Biochimica et Biophysica Acta 1847 (2015) 1004–1016

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbabio

Review

Large-scale genetic analysis of chloroplast biogenesis in maize☆

Susan Belcher [1], Rosalind Williams-Carrier [1], Nicholas Stiffler, Alice Barkan *

Institute of Molecular Biology, University of Oregon, Eugene, OR 97403, USA

ARTICLE INFO

Article history:
Received 30 January 2015
Accepted 16 February 2015
Available online 26 February 2015

Keywords:
Plastid
Photosynthesis
Maize
Arabidopsis
Mutator
Chloroplast

ABSTRACT

Background: Chloroplast biogenesis involves a collaboration between several thousand nuclear genes and ~100 genes in the chloroplast. Many of the nuclear genes are of cyanobacterial ancestry and continue to perform their ancestral function. However, many others evolved subsequently and comprise a diverse set of proteins found specifically in photosynthetic eucaryotes. Genetic approaches have been key to the discovery of nuclear genes that participate in chloroplast biogenesis, especially those lacking close homologs outside the plant kingdom. *Scope of Review:* This article summarizes contributions from a genetic resource in maize, the Photosynthetic Mutant Library (PML). The PML collection consists of ~2000 non-photosynthetic mutants induced by *Mu* transposons. We include a summary of mutant phenotypes for 20 previously unstudied maize genes, including genes encoding chloroplast ribosomal proteins, a PPR protein, tRNA synthetases, proteins involved in plastid transcription, a putative ribosome assembly factor, a chaperonin 60 isoform, and a NifU-domain protein required for Photosystem I biogenesis. *Major Conclusions:* Insertions in 94 maize genes have been linked thus far to visible and molecular phenotypes with the PML collection. The spectrum of chloroplast biogenesis genes that have been genetically characterized in maize is discussed in the context of related efforts in other organisms. This comparison shows how distinct organismal attributes facilitate the discovery of different gene classes, and reveals examples of functional divergence between monocot and dicot plants. *General Significance:* These findings elucidate the biology of an organelle whose activities are fundamental to agriculture and the biosphere. This article is part of a Special Issue entitled: Chloroplast Biogenesis.

1. Introduction

The chloroplast is a dynamic organelle whose ultrastructural and proteome complexity are comparable to those of free-living bacteria. Chloroplasts evolved from an endosymbiont of cyanobacterial ancestry, and were likely retained by their primordial host due to their photosynthetic capacity [reviewed in [1]]. Subsequently, massive gene transfer from the endosymbiont to the nucleus, integration of the organelle into host signaling and metabolic pathways, and coevolution of the nuclear and chloroplast genomes produced an organelle whose bacterial ancestry remains apparent but is embellished by numerous physiological and molecular novelties [reviewed in [2,3]].

The biogenesis of the photosynthetic apparatus in extant chloroplasts requires the coordinated expression of several thousand nuclear genes and ~100 chloroplast genes, followed by a complex series of protein targeting and assembly steps that lead to the elaboration of the thylakoid membrane system and the biogenesis of the multisubunit complexes that perform the core reactions of photosynthesis. Furthermore, the composition and activities of the photosynthetic apparatus adapt readily to environmental influences such as light and temperature. Mechanisms underlying the biogenesis and adaptation of the photosynthetic apparatus are, in general, poorly understood [reviewed in [4–6]]. An added layer of complexity arises in multicellular plants, in which chloroplasts belong to an organelle family, the plastids, that adopt different forms in different cell types [reviewed in [7]]. The differentiation of non-photosynthetic proplastids into chloroplasts occurs in conjunction with the differentiation of leaf cells from meristematic progenitors. Furthermore, two distinct photosynthetic cell types in C4 plants - mesophyll and bundle sheath - harbor chloroplasts with distinct morphologies, intracellular distribution, and enzymatic profiles [reviewed in [8]].

Chloroplast biogenesis and photosynthesis *per se* are nicely amenable to analysis by classical genetic approaches ("forward genetics") because defects in photosynthesis can be detected with simple screens and photosynthesis is dispensable when an alternative source of reduced carbon is provided. Despite the increasing ease of reverse-genetic approaches, forward genetics remains a powerful method for dissecting complex biological processes. For example, a screen based on chlorophyll fluorescence parameters yielded a rich harvest of nuclear

☆ This article is part of a Special Issue entitled: Chloroplast Biogenesis.

* Corresponding author. Tel.: +1 541 346 5145.
E-mail address: abarkan@uoregon.edu (A. Barkan).

[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.bbabio.2015.02.014
0005-2728/

genes functioning in the synthesis, assembly, and regulation of the thylakoid NADH dehydrogenase-like complex (NDH) [reviewed in [9]]. Screens for mutants with defects in plastid gene expression revealed functions of several "domains of unknown function", screens for the loss of specific photosynthetic complexes identified novel photosystem assembly factors, and screens for mutants with defects in non-photochemical quenching and state transitions screens provided important insights into mechanisms of photosynthetic acclimation [reviewed in [3,10–12]].

2. Maize as a model organism for the genetic analysis of chloroplast biogenesis

Phenotype-driven genetic analyses of chloroplast processes have concentrated on four organisms: the green alga *Chlamydomonas reinhardtii,* the C3 dicot *Arabidopsis thaliana*, the C3 monocot *Oryza sativa*, and the C4 monocot *Zea mays*. These organisms span a considerable phylogenetic distance and embed their chloroplasts in diverse developmental and physiological contexts. Each offers a distinct set of attributes that impacts which experimental approaches are most easily employed and the types of mutants that are recovered. *Chlamydomonas* is the only organism that allows the ready manipulation of both the nuclear and chloroplast genomes, but it cannot serve as a model for the many aspects of chloroplast biology that are specific to land plants. The expansive genomic tools available for *Arabidopsis* are unrivaled, and the ease of growing large numbers of mutants in a small space make *Chlamydomonas* and *Arabidopsis* well suited to high throughput metabolite and fluorescence-based screens [13]. On the other hand, the large seed reserves of maize and rice support rapid heterotrophic growth of non-photosynthetic mutants for several weeks without the need for specialized growth media. At eight days post germination, a non-photosynthetic maize seedling is typically 10 cm tall with a fresh weight of approximately 0.5 gm (see photographs in Figs. 1 through 5). This provides ready access to non-photosynthetic mutant tissue for biochemical analysis, and fosters comprehensive analyses of molecular phenotypes using methods that can be onerous in *Arabidopsis*. Maize and rice have proven to be particularly useful for studying mutations that either directly or indirectly cause the loss of plastid ribosomes: this condition causes embryo lethality in *Arabidopsis*, but typically yields albino seedlings in cereals whose molecular defects can be informative [reviewed in [14]]. This feature likely accounts for the fact that the nuclear gene complement involved in chloroplast RNA splicing in land plants has been elucidated primarily through genetic and biochemical approaches in maize (see below).

The use of maize for the genetic dissection of chloroplast processes was pioneered by Don Miles, who was the first to use "high chlorophyll fluorescence" (*hcf*) to screen for non-photosynthetic mutants in plants [15,16]. Miles initially screened EMS-mutagenized maize, but chemical mutagens were soon supplanted by the *Mutator* (*Mu*) transposon system as the mutagen of choice [17,18]. However, the high copy number of *Mu* transposons (~100 insertions per genome) hindered the assignment of causal insertions in *Mu* lines, and only one of the causal mutations in the Miles collection has been reported [17, 19]. This challenge of the *Mu* system was overcome with the recent development of high-throughput methods for sequencing *Mu* insertion sites and linking them to specific phenotypes [20,21].

3. Overview of the PML mutant collection

The PML collection was assembled as a tool to deeply sample the complement of nuclear genes required for the biogenesis of photosynthetically competent chloroplasts in plants. The collection consists of ~2000 independently arising mutants that were selected from *Mu*-active maize lines based on seedling chlorophyll deficiency (pale green, albino, yellow, virescent, etc.) or an *hcf* phenotype. The latter screen was abandoned early in the project because the vast majority of *hcf* mutants also have a visibly apparent reduction in chlorophyll. Most mutants in the PML collection are seedling lethal, and die at ~ three weeks post-germination when seed reserves are exhausted. The mutations are recessive and are propagated by crossing heterozygous plants. Genes that have been linked to phenotypes in the PML collection thus far are represented by an average of ~3.6 mutant alleles, suggesting that the collection is nearing saturation for the chlorophyll-deficiency screen that was used to assemble it. Based on allele frequencies to date, we estimate that the chloroplast phenotypes in the collection result from mutations in approximately 400 different genes.

The phenotypes used to assemble the PML collection are expected to capture genes involved in many aspects of chloroplast biogenesis and photosynthesis, including the synthesis, targeting, or assembly of subunits of Rubisco, Photosystem II (PSII), the cytochrome b_6f complex (cyt b_6f), Photosystem I (PSI), or the ATP Synthase (ATPase), the import of proteins into the chloroplast, the targeting of proteins to the thylakoid membrane, the synthesis of pigments, lipids, and prosthetic groups, and more. Certainly, the PML collection will not include every gene that functions in chloroplast biogenesis. For example, mutations that disrupt plastid division are unlikely to be represented, as they do not cause a detectable chlorophyll deficiency in *Arabidopsis* [22]. Also missing from the collection will be mutations in genes that are members of redundant gene pairs or pathways, genes that are essential for gametophyte or embryo viability, and genes whose function is superfluous under the growth conditions employed during screening.

Phenotypic data for mutants in the PML collection were collected in three stages:

(i) Pigmentation. Leaf pigmentation was qualitatively scored with a controlled vocabulary designed to capture nuances that provide clues about the underlying lesion. For example, mutants lacking the plastid translation machinery have "ivory" leaves, in contrast to the "white" phenotype caused by defects in carotenoid biosynthesis [reviewed in [18]]. In general, growth of non-photosynthetic mutants at high light intensities and low temperatures enhances chlorophyll loss. To maximize sensitivity, screens for new mutants were generally performed in a cool greenhouse during the winter. In some cases, pigment deficiencies are considerably more subtle when the same lines are grown under more moderate conditions.

(ii) Immunoblot analysis of photosynthetic complexes. Mutants that clearly accumulate some chlorophyll based on a qualitative visual inspection (approximately 60% of the mutants in the collection) were further analyzed by immunoblotting to quantify the abundance of core subunits of PSII, cyt b_6f, PSI, the thylakoid ATP synthase, and Rubisco. The abundance of a single "marker" subunit for each complex is generally a good predictor of the abundance of the complex as a whole due to the fact that improperly assembled subunits are generally unstable or inhibit the synthesis of partner subunits. Some mutants specifically lack a single complex, some have reduced levels of all of the monitored complexes, and others lack various subsets of complexes; these protein phenotypes provide initial clues about gene function that guide more detailed investigations. Albino mutants were not used for immunoblot analyses because they have pleiotropic losses of photosynthetic enzyme complexes that make the results uninformative.

(iii) RNA gel blot analysis of chloroplast RNAs. Approximately 500 of the mutants with protein deficiencies detected in the immunoblot assay were further analyzed by RNA gel blot hybridization, using probes to chloroplast genes encoding the missing proteins. Approximately 30 of these mutants were found to have aberrant chloroplast mRNA populations. This number under represents the proportion of mutants with underlying defects in chloroplast RNA metabolism because the probings were not comprehensive and because "ivory" mutants (a phenotype that often results from defects in plastid gene expression) were not examined in this way.

4. Linking phenotypes to genes in the PML collection

The PML collection is used for both phenotype-driven forward genetics and as a reverse-genetic resource to characterize genes suspected to be involved in particular chloroplast processes. The former approach initially involved laborious Southern-blot methods to identify *Mu* insertions that cosegregate with mutant phenotypes. Each such gene cloning project took approximately 1 person-year. However, this effort paid off in that it yielded the first proteins involved in chloroplast RNA splicing in plants (*crs1, crs2*) [23–25], the first molecular functions for a pentatricopeptide repeat (PPR) protein in plants (*crp1)* [26–28], and insights into the components and functions of the cpSec and cpTAT thylakoid targeting machineries (*hcf106, tha1, tha4*) [29–31]. Other groups employing the analogous approach to study chloroplast processes in maize were rewarded by the identification of the enigmatic ribosome-associated protein Iojap [32], a master transcriptional regulator of chloroplast development, GOLDEN2 [33], and the first assembly factor for Rubisco, BSD2 [34] (see Table 1).

The PML collection also provides an enriched source of mutations in genes relevant to chloroplast biology for use in reverse-genetic applications. Screens for insertions in genes-of-interest were initially performed one gene at a time, by PCR of pooled mutant DNAs. Mutants identified in this way were used to validate phenotype-driven gene identifications [25,35], establish functions for candidate splicing factors and thylakoid targeting components [36–39], and explore the functional repertoire of the PPR family [40–43].

