

A Study of New *Arabidopsis* Chloroplast RNA Editing Mutants Reveals General Features of Editing Factors and Their Target Sites ^{WJ|OA}

Kamel Hammani,^{a,b,1} Kenji Okuda,^{c,1} Sandra K. Tanz,^a Anne-Laure Chateigner-Boutin,^a Toshiharu Shikanai,^c and Ian Small^{a,2}

^a Australian Research Council Centre of Excellence in Plant Energy Biology, University of Western Australia, Crawley 6009 WA, Australia

^b Institut de Biologie Moléculaire des Plantes du Centre National de la Recherche Scientifique, Université de Strasbourg, 67084 Strasbourg Cedex, France

^c Department of Botany, Graduate School of Science, Kyoto University, Kyoto 606-8502 Japan

RNA editing in higher plant organelles results in the conversion of specific cytidine residues to uridine residues in RNA. The recognition of a specific target C site by the editing machinery involves *trans*-acting factors that bind to the RNA upstream of the C to be edited. In the last few years, analysis of mutants affected in chloroplast biogenesis has identified several pentatricopeptide repeat (PPR) proteins from the PLS subfamily that are essential for the editing of particular RNA transcripts. We selected other genes from the same subfamily and used a reverse genetics approach to identify six new chloroplast editing factors in *Arabidopsis thaliana* (OTP80, OTP81, OTP82, OTP84, OTP85, and OTP86). These six factors account for nine editing sites not previously assigned to an editing factor and, together with the nine PPR editing proteins previously described, explain more than half of the 34 editing events in *Arabidopsis* chloroplasts. OTP80, OTP81, OTP85, and OTP86 target only one editing site each, OTP82 two sites, and OTP84 three sites in different transcripts. An analysis of the target sites requiring the five editing factors involved in editing of multiple sites (CRR22, CRR28, CLB19, OTP82, and OTP84) suggests that editing factors can generally distinguish pyrimidines from purines and, at some positions, must be able to recognize specific bases.

INTRODUCTION

The term RNA editing covers a multitude of processes that lead to posttranscriptional sequence alterations in RNAs. In plants, RNA editing generally indicates highly specific cytidine-to-uridine conversions that are observed in both mitochondrial and plastid transcripts. Although similar editing events occur in other organisms (see review by Simpson and Emeson, 1996), it is clear that this form of RNA editing arose in an immediate ancestor of land plants and is not homologous to editing processes found in other phyla and may not be mechanistically related. It is likely that >500 sites are specifically altered in *Arabidopsis thaliana* mitochondrial transcripts (Giegé and Brennicke, 1999; Bentolila et al., 2008; Zehrmann et al., 2008), and at least 34 sites are known to be changed in *Arabidopsis* plastid transcripts (Chateigner-Boutin and Small, 2007). RNA editing typically affects the transcripts of protein-coding genes but has also been found to modify noncoding

transcribed regions, structural RNAs, and intron sequences. RNA editing is essential for correct gene expression: proteins translated from edited transcripts are usually different from the ones deduced from the gene sequence and usually present higher similarity to the corresponding nonplant homologs (reviewed in Shikanai, 2006).

A major question concerning the editing process is the manner by which the hundreds of editing sites are specifically targeted. In vivo experiments using transgenic chloroplasts have shown that mRNA editing sites are recognized via *cis*-acting elements that are generally located within ~30 nucleotides of the editing site (Chaudhuri et al., 1995; Bock et al., 1996, 1997; Chaudhuri and Maliga, 1996; Reed et al., 2001). This has been confirmed in vitro using synthetic RNA substrates to explore editing specificity in both chloroplast (Miyamoto et al., 2002; Hegeman et al., 2005; Hayes et al., 2006) and mitochondrial extracts (Neuwirt et al., 2005; Takenaka et al., 2004, 2008; Verbitskiy et al., 2008). In other systems with such pervasive RNA editing, target sites are recognized by guide RNAs complementary to the RNA strand to be edited (Simpson and Emeson, 1996). However, there is good biochemical evidence that the *trans*-factors that bind to the *cis*-elements neighboring editing sites in plant organelles are proteins (Chaudhuri et al., 1995; Bock and Koop, 1997; Hirose and Sugiura, 2001). UV cross-linking experiments in an in vitro tobacco (*Nicotiana tabacum*) editing system reveals that proteins with distinct molecular masses of 25, 56, 70, 91, and 93 kD specifically bind to the *cis*-acting elements required for editing in

