



# Pas de Trois: An Overview of Penta-, Tetra-, and Octo-Tricopeptide Repeat Proteins From *Chlamydomonas reinhardtii* and Their Role in Chloroplast Gene Expression

Karla S. Macedo-Osorio<sup>1,2,3\*</sup>, Agustino Martínez-Antonio<sup>2</sup> and Jesús A. Badillo-Corona<sup>1\*</sup>

## OPEN ACCESS

### Edited by:

Mamoru Sugita,  
Nagoya University, Japan

### Reviewed by:

Yves Choquet,  
Centre National de la Recherche  
Scientifique (CNRS), France  
Bruce Cahoon,  
University of Virginia's, United States

### \*Correspondence:

Karla S. Macedo-Osorio  
kmacedo@correo.xoc.uam.mx  
Jesús A. Badillo-Corona  
jbadillo@icpn.mx

### Specialty section:

This article was submitted to  
Plant Physiology,  
a section of the journal  
Frontiers in Plant Science

**Received:** 13 September 2021

**Accepted:** 26 October 2021

**Published:** 17 November 2021

### Citation:

Macedo-Osorio KS,  
Martínez-Antonio A and  
Badillo-Corona JA (2021) Pas de  
Trois: An Overview of Penta-, Tetra-,  
and Octo-Tricopeptide Repeat  
Proteins From *Chlamydomonas*  
*reinhardtii* and Their Role in  
Chloroplast Gene Expression.  
*Front. Plant Sci.* 12:775366.  
doi: 10.3389/fpls.2021.775366

<sup>1</sup>Instituto Politécnico Nacional, Unidad Profesional Interdisciplinaria de Biotecnología, México City, México, <sup>2</sup>Biological Engineering Laboratory, Genetic Engineering Department, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional-Unidad Irapuato, Irapuato, México, <sup>3</sup>División de Ciencias Biológicas y de la Salud, Universidad Autónoma Metropolitana-Xochimilco, México City, México

Penta-, Tetra-, and Octo-tricopeptide repeat (PPR, TPR, and OPR) proteins are nucleus-encoded proteins composed of tandem repeats of 35, 34, and 38–40 amino acids, respectively. They form helix-turn-helix structures that interact with mRNA or other proteins and participate in RNA stabilization, processing, maturation, and act as translation enhancers of chloroplast and mitochondrial mRNAs. These helical repeat proteins are unevenly present in plants and algae. While PPR proteins are more abundant in plants than in algae, OPR proteins are more abundant in algae. In *Arabidopsis*, maize, and rice there have been 450, 661, and 477 PPR proteins identified, respectively, which contrasts with only 14 PPR proteins identified in *Chlamydomonas reinhardtii*. Likewise, more than 120 OPR proteins members have been predicted from the nuclear genome of *C. reinhardtii* and only one has been identified in *Arabidopsis thaliana*. Due to their abundance in land plants, PPR proteins have been largely characterized making it possible to elucidate their RNA-binding code. This has even allowed researchers to generate engineered PPR proteins with defined affinity to a particular target, which has served as the basis to develop tools for gene expression in biotechnological applications. However, fine elucidation of the helical repeat proteins code in *Chlamydomonas* is a pending task. In this review, we summarize the current knowledge on the role PPR, TPR, and OPR proteins play in chloroplast gene expression in the green algae *C. reinhardtii*, pointing to relevant similarities and differences with their counterparts in plants. We also recapitulate on how these proteins have been engineered and shown to serve as mRNA regulatory factors for biotechnological applications in plants and how this could be used as a starting point for applications in algae.

**Keywords:** helical repeat proteins, PPR, OPR, TPR, HAT-TPR, chloroplast mRNA-processing, anterograde control

## INTRODUCTION

Evolutionary biologists estimate that it was 1.5 billion years ago when plastids started to develop as the result of the engulfment of a cyanobacterium by a heterotrophic eukaryotic cell (Yoon et al., 2004). Later, it is widely accepted, this endosymbiosis process gave rise to chloroplasts in plants and algae (reviewed by Stadnichuk and Kusnetsov, 2021). The evolution of this endosymbiosis resulted in a drastic genome reduction of the originally engulfed cyanobacterium; 0.15 Mbp presently in chloroplasts vs. 3 Mbp in free-living cyanobacteria (Martin et al., 2002; Yagi and Shiina, 2014). While thousands of genes disappeared, most likely as the result of being redundant with the ones already present in the host, others were transferred and integrated into the host nuclear genome (Bock, 2017; De Marchis et al., 2019). As a result of this genome reduction, 2–5% of the genes found in free-living cyanobacteria remain in modern chloroplasts (Eckardt, 2006; Stegemann and Bock, 2006; Yagi and Shiina, 2014). This gene interchange is also made evident by the fact that 3.5–6% of protein-coding genes in the nuclear genome of *Chlamydomonas* and other Plantae have a cyanobacterial origin (Moustafa and Bhattacharya, 2008). This genetic rearrangement ultimately led to the chloroplasts as we know them today: semi-autonomous organelles that contain their own genome and the expression machinery for many genes involved in protein synthesis (including tRNAs and rRNAs) and photosynthesis (including proteins involved in light capture, CO<sub>2</sub> fixation, and ATP synthesis).

It has been estimated that around 93–99% of proteins functioning inside chloroplasts are encoded in the nucleus (Woodson and Chory, 2008), translated in the cytosol, and then transported into the chloroplast (Nakai, 2018). In photosynthetic organisms, the subunits of all photosynthetic complexes, the translation apparatus, some proteins of the chloroplast envelope, and the key enzyme in carbon fixation, RuBisCo, are encoded by genes located in both the nuclear and chloroplast genomes (Börner, 2017). For example, in *Chlamydomonas reinhardtii*, Photosystem I (PSI) is composed of 14 proteins, four of which are encoded in the chloroplast (PsaA, PsaB, PsaC, and PsaJ subunits) and the other 10 are nucleus-encoded (Harris, 2009). The assembly of the protein complexes of the photosystems I and II, the ATP synthase, cytochrome *b<sub>6</sub>f* complex, and RuBisCo is regulated by the control by epistasis of synthesis (CES) process, dependent on the regulation of translation. The synthesis of one subunit becomes dependent on the synthesis of another: when certain core subunits are absent, synthesis of other chloroplast-encoded subunits from the same protein complex is reduced (Choquet et al., 1998; Choquet and Wollman, 2009). Similarly, the cofactors that are accumulated in excess of their binding proteins can regulate their own synthesis and that of the proteins (Pogson et al., 2008). The expression of the divided encoded genetic information requires coordinated communication between chloroplast and nucleus to maintain cell homeostasis.

Cellular processes such as chloroplast biogenesis (reviewed by Jarvis and López-Juez, 2013; Yoo et al., 2020), plastid differentiation (reviewed by Sadali et al., 2019), photosynthesis

including the assembly of photosynthetic apparatus and expression of photosynthetic genes (Lefebvre-Legendre et al., 2015, 2016; Douchi et al., 2016; Krupinska et al., 2019), pigment production (Terry and Bampton, 2019; Ganusova et al., 2020), and other metabolic processes such as starch and lipid biosynthesis, cell wall synthesis and/or modification, sugar transport, stress responses, and reactive oxygen species (ROS) responses are highly dependent on communication between the nucleus and the chloroplast (reviewed by Woodson and Chory, 2008; Börner, 2017; Piñas-Fernández and Strand, 2018). *Trans*-acting regulatory factors are generated in the nucleus and are essential for transcriptional, post-transcriptional, and translational regulation of chloroplast gene expression (Woodson and Chory, 2008; Piñas-Fernández and Strand, 2018). The regulation from the nucleus to the chloroplast is known as anterograde signaling. In contrast, retrograde signaling is when the signals originate from the organelles (chloroplasts and mitochondria) to influence the expression of nuclear genes, some of which, in turn, can alter the anterograde control (Hernández-Verdeja and Strand, 2018). Synchronization between anterograde and retrograde signaling allows the cell to survive against changes in the environment (Rochaix and Ramundo, 2017). These changes include temperature, light intensity, drought, and nutrient deficiency (Sun and Guo, 2016; Rochaix and Ramundo, 2017). As part of the anterograde signaling, in higher plants and microalgae, many nucleus-encoded proteins are RNA-binding factors for post-transcriptional regulation of gene expression, such as RNA stabilization, splicing, intercistronic processing of polycistronic RNAs, editing of chloroplast transcripts, and regulation of the translation (Lyska et al., 2013; Manavski et al., 2018; Zoschke and Bock, 2018; De Marchis et al., 2019).

RNA-binding proteins involved in post-transcriptional gene expression include the helical repeat protein family. This superfamily comprises the Tetra-tricopeptide repeat (TPR), Half-a-tetratricopeptide repeat (HAT), Penta-tricopeptide repeat (PPR), and Octo-tricopeptide repeat (OPR) proteins. TPR proteins are present in prokaryotic and eukaryotic organisms, where their principal function is the regulation of many cellular processes, including translation, by mediating protein–protein interactions (Bohne et al., 2016). PPR proteins are found in all eukaryotic organisms. In land plants, hundreds of PPR proteins have been identified, many of them even functionally characterized, determining that they are involved in RNA metabolism and translation regulation of organellar genes (Shikanai and Fujii, 2013; Barkan and Small, 2014; Rovira and Smith, 2019). OPR proteins are also found in prokaryotes and eukaryotes and similar functions have been attributed to them; however, although the number of OPR proteins is limited in land plants, in algae it is quite large (Bohne et al., 2009; Rahire et al., 2012).

Due to its ease of cultivation and its relative simplicity, the microalga *C. reinhardtii* is one of the best model organisms for the study of various cellular processes, including photosynthesis, flagellar function and pigment, biofuel, and recombinant proteins production (Salomé and Merchant, 2019). Furthermore, the study of the mechanisms regulating chloroplast

gene expression in *C. reinhardtii* is crucial for the development of tools for the improvement of the production of biomass and bioactive compounds. In this paper, we first review the role of PPR, TPR (including HAT), and OPR proteins in chloroplast gene expression in *C. reinhardtii* and then consider the emerging tools being developed for biotechnological applications.

## Chloroplast Gene Expression

In the last decade, the study of chloroplast gene function and regulation gained great attention as this organelle showed to be a low-cost platform for recombinant protein production (Economou et al., 2014; Shamriz and Ofoghi, 2017; Dyo and Purton, 2018). In the chloroplast of the unicellular microalga *C. reinhardtii* more than 100 recombinant proteins, including antibodies (Mayfield et al., 2003; Tran et al., 2009), immunotoxins (Tran et al., 2013a,b), antigens (Dreesen et al., 2010; Michelet et al., 2011; Jones et al., 2013), toxins (Kang et al., 2017), and growth factors (Rasala et al., 2010; Wannathong et al., 2016) have been produced. This has been possible thanks to advances in the study of chloroplast functioning and the development of molecular tools. The plastid genome was sequenced almost 2 decades ago and consists of 99 genes that include tRNAs, rRNAs, plastid-encoded RNA polymerase subunits, ribosomal and photosynthetic proteins (Maul et al., 2002; Gallaher et al., 2018). The synthesis of chloroplast proteins is a complex process highly regulated at translational level but also depending on changes in transcription rate and mRNA degradation (Herrin, 2009). Although, chloroplast mRNAs do not possess a 5' trimethylguanosine cap like cytoplasmic mRNAs, they have a relatively long half-life in the range of minutes to hours (Herrin, 2009). In many cases, the stability of these transcripts depends on nucleus-encoded proteins that protect and prevent mRNA degradation by nucleases.