The pace of linking genes to phenotypes in the PML collection was dramatically accelerated with the development of the "*Mu*-Illumina" method for sequencing all *Mu*-insertion sites in individual plants [20]. When used for phenotype-driven classical genetic analyses, causal mutations are identified by their cosegregation with the mutant phenotype after a simple two-generation crossing scheme. Eighteen such gene identifications are performed in a single Illumina lane, and the library preparation and data analysis take only a few weeks. The rate limiting steps in this process are the generation of the F2 ears from which the sequenced mutants are drawn, and subsequent validation by recovery of independent alleles. Thirteen genes identified in this manner have been reported since the development of the *Mu*-Illumina method; these are included in Table 1, along with all other published reports of maize genes whose mutation disrupts chloroplast biogenesis. The causal mutations underlying another set of mutants are reported here for the first time (Table 2), and document functions for an additional 20 maize genes. The genes identified in these forward genetic screens function in various aspects of chloroplast RNA metabolism, translation, DNA replication, protein targeting, transcription, and the assembly of photosynthetic complexes.

In a complementary approach, a reverse-genetic resource is being developed by using Illumina sequencing to catalog all *Mu* insertion sites in one individual representing each of the ~2000 mutants in the PML collection. This method does not identify which of the ~100 *Mu* insertions in the plant causes the chloroplast phenotype. However, a database of the sequenced insertions provides ready access to insertions in genes of interest for phenotypic analysis. Published studies describing genes whose functions have been elucidated by reverse genetics with the PML collection are summarized in Table 1. Additional mutants identified in this way are posted at http://pml.uoregon.edu/pml_table.php.

5. Previously unreported gene identifications made by forward-genetic analysis of PML mutants

Table 2 lists a subset of "new" maize genes whose phenotype-to-genotype assignments have been made by forward genetic *Mu*-Illumina analysis; genes that are under active study in our laboratory or collaborating laboratories are not included. The identity of the causal mutations were validated by complementation crosses between different mutant alleles, except in a few instances in which the molecular phenotypes match those reported for characterized orthologs in other organisms.

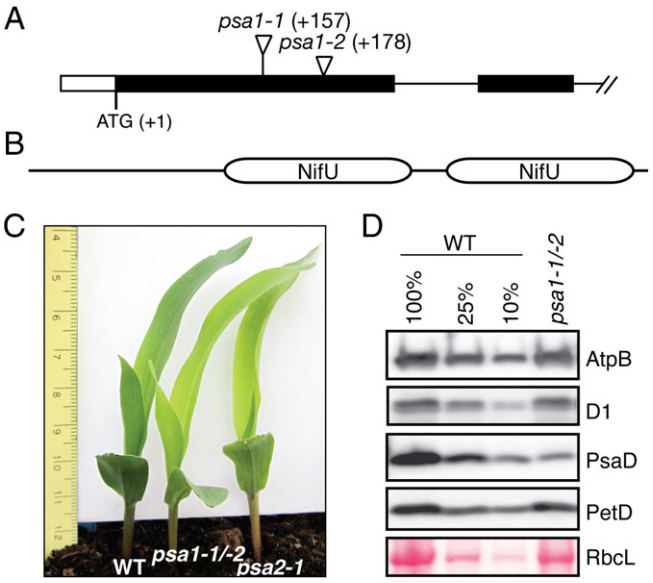

Fig. 1. Overview of *psa1* mutants. (A) Positions of the *Mu* insertions in the *psa1* gene. The nucleotide positions relative to the start codon are indicated. The insertion site sequences are provided in Supplementary Fig. 1. (B) Domain architecture of PSA1. (C) Phenotype of *psa1* mutants. A previously described mutant with a defect in PSI biogenesis, psa2, is shown for comparison. Plants were grown for nine days in soil. (D) Immunoblot profile of core subunits of photosynthetic enzyme complexes. A single blot was probed sequentially with antibodies to the indicated proteins. The Ponceau S stained blot below illustrates the abundance of RbcL, the large subunit of Rubisco.

This gene set includes genes encoding subunits of the photosynthetic apparatus (AtpC and PPDK), components of the chloroplast translation machinery (a cysteinyl and a leucyl tRNA synthetase), an RNA helicase involved in splicing and ribosome maturation (Zm-RH3) [44], proteins required for the activity of the plastid-encoded RNA polymerase (PEP) (Zm-PTAC14 and Zm-PRDA1) [45–47], the cytochrome biogenesis factors Zm-CCS1 and Zm-CCB4 [48–50], and proteins involved in plastoquinone synthesis (Zm-APG1) [51], porphyrin metabolism (Zm-Ycf54 and Zm-GUN4) [52–55]), assembly of photosynthetic enzyme complexes (Zm-Y3IP1 and chaperonin 60α1) [56,57], and chloroplast RNA stabilization/translation (Zm-PGR3) [58]. Molecular phenotypes of *cps1*, *cps2,* and *pet3* mutants had been reported long before the recent identification of their causal mutations. *cps1* mutants have a global defect in chloroplast translation [59], consistent with the finding that CPS1 encodes a cysteinyl-tRNA synthetase. *cps2* mutants specifically lack Rubisco [59], consistent with the finding that CPS2 encodes a Cpn60 isoform. *pet3* mutants specifically lack the *cyt* b_6f complex but synthesize its plastid-encoded subunits [60], consistent with the finding that PET3 encodes the maize ortholog of CCB4, which is required to attach heme to cytochrome b_6 [48,49]. The *Arabidopsis* orthologs of CPS1 and CPS6, a cysteinyl- and leucyl- tRNA synthetase, respectively, are dual targeted to chloroplasts and mitochondria [61]. However, the morphology and growth rate of the maize mutants are similar to those of other non-photosynthetic maize mutants, strongly suggesting that these tRNA synthetases in maize function solely, if not entirely, in chloroplasts.

Occasionally, phenotypes that do not relate to chloroplast biogenesis arise in our lines, and are included in the *Mu*-Illumina pipeline for gene identification. Two such gene identifications are included in Table 2. One of these, *prpo1,* conditions a lesion-mimic phenotype, as does knockown of its tobacco ortholog (*ppox)* [62]. The other results in adherant seedling leaves, similar to phenotypes observed for the *Arabidopsis* ortholog *abcg11* [63,64].

Several of the gene identifications summarized in Table 2 merit special attention because they provide the first functional information for the orthologous group or provide insight into gene function beyond that reported for orthologous proteins. These are discussed below.

5.1. PSA1, a PSI biogenesis factor with NifU domains

Our immunoblot phenotyping pipeline identified many non-allelic mutants that specifically lack PSI. The causal mutation in one of these, *psa2*, was recently reported: PSA2 is a lumenal protein with a DnaJ-related zinc finger domain and protein disulfide reductase activity [65], and was proposed to promote PSI assembly by mediating thiol transactions in the thylakoid lumen. Mutations in a different gene, *psa1*, condition pigment and protein phenotypes similar to those of *psa2* (Fig. 1). The causal insertions disrupt a gene encoding a protein with NifU domains (InterPro IPR001075), which have been implicated in the assembly of iron-sulfur clusters. The *psa1* ortholog belongs to a small gene family in *Arabidopsis* [66,67], one member of which (NFU2/AtCnfu-V) has been shown to be required for the accumulation of PSI and ferredoxin [66,68]. Some ortholog predictors place *psa1* in an orthologous group that has not been genetically characterized (NFU3/AtCNFUIVa/AT4G25910 - see http://cas-pogs.uoregon.edu/#/pog/19589). However, the orthology relationships in this family are ambiguous due to the gain and loss of paralogs since the divergence of monocots and dicots. In any case, the *psa1* phenotype in conjunction with the known activities of NifU proteins provides strong evidence that *psa1* promotes PSI accumulation by mediating the assembly of iron-sulfur clusters that associate with PSI.

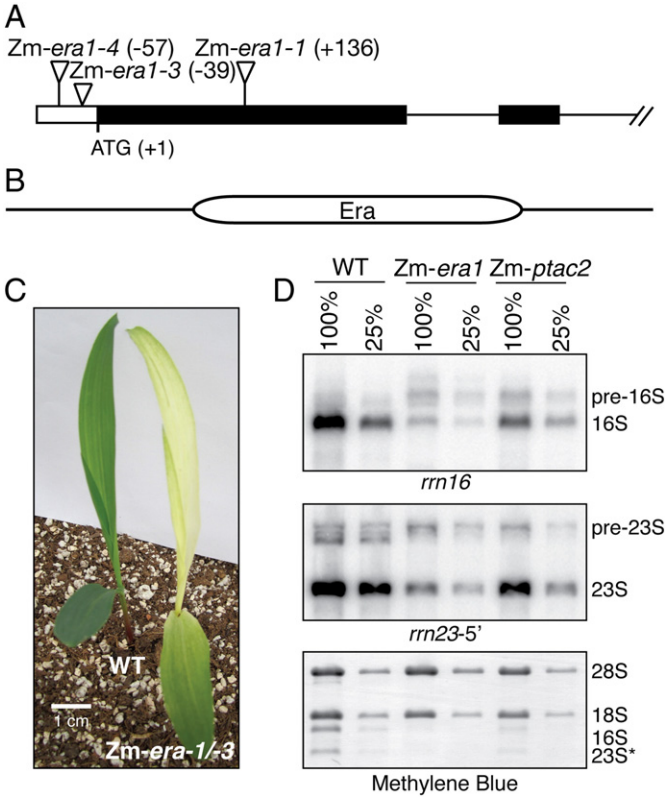

Fig. 2. Overview of Zm-*era1* mutants. (A) Positions of the *Mu* insertions in the Zm-*era1* gene. The nucleotide positions relative to the start codon are indicated. The insertion site sequences are provided in Supplementary Fig. 1. (B) Domain architecture of Zm-Era1. (C) Phenotype of Zm-*era1* mutants. The pictured individual is the heteroallelic progeny of a cross between a strong and weak allele. (D) RNA gel blot hybridizations showing defects in the processing and accumulation of chloroplast rRNAs. The upper and middle blots were probed with *rrn16* and *rrn23*, respectively. The methylene blue-stained membrane is shown below to illustrate relative loading.

5.2. Zm-Era1, a putative assembly factor for the chloroplast 30S ribosomal subunit

The causal insertions in three allelic mutations that condition an ivory or virescent phenotype (depending on allele strength) were mapped to a gene encoding a protein with an Era domain (IPR005662) (Fig. 2). The Era domain is related to a bacterial protein that promotes the assembly of the 30S ribosomal subunit [69,70]. The maize protein and its *Arabidopsis* ortholog localize to chloroplast nucleoids [71,72], as do several other proteins implicated in plastid ribosome assembly. Orthologs of this protein have not been functionally characterized in plants, but the *Arabidopsis* gene was assigned the name Era1 in a recent phylogenetic analysis [72]. Accordingly, we named the maize protein Zm-Era1.

To explore the possibility that Zm-Era1 is involved in the biogenesis of chloroplast ribosomes, chloroplast rRNAs in a hypomorphic Zm-*era1* mutant were analyzed by RNA gel blot hybridization (Fig. 2D). To control for secondary effects resulting from global defects in chloroplast gene expression, a Zm-*ptac2* mutant with a defect in chloroplast transcription and a similar pigment phenotype [73] was analyzed in parallel. The Zm-*era1* mutant had considerably lower levels of chloroplast rRNAs than did the Zm-*ptac2* control. Furthermore, the ratio of pre-16S rRNA to mature 16S rRNA was much higher in the Zm-*era1* mutant than in the Zm-*ptac2* mutant, consistent with the possibility that Zm-Era1 functions directly in the maturation of the chloroplast 30S ribosomal subunit. However, defects in rRNA processing and accumulation can result from a wide variety of primary defects, so validation of this hypothesis will require the detection of a physical association between Era1 and assembling 30S subunits.

5.3. A mutant lacking Zm-Ycf54, a component of the Mg-protoporphyrin monomethyl ester cyclase

A pair of allelic mutants with yellow-green pigmentation stood out during routine immunoblot phenotyping due to the loss of an abundant protein detected on stained blots that corresponds to a major Light Harvesting Chlorophyll a/b Binding Protein (LHCP) (Fig. 3). Core subunits of Rubisco, PSII, *cyt* b_6f, PSI, and the ATP synthase accumulate to near normal levels in these mutants (Fig. 3D and data not shown). The causal insertions were mapped to a gene encoding a homolog of the cyanobacterial protein Ycf54. This protein and its homologs in barley and tobacco are subunits of the Mg-protoporphyrin monomethyl ester cyclase involved in chlorophyll biosynthesis [54,55,74]. An antisense knockdown of the tobacco ortholog, LCAA, has been characterized [74] but heritable mutants in plants have not been reported previously. The Zm-*ycf54* mutant adds to the resources available for studying the roles of chlorophyll precursors in retrograde signaling, and the physiological consequences of decreasing the abundance of light harvesting complexes in a C4 plant.

