Post-Transcriptional Control of Chloroplast Gene Expression

Eva M. del Campo

Department of Plant Biology, University of Alcalá, Alcalá de Henares, 28871-Madrid, Spain.

Abstract: Chloroplasts contain their own genome, organized as operons, which are generally transcribed as polycistronic transcriptional units. These primary transcripts are processed into smaller RNAs, which are further modified to produce functional RNAs. The RNA processing mechanisms remain largely unknown and represent an important step in the control of chloroplast gene expression. Such mechanisms include RNA cleavage of pre-existing RNAs, RNA stabilization, intron splicing, and RNA editing. Recently, several nuclear-encoded proteins that participate in diverse plastid RNA processing events have been characterised. Many of them seem to belong to the pentatricopeptide repeat (PPR) protein family that is implicated in many crucial functions including organelle biogenesis and plant development. This review will provide an overview of current knowledge of the post-transcriptional processing in chloroplasts.

Keywords: Chloroplasts, primary transcripts, RNA processing, RNA stabilization, intron splicing, RNA editing, PPR protein

Introduction

Since the discovery of the existence of DNA\(^1\) and ribosomes\(^2\) in chloroplasts, many studies have been published about the structure of the chloroplast genome and its expression. These studies were facilitated by the development of cloning and sequencing techniques in the 1970s. The first physical map of plastid DNA was obtained from maize\(^3\) and the first plastid gene was cloned in 1977.\(^4\) A decade later, the complete chloroplast genome of tobacco,\(^5\) Marchantia polymorpha\(^6\) and rice\(^7\) was sequenced. These first approaches culminated in an emerging new field: gene organization and expression of the chloroplast genome. This field has subsequently become one of the most studied in plant molecular biology. The chloroplast genome has both prokaryotic and eukaryotic properties,\(^8\) but resembles prokaryotic systems since it has \(\sigma^{70}\) type promoters, a plastid encoded RNA polymerase, operons, “Shine-Dalgarno”-like sequences, and 70S ribosomes. The chloroplast genetic machinery also has characteristics of nuclear systems with the presence of introns and highly stable mRNAs. Consequently, the control of chloroplast gene expression includes several processes that are similar to those of prokaryotic and/or eukaryotic systems. These processes are: transcription, post-transcriptional processing, translation, and post-translational modifications. Generally, transcription rates and steady-state mRNA levels are not consistent suggesting that post-transcriptional RNA processing and stabilization are decisive steps in controlling plastid gene expression. This step principally includes RNA cleavage of pre-existing RNAs, RNA stabilization-degradation, intron splicing, and RNA editing (Fig. 1).

Plastid Transcriptional Machinery

Most of the genes encoded in higher plant chloroplasts, including genes involved in related functions, are organized as operons. However, they may also encode functionally unrelated genes.\(^9,10\) Plastid operons are transcribed as polycistronic units by at least two distinct RNA polymerase activities: the plastid-encoded (PEP) and the nuclear-encoded (NEP) RNA polymerases.\(^11,12\) PEP is a multisubunit complex which resembles eubacterial RNA polymerases and is the predominant transcriptional activity in mature chloroplasts. PEP recognizes \(E. \text{coli} \ \sigma^{70}\) type promoters whose typical TTGACA (−35) and TATAAT (−10) consensus elements are found upstream of most plastid transcriptional units. The PEP core enzyme is composed of four different subunits, \(\alpha, \beta, \beta'\) and \(\beta'\), which are encoded on the plastid genome by \(rpoA, rpoB, rpoC1\) and \(C2\) genes.\(^13\) The activity of the PEP core enzyme is regulated by sigma-like transcription factors (SLFs) that paly a role in promoter selection in a similar manner to the...
Six different sigma factors, SIG1–SIG6, have been described for Arabidopsis thaliana. The mRNAs of these SLFs are translated in the cytoplasm and the corresponding proteins are subsequently imported as precursor proteins into the plastids. Recently, several investigations have elucidated the role of sigma factors by analyzing Arabidopsis T-DNA insertion lines with disrupted SIG genes. SIG2 is known to specifically transcribe some of the tRNA genes and the psaI gene, SIG3 specifically transcribes the psbN gene in plastids, SIG4 is of specific importance for ndhF gene transcription, SIG5 has been shown to play an important role in the recognition of the blue-light dependent promoter of the psbD gene and SIG6 plays a more general role during early plastid differentiation and plant development.
There is a second nuclear-encoded transcription activity in chloroplasts (NEP, nuclear-encoded plastid RNA polymerase) supplementary to PEP. Most NEP promoters have a core sequence motif YRTA and are known as type-Ia, similar to plant mitochondria promoters. A subclass of NEP promoters, known as type-Ib, shares a GAA-box motif upstream of the YRTA-motif. Type-II NEP promoters lack these motifs and possess crucial sequences located downstream of the transcription initiation site, represented by dicot clpP promoters. Unlike PEP, NEP is a single subunit enzyme sharing homology with the RNA polymerases of phage T3 and T7. Initially, a gene encoding NEP was sequenced in several plants. Further isolation of functionally distinct NEP activities in spinach chloroplasts and the identification of two genes for NEP-like isozymes in Arabidopsis suggested the existence of additional NEP activities.

Recent evidence indicates that NEP is represented by two phage-type RNA polymerases (RpoTp and RpoTmp) that have overlapping as well as gene-specific functions in the transcription of plastidial genes in A. thaliana. RpoTp is localized in chloroplasts whereas RpoTmp, exclusively found in dicots, is presumably localized in both mitochondria and chloroplasts. In vitro transcription assays revealed no significant promoter specificity for RpoTmp and the accurate transcription initiation from overlapping subsets of mitochondrial and plastidial promoters without the aid of protein cofactors. RpoTp is a catalytic subunit of NEP involved in recognition of a distinct subset of type I NEP promoters. Mutational approaches indicated that the plastid RpoTp RNA Polymerase is required for chloroplast biogenesis and mesophyll cell proliferation in Arabidopsis. Evidence indicates that RpoTmp and RpoTp are involved in similar developmental events and that they are partially redundant.

However, in contrast to the role assigned to RpoTp in both early and late stages of vegetative development in Arabidopsis, RpoTmp is required in early seedling development. It has been shown that RPOTmp fulfills a specific function in the transcription of the rrn operon in proplasts/amyloplasts during seed imbibition/germination. In chloroplast, RpoTp is tightly associated with thylakoid membranes and interacts with a RING-H2 protein that in turn mediates intraplastidial trafficking of the RPOTmp RNA polymerase. The same research work presented a model in which light determines membrane association and functional switching of RPOTmp by triggering the synthesis of the RING protein. Interestingly, comparison of plastidial promoters from tobacco and Arabidopsis revealed a high diversity, which may also apply to other plants. The diversity in individual promoter usage in different plants suggests that there are species-specific ways of controlling gene expression in plastids.

**Chloroplast RNA Processing and Stability**

Evidence indicates that the control of chloroplast gene expression relies more on RNA processing and stability than on transcriptional regulation. In chloroplasts, polycistronic primary RNAs transcribed by PEP and/or NEP are generally processed into smaller transcripts which are further modified. RNA processing mechanisms remain largely unknown and sometimes controversial in spite of the diverse studies, focusing on several aspects of chloroplast gene expression reviewed in. Nowadays, the fact that post-transcriptional RNA processing of primary transcripts represents an important step in the control of chloroplast gene expression appears to be well accepted. In several cases, alternative processing of polycistronic primary transcripts may cause the simultaneous stabilization and degradation of alternative transcripts, resulting in the enhancement and inhibition of their translation, respectively. Whether or not transcript processing influences its translation into proteins remains controversial. In light of this, several investigations indicated that intercistronic processing is crucial for the translation of chloroplast operons and that the translation of monocistronic forms is more effective than translation of polycistronic forms. Nevertheless, in some cases it seems that translatable transcripts can be produced by both direct transcription from the promoter and intercistronic cleavage of pre-existing transcripts. Additionally, recent investigations with transgenic lines have demonstrated that processing into monocistrons is not required for over-expression of transgenes and that they are efficiently translated. Unlike higher plants, in the green alga Chlamydomonas reinhardtii, translation seems to be an essential step in the regulation of chloroplast gene expression. In these algae, transcript processing is less important in controlling plastid gene expression than in...
higher plants since nearly all genes appear to be transcribed as monocistronic RNAs.

In chloroplasts, transcript stability is mainly influenced by the presence of 5’ untranslated regions (5'-UTRs) and 3'-UTRs, that seem to be necessary to prevent the rapid degradation or low accumulation of primary transcripts. Deletions or mutating them destablizes the RNA, leading to reduced transcript accumulation and translation. Most of the plastid transcripts have short inverted repeat sequences (IR) that can potentially form a stem loop secondary structure (Fig. 1). In prokaryotic organisms, similar structures appear to play a crucial role in transcription termination of RNAs. However, in chloroplasts, transcription termination is very inefficient, resulting in considerable read-through transcription of downstream sequences. Therefore, the role of plastid 3'-UTRs differs from the role of its bacterial counterparts since they are more involved in transcript stability preventing 3’ to 5’ exonucleolytic degradation of transcripts than in the effective termination of transcription.