For many years, chloroplast gene arrangement was thought to result in the transcription of mono, bi, or tri-cistronic units. However, it was recently reported that 70% of genes from *C. reinhardtii* chloroplast are co-transcribed forming long polycistronic units, although it is still clear that in some cases, some of these genes are also regulated by single promoters (Cavaiuolo et al., 2017; Gallaher et al., 2018). Polycistronic mRNAs are integrated by two or more cistrons transcribed from the same promoter, however, unlike prokaryotes, cistron units in chloroplast are separated by intergenic regions rich in A-U (Pfalz et al., 2009). Before translation, polycistronic transcripts must be processed; intergenic regions serve as the target of endonucleases to adjust polycistronic mRNAs into translatable monocistrons. Once cistrons are cleaved, the resulting transcripts are exposed to exonucleases, which degrade them further from their ends until degradation is blocked by an internal robust mRNA-structure or by an mRNA-binding protein (RBP) bound to the transcript (Pfalz et al., 2009; Prikryl et al., 2011; Stoppel and Meurer, 2012). mRNA-binding proteins anchor to either the 5'-untranslated region (UTR) or 3'-UTR of the transcript protecting it from degradation during the trimming process by exonucleases (Figure 1A). In photosynthetic genes of plants and microalgae, these protective functions have been

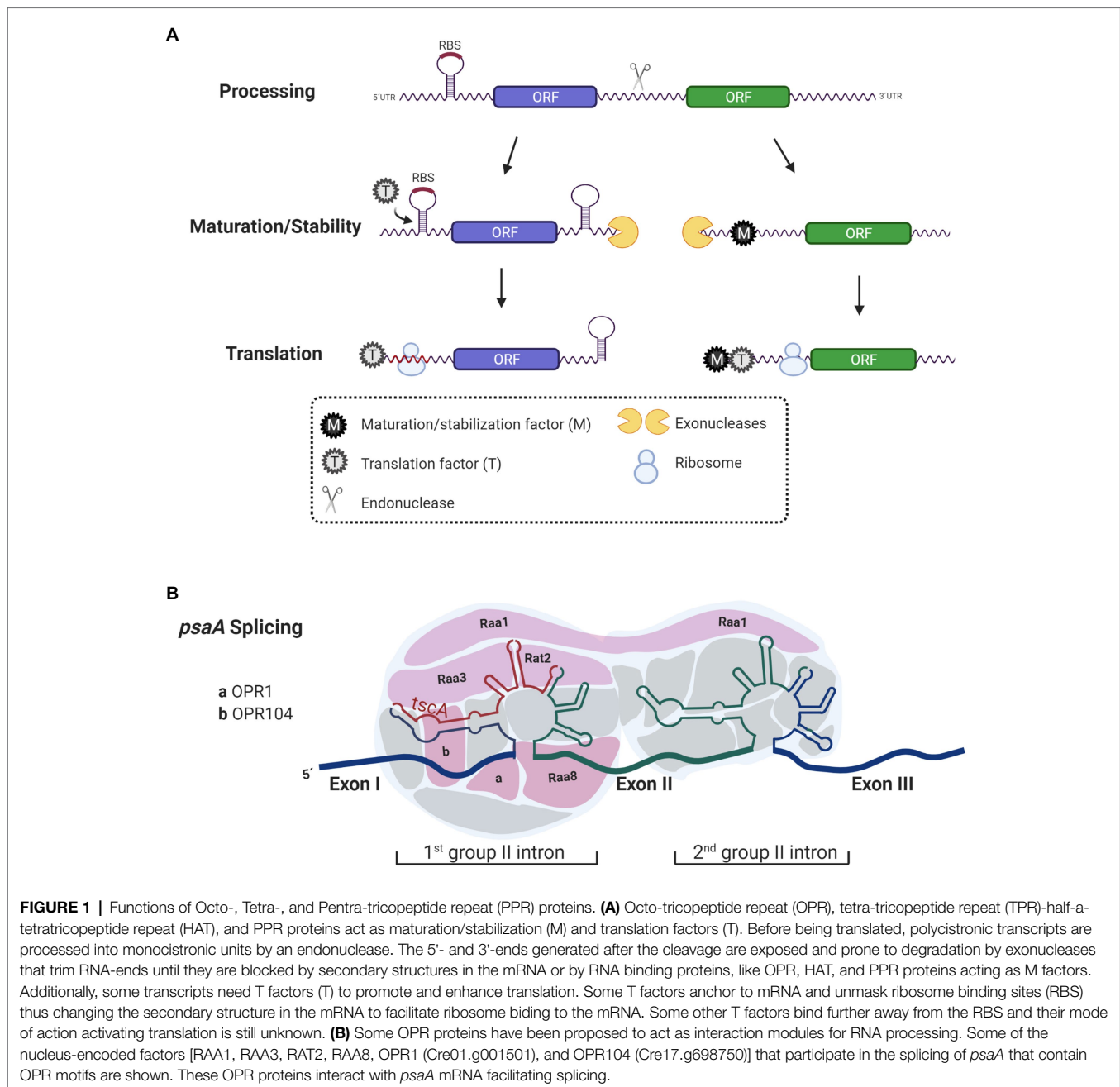
experimentally confirmed to be attributed to members of the helical repeat protein family: namely, PPR, OPR, or TPR proteins (Boudreau et al., 2000; Loiselay et al., 2008; Pfalz et al., 2009; Johnson et al., 2010; Wang et al., 2015; Viola et al., 2019). In plants and *Chlamydomonas*, PPR and OPR proteins act as post-transcriptional factors that participate in RNA stabilization and translation of chloroplast transcripts (Johnson et al., 2010; Prikryl et al., 2011; Wang et al., 2015; Viola et al., 2019). Although some TPR proteins also participate in these processes, their main function relies on the protein-protein interactions required for the assembly of photosynthesis complexes (Bohne et al., 2016). Additionally, in mitochondria and chloroplast of plants, PPR proteins also have editing functions (Andrés-Colás et al., 2017; Guillaumot et al., 2017).

P/O/TPR proteins are nucleus-encoded proteins characterized by the presence of tandem arrays of degenerated 35, 38–40, and 34 amino acid repeats, respectively (Rovira and Smith, 2019). On the one hand, PPR proteins are more abundant in plants than in microalgae. In *C. reinhardtii*, 14 PPR proteins have been identified (Tourasse et al., 2013). By contrast, in the genomes of *A. thaliana*, maize, and rice approximately 450, 661, and 477 members of the PPR proteins family have been identified, respectively.<sup>1</sup> Most of these PPR proteins have been predicted to be functional in plastids and mitochondria (Gutmann et al., 2019; Rovira and Smith, 2019). On the other hand, OPR proteins are disproportionately more abundant in microalgae where around >120 OPR have been estimated to be encoded in the nuclear genome of *C. reinhardtii* (Cavaiuolo et al., 2017; *Chlamydomonas* genome; [https://phytozome-next.jgi.doe.gov/info/Creinhardtii\\_v5\\_6](https://phytozome-next.jgi.doe.gov/info/Creinhardtii_v5_6)) vs. only one identified in *A. thaliana* (Kleinknecht et al., 2014). Although both PPR and OPR proteins seem to have similar functions associated with stability, maturation, and translation of chloroplast and mitochondria transcripts, the difference in their distribution in algae and plants may indicate different evolutionary routes in the regulation of gene expression in the organelles of these lineages (Fujii and Small, 2011; Rovira and Smith, 2019).

According to the nomenclature established for the Organelle *Trans*-Acting Factors (OTAFs) in *C. reinhardtii*, P/O/TPR proteins are named in their first letter according to their function: **M** if they are maturation or stability factors required for the stable accumulation of their target mRNA; or **T**, if they are factors required for translation of specific transcripts, respectively (Cavaiuolo et al., 2017). They are then assigned a letter A, B, C, D, or R if their targets are mRNAs for PSI, PSII, cytochrome *b<sub>6</sub>f*, ATP synthase, and RuBisCO subunits, respectively. A third letter corresponding to the last letter of the name of their target gene is required to be incorporated (Wang et al., 2015). For instance, MAC1 is a Maturation factor for the *psaC* mRNA that in turn is translated into the PsaC subunit from PSI. Similarly, TAB1 is required as a Translation factor for the production of PsaB protein of PSI.

As it was already observed, the regulation of chloroplast gene expression relies on several layers of control. One of these layers is the result of the coordinated action of P/O/

<sup>1</sup><https://ppr.plantenergy.uwa.edu.au/ppr/>



TPR proteins as we have glimpsed. In the following sections, we explore specific roles and mechanisms of action of P/O/TPR proteins in *C. reinhardtii*.

## PENTA-TRICOPEPTIDE REPEAT PROTEINS

Penta-tricopeptide Repeat proteins are distributed in all eucaryotes, and due to their abundance and relevant functions, these proteins have been extensively studied in plants (Gutmann et al., 2019). These were initially identified in nuclear mutants with non-photosynthetic phenotypes and altered

post-transcriptional processes (Fisk et al., 1999). Shortly after, the sequence of the *A. thaliana* genome revealed a new protein family that comprised hundreds of genes that until that moment had an unknown function (Aubourg et al., 2000). Later, Small and Peeters (2000) named them, PPR Proteins following their structural similarity with the previously described TPR motifs. Interestingly and surprisingly, Lurin et al. (2004) predicted that two-thirds of these proteins are targeted to mitochondria or chloroplasts in Arabidopsis.

These nucleus-encoded proteins are composed of degenerated motifs of 35 amino acids repeated in tandem (Small and Peeters, 2000). The PPR motif forms two antiparallel  $\alpha$ -helices, which interact to produce a helix-turn-helix motif (Figures 2A,B).





**TABLE 1** | Penta-tricopeptide repeat proteins identified and studied in *C. reinhardtii*.

Name	Number of PPR motifs	Target	Function	Reference
TCB1 (PPR1)	11*	<i>petB</i>	Translation and stabilization of <i>petB</i> mRNA	Tourasse et al., 2013 <sup>*</sup> ; Cavaiuolo et al., 2017
MRL1 (PPR2)	12/9*	<i>rbcL</i>	Stabilization of <i>rbcL</i> mRNA	Johnson et al., 2010; Johnson, 2011; Tourasse et al., 2013 <sup>*</sup>
PPR3	9	<i>rps4</i> , target suggested by the PPR code	Not studied yet	Tourasse et al., 2013; Cavaiuolo et al., 2017
PPR4	10	Unknown	Not studied yet	Tourasse et al., 2013
PPR5	7	Unknown	Not studied yet	Tourasse et al., 2013
PPR6	10	<i>psbF</i> , target suggested by the PPR code	Not studied yet	Tourasse et al., 2013; Cavaiuolo et al., 2017
PPR7	2 <sup>*</sup> /4	<i>rbcL</i> , <i>rpoC2</i> , <i>psbH</i> , <i>tscA</i> , 16S rRNA, <i>psbJ</i> , and <i>atpA</i> gene clusters	Stabilization of <i>rbcL</i> , <i>rpoC2</i> , <i>psbH</i> , and <i>tscA</i> transcripts. Maturation of polycistronic mRNAs <i>psbJ</i> and <i>atpA</i> gene clusters as well as of 16S rRNA.	Tourasse et al., 2013 <sup>*</sup> ; Jalal et al., 2015
PPR8	8	Unknown	Not studied yet	Tourasse et al., 2013
PPR9	4	Unknown	Not studied yet	Tourasse et al., 2013
PPR10	5	Unknown	Not studied yet	Tourasse et al., 2013
PPR11	2	Unknown	Not studied yet	Tourasse et al., 2013
PPR12	2	Unknown	Not studied yet	Tourasse et al., 2013
PPR13	5	Unknown	Not studied yet	Tourasse et al., 2013
MCA1 (PPR14)	14	<i>petA</i>	Stabilization of <i>petA</i> mRNA	Raynaud et al., 2007; Loiselay et al., 2008; Boulouis et al., 2011; Tourasse et al., 2013

Asterisks (\*) indicate the specific reference of that data.

profiling, it has been proposed that the role of PPR1 is mainly as a translation factor, hence the TCB1 name, and that in addition, it stabilizes *petB* (Cavaiuolo et al., 2017).

In contrast to MCA1 and MRL1, both of which have a single mRNA target, PPR7 protein, which is the most highly expressed PPR protein in *Chlamydomonas*, participates in regulating the expression of several chloroplast genes (Jalal et al., 2015). PPR7 is part of a ribonucleoprotein complex that contributes to the stabilization of *rbcL*, *rpoC2*, *psbH*, and *tscA* transcripts (Jalal et al., 2015). Additionally, PPR7 also participates in the processing and maturation of the following polycistronic transcripts: the gene cluster *atpA-psbI-cemA-atpH*, which encodes for the  $\alpha$ -subunit of the ATP synthase; a small PSII subunit, a putative envelope membrane protein involved in inorganic carbon uptake; and the subunit III of chloroplast ATP synthase, respectively (Jalal et al., 2015). PPR7 also participates in the maturation of the co-transcript *psbJ-atpI-psaJ-rps12*, which encodes for PsbJ (a PSII reaction center protein), the subunit IV of ATP synthase; PsaJ (a PSI reaction center subunit), and the ribosomal protein S12. Additionally, PPR7 participates in processing the 16S rRNA, which is part of the *rrn* operon that encodes for the 16S, 7S, 3S, 23S, 5S rRNAs, and two tRNAs (*trnI* and *trnA*; Jalal et al., 2015). Detailed mechanisms of PPR7 in processing all these polycistronic transcripts are still scarce and unclear, particularly when compared to the level of details known for PPR proteins in plants (Pfalz et al., 2009; Stoppel and Meurer, 2012). A hypothesis is that PPR7 could recruit certain endonucleases or modify large secondary structures in large mRNAs to unmask nuclease sensitive targets. In any case, given the diversity of functions of the regulated genes, PPR7 is essential for the proper functioning of chloroplasts (Jalal et al., 2015).