5.4. New insight into functions of genes that had been characterized in other organisms

Given the large community of researchers employing *Arabidopsis* for the genetic analysis of chloroplast processes, it is to be expected that many of the causal mutations identified in the PML collection map to orthologs of characterized *Arabidopsis* genes. Usually, the maize phenotypes are fully concordant with those reported for *Arabidopsis*, but sometimes distinct phenotypes reveal functional divergence. For example, the molecular phenotypes resulting from loss of the orthologous PPR proteins ATP4 (maize) and SVR7 (*Arabidopsis*) differ substantially [75–78], as do the molecular defects in maize and *Arabidopsis* mutants lacking the single-stranded nucleic acid binding protein WHY1 [79,80].

In other instances, the distinct anatomy and physiology of maize facilitate molecular analyses that provide insights beyond those that are readily obtained in *Arabidopsis*. This applies to the many genes that

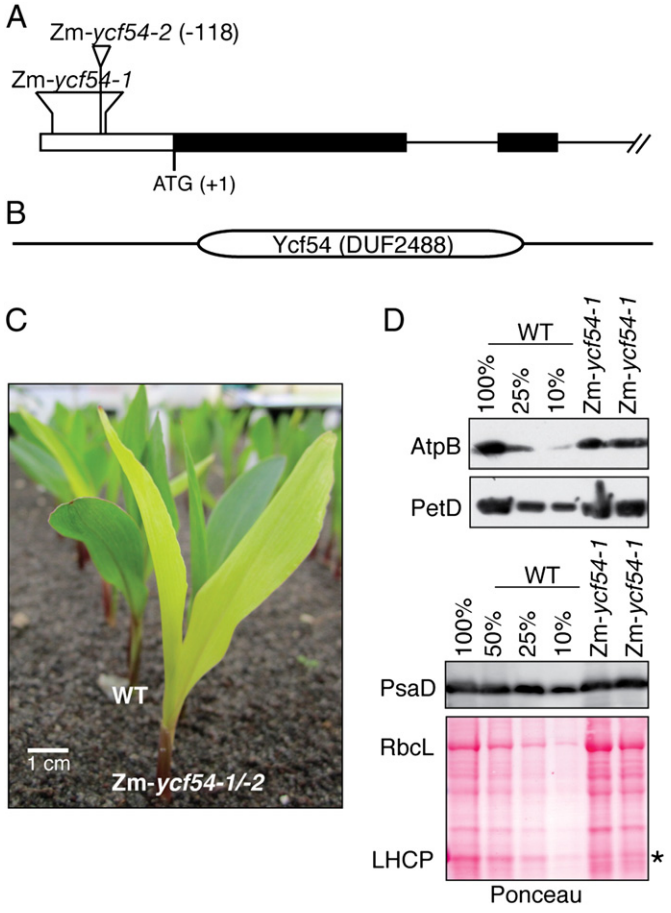

Fig. 3. Overview of Zm-*ycf54* mutants. (A) Positions of the *Mu* insertions in the Zm-*ycf54* gene. The nucleotide positions relative to the start codon are indicated. The insertion in Zm-*ycf54-1* was flanked by a deletion. The insertion site sequences are provided in Supplementary Fig. 1. (B) Domain architecture of Zm-YCF54. (C) Phenotype of a Zm-*ycf54* mutant. (D) Immunoblot profile of core subunits of photosynthetic enzyme complexes. The upper and lower panel involved identical samples but show different gels. The loss of an abundant stained protein corresponding to LHCP can be seen on the PonceauS-stained filter at the bottom (asterisk).

are essential for embryogenesis in *Arabidopsis* but not in maize, due to the species-specific requirement for a chloroplast translation product during embryogenesis [14]. The fact that such mutants survive through the seedling stage in maize has been particularly important for the characterization of the chloroplast splicing machinery (see section 6 below).

Newly identified maize mutants with insertions in the orthologs of *Arabidopsis ATAB2* and *PGR3* (see Table 2) provide additional examples of how phenotypic analyses in maize can clarify understanding of conserved functions while also revealing divergent functions.

5.4.1. Reevaluating the function of ATAB2 based on analysis of Zm-tab2 mutants

ATAB2 is the *Arabidopsis* ortholog of *Chlamydomonas Tab2*, which promotes PSI accumulation by activating translation of the chloroplast *psaB* mRNA [81,82]. The *Arabidopsis* mutants lack PSI but also are deficient for PSII, with translation defects reported for subunits of both complexes [82]. Based on this body of work, the current TAIR annotation states that ATAB2 "presumably functions as an activator of translation with targets at PS I and PS II."

The maize ortholog, Zm-*tab2*, was identified during *Mu*-Illumina analysis of two allelic mutants lacking PSI (Fig. 4). To evaluate chloroplast translation in these mutants, we employed a ribosome profiling method that quantitatively maps chloroplast ribosome footprints to specific mRNA regions at high resolution [76]. A genome-wide analysis of chloroplast ribosome footprints in Zm-*tab2* mutants did not detect reduced ribosome occupancy on photosystem mRNAs (Supplementary Fig. 2). Furthermore, slot-blot hybridization analysis of ribosome footprints prepared from Zm-*tab2* (maize) and *atab2* (*Arabidopsis*) mutants revealed that the ratio of ribosome footprints on the *psaB* relative to *rbcL* open reading frames is similar in the wild-type and mutant samples (Fig. 4D). These results strongly suggest that the mutants do not exhibit a specific defect in photosystem mRNA translation.

Based on these findings, it seems that the functions of *ATAB2* and its orthologs should be reevaluated. The presence of a *tab2* ortholog in cyanobacteria is also suggestive in this regard, as there is no precedent for conserved mRNA-specific translational regulators among chloroplasts and cyanobacteria.

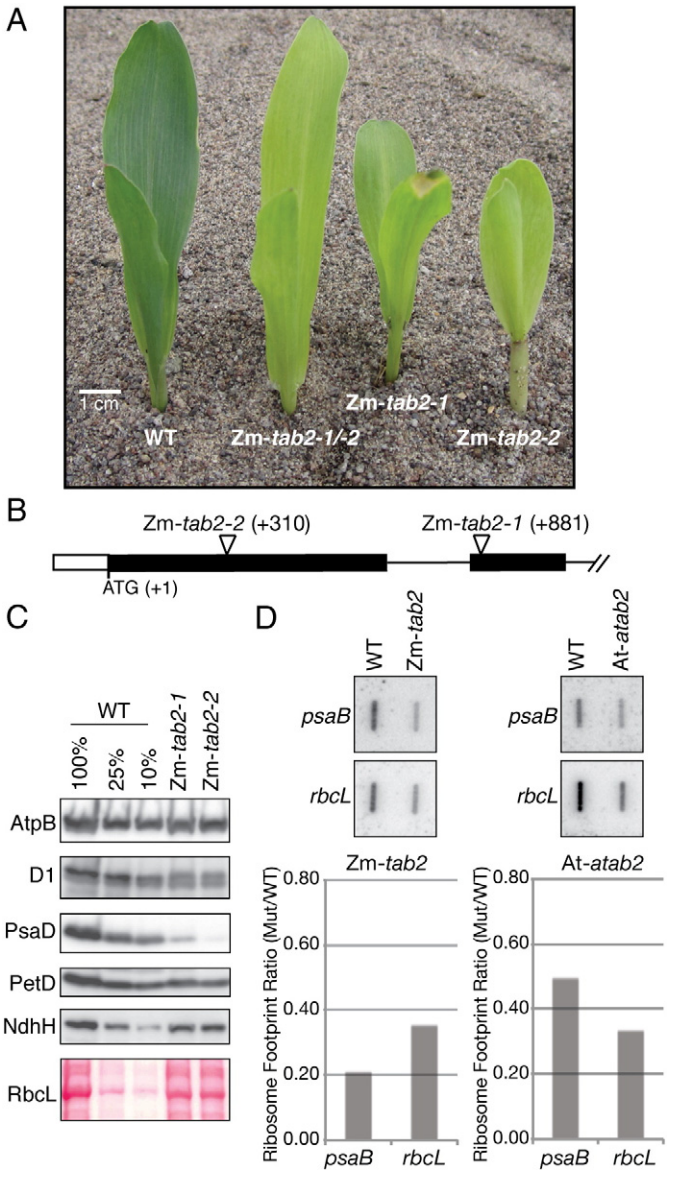

Fig. 4. Overview of Zm-*tab2* mutants. (A) Phenotype of Zm-*tab2* mutants. Plants were grown in soil for 8 days. (B) Positions of the *Mu* insertions in the Zm-*tab2* gene. The nucleotide positions relative to the start codon are indicated. The insertion site sequences are shown in Supplementary Fig. 1. (C) Immunoblot profile of core subunits of photosynthetic enzyme complexes. (D) Slot-blot analysis of ribosome footprints from the *psaB* and *rbcL* genes in maize and *Arabidopsis tab2* mutants. The data were quantified with a STORM phosphor-imager and are plotted below. The *Arabidopsis* mutant is SAIL line 803-D03, stock CS876967.

Table 1
Identified nuclear genes in maize required for chloroplast biogenesis: summary of published mutants.

Gene[1]	Maize locus ID	*Arabidopsis* ortholog	Method[2]	Seedling phenotype[3]	Protein phenotype[4]	Source	Gene function	Publications
Chloroplast gene expression and DNA metabolism								
crp1	GRMZM2G083950	At5g42310	F	pyg	Reduced PSI & cyt	PML	PPR protein; Activates *petA* and *psaC* translation; stabilizes processed *petB, petD* mRNAs; binds *petA, psaC, petD* 5′UTRs	[26,27,76,114]
ppr10/luteus15	GRMZM2G177169	At2g18940	R	vpyg	ATPase absent, PSI reduced	PML	PPR protein; activates *atpH* translation, stabilizes *atpH* and *psaJ* mRNAs, binds *atpI-atpH* and *psaJ-rpl33* intergenic regions	[41,76,115]
atp4	GRMZM2G128665	At4g16390	R	spg	ATPase specific defect	PML	PPR-SMR protein; activates *atpB* translation, binds *atpB* 5′UTR	[75,76]
tha8	AC217965.2_FG012	At3g27750	F	pyg	Global, reduced OEC23	PML, MTM	PPR protein; required for splicing *ycf3-int2* and *trnA* introns and associates with those introns *in vivo*	[90]
Zm-*otp51*	GRMZM2G325019	At2g15820	R	pyg	Strong PSI and cyt defect, mild global	PML	PPR-LAGLIDADG protein; required for splicing cp *ycf3-int2*	[90,93]
ppr5	GRMZM2G025409	At4g39620	R	Hypomorph-pg, null-iv	Global	PML	PPR protein; stabilizes unspliced trnG precursor and stimulates its splicing; binds the trnG intron	[40,116]
ppr2	GRMZM2G341621	At3g06430	R	iv	Global	PML	PPR protein required for plastid ribosome accumulation	[43]
ppr4	not in B73 v3; GenBank cDNA NP_001105869	At5g04810	R	Hypomorph-pg, null-iv runty	Global	PML	PPR protein required for *rps12* trans-splicing; binds *rps12* intron 1	[42]
crs1	GRMZM2G078412	At5g16180	F	Hypomorph-pyg, null-vpyg	Global, Atpase almost gone	PML	*cp RNA splicing* 1; founding member of CRM domain family. Binds *atpF* intron, promotes *atpF* splicing	[24,25,99, 114,117]
crs2	GRMZM2G132021	At5g16140, At5g38290	F	iv	Global	PML	*cp RNA splicing 2*; required for splicing of chloroplast subgroup IIB introns. peptidyl tRNA hydrolase homolog.	[23,24,37, 118,119]
caf1	Incorrect gene model. Gene spans GRMZM2G089222/ GRMZM2G173923 GenBank cDNA NP_001105220	At2g20020	R	iv	Global	PML	*CRS2-associated factor 1*;CRM domain protein, interacts with CRS2, promotes splicing of several cp introns	[37,99,114]
caf2	AC199526.5_FGT003	At1g23400	R	iv	Global	PML	*CRS2-associated factor 2*; CRM domain protein, interacts with CRS2, promotes splicing of several cp introns	[37,99,114,119]
rnc1	GRMZM2G035820	At4g37510	R	Hypomorph-pyg-vir, null -iv runty	Global	PML	RNAse III domain protein required for the splicing of many cp introns. Interacts with WTF1.	[36,38]
wtf1	GRMZM2G403797	At4g01037	R	Hypomorph-pyg, null-iv runty	Global	PML	cp RNA splicing factor. Founding member of PORR domain family. Interacts with RNC1.	[36,38]
Zm-*mterf4*	GRMZM2G029933	At4g02990	F	Hypomorph-vpyg, null-iv	Global	PML	cp mTERF domain protein. Binds *in vivo* to multiple group II introns and promotes their splicing.	[91]
crp4	GRMZM2G377761	At3g03710	F	pg	Global	PML, Schnable	cp polynucleotide phosphorylase, cpRNA processing and decay	[20,120]
Zm-*orrm1*	GRMZM5G899787	At3g20930	F	pg	Mild global, strong loss of cyt b_6f	PML	ORRM-domain protein, involved in cp RNA editing	[121]
Zm-*ptac12*	GRMZM5G897926	At2g34640	F	vpyg-vir	Global	PML	Maize ortholog of PTAC12; PEP associated protein required for PEP-mediated transcription	[46,73]
Zm-*ptac2*	GRMZM2G122116	At1g74850	F	Hypomorph-pyg, null-ivory	Global	PML	Maize ortholog of PTAC2; PEP associated protein required for PEP-mediated transcription	[46,73]
Zm-*murE*	GRMZM2G009070	At1g63680	F	Hypomorph-pyg, null-iv	Global	PML	Maize ortholog of MurE; PEP associated protein required for PEP-mediated transcription	[46,73,122]
Zm-*prin2*	GRMZM2G119906	At1g10522	F	Hypomorph- vpyg-vir, null-iv	Global	PML	Maize ortholog of PRIN2; PEP associated protein required for PEP-mediated transcription	[73,123,124]
Zm-*hcf107*	GRMZM2G121960	At3g17040	R	spg	Reduced PSII	D.Miles and W.Cook	HAT domain protein required to stabilize *psbH* 5′ end and promote *psbH* translation. Binds *psbH* 5′UTR	[19,125–127]
w2	GRMZM2G480171	At1g50840, At3g20540	F	Hypomorph-pyg, null-iv	Global	PML, Maize Genetics Coop	Organellar DNA polymerase gamma; Required for cpDNA replication.	[128]