Another post-transcriptional modification affecting transcript stability is RNA polyadenylation (Fig. 1). In chloroplasts, poly(A) tails are found in degradation intermediate 3’-ends that contain not only adenosine but also other residues, principally guanosine. In chloroplast extracts, polyadenylated RNAs are degraded faster than nonadenylated RNAs and are more abundant in vivo under specific conditions that promote RNA degradation. Thus, polyadenylation might promote plastid RNA turnover in vivo by targeting endonucleolytic cleavage products for degradation as described for bacteria and plant mitochondria. The molecular mechanism of RNA degradation in chloroplasts appears very similar to that of bacteria. The first step consists of endonucleolytic cleavage of the RNA molecule, followed by polyadenylation. The polyadenylated cleavage products, including mRNAs and released introns, are then directed to rapid exonucleolytic degradation by PNPase and possibly other exonucleases (Fig. 1). Recent studies have revealed that although this enzyme is essential for efficient 3'-end processing of mRNAs, it is insufficient to mediate transcript degradation revealing an additional function of this exonuclease in tRNA degradation in Arabidopsis thaliana. In the last few years, several nuclear-encoded proteins that participate in chloroplast transcript processing and stabilization have been characterised. Most of them have been studied in Arabidopsis mutants (see Table 1). CRS2 is a protein that is involved in the intercistronic processing of rps7-ndhB transcripts. Such RNA processing seems to be essential for ndhB translation. This protein was first described in Arabidopsis mutants known as “chlororespiratory reduction mutants,” with reduced chloroplast NDH activity. crr2-1 and crr2-2 are recessive mutant alleles responsible for the impaired accumulation of the NDH complex. HCF152, encoded by the gene hcf152, is a RNA-binding protein that is involved in the processing or stabilization of the petB transcripts within the psbB-psbT-psbH-petB-petD operon. This gene was first identified in the nonphotosynthetic mutant of Arabidopsis hcf152 which does not produce the cytochrome b6f. The P67 protein seems to participate in the processing and translation of specific chloroplast mRNAs in radish and Arabidopsis. PGR3 is a nuclear-encoded protein which might have different functions in conferring RNA stability to the primary tricistronic transcript of the petL operon. This regulatory protein was described in pgr3 (proton gradient regulation 3) mutants of Arabidopsis, which display high chlorophyll fluorescence (HCF) because of a reduced level of the cytochrome b6/f complex.

Mutants of several plant species other than Arabidopsis have revealed the existence of new nuclear-encoded proteins which participate in chloroplast RNA processing and/or stabilization. In maize, the CRP1 protein is required for the translation of the chloroplast petA and petD transcripts and for the processing of the petD mRNA from a polycistronic precursor. Analysis of double mutants that lack both chloroplast ribosomes and CRP1 function suggested that CRP1 activates a site-specific endoribonuclease independently of any role it plays in translation.

Zmppr5 is the maize ortholog of the embryo-essential Arabidopsis gene Atppr5. The protein product of this gene is bound in vivo to the unspliced precursor of trnG-UCC RNA. Null and hypomorphic Zmppr5 insertion mutants are embryo viable but show deficiency in chloroplast ribosomes and die as seedlings. In these mutants, transcription of trnG-UCC is unaffected but their encoded transcripts are dramatically decreased. This observation, in addition to biochemical data,
Table 1. List of pentaricopeptide repeat protein genes involved in the control of chloroplast gene expression.

| Organism       | Gene     | Subfamily | PPR repeats | Target           | Possible function                                                                 | Accession | Reference                                      |
|----------------|----------|-----------|-------------|------------------|----------------------------------------------------------------------------------|-----------|------------------------------------------------|
| A. thaliana    | CLB19    | P-subfamily | 9           | rpoA, clpP       | Editing of rpoA and clpP transcripts                                             | Q9MA50    | Chateigner-Boutin et al. | 169 |
|                | CRR2     | PCMP-H    | 9           | rps7/ndhB        | RNA processing between rps7 and ndhB                                              | NP_190263| Hashimoto et al.     | 78  |
|                | CRR4     | PCMP-E    | 11          | ndhD             | RNA editing                                                                      | NP_182060| Koter et al.; Okuda et al.; Okuda et al.   | 166, 167, 168 |
|                | CRR21    | P-subfamily | 13          | ndhD             | RNA editing                                                                      | NP_200385| Okuda et al.         | 168 |
|                | DG1      | P-subfamily | 7           | PEP              | Regulation of plastid-encoded polymerase-dependent chloroplast gene expression   | NP_201558| Chi et al.          | 184 |
|                | HCF152   | P-subfamily | 12          | psbH/petB        | Processing and/or stabilization of polycistronic transcripts of the operon       | NP_187576| Meierhoff et al.; Nakamura et al. | 179, 79 |
|                | OTP51    | P-subfamily | 4           | ycf3             | Splicing of ycf3 intron 2 and other group-IIa introns                            | NP_565382| de Longevialle et al. | 182 |
|                | P67      | P-subfamily | 2           | Not known        | Processing or the translation of RNAs                                            | NP_193372| Lahmy et al.        | 80  |
|                | PGR3     | P-subfamily | 27          | petL operon      | Stabilization of the primary tricistronic transcript of the petL operon          | NP_194913| Yamazaki et al.      | 81  |
|                | AtPPR4   | P-subfamily | 16          | Not known        | Ribosome biogenesis                                                              | NP_568141| Schmitz-Linneweber et al. | 113 |
| C. reinhardtii | MCA1     | P-subfamily | 10          | petA             | Required for stable accumulation of the petA transcript                          | AAK14341 | Raynaud et al.      | 86  |
| P. patens      | PPR531-11| P-subfamily | 11          | rps12, clpP      | Intergenic RNA cleavage between clpP and 5’-rps12 and the splicing of clpP pre-mRNA | BAF02664 | Hattori et al.      | 85  |
| O. sativa      | OsPPR1   | P-subfamily | 11          | Not known        | Chloroplast biogenesis                                                            | AAS93059 | Gothandam et al.     | 181 |
| Z. mays        | CRP1     | P-subfamily | 14          | petA, petD       | Processing of the petD and translation of the petA and petD RNAs                 | AAC25599 | Fisk et al.; Schmitz-Linneweber et al. | 82, 83 |
|                | PPR2     | P-subfamily | 11          | Not known        | Required for plastid ribosome accumulation                                        | AAP37977 | Williams and Barkan | 103 |
|                | PPR4     | P-subfamily | 16          | rps12            | Trans-splicing of rps12 RNA and ribosome biogenesis                              | ABF57644 | Schmitz-Linneweber et al. | 113 |
|                | PPR5     | P-subfamily | 8           | tmG-UCC          | Stabilization of the tmG-UCC tRNA precursor                                       | NP_001106062 | Beick et al. | 84  |
indicates that PPR5 stabilizes the trnG-UCC precursor by direct binding and protection of an endonuclease-sensitive site.

In the moss Physcomitrella patens, the ppr351-11-disrupted mutants display a significantly smaller protonemal colony, different chloroplast morphology, incomplete thylakoid membrane formation and a reduction of the quantum yield of photosystem II. Several analyses have demonstrated that PPR531-11 has a role in intergenic RNA cleavage between clpP and 5'-rps12 and in the splicing of clpP pre-mRNA affecting the steady-state level of ClpP, which regulates the formation and maintenance of thylakoid membranes in chloroplasts.

In the unicellular alga Chlamydomonas reinhardtii, expression of the chloroplast petA gene encoding cytochrome f, depends on two specific nucleus-encoded factors: MCA1, required for stable accumulation of the petA transcript, and TCA1, required for its translation. Mutants with tagged versions of MCA1 and TCA1 show limited petA mRNA accumulation and cytochrome f translation, respectively. It has been proposed that a rapid drop in MCA1 exhausts the pool of petA transcripts, and the progressive loss of TCA1 further prevents translation of cytochrome f where de novo biogenesis of cytochrome b(6)f complexes is not required.

**Intron Splicing**

Several chloroplast genes, encoding both structural RNAs and proteins, are interrupted by introns. In chloroplasts, introns are classified into two main groups according to their conserved primary and secondary structures as well as their different splicing pathways, these are termed group I and group II introns. Land plant chloroplast genomes contain c.a. 20 group II introns and a single group I intron (within the trnL-UAA gene). However, a relatively high number of group I introns have been reported for green algae within organellar LSU rDNAs. Group I introns are found more frequently in fungi, plants, and algae. In organellar DNAs, group I introns are found in genes encoding rRNAs, tRNAs, and proteins but they are limited to genes encoding rRNAs in the nucleus. Group I introns are located in functionally vital loci and they must be removed from transcripts by splicing, a process which occurs co-ordinately with ligation of RNA exons. The intron folds to form a secondary structure consisting of ten domains, P1 to P10, each with specific roles in the formation of a catalytic core responsible for carrying out the splicing and ligation. Most of the conserved nucleotides correspond to the four short sequences P, Q, R, and S. These sequences are located in the same 5' to 3' order at variable distances from each other (form c.a. 20 nt to many hundreds). All of the group I introns, from several genetic systems of diverse organisms identified to date including green algae chloroplasts, have a U at their 5'-end and a G at their 3'-end. Splicing proceeds through two transesterification reactions with the first reaction involving cleavage at the 5' splice site and simultaneous addition of guanosine to the 5' intron end. The second reaction involves cleavage at the 3' splice site with concomitant ligation of exons. Group I intron splicing may be autocatalytic (self-splicing) or maturase facilitated. Several proteins from fungal mitochondria encoded by group I introns promote their splicing in vivo. However, self-splicing has only been tested by an in vitro assay in mitochondrial group I introns from Aspergillus nidulans. In Chlamydomonas reinhardtii it has been demonstrated the existence of nuclear genes that promote splicing of group I introns in the chloroplast 23S rRNA and psbA genes.