## TETRA-TRICOPEPTIDE REPEAT PROTEINS

Like PPR proteins, TPR proteins are helical repeat proteins characterized by the presence of 1–26 tandem repeats, each one of 34 amino acids that form an antiparallel  $\alpha$ -helical hairpin (Bohne et al., 2016). The TPR motifs function as a scaffold for protein assembly in multi-subunit complexes and are involved in various processes ranging from transcriptional regulation and RNA metabolism to protein folding, chlorophyll synthesis, and transport (Table 2). Given the diversity of functions, it is not surprising that the location of TPR proteins is not limited to the chloroplast or the mitochondria. Two flagellar subunit proteins with TPR motifs, IFT139 and IFT88, have been identified in *C. reinhardtii* as part of the IFT-A and IFT-B subcomplex involved in intraflagellar transport (Pazour et al., 2000; Snell and Goodenough, 2009; Behal et al., 2012). IFT88-deficient *C. reinhardtii* mutants show a phenotype without flagella (Pazour et al., 2000). Because IFT88 is conserved in green algae, nematodes, and vertebrates, the analysis of *C. reinhardtii* IFT88 mutants has allowed the study of intra-flagellar transport in other species, for example, the study of its homologous genes in mice and humans called Tg737. Mice with Tg737 defects die shortly after birth from polycystic kidney disease (Pazour et al., 2000).

Tetra-tricopeptide repeat motifs can also act as sites for protein–protein interactions; in *Chlamydomonas* and tobacco, Ycf3, a chloroplasts-encoded TPR protein, is responsible for the initial assembly of PSI (Boudreau et al., 1997). The Ycf3 protein interacts with the PsaA and PsaB subunits forming the PSI reaction center subcomplex to which the Ycf4-module binds to provide stabilization and to facilitate the integration

**TABLE 2** | Tetra-tricopeptide repeat and HAT-TPR proteins identified and studied in *C. reinhardtii*.

Type of protein	Name	Number of TPR motifs	Target	Function	Reference
TPR	CGL71	3	Not apply	Assembly of the PSI complex. Involved in protecting the PSI assembly process from oxidative disruption	Heinrickel et al., 2016
	IFT139	4	Not apply	Involved in the intraflagellar transport. IFTA subcomplex.	Pazour et al., 2000; Behal et al., 2012
	IFT88	2	Not apply	Involved in the intraflagellar transport. IFTB subcomplex.	Pazour et al., 2000; Behal et al., 2012
	FAP70	2	Not apply	Involved in ciliary motility	Shamoto et al., 2018; Hou et al., 2021
	FLP	2	Not apply	Regulator of the chlorophyll synthesis pathway Factor for the <i>psbA</i> mRNA translation initiation and assembly of the nascent D1 protein.	Falciatore et al., 2005
	REP27	2	<i>psbA</i>	Activation of the bound D1 in the PSII reaction center during the PSII repair process.	Park et al., 2007; Dewez et al., 2009
	Tic40	1	Not apply	Component of the chloroplast protein import machinery.	Kalanon and McFadden, 2008
	Ycf3	3	PsaA, PsaD subunits	Assembly of the PSI complex.	Boudreau et al., 1997; Naver et al., 2001; Rochaix et al., 2004; Nellaepalli et al., 2018
HAT-TPR	MAC1	12	<i>psaC</i>	Stabilization of <i>psaC</i> 5'UTR.	Douchi et al., 2016
	MBB1	10	<i>psbB-psbT-psbH</i>	Processing of <i>psbB-psbT-psbH</i> operon. Stabilization of <i>psbB-psbT</i> and <i>psbH</i> 5'UTR.	Monod et al., 1992; Vaistij et al., 2000a,b; Loizeau et al., 2014
	MBD1 (NAC2)	9	<i>psbD</i>	Stabilization of <i>psbD</i> 5'UTR.	Kuchka et al., 1989; Boudreau et al., 2000; Schwarz et al., 2007, 2012

of the other central PSI subunits (PsaC, D, E, F, H, I, J, and L) to form the mature PSI-LHCI supercomplex (Naver et al., 2001; Nellaepalli et al., 2018, 2021). Ycf3 mutants display a light-sensitive phenotype, and they cannot grow photoautotrophically, demonstrating that Ycf3 is essential for photosynthesis. In addition to Ycf3, CGL71, another TPR protein, also participates in the biogenesis and stability of PSI, protecting it from oxidative disruption during its assembly (Heinrickel et al., 2016). FLP and Tic40 also participate in chloroplast biogenesis, the former in the regulation of chlorophyll biosynthesis and the latter as a component of the chloroplast import machinery (Falciatore et al., 2005; Kalanon and McFadden, 2008).

Another example is REP27, a TPR protein located on the thylakoid membrane, which participates in the PSII repair cycle (Park et al., 2007; Dewez et al., 2009). During oxygenic photosynthesis, the abundance of O<sub>2</sub> and the formation of oxidants lead to photo-oxidative damages causing irreversible

inactivation of the D1 protein, impairing electron transport, and inhibiting the function of PSII (Park et al., 2007). REP27 participates by facilitating and regulating different stages of D1 biosynthesis *de novo*. It first acts as a factor for the initiation of *psbA* translation (D1 protein). Then it facilitates the assembly of the nascent D1 peptide. Finally, REP27 activates the D1 protein through post-translational modifications conferring it its functional state in the PSII complex (Dewez et al., 2009). Thus, REP27 plays a triple role, in the regulation of PSII turnover, assembly, and activation.

Like PPR proteins, TPR proteins are also involved in the metabolism of RNA, and this function has been associated with the HAT (half-a-TPR) motif, a subclass of TPR motifs found in several proteins that stabilize and/or activate the translation of specific chloroplast RNAs (Bohne et al., 2016). Unlike the TPR motif, the HAT motifs are characterized by a distinct pattern of conserved amino acid positions and by the presence of highly conserved aromatic residues, mainly

tyrosine and tryptophan (Preker and Keller, 1998). In *C. reinhardtii* the existence of 10 HAT proteins has been estimated (Table 2; Bohne et al., 2016), from which three have been functionally characterized; NAC2 (also named MBD1; Kuchka et al., 1989; Boudreau et al., 2000; Schwarz et al., 2007), MAC1 (Douchi et al., 2016), and MBB1 (Vaistij et al., 2000a,b; Loizeau et al., 2014). NAC2 participates in the stabilization and processing of the 5'-UTRs of *psbD* (encoding the D2 subunit of the PSII reaction center), MAC1 for *psaC* (encoding the Photosystem I iron-sulfur protein), and MBB1 for *psbB* mRNA (encoding the PSII core antenna CP47). These factors bind to the 5'-UTR end of their target, protecting the transcript from exonucleases degradation. Additionally, MBB1, an orthologue of HCF107 (high chlorophyll fluorescence 107) protein in *A. thaliana*, participates in the processing of the *psbB-psbT-psbH* polycistronic transcript (Felder et al., 2001). Before translation, this mRNA must be divided into smaller transcripts, *psbB-psbT*, and *psbH* (Vaistij et al., 2000b). MBB1 binds to the 5'-UTR of *psbB* and *psbH* and it is speculated that this could recruit an endonuclease or modify the secondary structure of the mRNA exposing the sites for endonuclease cleavage. Once the polycistronic mRNA is cleaved, MBB1 stabilizes the monocistronic transcripts, protecting them from exonucleases (Loizeau et al., 2014).

## OCTOTRICOPEPTIDE REPEAT PROTEINS

Octo-tricopeptide repeat proteins are the most abundant helical repeat proteins in green algae. Their structure is characterized by the presence of motifs of 38–40 amino acids repeated in tandem, in the same way as in PPR and TPR proteins, where each motif forms a pair of antiparallel  $\alpha$ -helices (Eberhard et al., 2011; Wang et al., 2015). Besides green algae, OPR proteins are present in some pathogenic bacteria (*Coxiella burnetii*; Rahire et al., 2012), symbiotic protozoa (e. g. *Plasmodium falciparum*, and *Toxoplasma gondii*; Hillebrand et al., 2018; Hollin et al., 2021) and the cercozoan amoeba *Paulinella chromatophora* (Oberleitner et al., 2020); but are rare in land plants. In *A. thaliana*, the presence of only one OPR protein has been reported, called RAP, which is involved in the processing of the 16S ribosomal RNA chloroplast gene (Kleinknecht et al., 2014). In contrast, analysis *in silico* has revealed that there are more than 120 members (Goodstein et al., 2012; Boulouis et al., 2015). The abundance of OPR proteins in the nuclear genome of *Chlamydomonas* shows not only a high prevalence but also a complex regulation of chloroplast gene expression, where these factors play an essential role. In *C. reinhardtii*, this family of proteins is extensively involved in the post-transcriptional control of chloroplast gene expression (Table 3). The first OPR described in *C. reinhardtii* was the chloroplast translation factor TBC2 (Auchincloss et al., 2002), which together with the nuclear proteins TBC1 and TBC3 interact with the 5'-UTR end of the *psbC* mRNA favoring its translation (Zerges et al., 2003). Like TBC2, TAB1 (Stampacchia et al., 1997; Rahire et al., 2012), TDA1 (Eberhard et al., 2011), and TAA1 (Lefebvre-Legendre et al., 2015; Reifschneider et al., 2016) also function

as translation factors in the translation of *psaB*, *atpA*, and *psaA* mRNAs, respectively.

In addition to facilitating the translation of specific mRNAs, some OPR proteins can perform additional functions or even carry out dual functions (Eberhard et al., 2011; Ozawa et al., 2020). They can function as M (maturation) factors, protecting their mRNA target from degradation by exonucleases and/or participate in the processing of polycistronic mRNAs (Murakami et al., 2005; Wang et al., 2015; Viola et al., 2019; Ozawa et al., 2020), as well as acting as factors that participate in gene *trans*-splicing (Reifschneider et al., 2016). The coordinated interaction of these factors on maturation/stability and translation greatly affects organelle function and, consequently, photosynthesis, respiration, and environmental responses. Some of the OPR proteins functionally characterized as M factors comprise the proteins MBI1, MCG1, MTHI1, and MDA1, which stabilize the *psbI*, *petG*, *atpH/atpI*, and *atpA* transcripts, respectively (Wang et al., 2015; Viola et al., 2019; Ozawa et al., 2020). The binding of these proteins to the 5' UTR protects transcripts against 5'-3' exonucleolytic degradation by RNA exonucleases.

Interestingly, OPR proteins may also play a double role; in some cases, acting as T factors while also participating in the stabilization of their target transcript. The opposite may also take place, M factors contributing to translation of their target. Such is the case of TAA1 (Lefebvre-Legendre et al., 2015) a translation factor required for PsaA accumulation that also participates in preventing *psaA* degradation by exonucleases. Another example of a dual role is the MTHI1 protein; which, unlike TAA1, is involved in regulation of the expression of two different targets, *atpH* and *atpI* mRNAs (Ozawa et al., 2020). On the one hand, MTHI1 works as M factor for the *atpH* mRNA, allowing for the accumulation of a monocistronic transcript but also participating as translational enhancer. On the other hand, MTHI1 is required for translation activation of the *atpI* mRNA but not for its stabilization (Ozawa et al., 2020). Conversely, the expression of a single transcript can be regulated by more than one OPR protein; such is the case for the *atpA* transcript in the *atpA-psbI-cemA-atpH* gene cluster (Drapier et al., 1998). The accumulation of the AtpA protein depends on the maturation and translation activities of the MDA1 and TDA1 factors, respectively, both of which bind to the 5'-UTR of *atpA*. While MDA1 stabilizes and participates in processing the *atpA* transcript (Viola et al., 2019), TDA1 intervenes, facilitating the start of translation (Eberhard et al., 2011).

Another OPR protein is CCS2 which is required for the assembly of the cytochrome *c* complex. CCS2 contains an LWALAR consensus amino acids motif (Cline et al., 2017), which is different to the PPPEW motif, typical of other known OPR proteins such as TBC2, TDA1, and TAB1 (Auchincloss et al., 2002; Eberhard et al., 2011; Rahire et al., 2012). Due to its OPR motifs, CCS2 was initially suggested to participate in the maturation, stabilization, or translation of a chloroplast transcript involved in cytochrome *c* assembly. The transcript of the *ccsA* gene (a heme delivery factor required for cytochrome *c* maturation) has been suggested as the CCS2 target.



