The Octatricopeptide Repeat Protein Raa8 Is Required for Chloroplast trans Splicing

Christina Marx, Christiane Wünsch, Ulrich Kück
Lehrstuhl für Allgemeine und Molekulare Botanik, Ruhr-University Bochum, Bochum, Germany

The mRNA maturation of the tripartite chloroplast psaA gene from the green alga *Chlamydomonas reinhardtii* depends on various nucleus-encoded factors that participate in *trans* splicing of two group II introns. Recently, a multiprotein complex was identified that is involved in processing the *psaA* precursor mRNA. Using coupled tandem affinity purification (TAP) and mass spectrometry analyses with the *trans*-splicing factor Raa4 as a bait protein, we recently identified a multisubunit ribonucleoprotein (RNP) complex comprising the previously characterized *trans*-splicing factors Raa1, Raa3, Raa4, and Rat2 plus novel components. Raa1 and Rat2 share a structural motif, an octatricopeptide repeat (OPR), that presumably functions as an RNA interaction module. Two of the novel RNP complex components also exhibit a predicted OPR motif and were therefore considered potential *trans*-splicing factors. In this study, we selected bacterial artificial chromosome (BAC) clones encoding these OPR proteins and conducted functional complementation assays using previously generated *trans*-splicing mutants. Our assay revealed that the *trans*-splicing defect of mutant F19 was restored by a new factor we named *RAA8*; molecular characterization of complemented strains verified that Raa8 participates in splicing of the first *psaA* group II intron. Three of six OPR motifs are located in the C-terminal end of Raa8, which was shown to be essential for restoring *psaA* mRNA *trans* splicing. Our results support the important role played by OPR proteins in chloroplast RNA metabolism and also demonstrate that combining TAP and mass spectrometry with functional complementation studies represents a vigorous tool for identifying *trans*-splicing factors.
translation of chloroplast psaA mRNA are functions that are also mediated by the OPR protein Taal (translation of psaA mRNA), which regulates psaA translation in response to iron limitation (17).

Other OPR-containing proteins are required for the maturation of chloroplast psaA mRNA. The psaA gene encodes the P700 chlorophyll α-apoprotein of photosystem I and comprises three exons that are scattered around the chloroplast genome (18). Each exon is flanked by group II intron sequences, and thus two trans-splicing reactions are necessary to generate the mature psaA mRNA (19). Remarkably, the first group II intron is tripartite, comprising the two psaA primary transcripts of exons 1 and 2 as well as an additional small chloroplast RNA (tscA RNA) (20). Processing of the tscA RNA is a prerequisite for correct assembly and consecutive trans splicing of the first group II intron.

So far, two OPR proteins have been described that are required for psaA trans splicing. Rat2 (RNA maturation of psaA tscA RNA) and Raa1 (RNA maturation of psaA RNA) exhibit two and four OPR motifs, respectively (10, 14, 15). Mutants with defects in genes such as RAT2 or RAA1 not only are deficient in trans splicing but also have a defect in photosynthesis, since they lack the P700 chlorophyll α-apoprotein. Genetic analysis of photosystem I mutants revealed that at least 14 nucleus-encoded factors are required for psaA mRNA maturation (18, 21, 22).

Recently, we were able to demonstrate that both Raa1 and Rat2 are part of a chloroplast RNP complex involved in the first trans-splicing reaction (13). This complex was identified by an extensive proteomic analysis and contains four trans-splicing factors together with 19 as yet uncharacterized subunits. Of these 19 novel proteins, we identified two further OPR proteins that also might represent trans-splicing factors required for group II intron splicing (10, 12, 13).

Previously, trans-splicing factors were identified by analysis of mutants generated by conventional or insertional mutagenesis (22). For example, we identified mutant genes by ligation-mediated suppression (LMS) PCR and genome walking in combination with functional complementation analysis using bacterial artificial chromosome (BAC) clones (23). As an alternative approach to discover novel trans-splicing factors, we present here a complementation analysis of four as yet uncharacterized trans-splicing mutants (21, 24). The mutants were transformed with BAC clones carrying genes for the above-mentioned OPR proteins, which were previously detected solely by tandem affinity purification (TAP) followed by mass spectrometry. Evidence is provided that a 269-kDa protein carrying six copies of the OPR motif is required for trans splicing of the first psaA group II intron. Importantly, our results extend the list of OPR proteins from C. reinhardtii involved in splicing and processing of chloroplast RNAs.