A remarkable feature of group I introns is their ability to colonize new insertion sites resulting in their spread. Intron insertion can occur via two alternative processes: reverse splicing and intron homing. Reverse splicing involves the insertion of a free intron into the RNA and has been observed in mobile group I introns integrated into the small subunit rRNA of bacteria and yeast. Intron homing is the insertion of an intron into a homologous position within an intronless copy of DNA. Intron homing is catalyzed by endonucleases, and are called homing endonucleases (HEs). Encoded by open reading frames (ORF) within introns, they recognize and cleave the target gene. In eukaryotes, HEs are found within nuclear and organellar genomes including both mitochondria and chloroplasts. HEs comprise four families known as: LAGLIDADG, GIY-YIG, His-Cys box, and HNH. In chloroplasts, HEs belonging to the LAGLIDADG, GIY-YIG, and HNH families have been discovered. The most studied chloroplast HEs were found within green algae of the genus Chlamydomonas. The LAGLIDADG family includes
I-CreI and I-Ceu-I proteins from the chloroplasts of *C. reinhardtii* and *C. eugametos* respectively which have only one LAGLIDADG sequence motif and function as homodimers. X-ray crystallography has generated structural models for group I intron-encoded I-CreI HE [23S rRNA gene from *Chlamydomonas reinhardtii* chloroplast].

The LAGLIDADG motifs form structurally conserved alpha-helices located at the center of the interdomain. The DNA-binding interface forms a four-stranded beta-sheet located on either side of the LAGLIDADG alpha-helices. The last acidic residue of the LAGLIDADG motif participates in DNA cleavage by phosphodiester hydrolysis.100

The GIY-YIG family includes monomeric enzymes which are characterized by the conserved GIY-(X<sub>10–11</sub>)-YIG motif. In chloroplasts, the ORFs in introns 2 and 3 (Cr.psbA2 and Cr.psbA3) within the *psbA* gene of *C. reinhardtii* contain variants of the GIY-YIG motif.102 The I-CreII protein is an ORF within intron 4 (Cr.psbA4) of the *psbA* gene of *C. reinhardtii*. This HE contains an H–N–H and possibly a GIY–YIG motif.103 This protein is a double-strand-specific endonuclease that cleaves fused *psbA* exon 4–exon 5 DNA. Cleavage of heterologous *psbA* DNAs has been demonstrated indicating that the enzyme can tolerate multiple, but not all, substitutions in the recognition site.

Group II introns are broadly distributed in diverse genetic systems including the chloroplast genome. This intron group can be distinguished by its folding into a characteristic secondary structure consisting on six helical domains radiating from a central core. There are two exon binding sites (EBS1 and ENS2) located within domain I. These exon binding sites interact with two intron binding sites (EBS1 and IBS2) located within the first twelve nucleotides of the intron 5’ end and their splicing proceeds via two alternative pathways known as the “branching” and “hydrolytic” pathways. The branching pathway consists of two consecutive transesterification reactions. During the first reaction, the first nucleotide of the intron 5’ end establishes a temporary 2’–5’ bond with a bulging adenosine located within domain VI. After intron splicing, the 5’ and 3’ exons join and the intron is released in a lariat form. The alternative splicing pathway starts by the hydrolytic cleavage of the 5’–splice site instead of transesterification.

In chloroplasts, most group II introns have a bulging adenosine within their domain VI and the splicing seems to occur via the branching pathway except for the *trnV(UAC)* transcripts.106 In spite of the fact that plastid group II introns are large ribozymes, since they seem to be auto-spliced *in vitro*, experimental evidence indicates that proteins are required for the efficient splicing of many group II introns *in vivo*, but to date, few group II intron splicing factors have been identified. Some of the protein factors are encoded within certain plastid group II introns, which contain genes for maturase-like proteins involved in their own splicing as well as of other intron-containing plastid genes107,108 whereas others are nuclear encoded. Several nucleus-encoded proteins necessary for the splicing of various subsets of the c.a. 20 chloroplast group II introns in vascular plants have been reported. CRS1 is one of the first to be characterized and is necessary for the splicing of the group II intron in the chloroplast *atpF* gene.109,110 Further investigations have demonstrated the participation of additional proteins in *atpF* intron splicing. One such proteins is the ZmWHY1 that co-immunoprecipitates with CRS1. ZmWHY1 is the maize ortholog of WHY1 which acts as nuclear the transcription factors involved in pathogen-induced transcription in potato and *Arabidopsis* (StWHY1 and AtWHY1 respectively). Genome-wide co-immunoprecipitation assays have shown that ZmWHY1 in chloroplast extract is associated with DNA from throughout the plastid genome and with a subset of plastid RNAs that includes *atpF* transcripts.

Various genetic approaches allowed the identification of additional nucleus-encoded proteins that are required for the splicing of group II introns in maize (*Zea mays* chloroplasts: a CAF1/CRS2 complex, a CAF2/CRS2 complex, PPR4 and RNC1. Each of the afore mentioned nuclear-encoded factors is required for the splicing of distinct, but overlapping, subsets of the 17 group II introns in maize chloroplasts.83,109,111–113 CRS1, CAF1 and CAF2 harbor a CRM domain which is a RNA binding domain111,112,114 and their *Arabidopsis thaliana* orthologs conserve the splicing functions.110 CRS2 is related to peptidyl-tRNA hydrolase enzymes115–116 whereas PPR4 is a member of the pentatricopeptide repeat (PPR) family (see Table 1 and Fig. 2).113,117 RNC1 is a plant-specific protein that has been recovered in both CAF1 and CAF2 co-immunoprecipitates118 and has two ribonuclease III (RNase III) domains. RNC1 is found in complexes containing a subset of group II introns in the chloroplasts that include,
but are not limited to, CAF1- and CAF2-dependent introns. rnc1 mutants exhibit an inefficient splicing of many of the introns which are associated with RNC1 indicating that RNC1 promotes intron splicing in vivo. Despite its two RNase III domains, phylogenetic considerations and biochemical assays indicate that RNC1 lacks endonucleolytic activity. These and other results suggest that RNC1 promotes splicing via its RNA binding activity and that it is recruited to specific plastid introns via protein–protein interactions.

All of the investigations on nuclear-encoded splicing factors mentioned have contributed to the elucidation of the possible mechanisms by which they promote splicing. Nevertheless, the fate of introns after splicing remains an unresolved question. To this end, the analysis of the degradation products of ndhA, atpF, and petB transcripts in several plant species have demonstrated the existence of both incomplete introns and unspliced pre-mRNAs, which presumably correspond with their respective intermediate degradation products.76 Nucleotide sequencing of both 5' and 3' ends of such RNA species has shown that the cleavage affects specific intron domains and occurs within an unpaired bubble flanked by two-stem structures typical of prokaryotic RNase III processing sites. Degradation of both unspliced pre-mRNAs and lariat introns has also been proposed as an additional mechanism that controls the level of mature translatable mRNAs of chloroplast genes.

RNA Editing

In plants, with the exception of liverworts, RNA editing has been found in both mitochondria and chloroplasts.119 Generally, this post-transcriptional modification affects mRNAs but it can also affect structural RNAs. In chloroplasts, most editing events involve conversions of cytidine (C) to uridine (U), but “reverse” conversions of uridine to cytidine have also been noted in several studied hornworts and ferns.120–123 In the chloroplast of seed plants, about 30 different C to U transitions affecting mRNAs have been found.124–128 In bryophytes, the number of RNA editing sites in plastids range from zero in liverworts to almost 1,000 in hornworts.121,122,129 Editing often alters the amino acid identity and affects the amino acids that play a role in proper protein function.122,129,130–132 In some cases, editing creates new translation initiation codons, converting mRNAs into translatable messages or stop codons.122,129,133–136 The existence of these cryptic start codons created by RNA editing, led to the definition of open reading frames (ORFs). Editing sites have also been detected in the anticodon of tRNA (Leu) and within untranslated regions, including introns.122,137–139 However, it seems that the frequency of editing within non-coding regions is very low in comparison with the extent of editing within coding regions. The discovery that editing often leads to the conservation of certain amino acid residues in some proteins in both mitochondria and chloroplasts suggests that editing may act as a mechanism to prevent the deleterious effects of point mutations that have been maintained through evolution. The correspondence of 53 editing sites found in the fern Adiantum capillus-veneris to editing sites in hornworts, and some other land plants, suggests that a major component of RNA editing sites have been conserved for hundreds of millions of years.122 Editing has also been studied in transcript processing intermediates to elucidate possible connections between editing and other post-transcriptional processing events (Fig. 1). The first results indicate that editing is an early RNA processing step, which precedes splicing and cleavage of polycistronic transcripts.46,127,140–142 The complete editing of polycistronic transcripts before any processing event could prevent aberrant forms of the corresponding protein as a result of the translation of unedited transcripts. The editing process involves two consecutive events: site recognition and nucleotide modification. It seems that the modification process in C to U transitions occurs via deamination of the base in plant mitochondria143,144 although the factor(s) mediating this process in plant organelles have not yet been identified. The recognition process for both mitochondrial and chloroplast RNA editing remains unknown to date. In this line, it is very difficult to explain the extraordinary high specificity in the selection of bases to be edited. Several studies indicate that RNA flanking sequences or cis-elements typically located within 15–30 nucleotides are involved in editing site recognition.145–152

Computational analysis of the sequences within −30 to +10 nucleotides of RNA editing sites (neighbor sequences) within the genomic and cDNA sequences of chloroplast genes in the moss Takakia lepidzioides. allowed statistical analyses of chloroplast RNA editing sites to be performed.152 This study allowed the development of a prediction
RNA cleavage of pre-existing RNAs, RNA stabilization, intron splicing, and RNA editing in chloroplasts

Figure 2. (A) Structure of a hypothetical pentatricopeptide repeat protein. (B) Diagram depicting the PPR proteins listed in Table 1 showing the distribution of PPR motifs. PPR motifs are represented as blue boxes whereas other motifs are represented as red ovals (RRM: RNA recognition motif; LAGLIDADG: LAGLIDADG motif). Only the motifs identified by using Pfam v21.10 were represented. All depicted PPR proteins have a transit peptide (no represented) at their N-termini for their targeting to chloroplasts.