**TABLE 3** | Octo-tricopeptide repeat proteins identified and studied in *C. reinhardtii*.

Type of protein	Name	Number of OPR motifs	Target	Function	Reference
	CCS2	10	<i>Possibly ccsA</i>	Involved in assembly of cytochrome <i>c</i> complex.	Cline et al., 2017
	MBC1 (OPR56)	ND	<i>psbC</i>	Stabilization of <i>psbC</i> mRNA	Cavaiuolo et al., 2017
	MBI1	17	<i>psbI</i>	Stabilization of <i>psbI</i> 5'UTR	Wang et al., 2015
	MCG1	12	<i>petG</i>	Stabilization of <i>petG</i> 5'UTR	Wang et al., 2015
	MCD1	7	<i>petD</i>	Stabilization of the <i>petD</i> 5'UTR	Murakami et al., 2005
	MDA1	13	<i>atpA</i>	Stabilization and processing of the <i>atpA</i> transcript.	Drapier et al., 1998, 2002; Viola et al., 2019
	MTH11	9	<i>atpH, atpI</i>	Stabilization and translation of <i>atpH</i> mRNA. Translation activation and stabilization <i>atpI</i> mRNA.	Ozawa et al., 2020
OPR	RAA1	5	<i>psaA</i>	<i>Trans</i> -splicing of <i>psaA</i> mRNA	Rochaix et al., 2004; Merendino et al., 2006; Lefebvre-Legendre et al., 2016
	RAA3	5	<i>psaA</i>	<i>Trans</i> -splicing of <i>psaA</i> mRNA	Rivier et al., 2001; Reifschneider et al., 2016
	RAA8 (OPR120)	6	<i>psaA</i>	<i>Trans</i> -splicing of <i>psaA</i> mRNA	Marx et al., 2015; Lefebvre-Legendre et al., 2016; Reifschneider et al., 2016
	RAT2	2	<i>psaA</i>	<i>Trans</i> -splicing of <i>psaA</i> mRNA and processing of 3' end of <i>tscA</i> RNA	Eberhard et al., 2011; Marx et al., 2015; Lefebvre-Legendre et al., 2016; Reifschneider et al., 2016
	TAA1	7	<i>psaA</i>	Translation and stabilization of <i>psaA</i> mRNA.	Lefebvre-Legendre et al., 2015; Reifschneider et al., 2016
	TAB1	10	<i>psaB</i>	Translation of the <i>psaB</i> mRNA.	Stampacchia et al., 1997; Wostrikoff et al., 2004; Rahire et al., 2012
	TBC2	9	<i>psbC</i>	Translation of <i>psbC</i> mRNA	Auchincloss et al., 2002; Zerges et al., 2003; Rahim et al., 2016
	TDA1	8	<i>atpA</i>	Translation of <i>atpA</i> mRNA	Eberhard et al., 2011; Viola et al., 2019
	OPR1 (Cre01.g001501)	ND	<i>psaA</i>	<i>Trans</i> -splicing of <i>psaA</i> mRNA	Reifschneider et al., 2016
	OPR104 (Cre17.g698750)	ND	<i>psaA</i>	<i>Trans</i> -splicing of <i>psaA</i> mRNA	Reifschneider et al., 2016
OPR-NCL	NCC1	9	<i>atpA</i>	Destabilization of <i>atpA</i> transcripts	Boulouis et al., 2015
	NCC2	9	<i>petA</i>	Destabilization of <i>petA</i> transcripts	Boulouis et al., 2015

Experimental evidence has shown that CCS2 is not required for stabilization of *ccsA* thus its role as a **M** factor can be discarded. However, CCS2 acting as **T** factor is a matter of consideration. An additional proposed mode of action, although less plausible, is that CCS2 could stabilize the CCS complex (composed of CCS1 and CcsA) *via* protein–protein interactions or by facilitating the recruitment of additional components to this complex (Cline et al., 2017).

Another role attributed to OPR proteins is participation in splicing processes. A clear example of this is the maturation of *psaA* mRNA, which requires two steps of *trans*-splicing (Herrin and Schmidt, 1988; **Figure 1B**). The *psaA* gene is encoded in three separate exons that are scattered throughout the chloroplast genome. Each of the exons is transcribed separately and has flanking sequences that allow its assembly by forming structures belonging to the group II introns (Kück et al., 1987; Choquet et al., 1988). Additionally, the *trans*-splicing chloroplast A (*tscA*) locus codes for a short non-coding RNA that completes the structure of the first tripartite group II intron (Goldschmidt-Clermont et al., 1991). At least 14 factors encoded in the nucleus participate in the splicing of *psaA* (Reifschneider et al., 2016).

Although not all of them have been characterized, at least six of these factors contain OPR motifs: RAT2, RAA1, RAA3, RAA8 (also known as OPR120), OPR1 (Cre01.g001501), and OPR104 (Cre17.g698750; Merendino et al., 2006; Marx et al., 2015; Lefebvre-Legendre et al., 2016; Reifschneider et al., 2016; **Figure 1B**). All these factors are part of a post-transcriptional complex, presumably functioning as RNA interaction module (Lefebvre-Legendre et al., 2016; Reifschneider et al., 2016; Kück and Schmitt, 2021). On the one hand, RAT2 (RNA-maturation of *psaA-tscA*) is necessary to process *tscA* from a polycistronic precursor (Balczun et al., 2005). While RAA8 (RNA-maturation of *psaA*) participates in the splice junction of the first *psaA* intron (Marx et al., 2015), RAA1 (RNA-maturation of *psaA1*) is involved in the processing of the *tscA* mRNA and participates in the *trans*-splicing of introns 1 and 2 (Rochaix et al., 2004; Merendino et al., 2006). Like many of the nucleus-encoded proteins involved in the anterograde regulation, OPR proteins activity is regulated in response to environmental factors such as light induction (RAA1), light repression (RAA3), H<sub>2</sub>O<sub>2</sub> suppression (RAA1; Wang et al., 2015), proteolytic degradation in response to nitrogen (MCA1 and

TCA1; Wei et al., 2014), and iron deprivation (MAC1 and TAA1; Lefebvre-Legendre et al., 2015; Douchi et al., 2016).

Most functionally known OPR proteins have been identified by screening nuclear mutants with photosynthetic or respiratory defects after mutagenesis. However, Boulouis et al. (2015) characterized two nuclear spontaneous mutants with single amino acid substitutions in two distinct genes encoding OPR proteins: OPR87 (Cre15.g638950.t1) and OPR98 (Cre15.g640400.t1). These nuclear mutants, called *ncc1* and *ncc2*, have expanded the recognition capacity of the OPR proteins, targeting the coding regions of two chloroplast transcripts, leading to the destabilization of *atpA* or *petA* transcripts, respectively (Drapier et al., 2002; Boulouis et al., 2015). NCC1 and NCC2 may recruit an endonuclease or may themselves carry on endonucleolytic activity and degrade their new targets (Boulouis et al., 2015). These proteins were classified as NCL (NCC-Like) proteins, a new subfamily of paralogs encoding OPR-RAP proteins, that comprise a highly conserved central region, containing 7–12 OPR repeats and a C-terminal RAP domain (Boulouis et al., 2015). The deeper we study these mutants and other OPR, PPR, and TPR proteins the better we will understand the importance and influence of anterograde control in organellar gene expression regulation. From a biotechnological point of view, elucidating how these proteins bind to nucleic acid sequences and understanding how changes in their motifs, as in the case of NCC1 and NCC2, alter their specificity will have an impact in the development of molecular tools. This will also allow us to skillfully tune the regulation of chloroplast metabolism to obtain new products, change the balance of chloroplast metabolites, and perhaps even achieve the long-sought improvement of photosynthesis.

## BIOTECHNOLOGICAL APPLICATIONS OF PPR, TPR, AND OPR PROTEINS

The study of PPR proteins in plants has allowed us to elucidate the mechanism by which these proteins recognize and bind to their target mRNA. The level of details is such that it has been possible to postulate an RNA recognition code for PPR proteins (Barkan et al., 2012; Yagi et al., 2013, 2014). Considering that each PPR motif recognizes a ribonucleotide, it has been proposed that the amino acid residues at positions 1, 4, and ii (the penultimate residue) of each PPR motif are determinants for the recognition of their RNA bases. Being the residues 4 and ii, those with the greatest influence on ribonucleotide recognition (Barkan et al., 2012; Yagi et al., 2013; **Figures 2C,D**). Some of the recognition codes that have been proposed include the amino acids Val, Asn, and Ser at positions 1, 4, and ii, respectively, for cytosine recognition; Val, Asn, and Asp for uracil; Val, Thr, and Asp for guanine; and finally, Phe, Thr, and Asn for adenine (Barkan et al., 2012; Yagi and Shiina, 2014; **Figure 2D**). However, the bioinformatic analysis predicts that there are other combinations of the PPR code amino acids; in fact, a few of them have been explored (Barkan et al., 2012; Yin et al., 2013; Cheng et al., 2016; Yan et al., 2019). From a biotechnological point of view, elucidating the PPR

code is an opportunity area for the design of new RNA-binding proteins with a particular motif or additional domains bound to specific targets for the identification or control of mRNA metabolism.

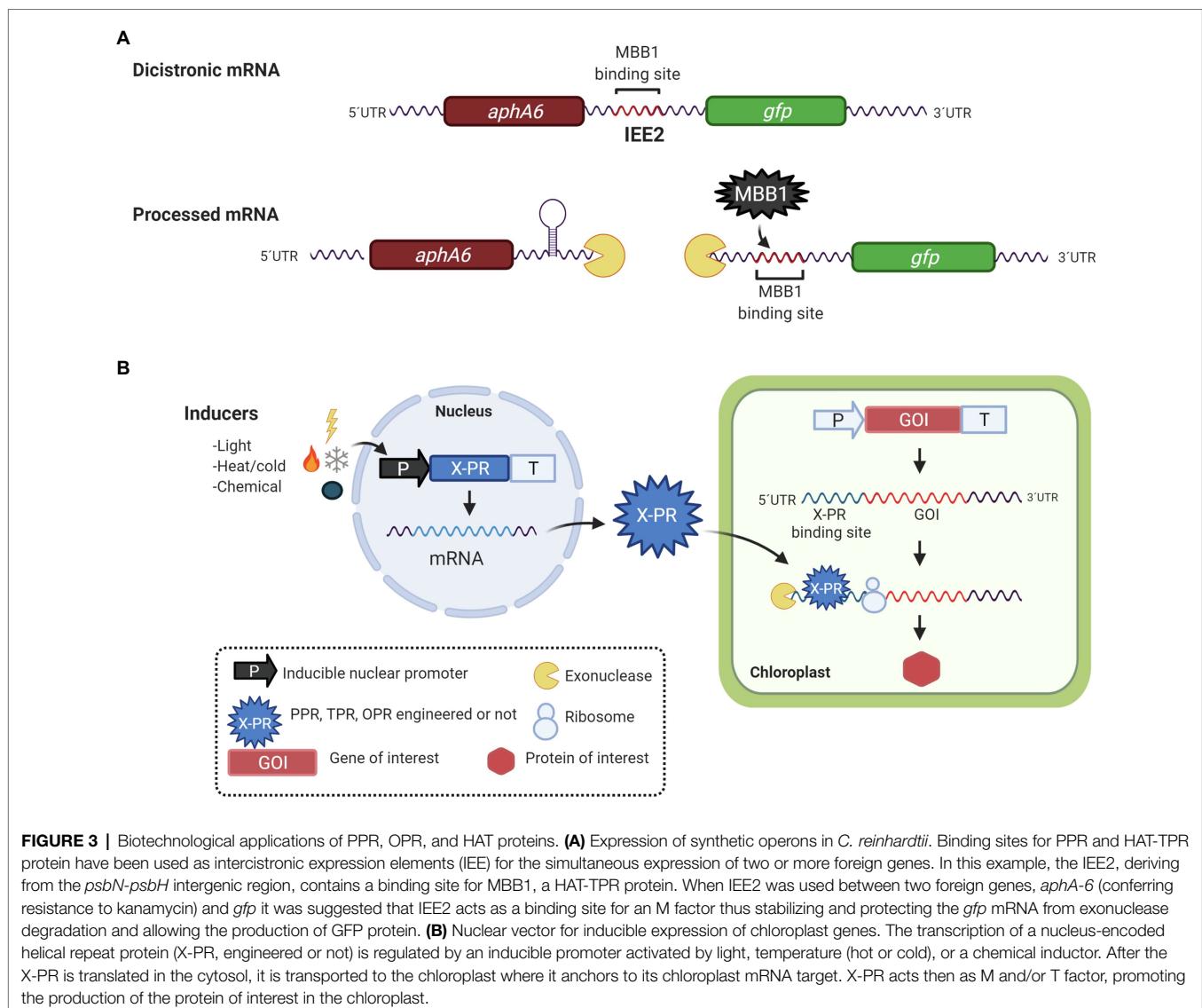
Imai et al. (2018) suggest using PPR motifs as scaffolds to develop designed RNA-binding proteins. The fusion of PPR motifs with other domains could allow the development of new tools for (a) RNA detection *in vivo* by assembling engineered-PPR (ePPR) to a reporter protein (e.g., GFP); (b) translation regulation by fusing ePPR motif to eIF4G (eukaryotic translation initiation factor 4G), allowing activation translation of the targeted mRNA; (c) development of a site-specific endonuclease, by fusing ePPR motif with an endonuclease, which allows recognizing and cleaving RNA in a specific sequence; (d) RNA editing, using an ePPR protein with cytidine deaminase domains to catalyze cytosine's reaction to uracil at a specific position; and (e) finally, the use of engineered motifs targeted to otherwise naked or susceptible to degradation mRNA. This, depending on the nature of each original P/O/TPR protein, could be used to increase the half-life of the transcript, enhance translation, or even tune down transcript translatability. The use of designed PPR proteins has great potential in developing tools that require RNA binding or tagging applications. McDermott et al. (2019) recently demonstrated that an artificial protein constructed with PPR motifs targeting a specific RNA-target could specifically bind to it *in vivo* and have low levels of off-target bindings. This work was the first approach to developing engineered PPR proteins as tools to study RNA metabolism in plants.