MATERIALS AND METHODS

Strains and growth conditions. C. reinhardtii wild-type strain 137C and previously described trans-splicing mutants F19, F15, M10, and F31 (21, 24) were grown in liquid Tris-acetate-phosphate or on Tris-acetate-phosphate agar plates (25). For photoautotrophic growth, test strains were spread onto HS (high-salt minimal medium) or Tris-acetate-phosphate agar plates (25). To generate autolysin, we used the high-efficiency mating strains CC-621 and CC-622 (25).

Transformation. Nuclear transformation was achieved by agitation with glass beads (26). To remove cell walls, cells were treated for at least 1 h with autolysin before transformation. Each transformation experiment was carried out at least three times, with 0.5 to 10 μg BAC DNA, and samples were spread on single Tris-acetate-phosphate agar plates. After 12 to 24 h of incubation under low light, plates were selected under high-light conditions for 2 to 3 weeks.

Construction of plasmids. Recombinant plasmids and oligonucleotides used for PCR and reverse transcription–PCR (RT-PCR) experiments or to generate transgenic algal strains are listed in Tables S1 and S2 in the supplemental material. BAC clones from the BAC library CRCCB8a were obtained from the Clemson University Genomics Institute (CUGI). After plasmid isolation, BAC DNA was verified by restriction enzyme hydrolysis and partial sequencing.

Derivatives of BAC 31O13 were generated to identify regions sufficient for rescue of mutant F19 by partial restriction of 31O13 with Sgfl (31O13-1 and 31O13-2) or BlnI (31O13-3 and 31O13-4). For subcloning of RAA8, 31O13 was first digested with Sful and Clal, resulting in plasmid 31O13-5. In the second step, 31O13-5 was partially digested with Milu (31O13-6). 31O13-6 comprises only RAA8 and Cre10.g4400050. To generate a plasmid that contains solely RAA8, 31O13-6 was partially digested with PvuI, resulting in 31O13-7.

Molecular genetic techniques. Standard molecular techniques were used as reported elsewhere (13, 23). C. reinhardtii total RNA and RNA blot experiments were prepared and performed, respectively, as described previously (15, 18, 23). After transfer of RNA samples to nylon membranes, hybridization was performed with a radioactively labeled probe. For isolation of genomic DNA, algal cells were treated as described elsewhere (27).

cDNA synthesis was conducted with the SuperScript III first-strand synthesis system for RT-PCR (Invitrogen) and purified using Amicon columns (Millipore). RT-PCR and PCR were performed using Taq DNA polymerase (Eppendorf). Phusion high-fidelity DNA polymerase (Thermo Scientific), GoTaq DNA polymerase (Promega), and FastStart Tag DNA polymerase (Roche).

For sequencing of the RAA8 gene locus, PCR fragments were directly sequenced or cloned into pDrive (Qiagen). For verification of point mutations, two independent PCRs were conducted with different primer combinations, followed by sequencing.

For PsA-A immunoblot analysis, total protein extracts were loaded onto 12% SDS-polyacrylamide gels with 6 M urea. After gel electrophoresis and blotting onto a polyvinylidene difluoride (PVDF) membrane (Roche), the membrane was blocked with 5% nonfat dry milk in washing buffer. After incubation with a rabbit polyclonal PsA antibody (Agrisera) at 4°C overnight, the membrane was decorated with peroxidase-linked anti-rabbit IgG for 1 h. Signals were detected using a Western enhanced chemiluminescence (ECL) substrate (Bio-Rad).

Bioinformatic analyses. Gene models and corresponding protein data were obtained from Phytozone v10.1. Protein motifs and domains were predicted by MEME Suite (28) and Motif Scan (29). Alignments of protein sequences were visualized using ClustalX (30). RNA binding regions were predicted by BindN (31). Secondary structure was analyzed using Phyre2 (32). Chloroplast target peptides were predicted by WoLF PSORT (33), TargetP (34), ChloroP (35), and PredAlgo (36).

RESULTS

Functional complementation analysis of class C mutants by use of BAC DNA. Recently, we identified a chloroplast RNP complex containing trans-splicing factors, intron RNA, and as yet uncharacterized protein subunits (13). This complex was shown to be involved in processing psaA exon 1 and 2 precursors. Two of these trans-splicing factors, Rat2 and Raa1, were described as OPR proteins, and two uncharacterized proteins were considered further OPR proteins of the splicing complex (10, 12). These are Cre17.g698750 and Cre10.g440000, and besides the OPR motifs, both have no conserved domains. To determine whether both OPR proteins are required for trans splicing of psaA mRNA, BAC
clones encoding these uncharacterized OPR components were selected for a complementation approach. BAC clones 13K19 and 31O13 were obtained from the CUGI BAC library and encode Cre17.g698750 and Cre10.g440000, respectively.