Algorithm which predicted c.a. 60% of true editing sites in *T. lepidozioides* transcripts. The success of this prediction algorithm suggests that the obtained patterns are indicative of key sites recognized by trans-factors around editing sites of *T. lepidozioides* chloroplast genes.

One of the latest identified editing cis-elements is the 5’ sequence GCCGUU, which is required for editing of tobacco *psbE* transcripts *in vitro*.

The analysis of *psbE* sequences from many plant species revealed that the GCCGUU sequence is present at a high frequency in plants that carry the
same editing event of psbE transcripts with the exception of Sciadopitys verticillata (Pinophyta). This plant species showed editing at this site despite the presence of nucleotides that differ from the conserved cis-element. Interestingly, chloroplast extracts from a species that has a difference in the motif and lacks the C target are incapable of editing tobacco psbE substrates, indicating that the necessary trans-acting factors were not retained without a C target. Conversely, several heterologous editing events have been reported in different plant species indicating the maintenance of plastid RNA editing activities independently of their target sites.\(^{154}\)

The transformation of the tobacco chloroplast genome has been extensively used to characterize cis-elements involved in editing site recognition in chloroplasts. These experiments, in combination with the introduction of point mutations, are very useful for identifying critical nucleotides that are targets for the editing apparatus.\(^ {146,147,155}\) By using these techniques many cis-acting elements required for the editing process have been discovered. Evidence indicate that they are not specific to an individual editing site allow recognition of a cluster of editing sites even in transcripts of different genes.\(^ {156}\) This finding is supported by the discovery of cis-elements of some 2–5 editing site clusters within different transcripts of various genes. Although lacking consensus sequences, they show some motifs in the 5’ region that are adjacent to editing sites that seem to be recognized by the same trans-elements.\(^ {157}\) Moreover, it has been shown that editing sites that share trans-elements are edited to an equal extent under similar physiological contexts.\(^ {158}\) In fact, in tobacco 34 editing events can be grouped into clusters of 2–5 editing sites according to sequence similarities immediately 5’ of the edited C. Analysis of transgenic tobacco plants with an over-expression of transcripts including each of the clusters showed impaired editing at these sites suggesting that the trans-factors are common to these editing sites therefore act as a limiting factor.\(^ {158–160}\) Moreover, the expression of transgenes bearing the sequences surrounding an specific editing site in sense and/or antisense orientation affected editing efficiency of both transgenic and endogenous transcripts.\(^ {161}\)

Nowadays, scant data exist on the trans-factors responsible for this recognition. Various indirect data indicate that each editing site, or in some cases a small set of sites, must be recognized by specific factors encoded in the plant nuclear genome since it seems that editing is not dependent on the chloroplast translational apparatus.\(^ {162,163}\) Nevertheless, recent studies on the influence of some physiological processes on editing revealed that treatments with antibiotics that inhibit translation in prokaryotes prevented certain C to U transitions.\(^ {164,165}\) Recently, several nuclear-encoded proteins have been identified as possible trans-acting factors essential for RNA editing (see Table 1).\(^ {148,166,168}\) CP31 is a RNA-binding protein required for the editing of two different tobacco sites in vitro.\(^ {148}\) CRR4 and CRR21 are PPR (Pentatricopeptide) proteins essential for editing of a specific site in the chloroplast ndhD mRNA of Arabidopsis thaliana.\(^ {166–168}\) Both CRR4 and CRR21 belong to the E+ subgroup of the PLS subfamily that is characterized by the presence of a conserved C-terminal region (the E/E+ domain). This E/E+ domain is highly conserved and exchangeable between CRR21 and CRR4, although it is not essential for RNA binding. It is possible that the E/E+ domain may have a common function in RNA editing rather than recognizing specific RNA sequences. CLB19 is a PPR protein similar to the editing specificity factors CRR4 and CRR21, but, unlike them, is implicated in the editing of two distinct target sites within the chloroplast, namely rpoA and clpP transcripts.\(^ {169}\) Further studies will be necessary to characterize the entire editing machinery.

The Role of PPR Protein in the Control of Chloroplast Gene Expression

A high number of nuclear mutants with non-photosynthetic phenotypes showing alterations in post-transcriptional steps have been isolated in higher plants\(^ {170,171}\) and in the green alga Chlamydomonas reinhardtii.\(^ {43}\) Generally, these mutants are affected in a single gene cluster or RNA. However, in some cases a single nuclear mutation simultaneously affects posttranscriptional processing of various operons.\(^ {48,109}\) The existence of these mutants suggests, on one hand, the existence of nuclear-encoded factors that control chloroplast RNA processing, and on the other hand, that such processing could play a crucial role in controlling chloroplast gene expression. Recently, several nuclear-encoded proteins that participate in chloroplast transcript processing and stabilization have been characterised (see Table 1).
Many of them seem to be pentatricopeptide repeat (PPR) proteins which are implicated in many crucial functions including organelle biogenesis and plant development.\(^{172}\)

The PPR protein family is characterized by a degenerate motif (PPR motif) consisting of around 35 amino acids that occurs in multiple tandem copies (Fig. 2).\(^{173}\) The structure of these proteins is similar to other proteins with a repeat motif known as the tetratricopeptide repeat (TPR) involved in protein-to-protein interactions.\(^{174}\) Both TRP and PPR proteins share several structural similarities: i) they have degenerate helical tandem repeat motifs, TPR and PPR, respectively; ii) these repeat units form a super helix to bind biomolecules; iii) each repeat consists of anti parallel alpha helix (A and B); and, iv) they have conserved tyrosine residues that facilitate intra helix packing. In spite of these similarities, significant differences between TPR and PPR proteins remain: i) PPR proteins are predominant in plants and generally absent in prokaryotes, whereas TPR proteins are mostly abundant in animals and lower plants and present in prokaryotes; ii) PPR proteins interact with nucleic acids binding to a single target molecule (mainly single stranded RNA) whereas TPR interacts mainly with other proteins and may bind to multiple target proteins forming a complex; ii) PPR proteins have a higher number of repeats (2 to 27) than TPR proteins (3 to 16); iii) they have repeat units of 35 and 34 amino acids for PPR and TPR proteins, respectively; and, iv) side chains of amino acids in the central groove are exclusively hydrophilic in PPR proteins whereas they vary considerably in TPR proteins. Another class of PPR proteins is the proteins commonly known as plant combinatorial and modular proteins or PCMPs. They have complex and variable arrangements of PPR motifs in different combinations.\(^{175,176}\) Apart from the predominant PPR repeat motifs, several other variable motifs have been found at the C-terminus in various PPR proteins. There are three different optional motifs in PPR proteins: E, Ep, and DYW.\(^{177}\) While E and Ep motifs are degenerate, the amino acid sequence of DYW motifs is well conserved, especially Cys and His.\(^{177}\) The occurrence of C-terminal motif is optional in classical PPR and has been implicated in the recruitment of catalytic factors for RNA processing.\(^{176}\)

Most of the known PPR proteins of land plants are nuclear-encoded and targeted to the mitochondria or chloroplasts since they contain a transit peptide at the N-terminus.\(^{175}\) Coordination of nuclear and organellar gene expression with organellar functions is essential to maintain cellular homeostasis, and to respond to changes in environmental conditions. Within this context of multiple regulatory signalling pathways, PPR proteins seem to play a significant role.\(^{178}\) PPR proteins seem to bind to specific chloroplast transcripts modulating their expression with other general factors.