Another opportunity for the biotechnological applications of PPR, TPR, and OPR proteins is their binding sites. Considering that natively these proteins bind to 5'- and 3'-UTR targets and function as M or T mRNA factors, it has been proposed to use them in the maturation-stabilization and translation of foreign genes with 5'-UTRs containing P/T/OPR proteins binding sites. Zhou et al. (2007), Legen et al. (2018), and Macedo-Osorio et al. (2018) have identified sequences called Intercistronic Expression Elements (IEEs), which contain TPR and PPR proteins binding sites, that have been used for the expression of heterologous polycistronic mRNA in tobacco (Zhou et al., 2007; Lu et al., 2013; Fuentes et al., 2016; Legen et al., 2018), tomato (Lu et al., 2013), and *C. reinhardtii* (Macedo-Osorio et al., 2018). IEEs were first described by Zhou et al. (2007), who identified a sequence of 50 bases on the 5'-UTR of the *psbH* gene in tobacco. It contains the binding site of the HAT protein HCF107; this element was recognized as an IEE. Under native conditions, HCF107 binds to the 5'-UTR of *psbH* and stabilizes the monocistronic transcript of *psbH* once it is cleaved from the bicistronic transcript *psbT-psbH* (Felder et al., 2001; Sane et al., 2005). When the IEE is inserted into the intergenic region of the bicistronic transcript *yfp-nptII* (yellow fluorescent protein and kanamycin resistance gene, respectively), it is able to stabilize and increase the translation of the monocistronic transcripts (Zhou et al., 2007). Its function has been attributed to the fact that it contains the binding site for HCF107 (Hammani et al., 2012). As expected, HCF107 joins to the 5'-UTR of the foreign gene, stabilizing and protecting the transcript from

exonucleases and generating higher translated product (Zhou et al., 2007; Lu et al., 2013; Fuentes et al., 2016). This IEE has been successfully used to express genes involved in the tocopherol biosynthesis pathway to produce vitamin E in tobacco and tomato (Lu et al., 2013) and for the expression of the core biosynthetic pathway of artemisinic acid production in tobacco chloroplasts (Fuentes et al., 2016). Although, these elements have great potential in metabolic pathway engineering, Legen et al. (2018) recently, demonstrated that when an IEE is used in the stabilization of foreign genes, there is a negative effect on the stabilization of the endogenous *psbH* transcript also, possibly due to a phenomenon of HCF107 titration for competition between the binding sites of the foreign transcript and the native *psbH* transcript. This problem can be solved by designing PPR and TPR proteins with specific binding sites to foreign genes. Additionally, Legen et al. (2018) showed that binding sites for PPR10, HCF152, and CCR2 can be used as IEEs when they were evaluated in a synthetic *neo-egfp* bicistron

in tobacco chloroplasts. Similar results were obtained by Macedo-Osorio et al. (2018) in *C. reinhardtii* when the intergenic region of *psbN-psbH* was used as IEE, which contains the anchoring site for the MBB1 protein, an ortholog of HCF107. This region, called IEE2, could stabilize the monocistronic transcript *gfp* on the bicistronic *aphA6-gfp* mRNA, allowing the translation and accumulation of the GFP protein. Similarly, in tobacco with HCF107, MBB1 could be binding to IEE2, functioning as an M-factor for the GFP monocistronic transcript in *C. reinhardtii* (Figure 3A).

Additionally, Rojas et al. (2019) developed an inducible switch to activate chloroplast transgenes' expression in tobacco using an engineered PPR proteins. In native conditions, PPR10 binds to the 5' end of the *atpH* mRNA protecting it from degradation by exonucleases and increasing its translational efficiency. In this case, Rojas et al. (2019) generated variants of PPR10 from maize, with modified amino acids that allow it to bind to a specific target sequence; the protein was expressed



in the nucleus under the regulation of an inducible promoter. The protein once produced could bind to a specific target sequence, located in 5'-UTR of the GFP mRNA expressed in the chloroplast. The inducible system allows the regulated expression of chloroplast genes from the nucleus (**Figure 3B**).

Similarly, Carrera-Pacheco et al. (2020) developed an inducible gene expression system in the *Chlamydomonas* chloroplast. The OPR protein, TDA1 of *Chlamydomonas* was nuclear expressed under the regulation of light, and heat-inducible promoters, HSP70A-RBCS2, the generated protein is capable of specifically binding to the 5'-UTR of *atpA* used to regulate the expression of *gfp* in the chloroplast, showing a 1.9-fold increase in the production of the recombinant protein. A reduced number of examples, albeit clear and concise, have shown that it is possible to engineer and manipulate PPR proteins to modulate and rewire gene expression, particularly with a biotechnological application, that is, to express foreign proteins in the chloroplast of plants. Application of engineered PPR, OPR, and TPR proteins in the chloroplast of *C. reinhardtii* has yet to see the light.

## FINAL REMARKS

Penta-tricopeptide repeat, TPR, and OPR proteins have been revealed as essential factors for the functioning of chloroplasts; they act as regulatory factors for maturation, stabilization, and translation of different transcripts coding for essential genes including those for photosynthesis. Although initial studies of PPR proteins in plants allowed the elucidation of their mechanisms of action, it is necessary to characterize additional PPR, TPR, and OPR proteins from *C. reinhardtii* to gain knowledge in their regulatory mechanisms in this important model organism. Understanding the way they control and

regulate the expression of genes, mainly in the chloroplast, opens the door to generate new tools to improve not only the production levels of recombinant proteins and food and relevant pharmaceutical compounds but also to perform metabolic engineering aimed at enhancing the production of lipids for biofuels. A more complex task, though not unrealistic would be to remodel cell metabolism in *Chlamydomonas* to enhance photosynthesis to make the CO<sub>2</sub> fixation rate more efficient and thus contribute to the reduction of atmospheric CO<sub>2</sub> levels, while producing sustainable value-added products.

## AUTHOR CONTRIBUTIONS

KM-O and JB-C: conceptualization. KM-O: investigation and writing of original draft. KM-O, AM-A, and JB-C: writing-reviewing and editing. AM-A: funding acquisition. All authors contributed to the article and approved the submitted version.

## FUNDING

KM-O was financially supported by CONACYT through a postdoctoral fellowship. Research in the laboratory of JB-C is financially supported by Instituto Politécnico Nacional and CONACYT through project 286497.

## ACKNOWLEDGMENTS

All figures were created with BioRender.com. KM-O would like to thank Oreo and Luneta for emotional support during the writing process.

## REFERENCES

- Andrés-Colás, N., Zhu, Q., Takenaka, M., De Rybel, B., Weijers, D., and Van Der Straeten, D. (2017). Multiple PPR protein interactions are involved in the RNA editing system in Arabidopsis mitochondria and plastids. *Proc. Natl. Acad. Sci.* 114, 8883–8888. doi: 10.1073/pnas.1705815114
- Aubourg, S., Boudet, N., Kreis, M., and Lecharny, A. (2000). In *Arabidopsis thaliana*, 1% of the genome codes for a novel protein family unique to plants. *Plant Mol. Biol.* 42, 603–613. doi: 10.1023/A:1006352315928
- Auchincloss, A. H., Zerges, W., Perron, K., Girard-Bascou, J., and Rochaix, J. D. (2002). Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of the chloroplast *psbC* mRNA in *Chlamydomonas reinhardtii*. *J. Cell Biol.* 157, 953–962. doi: 10.1083/jcb.200201060
- Balczun, C., Bunse, A., Hahn, D., Bennoun, P., Nickelsen, J., and Kück, U. (2005). Two adjacent nuclear genes are required for functional complementation of a chloroplast *trans*-splicing mutant from *Chlamydomonas reinhardtii*. *Plant J.* 43, 636–648. doi: 10.1111/j.1365-313X.2005.02478.x
- Barkan, A., Rojas, M., Fujii, S., Yap, A., Chong, Y. S., Bond, C. S., et al. (2012). A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet.* 8:e1002910. doi: 10.1371/journal.pgen.1002910
- Barkan, A., and Small, I. (2014). Pentatricopeptide repeat proteins in plants. *Annu. Rev. Biochem.* 65, 415–452. doi: 10.1146/annurev-arplant-050213-040159
- Behal, R. H., Miller, M. S., Qin, H., Lucker, B. F., Jones, A., and Cole, D. G. (2012). Subunit interactions and organization of the *Chlamydomonas reinhardtii* intraflagellar transport complex A proteins. *J. Biol. Chem.* 287, 11689–11703. doi: 10.1074/jbc.M111.287102
- Bock, R. (2017). Witnessing genome evolution: experimental reconstruction of endosymbiotic and horizontal gene transfer. *Annu. Rev. Genet.* 51, 1–22. doi: 10.1146/annurev-genet-120215-035329
- Bohne, A., Schwarz, C., Jalal, A., and Ossenbühl, F. (2009). Control of organellar gene expression in *Chlamydomonas reinhardtii* – future perspectives. *Endocytobiosis Cell Res.* 19, 70–80.
- Bohne, A., Schwenkert, S., Grimm, B., and Nickelsen, J. (2016). Roles of tetratricopeptide repeat proteins in biogenesis of the photosynthetic apparatus. *Int. Rev. Cell Mol. Biol.* 324, 187–227. doi: 10.1016/bs.ircmb.2016.01.005
- Börner, T. (2017). The discovery of plastid-to-nucleus retrograde signaling—a personal perspective. *Protoplasma* 254, 1845–1855. doi: 10.1007/s00709-017-1104-1
- Boudreau, E., Nickelsen, J., Lemaire, S. D., Ossenbühl, F., and Rochaix, J.-D. (2000). The *Nac2* gene of *Chlamydomonas reinhardtii* encodes a chloroplast TPR protein involved in *psbD* mRNA stability, processing, and/or translation. *EMBO J.* 19, 3366–3376. doi: 10.1093/emboj/19.13.3366
- Boudreau, E., Takahashi, Y., Lemieux, C., Turmel, M., and Rochaix, J. D. (1997). The chloroplast *ycf3* and *ycf4* open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem I complex. *EMBO J.* 16, 6095–6104. doi: 10.1093/emboj/16.20.6095
- Boulouis, A., Cecile, R., Bujaldon, S., Aznar, A., Wollman, F.-A., and Choquet, Y. (2011). The nucleus-encoded *trans*-acting factor MCA1 plays a critical role in the regulation of cytochrome *f* synthesis in *Chlamydomonas* chloroplasts. *Plant Cell* 23, 333–349. doi: 10.1105/tpc.110.078170