The BAC clones were used for complementation studies of photosynthesis mutants. These mutants were previously generated by random mutagenesis and further characterized to detect defects in trans splicing of the psaA mRNA (21, 24). Depending on the splicing defect, mutants can be divided into three distinct classes, namely, A, B, and C. Mutants unable to splice the second group II intron are designated class A, and class B mutants are blocked in both the first and second trans-splicing reactions (21). Class C mutants show a defect in the first trans-splicing reaction, and they either are unable to fuse exons 1 and 2 or, alternatively, have a defect in tscA RNA 3′-end processing. For this study, we chose a number of class C mutants from a type culture collection (ChlamyStation, Paris, France), namely, F19, M10, F31, and F5 (21, 24), for a complementation approach.

The four class C mutants were separately transformed with BAC clones 13K19 and 31O13, and transformants were selected under high-light conditions. Growth under these conditions indicated a restoration of photosystem I activity, and thus a complementation of the psaA trans-splicing defect. Transformation of M10, F5, and F31 with each of the two BACs revealed no transformants, while several transformants were obtained from transformation of F19 with BAC 31O13.

Photoautotrophic growth tests were performed on four selected transformants derived from complementation of F19 with BAC 31O13 (T404-2/6, T401-1/1, T401-2/1, and T401-2/2) (Fig. 1a). Tris-acetate-phosphate agar plates contained acetate as an alternative carbon source, whereas HS medium was used as a minimal medium. Therefore, photosystem I-deficient strains were unable to grow under high-light conditions both on Tris-acetate-phosphate plates and on HS medium without acetate, but they were able to grow under low-light conditions on Tris-acetate-phosphate agar plates (Fig. 1a; see Fig. S1 in the supplemental material). Wild-type strain 137C and mutant strain F19 served as controls for photoautotrophic growth tests.

The four selected transformants were able to grow under high-light conditions on HS and Tris-acetate-phosphate agar plates, thus demonstrating a restored photosystem I activity. For Northern analysis, RNAs were isolated from 137C, F19, and the four complementation strains. A psaA exon 1 probe detected a 2,800-nucleotide (nt) signal representing the mature psaA transcript in the wild-type strain 137C and all transformants, whereas a signal of about 400 nt corresponded to the psaA exon 1 primary transcript in mutant F19 (Fig. 1b). Furthermore, a tscA probe detected the 450-nt processed tscA transcript in all analyzed strains. Rehybridization with an RBSC1 probe served as a loading control. Northern analysis confirmed that mutant F19 has a defect in trans splicing of the first group II intron but not in 3′-end processing of tscA RNA. We previously showed that trans-splicing mutants with a defect in tscA processing generated precursor RNA molecules of 1,700 nt and 2,800 nt that are clearly distinguished from the processed, 450-nt tscA RNA (15, 37). All analyzed complementation strains exhibited restored photosystem I activity and regained the ability to generate mature psaA mRNA.

To further analyze the transformants, we performed immunodetection with an anti-PsaA antibody. As expected, since mutant F19 was unable to generate mature psaA mRNA transcripts, no PsaA protein was detected. In contrast, all tested transformants showed a signal for PsaA (Fig. 1c). Comparison with the wild-type signal revealed slightly reduced PsaA signal intensities for all transformants, presumably as a consequence of ectopic integration, and thus suboptimal gene expression, in the complemented strains.

These analyses revealed that recombinant BAC clone 31O13, carrying the Cre10.g440000 gene and 12 other genes, restored the trans-splicing defect of mutant F19. Since Cre10.g440000 is most
the lack of RAA8 in F19 is responsible for the psaA pre-mRNA splicing defect.

**Photosynthesis mutant F19 carries an in-frame stop codon in RAA8.** Since the BAC derivative 31O13-7 was sufficient to complement mutant F19, the F19 mutation must affect RAA8, the only gene carried by 31O13-7, with a predicted gene locus comprising nine exons (Fig. 3a).

Previously, F19 was generated by 5-fluorouracil mutagenesis, which is usually associated with DNA point mutations (38). To identify mutations responsible for the mutant phenotype, we performed PCR analysis and sequencing of the F19 DNA. By performing PCR analysis and sequencing of the F19 DNA, we detected a point mutation at base pair position 5,537, within exon 3. PCR analysis of wild-type 137C DNA and 31O13-7 DNA, following sequence alignment, revealed that the point mutation is unique to mutant F19 (Fig. 3b). The transversion of guanine to thymine leads to an in-frame stop codon instead of a glutamic acid codon (E1722). Whereas RAA8 encodes a protein of 2,745 residues, the stop codon in F19 leads to premature translation termination, resulting in a truncated protein of 1,720 residues. We concluded that the N-terminal part of Raa8 alone is not sufficient for complementation.