(continued on next page)

Table 1 (*continued*)

Gene[1]	Maize locus ID	*Arabidopsis* ortholog	Method[2]	Seedling phenotype[3]	Protein phenotype[4]	Source	Gene function	Publications
Zm-*why1*	GRMZM2G155662	At2g02740, At1g14410	R	Hypomorph-pyg, null-iv or embryo lethal, background dependent	Global	PML	single-stranded nucleic acid binding protein. Multifunctional: promotes *atpF* splicing, 23S rRNA maturation, stabilizes cp genome.	[79,80,129,130]
rlsb1	GRMZM2G087628	At1g71720	R	vpyg-vir	Global	PML & Schnable	cp RNA binding protein	[134]
Thylakoid protein targeting								
hcf106	GRMZM5G898735	At5g52440	F	hcf, pg	Reduced PSI, PSII, cyt & OEC23	Freeling Mu lines	thylakoid TAT transport	[29,131,132]
tha1	GRMZM2G090086 (bad gene model)	At4g01800	F	pyg	Reduced PSI, PSII, cyt & PC	PML	cp SecA homolog; required for thylakoid Sec-transport	[29,30]
tha4	GRMZM2G472651	At5g28750	F	spg	Reduced PSI, PSII, cyt & OEC23	PML	thylakoid TAT transport	[31]
tha5	GRMZM2G300408	At4g14870	F	pg	Reduced PSI, PSII, cyt & PC	PML	cp SecE; required for thylakoid Sec- transport	[20]
csr1	NP_001105732 (missing from B73 v3 genome)	At2g45770	R	Hypomorph-pyg, null-vpyg	Global	PML	CpFtsY. cp SRP receptor involved in thylakoid protein targeting	[39]
csy1	GRMZM5G809546	At2g18710	R	vpyg	Global, reduced PC, WT OEC23	Pioneer	CpSecY. thylakoid protein translocase	[133]
Assembly and homeostasis of photosynthetic apparatus								
pet2	GRMZM2G087063	At5g52110	F	pg	reduced cyt b_6f	PML	Ortholog of HCF208/CCB2; cyt *b6* heme attachment factor	[20,60]
psa2	GRMZM2G021687	At2g34860	F	spg	Reduced PSI	PML	DnaJ-type zinc finger domain protein required for PSI assembly	[65]
raf1	GRMZM2G457621	At3g04550, At5g28500	F	spg	Rubisco specific defect	PML	*Rubisco Accumulation Factor1*; promotes Rubisco assembly	[35]
raf2	GRMZM2G139123	At5g51110	F	spg	Rubisco specific defect	PML	*Rubisco Accumulation Factor2*; promotes Rubisco assembly	[100]
Zm-*clpP5*/ *vyl*	GRMZM2G121456	At1g02560	R	Hypomorph-vpyg, null-iv	Global	PML & Bailin Li	cp ATP-dependent Clp protease subunit 5	[135]
Chlorophyll deficient mutants cloned by other groups								
ij1	GRMZM2G004583	At3g12930	F	iv, striate, embryo defects; background dependent			*iojap1*. Required for plastid ribosome accumulation in seedlings. Associates with 50S ribosomal subunit.	[32,136]
bsd2	GRMZM2G062788	At3g47650	F	pg			*bundle sheath defective 2*. Rubisco assembly factor	[34]
g2/*bsd1*	GRMZM2G087804	At5g44190	F	pg			*golden2*. Nuclear transcription factor for chloroplast biogenesis genes	[33]
hcf60	GRMZM2G038013	At1g79850	F	hcf, pg			cp ribosomal protein S17	[137]
lpe1	GRMZM5G858417	ambiguous	F	pg			*leaf permease 1*; xanthine/uracil/ permease	[138]
Zm-*hcf136*/ *psb1**	GRMZM2G102838	At5g23120	F	spg, hcf	PSII-specific defect		PSII assembly factor	[139,140]
elm2	GRMZM2G101004	At1g58300	F	pg			*elongated mesocotyl2*. heme oxygenase4	[141]
cpx1/*nec4**	GRMZM5G870342	At1g03475	R	Yellow-necrotic	Global		coproporphyrinogen III oxidase	[142]
ppr8522/emb2	GRMZM5G884466	At5g67570	F	Embryo lethal; albino rescued seedling			cp PPR protein, plastid gene expression	[143]
zb7/*lw1**	GRMZM2G027059	At4g34350	F	Weak-pyg, strong-vpyg with bands	Global		ISPH enzyme involved in isoprenoid biosynthesis	[144]
*cf1**	GRMZM2G026117	AT5g08280	F	pyg necrotic			porphobilinogen deaminase	[145]
oy1	GRMZM2G419806	At5g45930	candidate	Yellow			Chlorophyll synthesis, Magnesium chelatase I	[146]

* Mu insertion alleles linked to phenotypes in PML collection but published previously by a different group.

[1] Alternate gene names separated by "/".

[2] F-forward genetics, phenotype-to-gene; R-reverse genetics, targeted screen.

[3] Seedling pigmentation: iv-ivory, pg-pale green, spg-slightly pale green, pyg-pale yellow green, vpyg- very pale yellow green, vir-virescent, hcf-high chlorophyll fluorescent, striate- longitudinal stripes.

[4] Protein data based on immunoblot survey of one core subunit of Rubisco, PSI, PSII, cyt b_6f, ATP synthase; "global"- reduction in all of these complexes. This data is provided only for PML mutants.

Table 2
New phenotype-driven gene identifications via *Mu*-Illumina analysis of PML mutants.[1]

Gene	Maize locus ID	*Arabidopsis* ortholog	Maize phenotype (pigment[2]; protein[3])	Gene/protein function	Closely related publications
Chloroplast gene expression					
cps1	GRMZM2G156565	At2g31170	pg (hypomorph); global	*Chloroplast protein synthesis 1*: cp cysteinyl-tRNA synthetase	[59,61,147]
cps5	GRMZM5G858471	At4g04350	iv; global	*Chloroplast protein synthesis 5*: cp leucyl-tRNA synthetase	[61]
Zm-*era1*	GRMZM2G158024	At5g66470	vpyg-vir; global	Ortholog of *Arabidopsis ERA1*. Required for cp ribosome accumulation. Putative cp ribosome assembly factor. ERA domain (IPR005662).	[72]
Zm-*prda1*	GRMZM2G079452	At5g48470	vpyg (hypomorph), iv (null); global	Ortholog of *Arabidopsis PRDA1*. cp nucleoid protein involved in PEP-mediated transcription; cpRNA defects in maize similar to other PEP-deficient mutants	[47]
Zm-*ptac14*	GRMZM5G807767	At4g20130	iv; global	Ortholog of *Arabidopsis PTAC14*. PEP-associated protein, required for PEP-mediated transcription; cpRNA defects in maize similar to other PEP-deficient mutants	[45,46,148]
Zm-*pgr3*	GRMZM2G372632	At4g31850	vpyg; global	Ortholog of *Arabidopsis PGR3*. PPR protein that stabilizes 5′-end of cp *petL* RNA and is required for NDH accumulation. Additional functions in maize.	[58,84]
Zm-*rh3b*	AC198418.3_FGP005	At5g26742	vpyg-vir (hypomorph); global	cp Dead box RNA helicase involved in cp intron splicing and assembly of the 50S ribosomal subunit	[44]
Biogenesis of photosynthetic enzymes/prosthetic groups					
Zm-*ccs1*	GRMZM2G038301	At1g49380	pg; specific loss of cyt b_6f	Ortholog of *CCS1*. Required for cytb_6f assembly	[149]
pet3	GRMZM2G177145	At1g59840	pg; specific loss of cyt b_6f	Ortholog of *CCB4*. Required for cyt b_6 assembly	[48,49]
Zm-*tab2*	GRMZM2G081955	At3g08010	pg; severe loss of PSI, mild global defect	Ortholog of *ATAB2* in *Arabidopsis* and *TAB2* in *Chlamydomonas*. Required for PSI accumulation. Activates translation of cp *psaB* RNA in *Chlamydomonas* but not in maize.	[81,82]
psa1	GRMZM2G100976	At4g25910	pg; specific loss of PSI	Required for PSI accumulation. Harbors NifU domain implicated in Fe-S cluster assembly. Ortholog of *Arabidopsis* Nfu-IVa.	[66,67]
cps2	AC215201.3_FGP005	At2g28000	pg; specific loss of Rubisco (hypomorph)	cpn60 α1. Molecular chaperone involved in Rubisco assembly. Also affects distribution of ribosomes on chloroplast mRNAs.	[35,57,59,100]
Zm-*y3IP1*	GRMZM2G002165	At5g44650	pg; specific loss of PSI	Ortholog of *Arabidopsis Y3IP1*. PSI assembly factor, cooperates with Ycf3	[56]
Zm-*ycf54*	GRMZM2G010196	At5g58250	pyg, reduced LHCP	Ortholog of barley/cyanobacterial *Ycf54*; subunit of Mg-protoporphyrin IX monomethyl ester cyclase	[54,55,74]
Zm-*apg1*	GRMZM2G082998	At3g63410	pyg; global	MPBQ/MSBQ methyltransferase involved in plastoquinone synthesis	[51]
Zm-*gun4*	GRMZM2G464328	At3g59400	pyg (hypomorph), yellow (null)	Ortholog of *Arabidopsis GUN4*; Mg-chelatase cofactor and porphyrin binding protein	[52,53]
prpo1/ Zm-*ppox*	GRMZM2G039396	At4g01690	pg with lesions	Ortholog of Arabidopsis *PPOX*/ *PPO1*. protophorphyrinogen oxidase	[62]
Structural component of photosynthetic enzyme					
Zm-*atpc*	GRMZM2G048907	At4g04640	spg; ATP-synthase specific defect	gamma subunit of chloroplast ATP synthase (ATPC1)	[150]
ppdk	GRMZM2G306345	At4g15530	pg; reduced PSI, PSII, cyt b_6f	Pyruvate orthophosphate dikinase	
Other					
Zm-*abcg11*	GRMZM2G177812	At1g17840	Adherant seedling leaves	Ortholog of *Arabidopsis* ABCG11/COF1/WBC11. Plasma membrane-localized ATP-binding cassette transporter; cutin transport to extracellular matrix	[63,64]

[1] Insertion sites are provided in Supplemental Fig. 1.
[2] pg-pale green; pyg – pale yellow green; vpyg – very pale yellow green; iv-ivory; vir –virescent.
[3] Based on immunoblots surveying one core subunit of PSI, PSII, cytb_6f, ATP synthase, Rubisco. "Global" – decrease in all surveyed complexes.

5.4.2. Divergent and conserved functions of the PPR protein PGR3

An insertion in the maize ortholog of *Arabidopsis* PGR3 was found to be the causal mutation in a mutant with "global" protein deficiencies and loss of specific RNAs from the plastid *petL* transcription unit (Fig. 5). PGR3 is a PPR protein that stabilizes the 5′ end of *petL* mRNA, enhances *petL* translation, and promotes the expression of one or more plastid gene encoding a subunit of the NDH complex [58,83,84].

The *petL* RNA defect in Zm-*pgr3* mutants (Fig. 5) shows the RNA stabilization function to be conserved between the maize and *Arabidopsis* PGR3 orthologs. However, the visible and molecular phenotypes of the maize mutant suggest additional functions beyond those in *Arabidopsis*. *Arabidopsis pgr3* mutants are not readily distinguishable from their normal siblings under standard laboratory growth conditions, but the maize mutants are strongly chlorotic (Fig. 5). Furthermore, the protein defects in *Arabidopsis* are limited to the cyt b_6f and NDH complexes, whereas the maize mutants exhibit a substantial reduction in core subunits of all complexes harboring plastid-encoded subunits (Fig. 5). Chloroplast rRNAs accumulate to reduced levels in the maize mutant but not in *Arabidopsis* (Fig. 5); this ribosome deficiency presumably contributes to the more global protein losses observed in maize.

