57. Choquet Y, Wollman F. The CES process. In: Stern D, Harris E, editors. The Chlamydomonas Sourcebook: Organellar and Metabolic Processes. 2 ed. Oxford: Academic Press; 2009. p. 1027–64.

58. Monde R, Zito F, Olive J, Wollman F, Stern D (2000) Post-transcriptional defects in tobacco chloroplast mutants lacking the cytochrome b6f complex. Plant J 21:61–72. PMID: 10652151

59. Scharff LB, Childs L, Walther D, Bock R (2011) Local Absence of Secondary Structure Permits Translation of mRNAs that Lack Ribosome-Binding Sites. PLoS Genet 7:e1002155. doi: 10.1371/journal.pgen.1002155 PGENETICS-D-11-00480 [pii]. PMID: 21731509

60. Barkan A, Walker M, Nolasco M, Johnson D (1994) A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. EMBO J 13:3170–81. PMID: 8039510

61. Hammani K, Cook W, Barkan A (2012) RNA binding and RNA remodeling activities of the Half-a-Tetratricopeptide (HAT) protein HCF107 underlie its effects on gene expression. PNAS 109:5651–6. doi: 10.1073/pnas.1200318109 PMID: 22451905

62. Prikryl J, Rojas M, Schuster G, Barkan A (2011) Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. Proc Natl Acad Sci USA 108:415–20. doi: 10.1073/pnas.1012076108 PMID: 21173259

63. Schmitz-Linneweber C, Williams-Carrier R, Barkan A (2005) RNA immunoprecipitation and microarray analysis show a chloroplast pentatricopeptide repeat protein to be associated with the 5'-region of mRNAs whose translation it activates. Plant Cell 17:2791–804. PMID: 16141451

64. Zoschke R, Kroeger T, Belcher S, Schottler MA, Barkan A, Schmitz-Linneweber C (2012) The Pentatricopeptide Repeat-SMR Protein ATP4 promotes translation of the chloroplast atpB/E mRNA. Plant J 72:547–58. doi: 10.1111/j.1365-313X.2012.05081.x PMID: 22708543

65. Felder S, Meierhoff K, Sane AP, Meurer J, Driemel C, Plucken H, et al. (2001) The nucleus-encoded HCF107 gene of Arabidopsis provides a link between intercistronic RNA processing and the accumulation of translation-competent psbH transcripts in chloroplasts. Plant Cell 13:2127–41. PMID: 11549768

66. Cai W, Okuda K, Peng L, Shikanai T (2011) PROTON GRADIENT REGULATION 3 recognizes multiple targets with limited similarity and mediates translation and RNA stabilization in plastids. Plant J 67:318–27. doi: 10.1111/j.1365-313X.2011.04593.x PMID: 21457370

67. Schult K, Meierhoff K, Paradies S, Toller T, Wolff P, Westhoff P (2007) The nuclear-encoded factor HCF173 is involved in the initiation of translation of the psbA mRNA in Arabidopsis thaliana. Plant Cell 19:1329–46. PMID: 17435084

68. Auchincloss A, Zerges W, Perron K, Girard-Bascou J, Rochaix J-D (2002) Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of the chloroplast psbC mRNA in Chlamydomonas reinhardtii. J Cell Biol 157:953–62. PMID: 12045185

69. Rahire M, Laroche F, Cerutti L, Rochaix JD (2012) Identification of an OPR protein involved in the translation initiation of the PsaB subunit of photosystem I. Plant J 72:652–61. doi: 10.1111/j.1365-313X. 2012.05111.x PMID: 22817760

70. Stampacchia O, Girard-Bascou J, Zanasco J-L, Zerges W, Bennoun P, Rochaix J-D (1997) A nuclear-encoded function essential for translation of the chloroplast *psaB* mRNA in Chlamydomonas. Plant Cell 9:773–82. PMID: 9165753

71. Schwarz C, Elles I, Kortmann J, Piotrowski M, Nickelsen J (2007) Synthesis of the D2 Protein of Photosystem II in Chlamydomonas Is Controlled by a High Molecular Mass Complex Containing the RNA Stabilization Factor Nac2 and the Translational Activator RBP40. Plant Cell 19:3627–39. PMID: 18055611

72. Boulouis A, Raynaud C, Bujaldon S, Aznar A, Wollman FA, Choquet Y (2011) The nucleus-encoded trans-acting factor MCA1 plays a critical role in the regulation of cytochrome f synthesis in Chlamydomonas chloroplasts. Plant Cell 23:333–49. tpc.110.078170 [pii] doi: 10.1105/tpc.110.078170 PMID: 21216944

73. Vaistij F, Boudreau E, Lemaire S, Goldschmidt-Clermont M, Rochaix J (2000) Characterization of Mbb1, a nucleus-encoded tetratricopeptide repeat protein required for expression of the chloroplast *psbB/psbT/psbH* gene cluster in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA 97:14813–8. PMID: 11121080

74. Zerges W, Girard-Bascou J, Rochaix JD (1997) Translation of the chloroplast psbC mRNA is controlled by interactions between its 5' leader and the nuclear loci TBC1 and TBC3 in Chlamydomonas reinhardtii. Mol Cell Biol 17:3440–8. PMID: 9154843

75. Zerges W, Rochaix J-D (1994) The 5' leader of a chloroplast mRNA mediates the translational requirements for two nucleus-encoded functions in *Chlamydomonas reinhardtii*. Mol Cell Biol 14:5268–77. PMID: 8035805