**Raa8 has a molecular mass of 269 kDa and exhibits six putative OPR motifs.** Raa8 has a predicted size of 269 kDa and is the largest trans-splicing factor described so far (Fig. 4a). Peptide counts obtained from mass spectrometry analysis (13) revealed a coverage of about 37.7% of the total Raa8 sequence. Using TargetP, ChloroP, PredAlgo, and WoLF PSORT, a chloroplast localization can be predicted for Raa8 (13, 33–36). In addition to the putative chloroplast signal peptide, MotifScan identified several stretches of low-complexity regions (29). These regions exhibit little diversity regarding their amino acid composition and are common in many eukaryotic proteins (39–41). Raa8 carries an alanine-rich region that covers nearly the entire protein, in addition to serine-, glycine-, glutamine-, and proline-rich stretches (Fig. 4a). The protein exhibits high alanine (24.4%), glycine (11.8%), and leucine (12.0%) contents. Moreover, BindN predicts seven putative RNA binding residues (31). Analysis of Raa8 together with several OPR proteins, such as Tab1, Tda1, Tbc2, and...

![FIG 2](image2.png)

**FIG 2** Complementation analysis of mutant F19. The genomic region inserted into BAC 31O13 is indicated at the top, and subclones are listed below. 31O13 carries 13 genes of chromosome 10, whose names in the figure are abbreviated to the final six characters (e.g., Cre10.g440050 = 440050). BAC vector sequences are indicated in black. Arrows represent the predicted open reading frames of the corresponding genes, whereas dashed lines indicate deleted regions. The red bar indicates the position of the detected point mutation in mutant F19. (b) Sequence alignment of the partial sequence of RAA8. A guanine-to-thymine transversion in F19 leads to a stop codon (*).

![FIG 3](image3.png)

**FIG 3** Map of the RAA8 locus. (a) Exon-intron organization of RAA8. Corresponding 5'- and 3'-untranslated regions (UTRs) at both ends are marked as open boxes. The red bar shows the position of the detected point mutation in mutant F19. (b) Sequence alignment of the partial sequence of RAA8 exon 3. Red shading indicates sequence deviations between the wild type and the mutants. A guanine-to-thymine transversion in F19 leads to a stop codon (*).
trans-splicing factors, by the program MEME Suite identified at least six OPR motifs (28).

Sequence conservation of the OPR motif is restricted to several amino acids, including the mostly conserved first proline and the tryptophan of the degenerate PPPEW pattern (Fig. 4b) (10, 11). Since OPR proteins belong to the α-solenoid superfamily, with an α-helical architecture, we performed a secondary structure prediction of partial Raa8 sequences by using Phyre2 (32). Raa8 contains several paired α-helices, with a high confidence value. Some of these α-helices match the six OPR motifs, and all are located in the C-terminal half of Raa8 (see Fig. S3 in the supplemental material).

Our functional complementation studies revealed that 31O13-4, carrying a truncated version of RAA8 (exon 1 to exon 5), was not able to rescue mutant F19. In silico analysis revealed that 31O13-4 encodes a 192-kDa instead of a 269-kDa protein. This indicates that the genomic region of exon 1 to exon 5 is not sufficient for functional complementation. Importantly, this truncated gene version does not encode the last three C-terminal OPR motifs. We therefore conclude that the complete set of OPR motifs is required for the trans-splicing function of Raa8, and thus for complementation.

DISCUSSION

We demonstrate here that functional complementation studies as a follow-up to a previous TAP and mass spectrometry approach enabled us to identify a gene defect in a previously generated trans-splicing class C mutant. Detection of the mutation in F19 revealed an uncharacterized OPR protein as a trans-splicing factor, which we named Raa8. This protein is part of an RNP complex that was previously identified by TAP and mass spectrometry and is involved in the first trans-splicing reaction of psaA mRNA. Hence, at least five trans-splicing factors, Rat2, Raa1, Raa3, Raa4, and the newly identified Raa8, are part of this complex.

Functional complementation studies using photosynthesis mutants to complement a combined TAP and mass spectrometry approach represent a vigorous tool for identifying trans-splicing factors. Homologous recombination is a prerequisite for generating mutant strains by insertion or deletion of distinct genes. In C. reinhardtii, the frequency of homologous recombination is exceptionally low, and most often, transforming DNA integrates randomly by the nonhomologous end-joining repair pathway (42). Until now, a rapid method to generate targeted mutations with a moderate success rate was still missing.