PPR proteins play essential roles in chloroplast gene expression, affecting transcription RNA processing and stabilization, intron splicing editing and translation (see Table 1 and Fig. 2). To date, only a few PPR proteins affecting chloroplast RNA processing and stabilization have been identified, mostly in Arabidopsis (see Table 1). CRR2 is a member of the plant combinatorial and modular protein (PCMP) family consisting of more than 200 genes in Arabidopsis. As mentioned earlier, CRR2 functions in the intergenic processing of chloroplast RNA between rps7 and ndhB, which is possibly essential for ndhB translation.\(^{78}\) CRP1 is a PPR protein with 14 tandem PPR motifs integrated in a multisubunit protein complex which is necessary for the accumulation of petB, petD, and petA chloroplast mRNAs in maize. The lack of the CRP1 protein results in the loss of the cytochrome b6f complex.\(^{48,82}\) The CRP1 protein is also directly associated with petA and psaC mRNAs in vivo, activating their translation.\(^{83}\) HCF152 is a PPR protein with 12 putative PPR motifs which binds certain areas of the petB transcript in Arabidopsis. This protein seems to exist in the chloroplast as a homodimer and is not associated with other proteins to form a high molecular mass complex.\(^{79,179,180}\) P67 is another PPR protein that could be involved in chloroplast RNA processing.\(^{80}\) Both HCF152 and P67 proteins show a significant similarity to the maize protein CRP1. PGR3 is a protein with 27 PPR motifs which appears to be involved not in the processing but in the stabilization and activation of the petL mRNA translation in Arabidopsis.\(^{81}\) ZmPPR5 is a protein with 8 PPR repeats ortholog of the embryo-essential Arabidopsis AtPPR5. This protein specifically binds the trnG-UCC group II intron and stabilizes the trnG-UCC precursor by directly protecting an endonuclease-sensitive site. These findings add to the evidence that chloroplast-localized PPR proteins that are embryo essential in Arabidopsis function in the biogenesis of the plastid translation apparatus. In rice, OSPPR1 is a
protein with 11 PPR repeats involved in the processing of chloroplast transcripts necessary in the early steps of plastid biogenesis.\textsuperscript{181}

PPR proteins are also involved in editing of specific chloroplast RNAs. The CRR4 protein belongs to the PCMP protein family with 11 PPR motifs and seems to be essential for RNA editing of \textit{ndhD} in chloroplasts of \textit{Arabidopsis}. It is speculated that CRR4 recognizes the target RNA and facilitates recruitment of general factors for RNA editing events in the plastid.\textsuperscript{166} It has been hypothesized that CRR4 protein functions as a trans-acting factor specifically interacting with a target sequence near the \textit{ndhD} editing site, affecting the start codon, and recruiting a putative editing enzyme such as cytidine deaminase, probably via the C-terminal E\textsubscript{b} domain.\textsuperscript{119,167} CRR21 is a PPR protein that is involved in the RNA editing of another editing site within \textit{ndhD} transcripts consisting of the conversion of the Ser-128 of NdhD protein to leucine.\textsuperscript{168} \textit{Arabidopsis crr21} mutants are specifically impaired in the RNA editing of this editing site and in the NDH complex suggesting that the Ser128Leu change has important consequences for the function of the NDH complex. Both CRR21 and CRR4 belong to the E+ subgroup of the PLS subfamily that is characterized by the presence of a conserved C-terminal region (the E/E+ domain). This E/E+ domain is highly conserved and exchangeable between CRR21 and CRR4 but it is not essential for RNA binding. Recent investigations suggest that the E/E+ domain has a common function in RNA editing rather than in recognizing specific RNA sequences. CLB19 is a PPR protein similar to the editing specificity factors CRR4 and CRR21, but, unlike them, is implicated in editing of two distinct target sites within the chloroplast, the \textit{rpoA} and \textit{clpP} transcripts. Mutants with a non-functional CLB19 protein show a yellow phenotype with impaired chloroplast development and early seedling lethality. In these mutants, transcript patterns are similar to a defect in the activity of the plastid-encoded RNA polymerase.

PPR proteins are also involved in chloroplast intron splicing. OTP51 is a PPR protein that is required for the splicing of \textit{ycf3} intron 2, and also influences the splicing of several other group-IIa introns. In \textit{Arabidopsis} mutants, the loss of OTP51 has consequences for photosystem-I and photosystem-II assembly, and for the photosynthetic fluorescence characteristics. This protein contains two LAGLIDADG motifs that are found in group-I intron maturases in other organisms. Interestingly, the amino acids reported to be important for maturase activity are conserved whereas amino acids thought to be important for the homing endonuclease activity of other LAGLIDADG proteins are missing in this protein. PPR4 is a chloroplast-targeted protein harbouring both a PPR tract and an RNA recognition motif. The association of PPR4 with the first intron of the plastid \textit{rps12} pre-mRNA and the fact that maize \textit{ppr4} mutants are defective for \textit{rps12} trans-splicing, indicates that this protein is an \textit{rps12} trans-splicing factor.\textsuperscript{181}

Thus far, PPRs have been considered exclusively eukaryotic, and they are greatly expanded in plants. However, the factors that underlie the expansion of this gene family in plants are not yet understood. Further studies are necessary to identify the diverse roles of the PPR family of proteins and to understand how PPR proteins help regulate the organelar gene expression and plant development.

**Concluding Remarks**

The control of chloroplast gene expression includes several processes that are similar to those of both prokaryotic and eukaryotic systems. These processes are: transcription, RNA processing, translation, and post-translational modifications (Fig. 1). Generally, transcription rates and steady-state mRNA levels are not comparable, suggesting that post-transcriptional RNA processing and stabilization are decisive steps in controlling gene expression in plastids. This step principally includes: RNA cleavage of pre-existing RNAs, RNA stabilization-degradation, intron splicing, and RNA editing. Recently, several nuclear-encoded proteins that participate in chloroplast transcript processing and stabilization have been characterised. Many of them seem to be pentatricopeptide repeat (PPR) proteins implicated in many crucial functions including organelle biogenesis and plant development (Table 1). PPR proteins seem to bind to specific chloroplast transcripts modulating their expression with other general factors and appear to be involved in the control of post-transcriptional gene expression in chloroplasts: in transcript processing, stabilization, editing, and translation. Although it is generally assumed that the PPR motifs form the RNA binding domain, the basis for RNA recognition remains unknown. To add clarity, point mutagenesis and crystal structure analysis studies are needed.
Moreover, the identification of interacting enzymes will be crucial to understanding the role of PPR proteins in the editing, splicing, stability and translation of diverse transcripts in chloroplasts. Finally, in spite of the increasing list of PPR proteins, as summarized in Table 1, there is little evidence of their involvement in the regulation of chloroplast metabolism in relation to plant development and in response to environmental changes. To reach this goal, further investigations focused on the behaviour of these newly described proteins in different developmental stages and in response to environmental conditions will be necessary.

**Disclosure**

The authors report no conflicts of interest.