These results indicate that Zm-PGR3 has functions beyond those of its *Arabidopsis* ortholog. In addition, the absence of the NDH complex that is observed in both species is likely to cause distinct physiological defects in maize due to the special role of the NDH complex in the bundle sheath chloroplasts of C4 plants [85].

6. Examples of chloroplast processes for which genetic analyses in maize have been particularly informative

Genetic analyses in maize have contributed to the understanding of many aspects of chloroplast biogenesis, but there are several instances in which work in maize has played a leading role. Two examples are summarized below.

6.1. Chloroplast RNA splicing and the discovery of plant-specific RNA binding domains that are dedicated to organellar RNA metabolism

Although introns are rare in bacteria, they are frequent in chloroplast genomes. Land plant chloroplast genomes typically encode ~18 introns, all of which are derived from self-splicing group I or group II intron ribozymes [reviewed in [86]]. These introns have lost the capacity to self-splice, and rely on a plethora of nucleus-encoded proteins that act in a combinatorial fashion to promote the splicing of specific introns. Most chloroplast splicing factors harbor non-canonical RNA binding

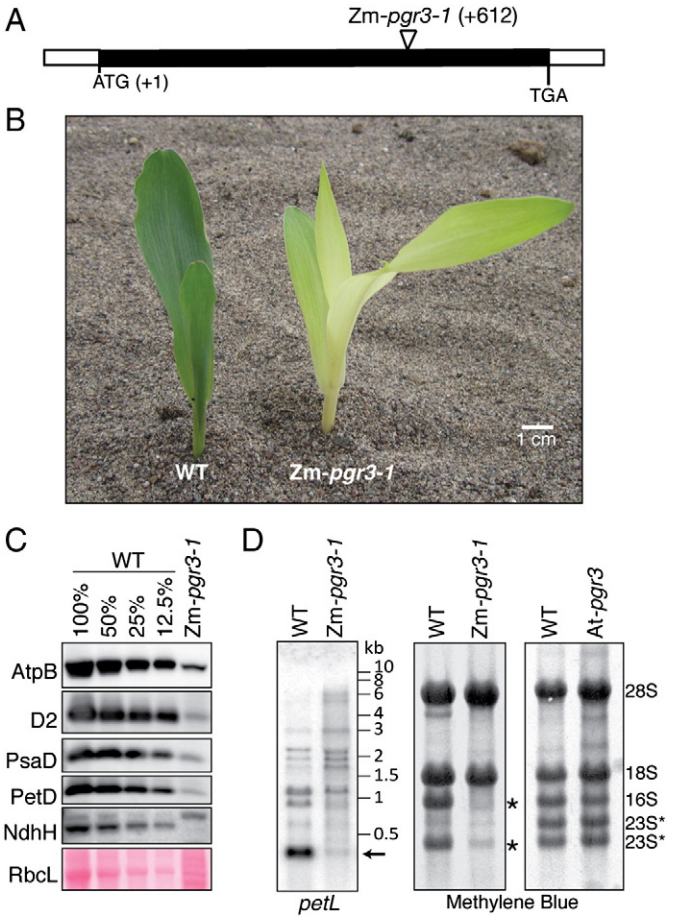

Fig. 5. Overview of Zm-*pgr3* mutants. (A) Position of the *Mu* insertion in the Zm-*pgr3* gene. The nucleotide position relative to the start codon is indicated. The insertion site sequence is shown in Supplementary Fig. 1. (B) Phenotype of a Zm-*pgr3* mutant. The plants were grown for 8 days in soil. (C) Immunoblot profile of core subunits of photosynthetic enzyme complexes. (D) RNA gel blot probed with a *petL* specific probe (left). The abundance of chloroplast rRNAs is shown on the methylene blue-stained blots to the right. The At-*pgr3* mutant has an insertion in the single exon of the *Arabidopsis PGR3* gene.

domains that are found almost exclusively in organelle-targeted proteins: the CRM, PORR, APO, PPR, and mTERF motifs [28,36,87–89]. Splicing occurs in the context of large intron ribonucleoprotein particles that generally include at least four different proteins [reviewed in [86]]. Thus, these formerly "self-splicing" introns splice in the context of particles that are as much protein as RNA.

The complexity of the chloroplast splicing machinery and the nature of the proteins involved were unanticipated. This understanding emerged largely from genetic analyses employing maize mutants in the PML collection, beginning with the identification of two mutants with distinct splicing defects, *chloroplast RNA splicing 1 (crs1)* and *crs2* [24]. Subsequent identification of the CRS2 Associated Factors CAF1 and CAF2 [37], followed by proteomic analysis of proteins that coimmunoprecipitate with CRS1, CAF1, and CAF2 implicated additional proteins in splicing (RNC1, WTF1, WHY1, APO1, RH3) [36,38,44,79,88], and the roles of these proteins were demonstrated by reverse-genetic and biochemical analyses. Additional chloroplast splicing factors were identified through independent investigations involving mutants in the PML collection (THA8, PPR4, PPR5, mTERF4) [40,42,90,91] and reverse-genetic analysis of members of the PPR family in *Arabidopsis* (OTP51 and OTP70) [92,93].

CRS1, CAF1, and CAF2 are the founding members of a plant-specific protein family harboring an ancient domain, the CRM domain, whose function as an RNA binding domain was revealed through this body of work [87]. Reverse genetic analysis of other members of the CRM domain family revealed two more chloroplast splicing factors, CFM2 and CFM3 [94,95]. WTF1 and APO1 are also the founding members of plant-specific gene families, each defined by domains that were shown through this work to be RNA binding domains: the PORR and APO domains, respectively. Most members of the CRM, PORR, and APO families are predicted to localize to chloroplasts or mitochondria [36, 87,96], and reverse genetic analysis of mitochondrial CRM and PORR domain proteins in *Arabidopsis* demonstrated these to be mitochondrial splicing factors [97,98]. Indeed, organellar splicing is the only molecular function ascribed to members of the CRM, PORR, and APO domain families, so it is likely that additional organellar splicing factors will be found among the uncharacterized members of these families.

Of the seventeen nucleus-encoded chloroplast splicing factors reported thus far in plants, only OTP51 and OTP70 were discovered independent of the body of work involving the maize PML collection. OTP51 and OTP70 are PPR proteins that are required for the splicing of *ycf3*-intron 2 and *rpoC1*-intron 1 in *Arabidopsis* [92,93]. The function of OTP51 is conserved in maize [90]. OTP70, however, facilitates the splicing of an intron that is not found in maize and, in fact, the current maize genome assembly lacks an OTP70 ortholog.

It is noteworthy that work in maize has dominated this field despite the much larger community of chloroplast researchers working in *Arabidopsis*. This is not due to a lack of functional conservation, as there is experimental evidence that the splicing functions of *crs1*, *caf1*, *caf2*, *otp51*, *tha8*, *cfm2*, *cfm3*, and *rh3* are conserved [44,90,94,95,99]. Instead, it seems likely that these genes have been elusive in *Arabidopsis* because most of them participate in the splicing of tRNAs and ribosomal protein mRNAs [reviewed in [86]] and are therefore essential for the function of the chloroplast translation machinery. The absence of chloroplast ribosomes conditions an embryo-lethal phenotype in *Arabidopsis*, but typically results in albino, seedling lethal mutants in maize whose molecular phenotypes can be informative [14]. In addition, RNA and protein-coimmunoprecipitation assays have been important for establishing protein-intron partnerships [36–38,40,42,44,79,88, 91,94,95] and these assays are more challenging in *Arabidopsis*, due to the greater difficulty of isolating large quantities of chloroplast extract.

6.2. Rubisco assembly

The assembly of the multimeric enzymes involved in photosynthesis is facilitated by general and complex-specific chaperones [reviewed in [12]]. The majority of known assembly factors for thylakoid membrane complexes have been discovered through genetic analysis in *Arabidopsis* [12]. By contrast, the genetic identification of assembly factors for Rubisco has come largely from work in maize. Three dedicated Rubisco assembly factors have been reported (BSD2, RAF1, and RAF2), all of which were discovered via forward genetic screens in maize [34,35, 100]. A fourth component of the Rubisco assembly network, Cpn60α1, proved to be the product of the maize *cps2* gene (see Table 2), whose mutant phenotype was reported two decades before the gene was identified [35,59]. Hypomorphic *cps2* alleles condition a severe loss of Rubisco with only a minor decrease in the abundance of other photosynthetic complexes [59]. Thus, although Cpn60s participate broadly in chloroplast protein folding/assembly [reviewed in [101]], the *cps2* mutant demonstrates that Rubisco assembly is particularly reliant on the Cpn60α1 isoform. Disruption of the rice *cps2* ortholog likewise causes a severe Rubisco deficiency, but the effects on other photosynthetic complexes were not reported [57]. Interestingly, the *cps2* mutant phenotype also revealed a role for Cpn60α1 in chloroplast translation: chloroplast mRNAs in *cps2* mutants exhibit an unusual bimodal distribution during polysome sedimentation experiments [59]. The basis for this is not known, but suggests that Cpn60α1 may bind nascent peptides cotranslationally to influence ribosome movement.

It isn't obvious why genetic analysis in maize has played such a prominent role in elucidating Rubisco assembly, as this process is likely

to be broadly conserved. Indeed, RAF1 and RAF2 orthologs are present in photosynthetic bacteria, where they also promote Rubisco assembly [102,103]. Perhaps the visible phenotype is more apparent in maize, or perhaps the sequestration of Rubisco in bundle sheath cells reduces pleiotropic consequences of its absence, such that the Rubisco-specific defect is more apparent in the C4 context.

7. Comparison of PML to related genetic resources in other organisms

The importance of delving more deeply into the nuclear gene complement relevant to photosynthesis is underscored by related efforts in other organisms. The Chloroplast 2010 Project [104,105] and Chloroplast Function Database [106] assembled *Arabidopsis* T-DNA lines for genes predicted to encode chloroplast-localized proteins and have collected some phenotypic data for these lines. The Chloroplast Function Database lists 72 genes whose disruption causes chlorophyll-deficient seedlings, many of which lack additional phenotypic information. The Chloroplast 2010 Project emphasized metabolic profiling of homozygous viable mutants [107], which largely excludes non-photosynthetic mutants analogous to those emphasized in the PML collection. The PhenoLeaf project [108] cataloged genes from the *Arabidopsis* Unimutant collection with visible leaf phenotypes, and includes 140 genes whose disruption gives a loss of leaf chlorophyll.

There is little overlap among the genes whose phenotypes have been documented in these projects and the 94 genes that have thus far been assigned a function via the PML collection in maize. Thus, these collections are complementary rather than redundant. Of course, there are numerous single-gene studies in *Arabidopsis*, rice, and *Chlamydomonas* that describe orthologs of genes also identified via PML. Thus far, approximately one-third of the maize genes whose functions have been defined with PML are orthologous to well-characterized genes in other organisms, one-third are orthologous to genes whose disruption had been linked only to very general phenotypes in other organisms (e.g. chloroplast development or embryogenesis), and one-third are novel genes that provided the first functional insight into their cognate orthologous groups.

In another related project, a phylogenomic approach identified a set of "GreenCut" genes found specifically in photosynthetic eukaryotes; this compilation provides a source for targeted study of plastid functions [109]. Interestingly, although all of the genes identified with the PML collection are conserved among angiosperms and many are plant-specific, only a small fraction are in the GreenCut. In some instances it is unclear why the gene was excluded from the GreenCut (e.g. the Rubisco assembly factor RAF1), but in many cases the answer is clear: Numerous genes whose functions have been assigned with mutants from the PML collection participate in chloroplast gene expression (RNA splicing, RNA stabilization, transcription, translational activation), and these processes and the nucleus-encoded proteins that mediate them evolved rapidly after the divergence of plants and algae.

8. Future prospects and community access

Several years back, a review entitled "*Chlamydomonas* and *Arabidopsis*: A Dynamic Duo" described the synergism of *Arabidopsis* and *Chlamydomonas* as model systems for studying the basic biology of photosynthetic eucaryotes: results from one species often inform understanding in the other, and differences highlight points of divergence [13]. As a C4 monocot with a well-developed genetic resource for dissecting chloroplast processes, maize adds a valuable member to the ensemble, creating a "Triumphant Trio". Functions of conserved genes are sometimes revealed more easily in maize than *Arabidopsis* due to differences in the role of plastid translation during embryogenesis, different partitioning of functions among paralogs, and the relative ease of obtaining mutant tissue and isolated chloroplasts for biochemical analyses. In addition, the physiological read-out of loss of conserved functions will sometimes differ in the C3 and C4 contexts, and there will also be examples of divergent functions. Analogous work in rice has also been informative [e.g. [57,110]], but the genetic resources for the study of chloroplast functions in rice are less developed.