C. reinhardtii mutants are commonly used as sources for forward genetic approaches to identify genes controlling specific processes. Usually, the screening of mutants relies on a distinct mutant phenotype that allows uncovering of specific gene functions based on complementation analysis. Commonly, insertional mutagenesis by random integration of a selection marker was used to generate mutants with a desired phenotype (42, 43). More recently, an alternative approach used whole-genome sequencing to identify gene mutations, but it was hampered by the mutational profile of the strains analyzed (44). Despite single nucleotide polymorphisms (indels) that lead to a large number of changes to the reference genome, additional chromosomal rearrangements or

FIG 4 The RAA8 gene encodes a 269-kDa protein that exhibits six predicted C-terminal OPR motifs. (a) Scheme of the primary structure of the Raa8 polypeptide and two truncated versions. In silico analyses predicted several low-complexity regions (Gly-rich, Pro-rich, Gln-rich, Ala-rich, and Ser-rich regions), RNA binding residues, a chloroplast transit peptide (cTP), and OPR motifs. Both F19 and 31O13-4 encode truncated proteins. (b) Six OPR motifs in Raa8 were predicted by MEME Suite (28). The alignment was visualized using ClustalX (30), with a default color scheme assigned for each residue that met the amino acid profile of the alignment at that position. The consensus sequence defined by MEME Suite is indicated at the bottom. Numbers in front of the OPR motifs indicate the positions of the first amino acids. Numbers behind the motifs represent the P values defined by MEME Suite, using characterized OPR proteins, such as Raa1 and Rat2.
suppressed recombination impedes the identification of mutations (44, 45).

In this study, we used an alternative approach to identify genes involved in chloroplast trans splicing by using data from extensive TAP and mass spectrometry analysis of an RNP complex (13). This RNP complex is involved in the first trans-splicing reaction and contains four characterized trans-splicing factors and additional, as yet uncharacterized protein subunits. We conducted functional complementation studies using previously generated trans-splicing mutants. More than 30 years ago, these trans-splicing mutants of C. reinhardtii were generated by chemical treatment or UV irradiation (21, 24, 46). Since the long-term propagation of strain collections leads to the accumulation of spontaneous mutations, identification of mutations by whole-genome sequencing is time-consuming and laborious, requiring numerous backcrosses and segregation analyses.

Here we took advantage of previously generated class C mutants and used BAC DNAs encoding either of the two uncharacterized OPR proteins, Cre10.g4400000 (Raa8) and Cre17.g698750, for transformation. Both proteins were already known to be part of the trans-splicing complex (13). Functional complementation studies and further molecular characterization of complemented strains enabled us to identify the novel trans-splicing factor Raa8. We unequivocally demonstrated that Raa8 is involved in trans splicing of the first psaA group II intron but is not crucial for tscA 3′-end processing. Thus, this is the first described OPR protein required exclusively for splicing of group II intron RNA. With the discovery of Raa8, we extend the rather short list of OPR proteins from C. reinhardtii that are involved in RNA splicing and/or processing.

Raa8 is a member of the OPR protein family that functions in chloroplast psaA mRNA processing. Mitochondria and chloroplasts originated from symbiotic incorporation of either alpha-proteobacteria or cyanobacteria (47, 48). Thus, both organelles show diverse similarities to bacteria, such as their circular genome, replication mode, RNA processing steps, and protein transport and translation machinery (49–52).

Endosymbiosis resulted in the reconstruction of genetic information mediated by gene transfer from the ancestral organelle genomes to the nucleus (53). The persistent compartmentation of genetic systems requires cross talk between different retrograde and anterograde signaling pathways to control plastid and mitochondrial biogenesis and their distinct biochemical processes (53–55). Organelle genes are often expressed constitutively and regulated at the posttranscriptional level. Therefore, nucleus-encoded factors are required to control the functionality of polycistronic transcripts by affecting their stability, editing, or maturation (56–58).

In algae and plants, chloroplast RNA metabolism is mostly controlled by PPR and OPR proteins, both of which are members of the α-solenoid superfamily (8, 10, 12, 58, 59). Numerous PPR proteins mediate translation, RNA stability, cleavage, editing, and splicing (8, 60). Whereas PPRs expanded in land plants, the sole OPR protein from Arabidopsis thaliana (61, 62) contains, in addition to four OPR motifs, an RAP (RNA binding domain abundant in apicomplexans) domain that is also present in the C. reinhardtii chloroplast protein Taa1 and the trans-splicing factor Raa3 (17, 63). Rap functions by binding to the 16S rRNA and mediates maturation of 16S rRNA via 5′-end processing (62). Similar to the case for the C. reinhardtii OPR proteins, such as Tda1 and Tbc2, the OPR motifs are localized mainly in the C-terminal half of Rap (10, 11, 62). Interestingly, the C-terminal part of Tda1 was shown to be sufficient to promote atpA mRNA translation. Moreover, complementation studies using truncated versions of both TAB1 and TDA1, as mentioned in the introduction, suggest that the entire set of OPR motifs is required for functionality of the proteins (10, 12).