**References**

1. Ris H, Plaut W. Ultrastructure of DNA-containing areas in the chloroplast of Chlamydomonas. J Cell Biol. 1962;13:383–391.
2. Lyttleton JW. Isolation of ribosomes from spinach chloroplasts. Exp Cell Res. 1962;26:312–317.
3. Bedbrook JR, Bogorad L. Endonuclease recognition sites mapped on Zea mays chloroplast DNA. Proc Natl Acad Sci U S A. 1976;73(12):4309–4313.
4. Bedbrook JR, Kolodner R, et al. Zea mays chloroplast ribosomal RNA genes are part of a 22,000 base pair inverted repeat. Cell. 1977;11(4):739–749.
5. Shinozaki K, Ohme M, et al. The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. EMBO J. 1986;5(9):2043–2049.
6. Ohyama K, Fukuzawa H, et al. Chloroplast gene organization deduced from complete sequence of liverwort Marchantia polymorpha chloroplast DNA. Nature. 1986;220(3):281–298.
7. Hiratsuka J, Shimada H, et al. The complete sequence of the rice (Oryza sativa) chloroplast genome. Nature. 1997;386(6624):570–574.
8. Mayfield SP, Yohn CB, et al. Regulation of chloroplast gene expression. Annu Rev Plant Physiol Plant Mol Biol. 1995;46:147–166.
9. Barkan A. Proteins encoded by a complex chloroplast transcription unit are each translated from both monocistronic and polycistronic mRNAs. EMBO J. 1988;7(9):2637–2644.
10. Westhoff P, Hermann RG. Complex RNA maturation in chloroplasts. The psbH operon from spinach. Eur J Biochem. 1988;171(3):551–564.
11. Lierer K, Börner T. Transcription of plastid genes. In Grasser KD (ed.), Regulation of Transcription in Plants, Blackwell Publishing, Oxford. 2007; p. 184–224.
12. Shrama T, Tsunoyama Y, et al. Plastid RNA polymerases, promoters, and transcription regulators in higher plants. Int Rev Cytol. 2005;244:1–68.
13. Little MC, Hallick RB. Chloroplast rpoA, rpoB, and rpoC genes specify at least three components of a chloroplast DNA-dependent RNA polymerase active in tRNA and mRNA transcription. J Biol Chem. 1998;263(28):14302–14307.
14. Allison LA. The role of sigma factors in plastid transcription. Biochimie. 2000;82(6–7):537–548.
15. Lerbs-Mach S. Regulation of tRNA transcription in plastids of higher plants. Biochimie. 2000;82(6–7):525–535.
16. Fujisawa M, Nagashima A, et al. Three new nuclear genes, sigD, sigE and sigF, encoding putative plastid RNA polymerase sigma factors in Arabidopsis thaliana. FEBS Lett. 2000;481(1):47–52.
17. Kanamaru K, Nagashima A, et al. An arabidopsis sigma factor (SIG2)-dependent expression of plastid-encoded tRNAs in chloroplasts. Plant Cell Physiol. 2001;42(10):1034–1043.
18. Nagashima A, Hanaoka M, et al. DNA microarray analysis of plastid gene expression in an arabidopsis mutant deficient in a plastid transcription factor sigma, SIG2. Biosci Biotechnol Biochem. 2004;68(3):694–704.
19. Zghidi W, Merendino L, et al. Nuclear-encoded plastid sigma factor SIG3 transcribes specifically the psbN gene in plastids. Nucleic Acids Res. 2007;35(2):455–464.
20. Favory JJ, Kobayushi M, et al. Specific function of a plastid sigma factor for ndhF gene transcription. Nucleic Acids Res. 2003;31(18):5991–5999.
21. Nagashima A, Hanaoka M, et al. The multiple-stress responsive plastid sigma factor, SIG5, directs activation of the psbD blue light-responsive promoter (BLRP) in Arabidopsis thaliana. Plant Cell Physiol. 2004;45(4):357–368.
22. Ishizaki Y, Tsunoyama Y, et al. A nuclear-encoded sigma factor, Arabidopsis SIG6, recognizes sigma-70 type plastid promoters and regulates early chloroplast development in cotyledons. Plant J. 2005;42(2):133–144.
23. Kühn K, Weihe A, et al. Multiple promoters are a common feature of mitochondrial genes in Arabidopsis. Nucleic Acids Res. 2005;33(1):337–346.
24. Kapoor S, Sugiyama M. Identification of two essential sequence elements in the nonconsensus type II PatPb-290 plastid promoter by using plastid transcription extracts from cultured tobacco BY-2 cells. Plant Cell. 1999;11(9):1799–1810.
25. Sirrman P, Silhavy D, et al. The phage-type PchP-P3 plastid promoter comprises sequences downstream of the transcription initiation site. Nucleic Acids Res. 1998;26(21):4874–4879.
26. Chang CC, Sheen J, et al. Functional analysis of two maize cDNAs encoding T7-like RNA polymerases. Plant Cell. 1999;11(5):911–926.
27. Biligny M, Courtois F, et al. Regulation of plastid rDNA transcription by interaction of CDF2 with two different RNA polymerases. EMBO J. 2000;19(8):1851–1860.
28. Hedcke B, Borner T, et al. Mitochondrial and chloroplast phage-type RNA polymerases in Arabidopsis. Science. 1997;277(5327):809–811.
29. Ikeda TM, Gray MW. Identification and characterization of T3/T7 bacteriophage-like plastid RNA polymerase sequences in wheat. Plant Mol Biol. 1999;40(4):567–578.
30. Hedike B, Borner T, et al. One RNA polymerase serving two genomes. EMBO Rep. 2000;1(5):435–440.
31. Kühn K, Bohne AV, et al. Arabidopsis phage-type RNA polymerases: Accurate in vitro transcription of organellar genes. Plant Cell. 2007;19(3):959–971.
32. Lierer K, Kaden D, et al. Overexpression of phage-type RNA polymerase RpoTp in tobacco demonstrates its role in chloroplast transcription by recognizing a distinct promoter type. Nucleic Acids Res. 2004;32(3):1159–1165.
33. Hricova A, Quesada V, et al. The SCABRA3 nuclear gene encodes the RpoTp RNA polymerase, which is required for chloroplast transcription by interaction of CDF2 with two different RNA polymerases. EMBO J. 2005;24(3):959–971.
34. Swiatecka-Hagenbruch M, Emanuel C, et al. Impaired function of the phage-type RNA polymerase RpoTp in transcription of chloroplast genes is compensated by a second phage-type RNA polymerase. Nucleic Acids Res. 2008;36(3):785–792.
35. Courtois F, Merendino L, et al. Phage-type RNA polymerase RPOTmp transcribes the rRN operon from the PC promoter at early developmental stages in Arabidopsis. Plant Physiol. 2007;145(3):712–721.
36. Azevedo J, Courtois F, et al. Intraplastidial trafficking of a phage-type RNA polymerase is mediated by a thylakoid RING-H2 protein. Proc Natl Acad Sci U S A. 2008;105(26):9123–9128.
37. Swiatecka-Hagenbruch M, Lierke K, et al. High diversity of plastidial promoters in Arabidopsis thaliana. Mol Genet Genomics. 2007;277(6):725–734.
38. Bollenbach TJ, Schuster G, et al. Cooperation of endo- and exoribonucleases in chloroplast mRNA turnover. Prog Nucleic Acid Res Mol Biol. 2004;78:305–337.
39. Jiang H, Heldt H, et al. Short dispersed repeats in the Chlamydomonas chloroplast genome are collocated with sites for mRNA 3′-end formation. Curr Genet. 2004;45(5):311–322.
40. Miyamoto T, Obokata J, et al. Posttranscriptional regulation and chloroplast mRNA turnover. Prog Nucleic Acid Res Mol Biol. 2007;277(6):725–734.
41. Maier UG, Bozarth A, et al. Complex chloroplast RNA metabolism: Just debugging the genetic programme? FEBS J. 2002;269(3):340–337.
42. Nickelsen J. Chloroplast RNA-binding proteins. Curr Genet. 2003;43(6):392–399.
43. Herrin DL, Nickelsen J. Chloroplast RNA processing and stability. Photosynth Res. 2004;82(3):301–314.
44. Reinhothe S, Reinhothe C, et al. A methyl jasmone-induced shift in the length of the 5′-untranslated region impairs translation of the plastid rbcL transcript in barley. EMBO J. 1993;12(4):1505–1512.
45. Hirose T, Sugimura M. Both RNA editing and RNA cleavage are required for translation of tobacco chloroplast ndhD mRNA: A possible regulatory mechanism for the expression of a chloroplast operon consisting of functionally unrelated genes. EMBO J. 1997;16(22):6804–6811.
46. Del Campo EM, Sabater B, et al. Post-transcriptional control of chloroplast gene expression. accumulation of stable psaC mRNA is due to downstream RNA cleavages in the ndhD gene. J Biol Chem. 2002;277(39):36457–36464.
47. Del Campo EM, Sabater B, et al. Characterization of the 5′- and 3′-ends of mRNAs of ndhH, ndhA and ndhG genes of the plastid ndhD operon. Biochimie. 2006;88:347–357.
48. Barkan A, Walker M, et al. 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. 1994;13(13):3170–3181.
49. Sturm NR, Kuras R, et al. The petD gene is transcribed by functionally redundant promoters in Chlamydomonas reinhardtii chloroplasts. Mol Cell Biol. 1994;14(9):6171–6179.
50. Quesada-Vargas T, Ruiz ON, et al. Characterization of heterologous multigene operons in transgenic chloroplasts: Transcription, processing, and translation. Plant Physiol. 2005;138(3):1746–1762.
51. Rochaix JD, Kuchka M, et al. Nuclear and chloroplast mutations affect the synthesis or stability of the chloroplast psbC gene product in chlamydomonas reinhardtii. EMBO J. 1989;8(4):1013–1021.
52. Zerges W, Rochaix JD. The 5′ leader of a chloroplast mRNA mediates the translational requirements for two nucleus-encoded functions in Chlamydomonas reinhardtii. Mol Cell Biol. 1994;14(8):5268–5277.
53. Anthonisen IL, Salvador ML, et al. Specific sequence elements in the 5′ untranslated regions of rbcL and atpB gene mRNAs stabilize transcripts in the chloroplast of Chlamydomonas reinhardtii. RNA. 2001;7(7):1024–1033.