To fully exploit the synergism of different model organisms, it is important for species-specific genome databases to display functional annotations obtained from study of orthologs in other species. This is currently all too rare. However, various online resources are available to facilitate cross-species functional inferences. For example, the POGs2 database http://cas-pogs.uoregon.edu/#/ [111] provides ready access to functional information for orthologous proteins in maize, rice, and *Arabidopsis*, and also includes links to *Chlamydomonas* orthologs through Plaza [112]. Complementary information can be found at the Plant Proteome Database (http://ppdb.tc.cornell.edu), which compiles proteome data from *Arabidopsis* and maize with easy navigation between the two species [113].

The PML collection is a community resource. Most of the mutants whose causal insertions have been identified are listed at http://pml.uoregon.edu/pml_table.php and seed is available at the Maize Genetics Coop Stock Center (http://maizecoop.cropsci.uiuc.edu) or directly from us. A search interface for publically-available lines with insertions in specified genes is available at http://teosinte.uoregon.edu/mu-illumina/ and a more extensive in-house database of insertions is available through collaboration.

Materials and methods

Immunoblots and RNA gel blot hybridizations were performed as described previously [18]. The following probes were used for RNA gel blot hybridizations (coordinates taken from GenBank accession NC_001666): *rrn16*- PCR fragment spanning residues 95593–96802 ; rrn235′- oligonculeotide spanning residues 99072–99131; *petL*-PCR fragment spanning residues 65168–65547. The antibodies for AtpB, D1, PsaD, and PetD were described previously [73]. The antibody to NdhH was a kind gift of Tsuyoshi Endo (Kyoto University).

The *Mu*-Illumina gene identifications were performed as described in [20], with minor modifications due to changes in the available Illumina instruments since the original publication. The genome-wide analysis of chloroplast ribosome occupancy in Zm-*tab2* mutants was performed with the microarray-based approach described in [76]. The slot-blot hybridization assay of *rbcL* and *psaB* ribosome occupancy in *atab2* and Zm-*tab2* mutants was performed analogously, except that the ribosome footprints were applied by slot-blotting to nylon membrane, and hybridized with probes spanning the *rbcL* or *psaB* open reading frames in the cognate species.

Transparency document

The Transparency document associated with this article can be found, in online version.

Acknowledgements

We would like to thank Tiffany Kroeger for her invaluable contributions to reverse-genetic screens and phenotypic analyses of PML mutants, Reimo Zoschke and Kenny Watkins for help with Zm-*tab2* ribosome profiling, Margarita Rojas for technical assistance, former members of David Stern's laboratory (Boyce Thompson Institute) for performing RNA gel blot analyses of mutants in the PML collection, and the many undergraduates in the Barkan lab who have contributed immunoblot data over the years. We would also like to thank Michael Freeling's laboratory (UC-Berkeley) for providing the founder lines that were used to generate the PML collection, Pat Schnable (Iowa State University) for providing several mutant alleles described here, Toshiharu Shikanai (Kyoto University) for providing the *pgr3* insertion line, and Tsuyoshi Endo for the gift of NdhH antibody. The development

of the PML resource has been supported by grants DBI-0421799, IOS-0922560, and IOS-1339130 from the National Science Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbabio.2015.02.014.

References

[1] J.N. Timmis, M.A. Ayliffe, C.Y. Huang, W. Martin, Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes, Nat. Rev. Genet. 5 (2004) 123–135.
[2] N. Rolland, G. Curien, G. Finazzi, M. Kuntz, E. Marechal, M. Matringe, S. Ravanel, D. Seigneurin-Berny, The biosynthetic capacities of the plastids and integration between cytoplasmic and chloroplast processes, Annu. Rev. Genet. 46 (2012) 233–264.
[3] A. Barkan, Expression of plastid genes: organelle-specific elaborations on a prokaryotic scaffold, Plant Physiol. 155 (2011) 1520–1532.
[4] J. Nickelsen, W. Zerges, Thylakoid biogenesis has joined the new era of bacterial cell biology, Front. Plant Sci. 4 (2013) 458.
[5] M. Pribil, M. Labs, D. Leister, Structure and dynamics of thylakoids in land plants, J. Exp. Bot. 65 (2014) 1955–1972.
[6] D. Lyska, K. Meierhoff, P. Westhoff, How to build functional thylakoid membranes: from plastid transcription to protein complex assembly, Planta 237 (2013) 413–428.
[7] P. Jarvis, E. Lopez-Juez, Biogenesis and homeostasis of chloroplasts and other plastids, Nat. Rev. Mol. Cell Biol. 14 (2013) 787–802.
[8] S. Weissmann, T.P. Brutnell, Engineering C4 photosynthetic regulatory networks, Curr. Opin. Biotechnol. 23 (2012) 298–304.
[9] L. Peng, H. Yamamoto, T. Shikanai, Structure and biogenesis of the chloroplast NAD(P)H dehydrogenase complex, Biochim. Biophys. Acta 1807 (2011) 945–953.
[10] P.E. Jensen, D. Leister, Chloroplast evolution, structure and functions, F1000prime Rep., 6, 2014, p. 40.
[11] J.D. Rochaix, Genetics of the biogenesis and dynamics of the photosynthetic machinery in eukaryotes, Plant Cell 16 (2004) 1650–1660.
[12] W. Chi, J. Ma, L. Zhang, Regulatory factors for the assembly of thylakoid membrane protein complexes, Phil. Trans. R. Soc. B 367 (2012) 3420–3429.
[13] B.L. Gutman, K.K. Niyogi, *Chlamydomonas* and *Arabidopsis*. A dynamic duo, Plant Physiol. 135 (2004) 607–610.
[14] N. Bryant, J. Lloyd, C. Sweeney, F. Myouga, D. Meinke, Identification of nuclear genes encoding chloroplast-localized proteins required for embryo development in *Arabidopsis thaliana*, Plant Physiol. 155 (2011) 1678–1689.
[15] C. Miles, D. Daniel, Chloroplast reactions of photosynthetic mutants in *Zea mays*, Plant Physiol. 53 (1974) 589–595.
[16] D. Miles, Mutants of higher plants: maize, Methods Enzymol. 69 (1980) 3–23.
[17] W. Cook, D. Miles, Transposon mutagenesis of nuclear photosynthetic genes in *Zea mays*, Photosyn. Res. 18 (1988) 33–59.
[18] A. Barkan, Approaches to investigating nuclear genes that function in chloroplast biogenesis in land plants, Methods Enzymol. 297 (1998) 38–57.
[19] K. Hammani, W. Cook, A. Barkan, RNA binding and RNA remodeling activities of the half-a-tetratricopeptide (HAT) protein HCF107 underlie its effects on gene expression, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 5651–5656.
[20] R. Williams-Carrier, N. Stiffler, S. Belcher, T. Kroeger, D.B. Stern, R.A. Monde, R. Coalter, A. Barkan, Use of illumina sequencing to identify transposon insertions underlying mutant phenotypes in high-copy mutator lines of maize, Plant J. 63 (2010) 167–177.
[21] C.T. Hunter, M. Suzuki, J. Saunders, S. Wu, A. Tasi, D.R. McCarty, K.E. Koch, Phenotype to genotype using forward-genetic Mu-seq for identification and functional classification of maize mutants, Front. Plant Sci. 4 (2014) 545.
[22] K.W. Osteryoung, K.D. Stokes, S.M. Rutherford, A.L. Percival, W.Y. Lee, Chloroplast division in higher plants requires members of two functionally divergent gene families with homology to bacterial *ftsZ*, Plant Cell 10 (1998) 1991–2004.
[23] B. Jenkins, A. Barkan, Recruitment of a peptidyl-tRNA hydrolase as a facilitator of group II intron splicing in chloroplasts, EMBO J. 20 (2001) 872–879.
[24] B. Jenkins, D. Kulhanek, A. Barkan, Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors, Plant Cell 9 (1997) 283–296.
[25] B. Till, C. Schmitz-Linneweber, R. Williams-Carrier, A. Barkan, CRS1 is a novel group II intron splicing factor that was derived from a domain of ancient origin, RNA 7 (2001) 1227–1238.
[26] A. Barkan, M. Walker, M. Nolasco, D. Johnson, 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 (1994) 3170–3181.
[27] D.G. Fisk, M.B. Walker, A. Barkan, Molecular cloning of the maize gene *crp1* reveals similarity between regulators of mitochondrial and chloroplast gene expression, EMBO J. 18 (1999) 2621–2630.
[28] I. Small, N. Peeters, The PPR motif - a TPR-related motif prevalent in plant organellar proteins, Trends Biochem. Sci. 25 (2000) 46–47.
[29] R. Voelker, A. Barkan, Two nuclear mutations disrupt distinct pathways for targeting proteins to the chloroplast thylakoid, EMBO J. 14 (1995) 3905–3914.
[30] R. Voelker, J. Mendel-Hartvig, A. Barkan, Transposon-disruption of a maize nuclear gene, *tha1*, encoding a chloroplast SecA homolog: *in vivo* role of cp-SecA in thylakoid protein targeting, Genetics 145 (1997) 467–478.
[31] M. Walker, L. Roy, E. Coleman, R. Voelker, A. Barkan, The maize *tha4* gene functions in sec-independent protein transport in chloroplasts and is related to *hcf106*, *tatA*, and *tatB*, J. Cell Biol. 147 (1999) 267–275.
[32] C.-d. Han, E.H.J. Coe, R.A. Martienssen, Molecular cloning and characterization of iojap (ij), a pattern striping gene of maize, EMBO J. 11 (1992) 4037–4046.
[33] L. Hall, L. Rossini, L. Cribb, J. Langdale, Golden 2: a novel transcriptional regulator of cellular differentiation in the maize leaf, Plant Cell 10 (1998) 925–936.
[34] T. Brutnell, R. Sawers, A. Mant, J. Langdale, BUNDLE SHEATH DEFECTIVE 2, a novel protein required for post-translational regulation of the *rbcL* gene in maize, Plant Cell 11 (1999) 849–864.
[35] L. Feiz, R. Williams-Carrier, K. Wostrikoff, S. Belcher, A. Barkan, D.B. Stern, Ribulose-1,5-bis-phosphate carboxylase/oxygenase accumulation factor1 is required for holoenzyme assembly in maize, Plant Cell 24 (2012) 3435–3446.
[36] T. Kroeger, K. Watkins, G. Friso, K.v. Wijk, A. Barkan, A plant-specific RNA binding domain revealed through analysis of chloroplast group II intron splicing. Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 4537–4542.
[37] G. Ostheimer, R. Williams-Carrier, S. Belcher, E. Osborne, J. Gierke, A. Barkan, Group II intron splicing factors derived by diversification of an ancient RNA binding module, EMBO J. 22 (2003) 3919–3929.
[38] K. Watkins, T. Kroeger, A. Cooke, R. Williams-Carrier, G. Friso, S. Belcher, K.v. Wijk, A. Barkan, A ribonuclease III domain protein functions in group II intron splicing in maize chloroplasts, Plant Cell 19 (2007) 2606–2623.
[39] Y. Asakura, T. Hirohashi, S. Kikuchi, S. Belcher, E. Osborne, S. Yano, I. Terashima, A. Barkan, M. Nakai, Maize mutants lacking chloroplast FtsY exhibit pleiotropic defects in the biogenesis of thylakoid membranes, Plant Cell 16 (2004) 201–214.
[40] S. Beick, C. Schmitz-Linneweber, R. Williams-Carrier, B. Jensen, A. Barkan, The pentatricopeptide repeat protein PPR5 stabilizes a specific tRNA precursor in maize chloroplasts, Mol. Cell. Biol. 28 (2008) 5337–5347.
[41] J. Pfalz, O. Bayraktar, J. Prikryl, A. Barkan, Site-specific binding of a PPR protein defines and stabilizes 5′ and 3′ mRNA termini in chloroplasts, EMBO J. 28 (2009) 2042–2052.
[42] C. Schmitz-Linneweber, R.E. Williams-Carrier, P.M. Williams-Voelker, T.S. Kroeger, A. Vichas, A. Barkan, A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast rps12 Pre-mRNA, Plant Cell 18 (2006) 2650–2663.
[43] P. Williams, A. Barkan, A chloroplast-localized PPR protein required for plastid ribosome accumulation, Plant J. 36 (2003) 675–686.
[44] Y. Asakura, E. Galarneau, K.P. Watkins, A. Barkan, K.J. van Wijk, Chloroplast RH3 DEAD box RNA helicases in maize and *Arabidopsis* function in splicing of specific group II introns and affect chloroplast ribosome biogenesis, Plant Physiol. 159 (2012) 961–974.
[45] Z.P. Gao, G.X. Chen, Z.N. Yang, Regulatory role of *Arabidopsis* pTAC14 in chloroplast development and plastid gene expression, Plant Signal. Behav. 7 (2012) 1354–1356.
[46] J. Pfalz, K. Liere, A. Kandlbinder, K.J. Dietz, R. Oelmuller, pTAC2, −6, and −12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression, Plant Cell 18 (2006) 176–197.
[47] J. Qiao, J. Li, W. Chu, M. Luo, PRDA1, a novel chloroplast nucleoid protein, is required for early chloroplast development and is involved in the regulation of plastid gene expression in *Arabidopsis*, Plant Cell Physiol. 54 (2013) 2071–2084.
[48] L. Lezhneva, R. Kuras, G. Ephritikhine, C. de Vitry, A novel pathway of cytochrome c biogenesis is involved in the assembly of the cytochrome b6f complex in *Arabidopsis* chloroplasts, J. Biol. Chem. 283 (2008) 24608–24616.
[49] R. Kuras, C. de Vitry, Y. Choquet, J. Girard-Bascou, D. Culler, S. Buschlen, S. Merchant, F.A. Wollman, Molecular genetic identification of a pathway for heme binding to cytochrome b6, J. Biol. Chem. 272 (1997) 32427–32435.
[50] Z. Xie, D. Culler, B.W. Dreyfuss, R. Kuras, F.A. Wollman, J. Girard-Bascou, S. Merchant, Genetic analysis of chloroplast c-type cytochrome assembly in *Chlamydomonas* reinhardtii: One chloroplast locus and at least four nuclear loci are required for heme attachment, Genetics 148 (1998) 681–692.
[51] R. Motohashi, T. Ito, M. Kobayashi, T. Taji, N. Nagata, T. Asami, S. Yoshida, K. Yamaguchi-Shinozaki, K. Shinozaki, Functional analysis of the 37 kDa inner envelope membrane polypeptide in chloroplast biogenesis using a Ds-tagged *Arabidopsis* pale-green mutant, Plant J. 34 (2003) 719–731.
[52] R.M. Larkin, J.M. Alonso, J.R. Ecker, J. Chory, GUN4, a regulator of chlorophyll synthesis and intracellular signaling, Science 299 (2003) 902–906.
[53] N.D. Adhikari, J.E. Froehlich, D.D. Strand, S.M. Buck, D.M. Kramer, R.M. Larkin, GUN4-porphyrin complexes bind the ChlH/GUN5 subunit of Mg-Chelatase and promote chlorophyll biosynthesis in *Arabidopsis*, Plant Cell 23 (2011) 1449–1467.
[54] D. Bollivar, I. Braumann, K. Berendt, S.P. Gough, M. Hansson, The Ycf54 protein is part of the membrane component of Mg-protoporphyrin IX monomethyl ester cyclase from barley (Hordeum vulgare L.), FEBS J. 281 (2014) 2377–2386.
[55] S. Hollingshead, J. Kopecna, P.J. Jackson, D.P. Canniffe, P.A. Davison, M.J. Dickman, R. Sobotka, C.N. Hunter, Conserved chloroplast open-reading frame ycf54 is required for activity of the magnesium protoporphyrin monomethylester oxidative cyclase in Synechocystis PCC 6803, J. Biol. Chem. 287 (2012) 27823–27833.
[56] C.A. Albus, S. Ruf, M.A. Schottler, W. Lein, J. Kehr, R. Bock, Y3IP1, a nucleus-encoded thylakoid protein, cooperates with the plastid-encoded Ycf3 protein in photosystem I assembly of tobacco and *Arabidopsis*, Plant Cell 22 (2010) 2838–2855.
[57] S.R. Kim, J.I. Yang, G. An, OsCpn60alpha1, encoding the plastid chaperonin 60alpha subunit, is essential for folding of rbcL, Mol. Cells 35 (2013) 402–409.
[58] W. Cai, K. Okuda, L. Peng, T. Shikanai, PROTON GRADIENT REGULATION 3 recognizes multiple targets with limited similarity and mediates translation and RNA stabilization in plastids, Plant J. 67 (2011) 318–327.
[59] A. Barkan, Nuclear mutants of maize with defects in chloroplast polysome assembly have altered chloroplast RNA metabolism, Plant Cell 5 (1993) 389–402.