- Boulouis, A., Drapier, D., Razafimanantsoa, H., Wostrikoff, K., Tourasse, N. J., Pascal, K., et al. (2015). Spontaneous dominant mutations in *Chlamydomonas* highlight ongoing evolution by gene diversification. *Plant Cell* 27, 984–1001. doi: 10.1105/tpc.15.00010
- Carrera-Pacheco, S. E., Hankamer, B., and Oey, M. (2020). Light and heat-shock mediated TDA1 overexpression as a tool for controlled high-yield recombinant protein production in *Chlamydomonas reinhardtii* chloroplasts. *Algal Res.* 48:101921. doi: 10.1016/j.algal.2020.101921
- Cavaiuolo, M., Kuras, R., Wollman, F. A., Choquet, Y., and Vallon, O. (2017). Small RNA profiling in *Chlamydomonas*: insights into chloroplast RNA metabolism. *Nucleic Acids Res.* 45, 10783–10799. doi: 10.1093/nar/gkx668
- Cheng, S., Gutmann, B., Zhong, X., Ye, Y., Fisher, M. F., Bai, F., et al. (2016). Redefining the structural motifs that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants. *Plant J.* 85, 532–547. doi: 10.1111/tpj.13121
- Choquet, Y., Goldschmidt-Clermont, M., Girard-Bascou, J., Kück, U., Bennoun, P., and Rochaix, J. D. (1988). Mutant phenotypes support a trans-splicing mechanism for the expression of the tripartite *psaA* gene in the *C. reinhardtii* chloroplast. *Cell* 52, 903–913. doi: 10.1016/0092-8674(88)90432-1
- Choquet, Y., Stern, D. B., Wostrikoff, K., Kuras, R., Girard-Bascou, J., and Wollman, F. A. (1998). Translation of cytochrome *f* is autoregulated through the 5' untranslated region of *petA* mRNA in *Chlamydomonas* chloroplasts. *Proc. Natl. Acad. Sci.* 95, 4380–4385. doi: 10.1073/pnas.95.8.4380
- Choquet, Y., and Wollman, F. A. (2009). “The CES process,” in *The Chlamydomonas Sourcebook*. eds. H. Elizabeth, S. David and W. George (San Diego, CA: Academic Press), 1027–1063.
- Cline, S. G., Laughbaum, I. A., and Hamel, P. P. (2017). CCS2, an octatricopeptide repeat protein, is required for plastid cytochrome *c* assembly in the green alga *Chlamydomonas reinhardtii*. *Front. Plant Sci.* 8:1306. doi: 10.3389/fpls.2017.0130
- De Marchis, F., Pompa, A., Ballottari, M., and Bellucci, M. (2019). Host-endosymbiont co-evolution shaped chloroplast translational regulation. *Bot. Lett.* 166, 309–325. doi: 10.1080/23818107.2019.1623716
- Dewez, D., Park, S., García-Cerdán, J. G., Lindberg, P., and Melis, A. (2009). Mechanism of REP27 protein action in the D1 protein turnover and photosystem II repair from photodamage. *Plant Physiol.* 151, 88–99. doi: 10.1104/pp.109.140798
- Douchi, D., Qu, Y., Longoni, P., Legendre-Lefebvre, L., Johnson, X., Schmitz-Linneweber, C., et al. (2016). A nucleus-encoded chloroplast Phosphoprotein governs expression of the photosystem I subunit PsaC in *Chlamydomonas reinhardtii*. *Plant Cell* 28, 1182–1199. doi: 10.1105/tpc.15.00725
- Drapier, D., Girard-Bascou, J., Stern, D. B., and Wollman, F. A. (2002). A dominant nuclear mutation in *Chlamydomonas* identifies a factor controlling chloroplast mRNA stability by acting on the coding region of the *atpA* transcript. *Plant J.* 31, 687–697. doi: 10.1046/j.1365-313X.2002.01387.x
- Drapier, D., Suzuki, H., Levy, H., Rimbault, B., Kindle, K. L., Stern, D. B., et al. (1998). The chloroplast *atpA* gene cluster in *Chlamydomonas reinhardtii*. Functional analysis of a polycistronic transcription unit. *Plant Physiol.* 117, 629–641. doi: 10.1104/pp.117.2.629
- Dreesen, I. A., Charpin-El, H. G., and Fussenegger, M. (2010). Heat-stable oral alga-based vaccine protects mice from *Staphylococcus aureus* infection. *J. Biotechnol.* 145, 273–280. doi: 10.1016/j.jbiotec.2009.12.006
- Dyo, Y. M., and Purton, S. (2018). The algal chloroplast is a synthetic biology platform for production of therapeutic proteins. *Microbiology* 164, 113–121. doi: 10.1099/mic.0.000599
- Eberhard, S., Loisel, C., Drapier, D., Bujaldon, S., Girard-Bascou, J., Kuras, R., et al. (2011). Dual functions of the nucleus-encoded factor TDA1 in trapping and translation activation of *atpA* transcripts in *Chlamydomonas reinhardtii* chloroplasts. *Plant J.* 67, 1055–1066. doi: 10.1111/j.1365-313X.2011.04657.x
- Eckardt, N. A. (2006). Genomic hopscotch: gene transfer from plastid to nucleus. *Plant Cell* 18, 2865–2867. doi: 10.1105/tpc.106.049031
- Economou, C. K., Wannathong, T., Szaub, J., and Purton, S. (2014). “A simple, low-cost method for chloroplast transformation of the green alga *Chlamydomonas reinhardtii*,” in *Chloroplast Biotechnology. Methods and Protocols*. Vol. 1132. ed. P. Maliga (New York, NY: Humana Press), 401–411.
- Falciatore, A., Merendino, L., Barneche, F., Ceol, M., Meskauskiene, R., Apel, K., et al. (2005). The FLP proteins act as regulators of chlorophyll synthesis in response to light and plastid signals in *Chlamydomonas*. *Genes Dev.* 19, 176–187. doi: 10.1101/gad.321305
- Felder, S., Meierhoff, K., Sane, A. P., Meurer, J., Driemel, C., Plücker, H., et al. (2001). The nucleus-encoded HCF107 gene of *Arabidopsis* provides a link between intergenic RNA processing and the accumulation of translation-competent *psbH* transcripts in chloroplasts. *Plant Cell* 13, 2127–2141. doi: 10.1105/TPC.010090
- Fisk, D. G., Walker, M. B., and Barkan, A. (1999). Molecular cloning of the maize gene *crp1* reveals similarity between regulators of mitochondrial and chloroplast gene expression. *EMBO J.* 18, 2621–2630. doi: 10.1093/emboj/18.9.2621
- Fuentes, P., Zhou, F., Erban, A., Karcher, D., Kopka, J., and Bock, R. (2016). A new synthetic biology approach allows transfer of an entire metabolic pathway from a medicinal plant to a biomass crop. *elife* 5:e13664. doi: 10.7554/eLife.13664
- Fujii, S., and Small, I. (2011). The evolution of RNA editing and pentatricopeptide repeat genes. *New Phytol.* 191, 37–47. doi: 10.1111/j.1469-8137.2011.03746.x
- Gallaher, S. D., Fitz-Gibbon, S. T., Strenkert, D., Purvine, S. O., Pellegrini, M., and Merchant, S. S. (2018). High-throughput sequencing of the chloroplast and mitochondrion of *Chlamydomonas reinhardtii* to generate improved de novo assemblies, analyze expression patterns and transcript speciation, and evaluate diversity among laboratory strains and wild isolates. *Plant J.* 93, 545–565. doi: 10.1111/tpj.13788
- Ganusova, E. E., Reagan, B. C., Fernandez, J. C., Azim, M. F., Sankoh, A. F., Freeman, K. M., et al. (2020). Chloroplast-to-nucleus retrograde signaling controls intercellular trafficking via plasmodesmata formation. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.* 375:20190408. doi: 10.1098/rstb.2019.0408
- Goldschmidt-Clermont, M., Choquet, Y., Girard-Bascou, J., and Michel, F. (1991). A small chloroplast RNA may be required for trans-splicing in *Chlamydomonas reinhardtii*. *Cell* 65, 135–143. doi: 10.1016/0092-8674(91)90415-U
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186. doi: 10.1093/nar/gkr944
- Guillaumot, D., Lopez-Obando, M., Baudry, K., Avon, A., Rigail, G., de Longevialle, A. F., et al. (2017). Two interacting PPR proteins are major *Arabidopsis* editing factors in plastid and mitochondria. *Proc. Natl. Acad. Sci.* 114, 8877–8882. doi: 10.1073/pnas.1705780114
- Gutmann, B., Royan, S., Schallenberg-Rüdinger, M., Lenz, H., Castleden, I. R., McDowell, R., et al. (2019). The expansion and diversification of pentatricopeptide repeat RNA editing factors in plants. *Mol. Plant* 13, 215–230. doi: 10.1016/j.molp.2019.11.002
- Hammani, K., Cook, W. B., and Barkan, A. (2012). RNA binding and RNA remodeling activities of the half-a-tetratricopeptide (HAT) protein HCF107 underlie its effects on gene expression. *Proc. Natl. Acad. Sci.* 109, 5651–5656. doi: 10.1073/pnas.1200318109
- Harris, E. H. (2009). *The Chlamydomonas Sourcebook: Introduction to Chlamydomonas and Its Laboratory Uses*. 2nd Edn. San Diego, CA: Academic Press.
- Heinrich, M., Kim, R. G., Wittkopp, T. M., Yang, W., Walters, K. A., Herbert, S. K., et al. (2016). Tetratricopeptide repeat protein protects photosystem I from oxidative disruption during assembly. *Proc. Natl. Acad. Sci.* 113, 2774–2779. doi: 10.1073/pnas.1524040113
- Hernández-Verdeja, T., and Strand, A. (2018). Retrograde signals navigate the path to chloroplast development. *Plant Physiol.* 176, 967–976. doi: 10.1104/pp.17.01299
- Herrin, D. L. (2009). “Chapter 27 – chloroplast RNA processing and stability,” in *The Chlamydomonas Sourcebook*. eds. H. Elizabeth, S. David and W. George (San Diego, CA), 937–965.
- Herrin, D. L., and Schmidt, G. W. (1988). Trans-splicing of transcripts for the chloroplast *psaA1* gene. In vivo requirement for nuclear gene products. *J. Biol. Chem.* 263, 14601–14604.
- Hillebrand, A., Matz, J. M., Almendinger, M., Müller, K., Matuschewski, K., and Schmitz-Linneweber, C. (2018). Identification of clustered organellar short (cos) RNAs and of a conserved family of organellar RNA-binding proteins, the heptatricopeptide repeat proteins, in the malaria parasite. *Nucleic Acids Res.* 46, 10417–10431. doi: 10.1093/nar/gky710
- Hollin, T., Jaroszewski, L., Stajich, J. E., Godzik, A., and Le Roch, K. G. (2021). Identification and phylogenetic analysis of RNA binding domain abundant in apicomplexans or RAP proteins. *Microb. Genom.* 7:mgen000541. doi: 10.1099/mgen.0.000541