54. Zou Z, Ebel C, et al. The stem-loop region of the tobacco psbA 5′UTR is an important determinant of mRNA stability and translation efficiency. Mol Genet Genomics. 2003;269(3):340–349.
55. Barnes D, Franklin S, et al. Contribution of 5′- and 3′-untranslated regions of plastid mRNAs to the expression of Chlamydomonas reinhardtii chloroplast genes. Mol Genet Genomics. 2005;274(6):625–636.
56. Zicker AA, Kadakia CS, et al. Distinct roles for the 5′ and 3′ untranslated regions in the degradation and accumulation of chloroplast tufA mRNA: Identification of an early intermediate in the in vivo degradation pathway. Plant Mol Biol. 2007;63(5):689–702.
57. Goldschmidt-Clermont M, Rahire M, et al. Redundant cis-acting determinants of 3′ processing and RNA stability in the chloroplast rbcL mRNA of Chlamydomonas. Plant J. 2008;53(3):566–577.
58. Salvador ML, Suay L, et al. Changes in the 5′-untranslated region of the rbcL gene accelerate transcript degradation more than 50-fold in the chloroplast of Chlamydomonas reinhardtii. Curr Genet. 2004;45(3):176–182.
59. Suay L, Salvador ML, et al. Specific roles of 5′ RNA secondary structures in stabilizing transcripts in chloroplasts. Nucleic Acids Res. 2005;33(15):4754–4761.
60. Stern DB, Gruissem W. Control of plastid gene expression. 3′-inverted repeats act as mRNA processing and stabilizing elements, but do not terminate transcription. Cell. 1987;51(6):1145–1157.
61. Rott R, Drager RG, et al. The 3′ untranslated regions of chloroplast genes in Chlamydomonas reinhardtii do not serve as efficient transcriptional terminators. Mol Gen Genet. 1996;252(6):676–683.
62. Kudla J, Hayes R, et al. Polyadenylation accelerates degradation of chloroplast mRNA. EMBO J. 1996;15(24):7137–7146.
63. Schuster G, Lisitisky I, et al. Polyadenylation and degradation of mRNA in the chloroplast. Plant Physiol. 1999;120(4):937–944.
64. Hayes R, Kudla J, et al. Degradation chloroplast mRNA: The role of polyadenylation. Trends Biochem Sci. 1999;24(5):199–202.
65. Yehudai-Reshof S, Portnoy V, et al. Domain analysis of the chloroplast polyuridylate phosphorylase reveals discrete functions in RNA degradation, polyadenylation, and sequence homology with exosome proteins. Plant Cell. 2003;15(9):2003–2019.
66. Nishimura Y, Kikis EA, et al. Antisense transcript and RNA processing alterations suppress instability of polyadenylated mRNA in Chlamydomonas chloroplasts. Plant Cell. 2004;16(11):2849–2869.
67. Sarkar N. Polyadenylation of mRNA in prokaryotes. Annu Rev Biochem. 1997;66:173–197.
68. Blum EA, Carposius J, et al. Polyadenylation promotes degradation of 3′-structured RNA by the Escherichia coli mRNA degradation system in vitro. J Biol Chem. 1999;274(7):4009–4016.
69. Carposius AJ, Vanzo NF, et al. mRNA degradation. A tale of poly(A) and multiprotein machines. Trends Genet. 1999;15(1):24–28.
70. Coburn GA, Mackie GA. Degradation of mRNA in Escherichia coli: An old problem with some new twists. Prog Nucleic Acid Res Mol Biol. 1999;62:55–108.
71. Lupold DS, Casile AG, et al. Polyadenylation occurs at multiple sites in maize mitochondrial cos2 mRNA and is independent of editing status. Plant Cell. 1999;11(8):1565–1578.
72. Gagliardi D, Perrin R, et al. Plant mitochondrial polyadenylated mRNAs are degraded by a 3′- to 5′-exoribonuclease activity, which proceeds unimpeded by stable secondary structures. J Biol Chem. 2001;276(47):43541–43547.
73. Monde RA, Schuster G, et al. Processing and degradation of chloroplast mRNA. Biochimie. 2000;82(6–7):573–582.
74. Lisitisky I, Schuster G. Preferential degradation of polyadenylated and polyuridylated RNAs by the bacterial exoribonuclease polyuridylate phosphorylase. Eur J Biochem. 1999;261(2):468–474.
75. Komine Y, Kwong L, et al. Polyadenylation of three classes of chloroplast mRNA in Chlamydomonas reinhardtii mRNA. RNA. 2000;6(4):598–607.
76. Del Campo EM, Casano LM. Degradation of plastid unspliced transcripts and lariat group II introns. Biochimie. 2000;82(6–7):573–582.
77. Walter M, Kilian J, et al. PNPase activity determines the efficiency of mRNA 3′-end processing, the degradation of tRNA and the extent of polyadenylation in chloroplasts. EMBO J. 2002;21(24):6905–6914.
78. Hashimoto M, Endo T, et al. A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast ndhB in arabidopsis. Plant J. 2003;36(4):541–549.
79. Nakamura T, Meierhoff K, et al. RNA-binding properties of HCF152, an arabidopsis PPR protein involved in the processing of chloroplast RNA. Eur J Biochem. 2003;270(20):4070–4081.
80. Lahmy S, Barneche F, et al. A chloroplastic RNA-binding protein is a new member of the PPR family. FEBS Lett. 2000;480(2–3):255–260.
102. Holloway SP, Deshpande NN, et al. The catalytic group-I introns of
88. Lucas P, Otis C, et al. Papid evolution of the DNA-binding site LAGL-
103. Kim HH, Corina LE, et al. Expression, purifi
cations for group II introns and RNA splicing in chloroplasts.
89. Pombert JF, Lemieux C, et al. The complete chloroplast DNA sequence
96. Li F, Holloway SP, et al. Nuclear genes that promote splicing of group I
93. Cech TR. Conserved sequences and structures of group I introns: build-
92. Cech TR, Damberger SH, et al. Representation of the secondary and ter-
83. Schmitz-Linneweber C, Williams-Carrier R, et al. RNA immunoprecipitation
Gene Regulation and Systems Biology 2009:3
97. Birgisdottir AB, Johansen SD. Reverse splicing of a mobile twin-
85. Hattori M, Miyake H, et al. A pentatricopeptide repeat protein is
86. Raynaud C, Loiselay C, et al. Evidence for regulatory function of
84. Raynaud C, Loiselay C, et al. Evidence for regulatory function of
81. Fisk DG, Walker MB, et al. Molecular cloning of the maize gene crp1
82. Schmitz-Linneweber C, Williams-Carrier R, et al. 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. 2005;17(10):2791–2804.
87. Beick S, Schmitz-Linneweber C, et al. The pentatricopeptide repeat protein SSR5 stabilizes a specific tRNA precursor in maize chloroplasts. Mol Cell Biol. 2008;28(17):5317–5347.
85. Hattori M, Miyake H, et al. A pentatricopeptide repeat protein is required for RNA processing of cplP pre-mRNA in moss chloroplasts. J Biol Chem. 2007;282(14):10773–10782.
86. Raynaud C, Loiselay C, et al. Evidence for regulatory function of
nucleus-encoded factors on mRNA splicing and translation in the chloroplast. Proc Natl Acad Sci U S A. 2007;104(21):9093–9098.
87. Turmel M, Mercier JP, et al. The site-specific DNA endonuclease encoded by a group I intron in the Chlamydomonas pallidostigmae chloroplast small subunit rRNA gene introduces a single-strand break at low concentrations of Mg2+. Nucleic Acids Res. 1995;23(13):2519–2525.
88. Lucas P, Otis C, et al. Papid evolution of the DNA-binding site LAGL-
101. Asakura Y, Barkan A. Arabidopsis orthologs of maize chloroplast
102. Schmitz-Linneweber C, Williams-Carrier R, et al. CRS1 is a novel group II intron
107. Liere K, Link G. RNA-binding activity of the
105. Chu VT, Liu Q, et al. More than one way to splice an RNA: Branching without a bulge and splicing without branching in group II introns. RNA. 1998;4(10):1186–1202.
104. Vogel J, Borner T. Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. EMBO J. 2002;21(14):3794–3803.
103. Li F, Holloway SP, et al. Nuclear genes that promote splicing of group I
96. Li F, Holloway SP, et al. Nuclear genes that promote splicing of group I
93. Cech TR. Conserved sequences and structures of group I introns: build-
92. Cech TR, Damberger SH, et al. Representation of the secondary and ter-
83. Schmitz-Linneweber C, Williams-Carrier R, et al. RNA immunoprecipitation
Gene Regulation and Systems Biology 2009:3
97. Birgisdottir AB, Johansen SD. Reverse splicing of a mobile twin-
85. Hattori M, Miyake H, et al. A pentatricopeptide repeat protein is
86. Raynaud C, Loiselay C, et al. Evidence for regulatory function of
nucleus-encoded factors on mRNA splicing and translation in the chloroplast. Proc Natl Acad Sci U S A. 2007;104(21):9093–9098.
87. Turmel M, Mercier JP, et al. The site-specific DNA endonuclease encoded by a group I intron in the Chlamydomonas pallidostigmae chloroplast small subunit rRNA gene introduces a single-strand break at low concentrations of Mg2+. Nucleic Acids Res. 1995;23(13):2519–2525.
88. Lucas P, Otis C, et al. Papid evolution of the DNA-binding site LAGL-
101. Asakura Y, Barkan A. Arabidopsis orthologs of maize chloroplast
102. Schmitz-Linneweber C, Williams-Carrier R, et al. CRS1 is a novel group II intron
107. Liere K, Link G. RNA-binding activity of the
105. Chu VT, Liu Q, et al. More than one way to splice an RNA: Branching without a bulge and splicing without branching in group II introns. RNA. 1998;4(10):1186–1202.
128. Tillich M, Funk HT, et al. Editing of plastid RNA in Arabidopsis thaliana ecotypes. Plant J. 2005;43(5):708–715.