The striking role of OPR-containing protein regions in chloroplast RNA metabolism is also supported by the results of this work. Similar to the case for other OPR proteins, such as Tda1 and Tbc2, our analysis revealed that the six OPR domains of Raa8 are localized in the C-terminal half of the protein’s primary structure. Characterization of mutant F19 revealed that a stop codon in exon 3 of RAA8 results in a truncated protein. Thus, the mutant encodes a 167-kDa protein lacking five of the six OPR motifs. Moreover, restoration of photosynthesis in F19 failed when a truncated, 192-kDa protein (expressed by derivative 31O13-4) exhibiting only three of the six OPRs was used for complementation. Our results show that the C-terminal region of Raa8 is required for trans-splicing function, suggesting that the missing OPRs in the truncated versions provide essential functions for group II intron splicing.

In vivo, trans splicing of psaA mRNA requires trans-splicing factors, which mostly share structural features such as low-complexity regions and OPR motifs. Accurate splicing of the group II introns is a prerequisite for successful translation of Psaa. Therefore, OPR proteins may adjust the tripartite intron fragments of the first group II intron through nucleotide binding to OPR motifs. The common α-helical structure of PPR and OPR motifs suggests that OPR motifs, similarly to PPR motifs, bind single nucleotides of the target RNA (6, 7).

Besides a direct function of OPR proteins, such as Raa8, in trans splicing, other OPR proteins, such as Raa1 and Rats2, are involved in 3′-end processing of tscA RNA. Presumably, these proteins enhance proper processing of the precursor transcript through sequence-specific binding of the OPR motif. Conversely, proteins without OPR motifs, such as Raa4 and Raa3, may act as chaperones. Binding of Raa4 to RNA was demonstrated previously, thus suggesting that Raa4 promotes folding and stability of the intron RNA (23).

The results of our work expand the roles of OPR proteins in mRNA splicing. The characterization of chloroplast-localized OPR proteins supports their specific functioning in various roles of chloroplast RNA metabolism, which is presumably not restricted to algal organisms but will most probably also be discovered in chloroplasts of higher plants.

ACKNOWLEDGMENTS

This work was funded by the Deutsche Forschungsgemeinschaft (grant KU 517/13-1).

We thank Ingeborg Godehardt for excellent technical assistance and ChlamiStation (France) for providing algal strains. We are grateful to J. Jacobs for helpful comments on the manuscript.

REFERENCES

1. Kohse B, Kajava AV. 2000. When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. Trends Biochem Sci 25:509–515. http://dx.doi.org/10.1016/S0968-0004(00)01667-4.
2. Deng D, Yan C, Wu J, Pan X, Yan N. 2014. Revisiting the TALE repeat. Protein Cell 5:297–306. http://dx.doi.org/10.1007/s13238-014-0035-2.
3. Blatch GL, Lässle M. 1999. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Bioessays 21:932–939.
4. Small ID, Peeters N. 2000. The PPR motif—a TPR-related motif prevalent in plant organelar proteins. Trends Biochem Sci 25:46–47.

5. Manova S. 2015. An overview of pentatricopeptide repeat proteins and their applications. Biochimie 113:93–99. http://dx.doi.org/10.1016/j.biochi.2015.04.004.

6. Li Q, Li Y-C, Liu Y, Liu J, Yu F, Wang Z, Long J, He J, Wang HW, Wang J, Zhu JK, Shi Y, Yan N. 2013. Structural basis for the modular recognition of single-stranded RNA by PPR proteins. Nature 504:168–171. http://dx.doi.org/10.1038/nature12651.

7. Barkan A, Rojas M, Fujii S, Yap A, Chong YS, Bond CS, Small I. 2012. A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. PLoS Genet 8:e1002910. http://dx.doi.org/10.1371/journal.pgen.1002910.

8. Barkan A, Small I. 2014. Pentatricopeptide repeat proteins. Annu Rev Plant Biol 65:415–442. http://dx.doi.org/10.1146/annurev-plant-050213-040159.

9. Tousarre NJ, Choquet Y, Vallon O. 2013. PPR proteins of green algae. RNA Biol 10:1526–1542. http://dx.doi.org/10.4161/rna.26127.

10. Eberhard S, Loiselay C, Drapier D, Bujaldon S, Girard-Bascou J, Kuras R, Choquet Y, Wollman FA. 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. http://dx.doi.org/10.1111/j.1365-313X.2011.04657.x.