76. Eberhard S, Drapier D, Wollman F (2002) Searching limiting steps in the expression of chloroplast-encoded proteins: relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of *Chlamydomonas reinhardtii*. Plant J 31:149–60. PMID: 12121445

77. Udy DB, Belcher S, Williams-Carrier R, Gualberto JM, Barkan A (2012) Effects of reduced chloroplast gene copy number on chloroplast gene expression in maize. Plant Physiol 160:1420–31. doi: 10.1104/pp.112.204198 PMID: 22977281

78. Zoschke R, Liere K, Borner T (2007) From seedling to mature plant: arabidopsis plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. Plant J 50:710–22. TPJ3084 [pii] doi: 10.1111/j.1365-313X.2007.03084.x PMID: 17425718

79. Baumgartner BJ, Rapp JC, Mullet JE (1993) Plastid genes encoding the transcription/translation apparatus are differentially transcribed early in barley chloroplast development. Plant Physiol 101:781–91. PMID: 12231729

80. Emanuel C, Weihe A, Graner A, Hess WR, Borner T (2004) Chloroplast development affects expression of phage-type RNA polymerases in barley leaves. Plant J 38:460–72. doi: 10.1111/j.0960-7412. 2004.02060.x TPJ2060 [pii]. PMID: 15086795

81. Cahoon AB, Harris FM, Stern DB (2004) Analysis of developing maize plastids reveals two mRNA stability classes correlating with RNA polymerase type. EMBO Rep 5:801–6. doi: 10.1038/sj.embor. 7400202 PMID: 15258614

82. Zhelyazkova P, Sharma CM, Forstner KU, Liere K, Vogel J, Borner T (2012) The primary transcriptome of barley chloroplasts: numerous noncoding RNAs and the dominating role of the plastid-encoded RNA polymerase. Plant Cell 24:123–36. tpc.111.089441 [pii] doi: 10.1105/tpc.111.089441 PMID: 22267485

83. Chi W, He B, Mao J, Jiang J, Zhang L (2015) Plastid sigma factors: Their individual functions and regulation in transcription. Biochim Biophys Acta 1847:770–8. doi: 10.1016/j.bbabio.2015.01.001 PMID: 25596450

84. Baumgartner BJ, Rapp JC, Mullet JE (1989) Plastid Transcription Activity and DNA Copy Number Increase Early in Barley Chloroplast Development. Plant Physiol 89:1011–8. PMID: 16666609

85. Klaff P, Gruissem W (1991) Changes in chloroplast mRNA stability during leaf development. Plant Cell 3:517–29. PMID: 12324602

86. Zhelyazkova P, Hammani K, Rojas M, Voelker R, Vargas-Suarez M, Borner T, et al. (2012) Protein-mediated protection as the predominant mechanism for defining processed mRNA termini in land plant chloroplasts. Nucleic Acids Res 40:3092–105. gkr1137 [pii] doi: 10.1093/nar/gkr1137 PMID: 22156165

87. Barkan A, Small I (2014) Pentatricopeptide Repeat Proteins in Plants. Annu Rev Plant Biol 65:415–42. doi: 10.1146/annurev-arplant-050213-040159 PMID: 24471833

88. Nakamura T, Ohta M, Sugiura M, Sugita M (2001) Chloroplast ribonucleoproteins function as a stabilizing factor of ribosome-free mRNAs in the stroma. J Biol Chem 276:147–52. doi: 10.1074/jbc.M008817200 PMID: 11038367

89. Tillich M, Hardel SL, Kupsch C, Armbruster U, Delannoy E, Gualberto JM, et al. (2009) Chloroplast ribonucleoprotein CP31A is required for editing and stability of specific chloroplast mRNAs. Proc Natl Acad Sci U S A 106:6002–7. doi: 10.1073/pnas.0808529106 PMID: 19297624

90. Fisk DG, Walker MB, Barkan A (1999) Molecular cloning of the maize gene crp1 reveals similarity between regulators of mitochondrial and chloroplast gene expression. EMBO J 18:2621–30. PMID: 10228173

91. Sane AP, Stein B, Westhoff P (2005) The nuclear gene HCF107 encodes a membrane-associated R-TPR (RNA tetratricopeptide repeat)-containing protein involved in expression of the plastidial psbH gene in Arabidopsis. Plant J 42:720–30. TPJ2409 [pii] doi: 10.1111/j.1365-313X.2005.02409.x PMID: 15918885

92. Fujii S, Sato N, Shikanai T (2013) Mutagenesis of individual pentatricopeptide repeat motifs affects RNA binding activity and reveals functional partitioning of Arabidopsis PROTON gradient regulation3. Plant Cell 25:3079–88. doi: 10.1105/tpc.113.112193 PMID: 23975900

93. Yamazaki H, Tasaka M, Shikanai T (2004) PPR motifs of the nucleus-encoded factor, PGR3, function in the selective and distinct steps of chloroplast gene expression in Arabidopsis. Plant J 38:152–63. PMID: 15053768

94. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal 17:10–2.

95. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–9. doi: 10.1038/nmeth.1923 PMID: 22388286

96. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–9. doi: 10.1093/bioinformatics/btp352 PMID: 19505943

97. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25:1105–11. doi: 10.1093/bioinformatics/btp120 PMID: 19289445

98. Xie Z, Merchant S (1996) The plastid-encoded ccsA gene is required for heme attachment to chloroplast c-type cytochromes. J Biol Chem 271:4632–9. PMID: 8617725