- Hou, Y., Zhao, L., Kubo, T., Cheng, X., McNeill, N., Oda, T., et al. (2021). Chlamydomonas FAP70 is a component of the previously uncharacterized ciliary central apparatus projection C2a. *J. Cell Sci.* 134:JCS258540. doi: 10.1242/jcs.258540
- Imai, T., Yagi, Y., and Nakamura, T. (2018). "Recent progress toward RNA manipulation with engineered pentatricopeptide repeat proteins," in *Applied RNA Bioscience*. eds. S. Masuda and S. Izawa (Singapore: Springer), 151–160.
- Jalal, A., Schwarz, C., Schmitz-Linneweber, C., Vallon, O., Nickelsen, J., and Bohne, A. V. (2015). A small multifunctional pentatricopeptide repeat protein in the chloroplast of *Chlamydomonas reinhardtii*. *Mol. Plant* 8, 412–426. doi: 10.1016/j.molp.2014.11.019
- Jarvis, P., and López-Juez, E. (2013). Biogenesis and homeostasis of chloroplasts and other plastids. *Nat. Rev. Mol. Cell Biol.* 14, 787–802. doi: 10.1038/nrm3702
- Johnson, X. (2011). Manipulating RuBisCO accumulation in the green alga, *Chlamydomonas reinhardtii*. *Plant Mol. Biol.* 76, 397–405. doi: 10.1007/s11103-011-9783-z
- Johnson, X., Wostrikoff, K., Finazzi, G., Kuras, R., Schwarz, C., Bujaldon, S., et al. (2010). MRL1, a conserved pentatricopeptide repeat protein, is required for stabilization of *rbcl* mRNA in *Chlamydomonas* and *Arabidopsis*. *Plant Cell* 22, 234–248. doi: 10.1105/tpc.109.066266
- Jones, C. S., Luong, T., Hannon, M., Tran, M., Gregory, J. A., Shen, Z., et al. (2013). Heterologous expression of the C-terminal antigenic domain of the malaria vaccine candidate Pfs48/45 in the green algae *Chlamydomonas reinhardtii*. *Appl. Microbiol. Biotechnol.* 97, 1987–1995. doi: 10.1007/s00253-012-4071-7
- Kalanon, M., and McFadden, G. I. (2008). The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*: A bioinformatic comparison of Toc and Tic components in plants, green algae, and red algae. *Genetics* 179, 95–112. doi: 10.1534/genetics.107.085704
- Kang, S., Odom, O. W., Thangamani, S., and Herrin, D. L. (2017). Toward mosquito control with a green alga: expression of cry toxins of *Bacillus thuringiensis* subsp. *israelensis* (Bti) in the chloroplast of *Chlamydomonas*. *J. Appl. Phycol.* 29, 1377–1389. doi: 10.1007/s10811-016-1008-z
- Kleinknecht, L., Wang, F., Stübe, R., Philippar, K., Nickelsen, J., and Bohne, A. (2014). RAP, the sole octatricopeptide repeat protein in *Arabidopsis*, is required for chloroplast 16S rRNA maturation. *Plant Cell* 26, 777–787. doi: 10.1105/tpc.114.122853
- Krupinska, K., Braun, S., Saeid, M., Anke, N., Götz, S., and Wolfgang, H. (2019). The nucleoid-associated protein WHIRLY1 is required for the coordinate assembly of plastid and nucleus-encoded proteins during chloroplast development. *Planta* 249, 1337–1347. doi: 10.1007/s00425-018-03085-z
- Kuchka, M. R., Goldschmidt-Clermont, M., van Dillewijn, J., and Rochaix, J. D. (1989). Mutation at the *Chlamydomonas* nuclear NAC2 locus specifically affects stability of the chloroplast *psbD* transcript encoding polypeptide D2 of PS II. *Cell* 58, 869–876. doi: 10.1016/0092-8674(89)90939-2
- Kück, U., Choquet, Y., Schneider, M., Dron, M., and Bennis, P. (1987). Structural and transcription analysis of two homologous genes for the P700 chlorophyll a-apoproteins in *Chlamydomonas reinhardtii*: evidence for *in vivo* trans-splicing. *EMBO J.* 6, 2185–2195. doi: 10.1002/j.1460-2075.1987.tb02489.x
- Kück, U., and Schmitt, O. (2021). The chloroplast trans-splicing RNA-protein supercomplex from the green alga *Chlamydomonas reinhardtii*. *Cell* 10:290. doi: 10.3390/cells10020290
- Lefebvre-Legendre, L., Choquet, Y., Kuras, R., Loubéry, S., Douchi, D., and Goldschmidt-Clermont, M. (2015). A nucleus-encoded chloroplast protein regulated by iron availability governs expression of the photosystem I subunit PsaA in *Chlamydomonas reinhardtii*. *Plant Physiol.* 167, 1527–1540. doi: 10.1104/pp.114.253906
- Lefebvre-Legendre, L., Reifschneider, O., Kollipara, L., Sickmann, A., Wolters, D., Kück, U., et al. (2016). A pioneer protein is part of a large complex involved in trans-splicing of a group II intron in the chloroplast of *Chlamydomonas reinhardtii*. *Plant J.* 85, 57–69. doi: 10.1111/tpj.13089
- Legen, J., Ruf, S., Kroop, X., Wang, G., Barkan, A., Bock, R., et al. (2018). Stabilization and translation of synthetic operon-derived mRNAs in chloroplasts by sequences representing PPR protein binding sites. *Plant J.* 94, 8–21. doi: 10.1111/tpj.13863
- Loiselay, C., Gumpel, N. J., Girard-Bascou, J., Watson, A. T., Purton, S., Wollman, F.-A., et al. (2008). Molecular identification and function of *cis*- and *trans*-acting determinants for *petA* transcript stability in *Chlamydomonas reinhardtii* chloroplasts. *Mol. Cell Biol.* 28, 5529–5542. doi: 10.1128/MCB.02056-07
- Loizeau, K., Qu, Y., Depp, S., Fiechter, V., Ruwe, H., Lefebvre-Legendre, L., et al. (2014). Small RNAs reveal two target sites of the RNA-maturation factor Mbb1 in the chloroplast of *Chlamydomonas*. *Nucleic Acids Res.* 42, 3286–3297. doi: 10.1093/nar/gkt1272
- Lu, Y., Rijzaani, H., Karcher, D., Ruf, S., and Bock, R. (2013). Efficient metabolic pathway engineering in transgenic tobacco and tomato plastids with synthetic multigene operons. *Proc. Natl. Acad. Sci.* 110, E623–E632. doi: 10.1073/pnas.1216898110
- Lurin, C., Andrés, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyère, C., et al. (2004). Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* 16, 2089–2103. doi: 10.1105/tpc.104.022236
- Lyska, D., Meierhoff, K., and Westhoff, P. (2013). How to build functional thylakoid membranes: From plastid transcription to protein complex assembly. *Planta* 237, 413–428. doi: 10.1007/s00425-012-1752-5
- Macedo-Osorio, K. S., Pérez-España, V. H., Garibay-Orijel, C., Daniel, G.-Z., Durán-Figueroa, N. V., and Badillo-Corona, J. A. (2018). Intercistronic expression elements (IEE) from the chloroplast of *Chlamydomonas reinhardtii* can be used for the expression of foreign genes in synthetic operons. *Plant Mol. Biol.* 98, 303–317. doi: 10.1007/s11103-018-0776-z
- Manavski, N., Schmid, L. M., and Meurer, J. (2018). RNA-stabilization factors in chloroplasts of vascular plants. *Essays Biochem.* 62, 51–64. doi: 10.1042/EBC20170061
- Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., et al. (2002). Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci.* 99, 12246–12251. doi: 10.1073/pnas.182432999
- Marx, C., Wünsch, C., and Kück, U. (2015). The octatricopeptide repeat protein RAA8 is required for chloroplast trans-splicing. *Eukaryot. Cell* 14, 998–1005. doi: 10.1128/EC.00096-15
- Maul, J. E., Lilly, J. W., Cui, L., Claude, W., Miller, W., Harris, E. H., et al. (2002). The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats. *Plant Cell* 14, 2659–2679. doi: 10.1105/tpc.006155
- Mayfield, S. P., Franklin, S. E., and Lerner, R. A. (2003). Expression and assembly of a fully active antibody in algae. *Proc. Natl. Acad. Sci.* 100, 438–442. doi: 10.1073/pnas.0237108100
- McDermott, J. J., Watkins, K. P., Williams-Carrier, R., and Barkan, A. (2019). Ribonucleoprotein capture by *in vivo* expression of a designer pentatricopeptide repeat protein. *Plant Cell* 31, 1723–1733. doi: 10.1105/tpc.19.00177
- Merendino, L., Perron, K., Rahire, M., Howald, I., Rochaix, J. D., and Goldschmidt-Clermont, M. (2006). A novel multifunctional factor involved in trans-splicing of chloroplast introns in *Chlamydomonas*. *Nucleic Acids Res.* 34, 262–274. doi: 10.1093/nar/gkj429
- Michelet, L., Lefebvre-Legendre, L., Burr, S. E., Rochaix, J.-D., and Goldschmidt-Clermont, M. (2011). Enhanced chloroplast transgene expression in a nuclear mutant of *Chlamydomonas*. *Plant Biotechnol. J.* 9, 565–574. doi: 10.1111/j.1467-7652.2010.00564.x
- Monod, C., Goldschmidt-Clermont, M., and Rochaix, J.-D. (1992). Accumulation of chloroplast *psbB* RNA requires a nuclear factor in *Chlamydomonas reinhardtii*. *Mol. Gen. Genet.* 231, 449–459. doi: 10.1007/BF00292715
- Moustafa, A., and Bhattacharya, D. (2008). PhyloSort: A user-friendly phylogenetic sorting tool and its application to estimating the cyanobacterial contribution to the nuclear genome of *Chlamydomonas*. *BMC Evol. Biol.* 8:6. doi: 10.1186/1471-2148-8-6
- Murakami, S., Kuehne, K., and Stern, D. B. (2005). A spontaneous tRNA suppressor of a mutation in the *Chlamydomonas reinhardtii* nuclear MCD1 gene required for stability of the chloroplast *petD* mRNA. *Nucleic Acids Res.* 33, 3372–3380. doi: 10.1093/nar/gki651
- Nakai, M. (2018). New perspectives on chloroplast protein import. *Plant Cell Physiol.* 59, 1111–1119. doi: 10.1093/pcp/pcy083
- Naver, H., Boudreau, E., and Rochaix, J. (2001). Functional studies of Ycf3: its role in assembly of photosystem I and interactions with some of its subunits. *Plant Cell* 13, 2731–2745. doi: 10.1105/tpc.010253
- Nellaepalli, S., Kim, R. G., Grossman, A. R., and Takahashi, Y. (2021). Interplay of four auxiliary factors is required for the assembly of photosystem I reaction center subcomplex. *Plant J.* 106, 1075–1086. doi: 10.1111/tpj.15220