11. Auchenium AH, Zerges W, Perron K, Girard-Bascou J, Rochaix JD. 2002. Characterization of TBI2, a nucleus-encoded factor specifically required for transcription of the chloroplast psbC mRNA in Chlamydomonas reinhardtii. J Cell Biol 157:953–962. http://dx.doi.org/10.1083/jcb.200201060.

12. Rahire M, Laroche F, Cerutti L, Rochaix JD. 2012. Identification of an OPR protein involved in the translation initiation of the Psb1 subunit of photosystem I. Plant J 72:652–661. http://dx.doi.org/10.1111/j.1365-313X.2012.05111.x.

13. Jacobs J, Marx C, Kock V, Reifschneider O, Fränzel B, Krisp C, Wolters H, Marx et al. 2009. Two adjacent nuclear genes are required for functional complementation of trans-splicing mutant from Chlamydomonas reinhardtii. Proc Natl Acad Sci U S A 87:1228–1229. http://dx.doi.org/10.1073/pnas.87.3.1228.

14. Newman SM, Boynton JE, Gillham NW, Randolph-Anderson BL, Johnson AM, Harris EH. 1998. Transformation of chloroplast ribosomal RNA genes of Chlamydomonas reinhardtii by gene transfer and genetic characterization of integration events. Genetics 126:875–888.

15. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Cinderoni L, Jensen, R, Li WW, Noble WS. 2009. MEME Suite: tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208. http://dx.doi.org/10.1093/nar/gkp335.

16. Pagani M, Ioannidis V, Cerutti L, Zahn-Zabal M, Jongneel CV, Falquet L. 2004. MyHits: a new interactive resource for protein annotation and domain identification. Nucleic Acids Res 32:W332–W335.

17. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948. http://dx.doi.org/10.1093/bioinformatics/btm404.

18. Wang L, Brown SJ. 2006. BindN: A Web-based tool for efficient prediction of DNA and RNA binding sites in amino acid sequences. Nucleic Acids Res 34:242–248. http://dx.doi.org/10.1093/nar/gkj425.

19. Kelley LA, Sternberg MJ. 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc 4:363–371. http://dx.doi.org/10.1038/nprot.2009.2.

20. Horton P, Park JK, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. 2007. WolF PSORT: protein localization predictor. Nucleic Acids Res 35:W585–W587.

21. Emanuelsson O, Brunak S, von Heijne G, Nielsen H. 2007. Locating proteins in the cell using TargetP, SignalP, and related tools. Nat Protoc 2:953–971. http://dx.doi.org/10.1038/nprot.2007.131.

22. Emanuelsson O, Nielsen H, von Heijne G. 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci 8:978–984. http://dx.doi.org/10.1111/j.1097-0134.1999.tb08310.x.

23. Tardif M, Attea A, Specht M, Cogne G, Rolland S, Hippler M, Merro F, Bruley C, Petlier G, Vallon O, Cournac L. 2012. PredAlgo: a new subcellular localization prediction tool dedicated to green algae. Mol Biol Evol 29:3625–3639. http://dx.doi.org/10.1093/molbev/mss178.

24. Hahn D, Nickelsen J, Hackert A, Kück U. 1998. A single nuclear locus is involved in both chloroplast RNA trans-splicing and 3’ end processing. Plant J 15:575–581. http://dx.doi.org/10.1046/j.1365-313X.1998.00324.x.

25. Baumgarten C, Kuehne K, Stern DB. 2005. A spontaneous tRNA suppressor of a mutation in the Chlamydomonas reinhardtii nucleus MCD1 gene required for stability of the chloroplast petD mRNA. Nucleic Acids Res 33:3372–3380. http://dx.doi.org/10.1093/nar/gki651.

26. Golding GB. 1999. Simple sequence is abundant in eukaryotic proteins. Protein Sci 8:1358–1361. http://dx.doi.org/10.1111/j.1046-8199.1999.7937.x.

27. Marciotte EM, Pellegrini M, Yeates TO, Eisenberg D. 1999. A census of protein repeats. J Mol Biol 293:151–160. http://dx.doi.org/10.1006/jmbi.1999.3136.

28. DePristo MA, Ziyersmit MM, Hartl DL. 2006. On the abundance, amino acid composition, and evolutionary dynamics of low-complexity regions in proteins. Gene 378:19–30. http://dx.doi.org/10.1016/j.gene.2006.03.023.

29. Jinkerson RE, Jonikas MC. 2015. Molecular techniques to interrogate and edit the Chlamydomonas nucleus genome. Plant J 82:393–412. http://dx.doi.org/10.1111/tpj.12801.