[60] R. Voelker, A. Barkan, Nuclear genes required for post-translational steps in the biogensis of the chloroplast cytochrome *b6f* complex, Mol. Gen. Genet. 249 (1995) 507–514.

[61] A.M. Duchene, A. Giritch, B. Hoffmann, V. Cognat, D. Lancelin, N.M. Peeters, M. Zaepfel, L. Marechal-Drouard, I.D. Small, Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in *Arabidopsis thaliana*, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 16484–16489.

[62] I. Lermontova, B. Grimm, Reduced activity of plastid protoporphyrinogen oxidase causes attenuated photodynamic damage during high-light compared to low-light exposure, Plant J. 48 (2006) 499–510.

[63] D. Panikashvili, S. Savaldi-Goldstein, T. Mandel, T. Yifhar, R.B. Franke, R. Hofer, L. Schreiber, J. Chory, A. Aharoni, The *Arabidopsis* DESPERADO/AtWBC11 transporter is required for cutin and wax secretion, Plant Physiol. 145 (2007) 1345–1360.

[64] D. Bird, F. Beisson, A. Brigham, J. Shin, S. Greer, R. Jetter, L. Kunst, X. Wu, A. Yephremov, L. Samuels, Characterization of *Arabidopsis* ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion, Plant J. 52 (2007) 485–498.

[65] R. Fristedt, R. Williams-Carrier, S.S. Merchant, A. Barkan, A thylakoid membrane protein harboring a DnaJ-type zinc finger domain is required for photosystem I accumulation in plants, J. Biol. Chem. 289 (2014) 30657–30667.

[66] T. Yabe, K. Morimoto, S. Kikuchi, K. Nishio, I. Terashima, M. Nakai, The *Arabidopsis* chloroplastic NifU-like protein CnfU, which can act as an iron-sulfur cluster scaffold protein, is required for biogenesis of ferredoxin and photosystem I, Plant Cell 16 (2004) 993–1007.

[67] S. Leon, B. Touraine, C. Ribot, J.F. Briat, S. Lobreaux, Iron-sulphur cluster assembly in plants: distinct NFU proteins in mitochondria and plastids from *Arabidopsis thaliana*, Biochem. J. 371 (2003) 823–830.

[68] B. Touraine, J.P. Boutin, A. Marion-Poll, J.F. Briat, G. Peltier, S. Lobreaux, Nfu2: a scaffold protein required for [4Fe-4S] and ferredoxin iron-sulphur cluster assembly in *Arabidopsis* chloroplasts, Plant J. 40 (2004) 101–111.

[69] P.C. Loh, T. Morimoto, Y. Matsuo, T. Oshima, N. Ogasawara, The GTP-binding protein YqeH participates in biogenesis of the 30S ribosome subunit in Bacillus subtilis, Genes Genet. Syst. 82 (2007) 281–289.

[70] A. Sayed, S. Matsuyama, M. Inouye, Era, an essential Escherichia coli small G-protein, binds to the 30S ribosomal subunit, Biochem. Biophys. Res. Commun. 264 (1999) 51–54.

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

[72] I.N. Suwastika, M. Denawa, S. Yomogihara, C.H. Im, W.Y. Bang, R.L. Ohniwa, J.D. Bahk, K. Takeyasu, T. Shiina, Evidence for lateral gene transfer (LGT) in the evolution of eubacteria-derived small GTPases in plant organelles, Front. Plant Sci. 5 (2014) 678.

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

[74] C.A. Albus, A. Salinas, O. Czarnecki, S. Kahlau, M. Rothbart, W. Thiele, W. Lein, R. Bock, B. Grimm, M.A. Schottler, LCAA, a novel factor required for magnesium protoporphyrin monomethylester cyclase accumulation and feedback control of aminolevulinic acid biosynthesis in tobacco, Plant Physiol. 160 (2012) 1923–1939.

[75] R. Zoschke, T. Kroeger, S. Belcher, M.A. Schottler, A. Barkan, C. Schmitz-Linneweber, The pentatricopeptide repeat-SMR protein ATP4 promotes translation of the chloroplast atpB/E mRNA, Plant J. 72 (2012) 547–558.

[76] R. Zoschke, K. Watkins, A. Barkan, A rapid microarray-based ribosome profiling method elucidates chloroplast ribosome behavior in vivo, Plant Cell 25 (2013) 2265–2275.

[77] X. Liu, F. Yu, S. Rodermel, An *Arabidopsis* pentatricopeptide repeat protein, SUPPRESSOR OF VARIEGATION7, is required for FtsH-mediated chloroplast biogenesis, Plant Physiol. 154 (2010) 1588–1601.

[78] R. Zoschke, Y. Qu, Y.O. Zubo, T. Borner, C. Schmitz-Linneweber, Mutation of the pentatricopeptide repeat-SMR protein SVR7 impairs accumulation and translation of chloroplast ATP synthase subunits in *Arabidopsis thaliana*, J. Plant Res. 126 (2013) 403–414.

[79] J. Prikryl, K.P. Watkins, G. Friso, K.J. Wijk, A. Barkan, A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis, Nucleic Acids Res. 36 (2008) 5152–5165.

[80] A. Marechal, J.S. Parent, F. Veronneau-Lafortune, A. Joyeux, B.F. Lang, N. Brisson, Whirly proteins maintain plastid genome stability in *Arabidopsis*, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 14693–14698.

[81] D. Dauvillee, O. Stampacchia, J. Girard-Bascou, J.D. Rochaix, Tab2 is a novel conserved RNA binding protein required for translation of the chloroplast psaB mRNA, EMBO J. 22 (2003) 6378–6388.

[82] F. Barneche, V. Winter, M. Crevecoeur, J.D. Rochaix, ATAB2 is a novel factor in the signalling pathway of light-controlled synthesis of photosystem proteins, EMBO J. 25 (2006) 5907–5918.

[83] S. Fujii, N. Sato, T. Shikanai, Mutagenesis of individual pentatricopeptide repeat motifs affects RNA binding activity and reveals functional partitioning of *Arabidopsis* PROTON gradient regulation3, Plant Cell 25 (2013) 3079–3088.

[84] H. Yamazaki, M. Tasaka, T. Shikanai, PPR motifs of the nucleus-encoded factor, PGR3, function in the selective and distinct steps of chloroplast gene expression in *Arabidopsis*, Plant J. 38 (2004) 152–163.

[85] T. Shikanai, Central role of cyclic electron transport around photosystem I in the regulation of photosynthesis, Curr. Opin. Biotechnol. 26 (2014) 25–30.

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

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

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

[89] T. Kleine, *Arabidopsis thaliana* mTERF proteins: evolution and functional classification, Front. Plant Sci. 3 (2012) 233.

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

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

[92] A.L. Chateigner-Boutin, C. Colas des Francs-Small, E. Delannoy, S. Kahlau, S.K. Tanz, A. Falcon de Longevialle, S. Fujii, I. Small, OTP70 is a pentatricopeptide repeat protein of the E subgroup involved in splicing of the plastid transcript rpoC1, Plant J. 65 (2011) 532–542.

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

[94] Y. Asakura, A. Barkan, A CRM domain protein functions dually in group I and group II intron splicing in land plant chloroplasts, Plant Cell 19 (2007) 3864–3875.

[95] Y. Asakura, O. Bayraktar, A. Barkan, Two CRM protein subfamilies cooperate in the splicing of group IIB introns in chloroplasts, RNA 14 (2008) 2319–2332.

[96] K. Amann, L. Lezhneva, G. Wanner, R.G. Herrmann, J. Meurer, ACCUMULATION OF PHOTOSYSTEM ONE1, a member of a novel gene family, is required for accumulation of [4Fe-4S] cluster-containing chloroplast complexes and antenna proteins, Plant Cell 16 (2004) 3084–3097.

[97] M. Zmudjak, C. Colas des Francs-Small, I. Keren, F. Shaya, E. Belausov, I. Small, O. Ostersetzer-Biran, mCSF1, a nucleus-encoded CRM protein required for the processing of many mitochondrial introns, is involved in the biogenesis of respiratory complexes I and IV in *Arabidopsis*, New Phytol. 199 (2013) 379–394.

[98] C.C. Francs-Small, T. Kroeger, M. Zmudjak, O. Ostersetzer-Biran, N. Rahimi, I. Small, A. Barkan, A PORR domain protein required for rpl2 and ccmF(C) intron splicing and for the biogenesis of c-type cytochromes in *Arabidopsis mitochondria*, Plant J. 69 (2012) 996–1005.

[99] Y. Asakura, A. Barkan, *Arabidopsis* orthologs of maize chloroplast splicing factors promote splicing of orthologous and species-specific group II introns, Plant Physiol. 142 (2006) 1656–1663.

[100] L. Feiz, R. Williams-Carrier, S. Belcher, M. Montano, A. Barkan, D.B. Stern, A protein with an inactive pterin-4a-carbinolamine dehydratase domain is required for Rubisco biogenesis in plants, Plant J. 80 (2014) 862–869.

[101] R. Trosch, T. Muhlhaus, M. Schroda, F. Willmund, ATP-dependent molecular chaperones in plastids – More complex than expected, Biochim. Biophys. Acta (2015).