- Nellaepalli, S., Ozawa, S.-I., Kuroda, H., and Takahashi, Y. (2018). The photosystem I assembly apparatus consisting of Ycf3-Y3IP1 and Ycf4 modules. *Nat. Commun.* 9:2439. doi: 10.1038/s41467-018-04823-3
- Oberleitner, L., Poschmann, G., Macorano, L., Schott-Verdugo, S., Gohlke, H., Stühler, K., et al. (2020). The puzzle of metabolite exchange and identification of putative octotricopeptide repeat expression regulators in the nascent photosynthetic organelles of *Paulinella chromatophora*. *Front. Microbiol.* 11:607182. doi: 10.3389/fmicb.2020.607182
- Ozawa, S. I., Cavaiuolo, M., Jarrige, D., Kuras, R., Rutgers, M., Eberhard, S., et al. (2020). The OPR protein MTH1 controls the expression of two different subunits of ATP synthase CFo in *Chlamydomonas reinhardtii*. *Plant Cell* 32, 1179–1203. doi: 10.1105/tpc.19.00770
- Park, S., Khamai, P., Garcia-Cerdan, J. G., and Melis, A. (2007). REP27, a Tetratricopeptide repeat nuclear-encoded and chloroplast-localized protein, functions in D1/32-kD reaction center protein turnover and photosystem II repair from photodamage. *Plant Physiol.* 143, 1547–1560. doi: 10.1104/pp.107.096396
- Pazour, G. J., Dickert, B. L., Vucica, Y., Seeley, E. S., Rosenbaum, J. L., Witman, G. B., et al. (2000). Chlamydomonas IFT 88 and its mouse homologue, polycystic kidney disease gene *Tg737*, are required for assembly of cilia and flagella. *J. Cell Biol.* 151, 709–718. doi: 10.1083/jcb.151.3.709
- Pfalz, J., Bayraktar, O. A., Prikryl, J., and Barkan, A. (2009). Site-specific binding of a PPR protein defines and stabilizes 5' and 3' mRNA termini in chloroplasts. *EMBO J.* 28, 2042–2052. doi: 10.1038/emboj.2009.121
- Piñas-Fernández, A., and Strand, A. (2018). Signaling between the organelles and the nucleus. *Annu. Plant Rev.* 33, 307–335. doi: 10.1002/9781119312994.apr0357
- Pogson, B. J., Woo, N., Fo, B., and Small, I. (2008). Plastid signalling to the nucleus and beyond. *Trends Plant Sci.* 13, 602–609. doi: 10.1016/j.tplants.2008.08.008
- Preker, P. J., and Keller, W. (1998). The HAT helix, a repetitive motif implicated in RNA processing. *Trends Biochem. Sci.* 23, 15–16. doi: 10.1016/S0968-0004(97)01156-0
- Prikryl, J., Rojas, M., Schuster, G., and Barkan, A. (2011). Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. *Proc. Natl. Acad. Sci.* 108, 415–420. doi: 10.1073/pnas.1012076108
- Rahim, M., Vigneault, F., and Zerges, W. (2016). The RNA structure of cis-acting translational elements of the chloroplast *psbC* mRNA in *Chlamydomonas reinhardtii*. *Front. Plant Sci.* 7:828. doi: 10.3389/fpls.2016.00828
- Rahire, M., Laroche, F., Cerutti, L., and Rochaix, J. D. (2012). Identification of an OPR protein involved in the translation initiation of the PsaB subunit of photosystem I. *Plant J.* 72, 652–661. doi: 10.1111/j.1365-313X.2012.05111.x
- Rasala, B. A., Muto, M., Lee, P. A., Jager, M., Cardoso, R. M., Behnke, A., et al. (2010). Production of therapeutic proteins in algae, analysis of expression of seven human proteins in the chloroplast of *Chlamydomonas reinhardtii*. *Plant Biotechnol. J.* 8, 719–733. doi: 10.1111/j.1467-7652.2010.00503.x
- Raynaud, C., Loiselay, C., Wostrikoff, K., Kuras, R., Girard-bascou, J., and Choquet, Y. (2007). Evidence for regulatory function of nucleus-encoded factors on mRNA stabilization and translation in the chloroplast. *Proc. Natl. Acad. Sci.* 104, 9093–9098. doi: 10.1073/pnas.0703162104
- Reifschneider, O., Marx, C., Jacobs, J., Kollipara, L., Sickmann, A., Wolters, D., et al. (2016). A Ribonucleoprotein supercomplex involved in trans-splicing of organelle group II introns. *J. Biol. Chem.* 291, 23330–23342. doi: 10.1074/jbc.M116.750570
- Rivier, C., Goldschmidt-Clermont, M., and Rochaix, J. D. (2001). Identification of an RNA-protein complex involved in chloroplast group II intron trans-splicing in *Chlamydomonas reinhardtii*. *EMBO J.* 20, 1765–1773. doi: 10.1093/emboj/20.7.1765
- Rochaix, J. D., Perron, K., Dauvillée, D., Laroche, F., Takahashi, Y., and Goldschmidt-Clermont, M. (2004). Post-transcriptional steps involved in the assembly of photosystem I in *Chlamydomonas*. *Biochem. Soc. Trans.* 32, 567–570. doi: 10.1042/BST0320567
- Rochaix, J., and Ramundo, S. (2017). Chloroplast signaling and quality control. *Essays Biochem.* 62, 13–20. doi: 10.1042/EBC20170048
- Rojas, M., Yu, Q., Williams-Carrier, R., Maliga, P., and Barkan, A. (2019). Engineered PPR proteins as inducible switches to activate the expression of chloroplast transgenes. *Nat. Plants* 5, 501–511. doi: 10.1038/s41477-019-0412-1
- Rovira, A. G., and Smith, A. G. (2019). PPR proteins – orchestrators of organelle RNA metabolism. *Physiol. Plant.* 166, 451–459. doi: 10.1111/ppl.12950
- Sadali, N. M., Sowden, R. G., Ling, Q., and Jarvis, R. P. (2019). Differentiation of chromoplasts and other plastids in plants. *Plant Cell Rep.* 38, 803–818. doi: 10.1007/s00299-019-02420-2
- Salomé, P. A., and Merchant, S. S. (2019). A series of fortunate events: introducing Chlamydomonas as a reference organism. *Plant Cell* 31, 1682–1707. doi: 10.1105/tpc.18.00952
- Sane, A. P., Stein, B., and Westhoff, P. (2005). The nuclear gene HCF107 encodes a membrane-associated R-TPR (RNA tetratricopeptide repeat)-containing protein involved in expression of the plastidial *psbH* gene in Arabidopsis. *Plant J.* 42, 720–730. doi: 10.1111/j.1365-313X.2005.02409.x
- Schwarz, C., Bohne, A. V., Wang, F., Cejudo, F. J., and Nickelsen, J. (2012). An intermolecular disulfide-based light switch for chloroplast *psbD* gene expression in *Chlamydomonas reinhardtii*. *Plant J.* 72, 378–389. doi: 10.1111/j.1365-313X.2012.05083.x
- Schwarz, C., Elles, I., Kortmann, J., Piotrowski, M., and Nickelsen, J. (2007). Synthesis of the D2 protein of photosystem II in *Chlamydomonas* is controlled by a high molecular mass complex containing the RNA stabilization factor Nac2 and the translational activator RBP40. *Plant Cell* 19, 3627–3639. doi: 10.1105/tpc.107.051722
- Shamoto, N., Narita, K., Kubo, T., Oda, T., and Takeda, S. (2018). CFAP70 is a novel axoneme-binding protein that localizes at the base of the outer dynein arm and regulates ciliary motility. *Cell* 7:124. doi: 10.3390/cells7090124
- Shamriz, S., and Ofoghi, H. (2017). Outlook in the application of *Chlamydomonas reinhardtii* chloroplast as a platform for recombinant protein production. *Biotechnol. Genet. Eng. Rev.* 32, 92–106. doi: 10.1080/02648725.2017.1307673
- Shen, C., Zhang, D., Guan, Z., Liu, Y., Yang, Z., Yang, Y., et al. (2016). Structural basis for specific single-stranded RNA recognition by designer pentatricopeptide repeat proteins. *Nat. Commun.* 7:11285. doi: 10.1038/ncomms11285
- Shikanai, T., and Fujii, S. (2013). Function of PPR proteins in plastid gene expression. *RNA Biol.* 10, 1446–1456. doi: 10.4161/rna.25207
- Small, I. D., and Peeters, N. (2000). The PPR motif – a TPR-related motif prevalent in plants organellar proteins. *Trends Biochem. Sci.* 25, 45–47. doi: 10.1016/S0968-0004(99)01520-0
- Snell, W. J., and Goodenough, U. W. (2009). “Chapter 12 – Flagellar adhesion, flagellar-generated signalling, and gamete fusion during mating,” in *The Chlamydomonas Sourcebook*. eds. H. Elizabeth, S. David and W. George (San Diego, CA), 369–394.
- Stadnichuk, I. N., and Kusnetsov, V. V. (2021). Endosymbiotic origin of chloroplasts in plant cells' evolution. *Russ. J. Plant Physiol.* 68, 1–16. doi: 10.1134/S1021443721010179
- Stampacchia, O., Girard-Bascou, J., Zanasco, J. L., Zerges, W., Bennoun, P., and Rochaix, J. D. (1997). A nuclear-encoded function essential for translation of the chloroplast *psaB* mRNA in *Chlamydomonas*. *Plant Cell* 9, 773–782. doi: 10.1105/tpc.9.5.773
- Stegemann, S., and Bock, R. (2006). Experimental reconstruction of functional gene transfer from the tobacco plastid genome to the nucleus. *Plant Cell* 18, 2869–2878. doi: 10.1105/tpc.106.046466
- Stoppel, R., and Meurer, J. (2012). The cutting crew – ribonucleases are key players in the control of plastid gene expression. *J. Exp. Bot.* 63, 1663–1673. doi: 10.1093/jxb/err401
- Sun, A. Z., and Guo, F. Q. (2016). Chloroplast retrograde regulation of heat stress responses in plants. *Front. Plant Sci.* 7:398. doi: 10.3389/fpls.2016.00398
- Terry, M. J., and Bampton, J. (2019). “The role of tetrapyrroles in chloroplast-to-nucleus retrograde signaling,” in *Metabolism, Structure and Function of Plant Tetrapyrroles: Control Mechanisms of Chlorophyll Biosynthesis and Analysis of Chlorophyll-Binding Proteins*. ed. B. Grimm (San Diego, CA: Advances in Botanical Research, Elsevier Ltd), 225–246.
- Tourasse, N. J., Choquet, Y., and Vallon, O. (2013). PPR proteins of green algae. *RNA Biol.* 10, 1526–1542. doi: 10.4161/rna.26127
- Tran, M., Henry, R., Siefker, D. T., Van, C., Newkirk, G., Kim, J., et al. (2013b). Production of anti-cancer immunotoxins in algae: ribosome inactivating proteins fusion partners. *Biotechnol. Bioeng.* 110, 2826–2835. doi: 10.1002/bit.24966
- Tran, M., Van, C., Barrera, D. J., Pettersson, P. L., Peinado, C., Bui, J., et al. (2013a). Production of unique immunotoxin cancer therapeutics in algal chloroplasts. *Proc. Natl. Acad. Sci.* 110, E15–E22. doi: 10.1073/pnas.1214638110
- Tran, M., Zhou, B., Pettersson, P. L., Gonzalez, M. J., and Mayfield, S. P. (2009). Synthesis and assembly of a full-length human monoclonal antibody in algal chloroplasts. *Biotechnol. Bioeng.* 104, 663–673. doi: 10.1002/bit.22446

- Vaistij, F. E., Boudreau, E., Lemaire, S. D., Goldschmidt-Clermont, M., and Rochaix, J. D. (2000a). Characterization of Mbb1, a nucleus-encoded tetratricopeptide-like repeat protein required for expression of the chloroplast *psbB/psbT/psbH* gene cluster in *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci.* 97, 14813–14818. doi: 10.1073/pnas.97.26.14813
- Vaistij, F. E., Goldschmidt-Clermont, M., Wostrikoff, K., and Rochaix, J. D. (2000b). Stability determinants in the chloroplast *psbB/T/H* mRNAs of *Chlamydomonas reinhardtii*. *Plant J.* 21, 469–482. doi: 10.1046/j.1365-313x.2000.00700.x
- Viola, S., Cavaiuolo, M., Drapier, D., Eberhard, S., Vallon, O., Wollman, F. A., et al. (2019). MDA1, a nucleus-encoded factor involved in the stabilization and processing of the *atpA* transcript in the chloroplast of *Chlamydomonas*. *Plant J.* 98, 1033–1047. doi: 10.1111/tbj.14300
- Wang, F., Johnson, X., Cavaiuolo, M., Bohne, A. V., Nickelsen, J., and Vallon, O. (2015). Two *Chlamydomonas* OPR proteins stabilize chloroplast mRNAs encoding small subunits of photosystem II and cytochrome *b<sub>6</sub>*. *Plant J.* 82, 861–873. doi: 10.1111/tbj.12858
- Wannathong, T., Waterhouse, J. C., Young, R. E., Economou, C. K., and Purton, S. (2016). New tools for chloroplast genetic engineering allow the synthesis of human growth hormone in the green alga *Chlamydomonas reinhardtii*. *Appl. Microbiol. Biotechnol.* 100, 5467–5477. doi: 10.1007/s00253-016-7354-6
- Wei, L., Derrien, B., Gautier, A., Houille-Vernes, L., Boulouis, A., Saint-Marcoux, D., et al. (2014). Nitric oxide-triggered remodeling of chloroplast bioenergetics and thylakoid proteins upon nitrogen starvation in *Chlamydomonas reinhardtii*. *Plant Cell* 26, 353–372. doi: 10.1105/tpc.113.120121
- Woodson, J. D., and Chory, J. (2008). Coordination of gene expression between organellar and nuclear genomes. *Nat. Rev. Genet.* 9, 383–395. doi: 10.1038/nrg2348
- Wostrikoff, K., Girard-Bascou, J., Wollman, F. A., and Choquet, Y. (2004). Biogenesis of PSI involves a cascade of translational autoregulation in the chloroplast of *Chlamydomonas*. *EMBO J.* 23, 2696–2705. doi: 10.1038/sj.emboj.7600266
- Yagi, Y., Hayashi, S., Kobayashi, K., Hirayama, T., and Nakamura, T. (2013). Elucidation of the RNA recognition code for Pentatricopeptide repeat proteins involved in organelle RNA editing in plants. *PLoS One* 8:e57286. doi: 10.1371/journal.pone.0057286
- Yagi, Y., Nakamura, T., and Small, I. (2014). The potential for manipulating RNA with pentatricopeptide repeat proteins. *Plant J.* 78, 772–782.
- Yagi, Y., and Shiina, T. (2014). Recent advances in the study of chloroplast gene expression and its evolution. *Front. Plant Sci.* 5:61. doi: 10.3389/fpls.2014.00061
- Yan, J., Yao, Y., Hong, S., Yang, Y., Shen, C., Zhang, Q., et al. (2019). Delineation of pentatricopeptide repeat codes for target RNA prediction. *Nucleic Acids Res.* 47, 3728–3738. doi: 10.1093/nar/gkz075
- Yin, P., Li, Q., Yan, C., Liu, Y., Liu, J., Yu, F., et al. (2013). Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* 504, 168–171. doi: 10.1038/nature12651
- Yoo, C., Han, S., and Chen, M. (2020). Nucleus-to-plastid phytochrome signalling in controlling chloroplast biogenesis. *Annu. Plant Rev.* 3, 251–280. doi: 10.1002/9781119312994.apr0615
- Yoon, H. S., Hackett, J. D., Ciniglia, C., Pinto, G., and Bhattacharya, D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* 21, 809–818. doi: 10.1093/molbev/msh075
- Zerges, W., Auchincloss, A. H., and Rochaix, J. D. (2003). Multiple translational control sequences in the 5' leader of the chloroplast *psbC* mRNA interact with nuclear gene products in *Chlamydomonas reinhardtii*. *Genetics* 163, 895–904. doi: 10.1093/genetics/163.3.895
- Zhou, F., Karcher, D., and Bock, R. (2007). Identification of a plastid intercistronic expression element (IEE) facilitating the expression of stable translatable monocistronic mRNAs from operons. *Plant J.* 52, 961–972. doi: 10.1111/j.1365-313X.2007.03261.x
- Zoschke, R., and Bock, R. (2018). Chloroplast translation: structural and functional organization, operational control and regulation. *Plant J.* 30, 745–770. doi: 10.1105/tpc.18.00016

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Macedo-Osorio, Martínez-Antonio and Badillo-Corona. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.