30. Nickelsen J, Kück U. 2000. The unicellular green alga Chlamydomonas reinhardtii as an experimental system to study chloroplast RNA metabolism. Naturwissenschaften 87:97–107. http://dx.doi.org/10.1007/s001140050686.

31. Dutcher SK, Li I, Lin H, Meyer L, Giddings TH, Jr, Kwan AL, Lewis BL. 2012. Whole-genome sequencing to identify mutants and polymorphisms
in Chlamydomonas reinhardtii. G3 (Bethesda) 2:15–22. http://dx.doi.org/10.1534/g3.111.000919.

45. Lin H, Miller ML, Granas DM, Dutcher SK. 2013. Whole genome sequencing identifies a deletion in protein phosphatase 2A that affects its stability and localization in Chlamydomonas reinhardtii. PLoS Genet 9:e1003841. http://dx.doi.org/10.1371/journal.pgen.1003841.

46. Garnier J, Guyon D, Picaud A. 1979. Characterization of new strains of nonphotosynthetic mutants of Chlamydomonas reinhardtii. I. Fluorescence, photochemical activities, chlorophyll-protein complexes. Plant Cell Physiol 20:1013–1027.

47. Gray MW, Burger G, Lang BF. 2001. The origin and early evolution of mitochondria. Genome Biol 2:116. http://dx.doi.org/10.1186/g418.

48. Gould SB, Waller RF, McFadden GI. 2008. Plastid evolution. Annu Rev Plant Biol 59:491–517. http://dx.doi.org/10.1146/annurev.arplant.59.032607.092915.

49. Wakasugi T, Tsudzuki T, Sugiura M. 2001. The genomics of land plant chloroplasts: gene content and alteration of genomic information by RNA editing. Photosynth Res 70:107–118. http://dx.doi.org/10.1023/A:1013892009589.

50. Small ID, Rackham O, Filipovska A. 2013. Organelle transcriptomes: products of a deconstructed genome. Curr Opin Microbiol 16:652–658. http://dx.doi.org/10.1016/j.mib.2013.07.011.

51. Wang P, Dalbey RE. 2011. Inserting membrane proteins: the YidC/Oxa1/Alb3 machinery in bacteria, mitochondria, and chloroplasts. Biochim Biophys Acta 1808:866–875. http://dx.doi.org/10.1016/j.bbamem.2010.08.014.

52. Gray MW. 6 April 2015. Mosaic nature of the mitochondrial proteome: implications for the origin and evolution of mitochondria. Proc Natl Acad Sci U S A. http://dx.doi.org/10.1073/pnas.1421379112.

53. Timmis JN, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet 5:123–135. http://dx.doi.org/10.1038/nrg1271.

54. Woodson JD, Chory J. 2008. Coordination of gene expression between organellar and nuclear genomes. Nat Rev Genet 9:383–395. http://dx.doi.org/10.1038/nrg2348.

55. Martin W, Koonin EV. 2006. Introns and the origin of nucleus-cytoplasmic compartmentalization. Nature 440:41–45. http://dx.doi.org/10.1038/nature04531.

56. Stern DB, Goldschmidt-Clermont M, Hansson MR. 2010. Chloroplast RNA metabolism. Annu Rev Plant Biol 61:125–155. http://dx.doi.org/10.1146/annurev-arplant-042809-112242.

57. Binder S, Brennicke A. 2003. Gene expression in plant mitochondria: transcriptional and post-transcriptional control. Philos Trans R Soc Lond B Biol Sci 358:181–189. http://dx.doi.org/10.1098/rstb.2002.1179.

58. de Longevialle AF, Small ID, Lurin C. 2010. Nuclearily encoded splicing factors implicated in RNA splicing in higher plant organelles. Mol Plant 3:691–705. http://dx.doi.org/10.1093/mp/ssq025.

59. Katiyar-Agarwal S, Gao S, Vivian-Smith A, Jin H. 2007. A novel class of bacteria-induced small RNAs in Arabidopsis. Genes Dev 21:3123–3134. http://dx.doi.org/10.1101/gad.1595107.

60. Kleinkecht L, Wang F, Stöbe R, Philippar K, Nickelsen J, Bohne AV. 2014. RAP, the sole octotricopeptide repeat protein in Arabidopsis, is required for chloroplast 16S rRNA maturation. Plant Cell 26:777–787. http://dx.doi.org/10.1105/tpc.114.122835.

61. Lee I, Hong W. 2004. RAP—a putative RNA-binding domain. Trends Biochem Sci 29:567–570. http://dx.doi.org/10.1016/j.tibs.2004.09.005.