[102] P. Kolesinski, I. Belusiak, M. Czarnocki-Cieciura, A. Szczepaniak, Rubisco Accumulation Factor 1 from *Thermosynechococcus elongatus* participates in the final stages of ribulose-1,5-bisphosphate carboxylase/oxygenase assembly in Escherichia coli cells and in vitro, FEBS J. 281 (2014) 3920–3932.

[103] N.M. Wheatley, C.D. Sundberg, S.D. Gidaniyan, D. Cascio, T.O. Yeates, Structure and identification of a pterin dehydratase-like protein as a ribulose-bisphosphate carboxylase/oxygenase (RuBisCO) assembly factor in the alpha-carboxysome, J. Biol. Chem. 289 (2014) 7973–7981.

[104] Y. Lu, L. Savage, M. Larson, C. Wilkerson, R.L. Last, Chloroplast 2010: a database for large-scale phenotypic screening of Arabidopsis mutants, Plant Physiol. 15 (2011) 1589–1600.

[105] L.J. Savage, K.M. Imre, D.A. Hall, R.L. Last, Analysis of essential *Arabidopsis* nuclear genes encoding plastid-targeted proteins, PLoS One 8 (2013) e73291.

[106] F. Myouga, K. Akiyama, Y. Tomonaga, A. Kato, Y. Sato, M. Kobayashi, N. Nagata, T. Sakurai, K. Shinozaki, The chloroplast function database II: a comprehensive collection of homozygous mutants and their phenotypic/genotypic traits for nuclear-encoded chloroplast proteins, Plant Cell Physiol. 54 (2013) e2.

[107] I. Ajjawi, Y. Lu, L.J. Savage, S.M. Bell, R.L. Last, Large-scale reverse genetics in *Arabidopsis*: case studies from the Chloroplast 2010 Project, Plant Physiol. 152 (2010) 529–540.

[108] D. Wilson-Sanchez, S. Rubio-Diaz, R. Munoz-Viana, J.M. Perez-Perez, S. Jover-Gil, M.R. Ponce, J.L. Micol, Leaf phenomics: a systematic reverse genetic screen for *Arabidopsis* leaf mutants, Plant J. 79 (2014) 878–891.

[109] S.J. Karpowicz, S.E. Prochnik, A.R. Grossman, S.S. Merchant, The GreenCut2 resource, a phylogenomically derived inventory of proteins specific to the plant lineage, J. Biol. Chem. 286 (2011) 21427–21439.

[110] T. Long, D. Guo, D. He, W. Shen, X. Li, The tRNA 3′-end processing enzyme tRNase Z2 contributes to chloroplast biogenesis in rice, J. Integr. Plant Biol. 55 (2013) 1104–1118.

[111] M. Tomcal, N. Stiffler, A. Barkan, POGs2: a web portal to facilitate cross-species inferences about protein architecture and function in plants, PLoS One 8 (2013) e82569.

[112] M. Van Bel, S. Proost, E. Wischnitzki, S. Movahedi, C. Scheerlinck, Y. Van de Peer, K. Vandepoele, Dissecting plant genomes with the PLAZA comparative genomics platform, Plant Physiol. 158 (2012) 590–600.

[113] M. Huang, G. Friso, K. Nishimura, X. Qu, P.D. Olinares, W. Majeran, Q. Sun, K.J. van Wijk, Construction of plastid reference proteomes for maize and *Arabidopsis* and evaluation of their orthologous relationships; the concept of orthoproteomics, J. Proteome Res. 12 (2013) 491–504.

[114] C. Schmitz-Linneweber, R. Williams-Carrier, A. Barkan, 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 (2005) 2791–2804.
[115] J. Prikryl, M. Rojas, G. Schuster, A. Barkan, Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 415–420.
[116] R. Williams-Carrier, T. Kroeger, A. Barkan, Sequence-specific binding of a chloroplast pentatricopeptide repeat protein to its native group II intron ligand, RNA 14 (2008) 1930–1941.
[117] O. Ostersetzer, K. Watkins, A. Cooke, A. Barkan, CRS1, a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains, Plant Cell 17 (2005) 241–255.
[118] G.J. Ostheimer, H. Hadjivassiliou, D.P. Kloer, A. Barkan, B.W. Matthews, Structural analysis of the group II intron splicing factor CRS2 yields insights into its protein and RNA interaction surfaces, J. Mol. Biol. 345 (2005) 51–68.
[119] G.J. Ostheimer, M. Rojas, H. Hadjivassiliou, A. Barkan, Formation of the CRS2-CAF2 group ll intron splicing complex is mediated by a 22 amino acid motif in the C-terminal region of CAF2, J. Biol. Chem. 281 (2006) 4732–4738.
[120] A. Germain, S. Herlich, S. Larom, S.H. Kim, G. Schuster, D.B. Stern, Mutational analysis of *Arabidopsis* chloroplast polynucleotide phosphorylase reveals roles for both RNase PH core domains in polyadenylation, RNA 3′-end maturation and intron degradation, Plant J. 67 (2011) 381–394.
[121] T. Sun, A. Germain, K. Hammani, A. Barkan, M. Hanson, S. Bentolila, An RNA recognition motif-containing protein is required for plastid RNA editing in *Arabidopsis* and maize, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) E1169–E1178.
[122] M. Garcia, F. Myouga, K. Takechi, H. Sato, K. Nabeshima, N. Nagata, S. Takio, K. Shinozaki, H. Takano, An *Arabidopsis* homolog of the bacterial peptidoglycan synthesis enzyme MurE has an essential role in chloroplast development, Plant J. 53 (2008) 924–934.
[123] P. Kindgren, D. Kremnev, N.E. Blanco, J. de Dios Barajas Lopez, A.P. Fernandez, C. Tellgren-Roth, T. Kleine, I. Small, A. Strand, The plastid redox insensitive 2 mutant of *Arabidopsis* is impaired in PEP activity and high light-dependent plastid redox signalling to the nucleus, Plant J. 70 (2012) 279–291.
[124] D. Kremnev, A. Strand, Plastid encoded RNA polymerase activity and expression of photosynthesis genes required for embryo and seed development in *Arabidopsis*, Front. Plant Sci. 5 (2014) 385.
[125] S. Felder, K. Meierhoff, A.P. Sane, J. Meurer, C. Driemel, H. Plucken, P. Klaff, B. Stein, N. Bechtold, P. Westhoff, 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 (2001) 2127–2141.
[126] A.P. Sane, B. Stein, P. Westhoff, 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 (2005) 720–730.
[127] T. Levey, P. Westhoff, K. Meierhoff, Expression of a nuclear-encoded psbH gene complements the plastidic RNA processing defect in the PSII mutant hcf107 in *Arabidopsis thaliana*, Plant J. 80 (2014) 292–304.
[128] D.B. Udy, S. Belcher, R. Williams-Carrier, J.M. Gualberto, A. Barkan, Effects of reduced chloroplast gene copy number on chloroplast gene expression in maize, Plant Physiol. 160 (2012) 1420–1431.
[129] Y.F. Zhang, M.M. Hou, B.C. Tan, The requirement of WHIRLY1 for embryogenesis is dependent on genetic background in maize, PLoS One 8 (2013) e67369.
[130] J. Melonek, M. Mulisch, C. Schmitz-Linneweber, E. Grabowski, G. Hensel, K. Krupinska, Whirly1 in chloroplasts associates with intron containing RNAs and rarely co-localizes with nucleoids, Planta 232 (2010) 471–481.
[131] R. Martienssen, A. Barkan, M. Freeling, W. Taylor, Molecular cloning of a maize gene involved in photosynthetic membrane organization that is regulated by Robertson's Mutator, EMBO J. 8 (1989) 1633–1639.
[132] A.M. Settles, A. Yonetani, A. Baron, D.R. Bush, K. Cline, R. Martienssen, Sec-independent protein translocation by the maize Hcf106 protein, Science 278 (1997) 1467–1470.
[133] L.M. Roy, A. Barkan, A secY homologue is required for the elaboration of the chloroplast thylakoid membrane and for normal chloroplast gene expression, J. Cell Biol. 141 (1998) 385–395.
[134] S.M. Bowman, M. Patel, P. Yerramsetty, C.M. Mure, A.M. Zielinski, J.A. Bruenn, J.O. Berry, A novel RNA binding protein affects rbcL gene expression and is specific to bundle sheath chloroplasts in C4 plants, BMC Plant Biol. 13 (2013) 138.
[135] A. Xing, M.E. Williams, T.M. Bourett, W. Hu, Z. Hou, R.B. Meeley, J. Jaqueth, T. Dam, B. Li, A pair of homoeolog ClpP5 genes underlies a virescent yellow-like mutant and its modifier in maize, Plant J. 79 (2014) 192–205.
[136] V. Walbot, E.H. Coe, Nuclear gene *iojap* conditions a programmed change to ribosome-less plastids in *Zea mays*, Proc. Natl. Acad. Sci. U. S. A. 76 (1979) 2760–2764.
[137] N.P. Schultes, R.J. Sawers, T.P. Brutnell, R.W. Krueger, Maize high chlorophyll fluorescent 60 mutation is caused by an Ac disruption of the gene encoding the chloroplast ribosomal small subunit protein 17, Plant J. 21 (2000) 317–327.
[138] N.P. Schultes, T.P. Brutnell, A. Allen, S.L. Dellaporta, T. Nelson, J. Chen, Leaf permease1 gene of maize is required for chloroplast development, Plant Cell 8 (1996) 463–475.
[139] S. Covshoff, W. Majeran, P. Liu, J.M. Kolkman, K.J. van Wijk, T.P. Brutnell, Deregulation of maize C4 photosynthetic development in a mesophyll cell-defective mutant, Plant Physiol. 146 (2008) 1469–1481.
[140] J. Meurer, H. Plucken, K. Kowallik, P. Westhoff, A nuclear-encoded protein of prokaryotic origin is essential for the stability of photosystem II in *Arabidopsis thaliana*, EMBO J. 17 (1998) 5286–5297.
[141] D. Shi, X. Zheng, L. Li, W. Lin, W. Xie, J. Yang, S. Chen, W. Jin, Chlorophyll deficiency in the maize elongated mesocotyl2 mutant is caused by a defective heme oxygenase and delaying grana stacking, PLoS One 8 (2013) e80107.
[142] P. Williams, K. Hardeman, J. Fowler, C. Rivin, Divergence of duplicated genes in maize: evolution of contrasting targeting information for enzymes in the porphyrin pathway, Plant J. 45 (2006) 727–739.
[143] D. Sosso, M. Canut, G. Gendrot, A. Dedieu, P. Chambrier, A. Barkan, G. Consonni, P.M. Rogowsky, PPR8522 encodes a chloroplast-targeted pentatricopeptide repeat protein necessary for maize embryogenesis and vegetative development, J. Exp. Bot. 63 (2012) 5843–5857.
[144] X.M. Lu, X.J. Hu, Y.Z. Zhao, W.B. Song, M. Zhang, Z.L. Chen, W. Chen, Y.B. Dong, Z.H. Wang, J.S. Lai, Map-based cloning of zb7 encoding an IPP and DMAPP synthase in the MEP pathway of maize, Mol. Plant 5 (2012) 1100–1112.
[145] M. Huang, T.L. Slewinski, R.F. Baker, D. Janick-Buckner, B. Buckner, G.S. Johal, D.M. Braun, Camouflage patterning in maize leaves results from a defect in porphobilinogen deaminase, Mol. Plant 2 (2009) 773–789.
[146] R.J. Sawers, J. Viney, P.R. Farmer, R.R. Bussey, G. Olsefski, K. Anufrikova, C.N. Hunter, T.P. Brutnell, The maize Oil yellow1 (Oy1) gene encodes the I subunit of magnesium chelatase, Plant Mol. Biol. 60 (2006) 95–106.
[147] C. Kagi, N. Baumann, N. Nielsen, Y.D. Stierhof, R. Gross-Hardt, The gametic central cell of *Arabidopsis* determines the lifespan of adjacent accessory cells, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 22350–22355.
[148] Z.P. Gao, Q.B. Yu, T.T. Zhao, Q. Ma, G.X. Chen, Z.N. Yang, A functional component of the transcriptionally active chromosome complex, *Arabidopsis* pTAC14, interacts with pTAC12/HEMERA and regulates plastid gene expression, Plant Physiol. 157 (2011) 1733–1745.
[149] Z. Xie, S. Merchant, A novel pathway for cytochromes c biogenesis in chloroplasts, Biochim. Biophys. Acta 1365 (1998) 309–318.
[150] C. Dal Bosco, L. Lezhneva, A. Biehl, D. Leister, H. Strotmann, G. Wanner, J. Meurer, Inactivation of the chloroplast ATP synthase gamma subunit results in high non-photochemical fluorescence quenching and altered nuclear gene expression in *Arabidopsis thaliana*, J. Biol. Chem. 279 (2004) 1060–1069.