129. Sugita M, Miyata Y, et al. Extensive RNA editing in transcripts from the PsbH operon and RpoA gene of plastids from the enigmatic moss Takakia lepidodzioides. Biosci Biotechnol Biochem. 2006;70(9):2268–2274.

130. Bock R, Kossel H, et al. Introduction of a heterologous editing site into the tobacco plastid genome: The lack of RNA editing leads to a mutant phenotype. EMBO J. 1994;13(19):4623–4628.

131. Zito F, Kuras R, et al. Mutations of cytochrome b6 in Chlamydomonas reinhardtii disclose the functional significance for a proline to leucine conversion by petB editing in maize and tobacco. Plant Mol Biol. 1997;33(1):79–86.

132. Sasaki Y, Kozaki A, et al. Chloroplast RNA editing required for primary transcripts of ndhA mRNA by creation of an initiation codon. Nature. 1991;353(6340):178–180.

133. Hoch B, Maier RM, et al. Editing of a chloroplast mRNA by creation of an initiation codon. J Exp Bot. 1997;276(6):3937–3940.

134. Lopez-Serrano M, Del Campo EM, et al. Introduction of a heterologous editing site into the tobacco plastid genome: The lack of RNA editing leads to a mutant phenotype. EMBO J. 1994;13(19):4623–4628.

135. Zito F, Kuras R, et al. Mutations of cytochrome b6 in Chlamydomonas reinhardtii disclose the functional significance for a proline to leucine conversion by petB editing in maize and tobacco. Plant Mol Biol. 1997;33(1):79–86.

136. Sasaki Y, Kozaki A, et al. Chloroplast RNA editing required for primary transcripts of ndhA mRNA by creation of an initiation codon. Nature. 1991;353(6340):178–180.

137. Hoch B, Maier RM, et al. Editing of a chloroplast mRNA by creation of an initiation codon. J Exp Bot. 1997;276(6):3937–3940.

138. Kudla J, Igloi GL, et al. RNA editing in tobacco chloroplasts leads to functional acetyl-CoA carboxylase in plants. J Biol Chem. 2001;276(6):3937–3940.

139. Drescher A, Hupfer H, et al. C-to-U conversion in the intercistronic trnK-matK intron. Plant Mol Biol. 2002;41(1):48–52.

140. Freyer R, Obokata J, et al. A site-specific factor interacts directly with its cognate RNA editing site in chloroplast transcripts. Proc Natl Acad Sci U S A. 2004;101(1):48–52.

141. Miyamoto T, Obokata J, et al. A site-specific factor interacts directly with its cognate RNA editing site in chloroplast transcripts. Proc Natl Acad Sci U S A. 2004;101(1):48–52.

142. Yura K, Miyata Y, et al. Characteristics and prediction of RNA editing sites in transcripts of the moss Takakia lepidodzioides chloroplast. DNA Res. 2008.

143. Hayes ML, Hanson MR. High conservation of a 5’ endonuclease required for RNA editing of a C target in chloroplast psbE transcripts. J Mol Evol. 2008;67(3):233–245.

144. Tillich M, Poltnigg P, et al. Maintenance of plastid RNA editing activities independently of their target sites. EMBO Rep. 2006;7(3):308–313.

145. Herrmann M, Bock R. Transfer of plastid RNA-editing activity to novel sites suggests a critical role for spacing in editing-site recognition. Proc Natl Acad Sci U S A. 1999;96(9):4856–4861.

146. Peeters NM, Hanson MR. Transcript abundance supercedes editing efficiency as a factor in developmental variation of chloroplast gene expression. RNA. 2002;8(4):497–511.

147. Chateigner-Boutin AL, Hanson MR. Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared cis elements. Mol Cell Biol. 2002;22(8):4484–4486.

148. Chateigner-Boutin AL, Hanson MR. Developmental co-variation of RNA editing extent of plastid editing sites exhibiting similar cis-elements. Nucleic Acids Res. 2003;31(10):2586–2594.

149. Hayes ML, Reed ML, et al. Sequence elements critical for efficient RNA editing at a tobacco chloroplast transcript in vivo and in vitro. Nucleic Acids Res. 2006;34(13):3742–3754.

150. Keller WP, Hayes ML, et al. Cross-competition in editing of chloroplast RNA transcripts in vitro implicates sharing of trans-factors between different C targets. J Biol Chem. 2008;283(12):7314–7319.

151. Hegeman CE, Halter CP, et al. Expression of complementary RNA from chloroplast transgenes affects editing efficiency of transgene and endogenous chloroplast transcripts. Nucleic Acids Res. 2005;33(5):1454–1464.

152. Zeltz P, Hess WR, et al. Editing of the chloroplast rpoB transcript is independent of chloroplast translation and shows different patterns in barley and maize. EMBO J. 1993;12(11):4291–4296.

153. Hess WR, Hoch B, et al. Inefficient rpl2 splicing in barley mutants with ribosome-deficient plastids. Plant Cell. 1994;6(10):1455–1465.

154. Karcher D, Bock R. Site-selective inhibition of plastid RNA editing by heat shock and antibiotics: A role for plastid translation in RNA editing. Nucleic Acids Res. 1998;26(5):1185–1190.

155. Karcher D, Bock R. Temperature sensitivity of RNA editing by heat shock and antibiotics: A role for plastid translation in RNA editing. Nucleic Acids Res. 1998;26(5):1185–1190.

156. Chateigner-Boutin AL, Hanson MR. Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared cis elements. Mol Cell Biol. 2002;22(8):4484–4486.

157. Chateigner-Boutin AL, Hanson MR. Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared cis elements. Mol Cell Biol. 2002;22(8):4484–4486.

158. Chateigner-Boutin AL, Hanson MR. Developmental co-variation of RNA editing extent of plastid editing sites exhibiting similar cis-elements. Nucleic Acids Res. 2003;31(10):2586–2594.

159. Hayes ML, Reed ML, et al. Sequence elements critical for efficient RNA editing at a tobacco chloroplast transcript in vivo and in vitro. Nucleic Acids Res. 2006;34(13):3742–3754.

160. Chateigner-Boutin AL, Hanson MR. Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared cis elements. Mol Cell Biol. 2002;22(8):4484–4486.

161. Chateigner-Boutin AL, Hanson MR. Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared cis elements. Mol Cell Biol. 2002;22(8):4484–4486.

162. Zeltz P, Hess WR, et al. Editing of the chloroplast rpoB transcript is independent of chloroplast translation and shows different patterns in barley and maize. EMBO J. 1993;12(11):4291–4296.

163. Hess WR, Hoch B, et al. Inefficient rpl2 splicing in barley mutants with ribosome-deficient plastids. Plant Cell. 1994;6(10):1455–1465.

164. Karcher D, Bock R. Site-selective inhibition of plastid RNA editing by heat shock and antibiotics: A role for plastid translation in RNA editing. Nucleic Acids Res. 1998;26(5):1185–1190.

165. Karcher D, Bock R. Temperature sensitivity of RNA editing by heat shock and antibiotics: A role for plastid translation in RNA editing. Nucleic Acids Res. 1998;26(5):1185–1190.

166. Chateigner-Boutin AL, Hanson MR. Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared cis elements. Mol Cell Biol. 2002;22(8):4484–4486.

167. Chateigner-Boutin AL, Hanson MR. Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared cis elements. Mol Cell Biol. 2002;22(8):4484–4486.

168. Okuda K, Miyata Y, et al. Characteristics and prediction of RNA editing sites in transcripts of the moss Takakia lepidodzioides chloroplast. DNA Res. 2008.

169. Kudla J, Igloi GL, et al. RNA editing in tobacco chloroplasts leads to functional acetyl-CoA carboxylase in plants. J Biol Chem. 2001;276(6):3937–3940.

170. Drescher A, Hupfer H, et al. C-to-U conversion in the intercistronic trnK-matK intron. Plant Mol Biol. 2002;41(1):48–52.

171. Verheul EL, Hedden P, et al. The editing activity of the maize small RNA molecule is not affected by changes in the levels of the corresponding mRNA. Nucleic Acids Res. 1994;22(3):661–668.

172. Schmitz-Linneweber C, Small I. Pentatricopeptide repeat proteins: A socket set for organelle gene expression. Gene Regulation and Systems Biology 2009:3.
RNA cleavage of pre-existing RNAs, RNA stabilization, intron splicing, and RNA editing in chloroplasts

174. Das AK, Cohen PW, et al. The structure of the tetratricopeptide repeats of protein phosphatase 5: Implications for TPR-mediated protein-protein interactions. *EMBO J.* 1998;17(5):1192–1199.

175. Lurin C, Andres C, et al. Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell.* 2004;16(8):2089–2103.

176. Rivals E, Bruyere C, et al. Formation of the *Arabidopsis* pentatricopeptide repeat family. *Plant Physiol.* 2006;141(3):825–839.

177. Aubourg S, Boudet N, et al. In *Arabidopsis thaliana*, 1% of the genome codes for a novel protein family unique to plants. *Plant Mol Biol.* 2000;42(4):603–613.

178. Koussevitzky S, Nott A, et al. Signals from chloroplasts converge to regulate nuclear gene expression. *Science.* 2007;316(5825):715–719.

179. Meierhoff K, Felder S, et al. HCF152, an arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast *psbB*-*psbT*-petH-petB-petD RNAs. *Plant Cell.* 2003;15(6):1480–1495.

180. Nakamura, T, Schuster G, et al. Chloroplast RNA-binding and pentatricopeptide repeat proteins. *Biochem Soc Trans.* 2004;32 (Pt 4):571–574.

181. Gothandam KM, Kim ES, et al. OsPPR1, a pentatricopeptide repeat protein of rice is essential for the chloroplast biogenesis. *Plant Mol Biol.* 2005;58(3):421–433.

182. De Longevialle AF, Hendrickson L, et al. 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.* 2008;56(1):157–168.

183. Finn RD, Mistry J, et al. Pfam: Clans, web tools and services. *Nucleic Acids Res.* 2006;34(Database issue):D247–51.

184. Chi W, Ma J, et al. The pentatricopeptide repeat protein DELAYED GREENING1 is involved in the regulation of early chloroplast development and chloroplast gene expression in *Arabidopsis*. *Plant Physiol.* 2008;147(2):573–584.

185. Williams PM, Barkan A. A chloroplast-localized PPR protein required for plastid ribosome accumulation. *Plant J.* 2003;36(5):675–686.