¹ These authors contributed equally to this work.

² Address correspondence to iansmall@cylle.uwa.edu.au.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Ian Small (iansmall@cylle.uwa.edu.au).

^{WJ} Online version contains Web-only data.

^{OA} Open Access articles can be viewed online without a subscription. www.plantcell.org/cgi/doi/10.1105/tpc.109.071472

psbL, *psbE*, *petB*, *rpoB*, and *rpoA*, respectively (Hirose and Sugiura, 2001; Miyamoto et al., 2002, 2004; Kobayashi et al., 2008). The same protein of 95 kD specifically binds *cis*-acting elements of two editing sites, *ndhB*-9 and *ndhF*-2, showing that the same *trans*-factor can recognize more than one site (Kobayashi et al., 2008). However, none of these biochemically investigated factors has been identified.

A lack of editing of many of the sites identified in plant organelles would be expected to lead to striking phenotypes due to mis-expression of proteins involved in photosynthesis or respiration. Investigation of *Arabidopsis* mutants with such phenotypes has identified a number of nuclear genes suspected to encode RNA-editing specificity factors (Kotera et al., 2005; Okuda et al., 2007, 2009; Chateigner-Boutin et al., 2008; Cai et al., 2009; Kim et al., 2009; Robbins et al., 2009; Yu et al., 2009; Zehrmann et al., 2009; Zhou et al., 2009). In each case, loss of one of these factors leads to a specific loss of editing of one or at most a few sites. Proof that any of these genetically identified factors directly targets its editing site is limited to a study of the protein CRR4. This factor is required for editing of the initiation codon of plastid *ndhD* transcripts and binds specifically within the region $-25/+10$ spanning the editing site (Okuda et al., 2006). CRR4, like of all of the editing factors described to date, is a pentatricopeptide repeat (PPR) protein (Schmitz-Linneweber and Small, 2008), characterized by tandem arrays of the 35-amino acid motif (Small and Peeters, 2000) for which this family of proteins is named. PPR proteins form a large family of >450 RNA binding proteins in *Arabidopsis*, the majority of which are thought to be targeted to mitochondria or chloroplasts (Lurin et al., 2004). More precisely, the reported organelle editing factors are all members of the plant-specific E and DYW subclasses of the PPR family (Table 1), characterized by distinctive C-terminal domains (Lurin et al., 2004; O'Toole et al., 2008). The E (extended) domain is a degenerate motif with some similarities to PPR motifs, while the DYW domain, named for its typical Asp-Tyr-Trp C-terminal tripeptide, is much more highly conserved, and its presence correlates phylogenetically with plant organelle RNA editing (Salone et al., 2007; Rudinger et al., 2008).

Most of the editing factors described so far were identified through forward genetic screens, generally via the distinctive phenotype caused by loss of editing. However, the similarity among all of them suggested to us that a reverse genetics approach concentrating on genes encoding similar proteins might identify new editing specificity factors. Using a high-resolution melting of amplicons technique for screening for editing defects (Chateigner-Boutin and Small, 2007), we identified six novel *Arabidopsis* mutants affected in editing of plastid mRNAs. An analysis of these mutants and comparison with the previously described mutants allows us to make some progress toward understanding the recognition of editing sites in plant organelles.

RESULTS

Identification of PPR Genes Required for RNA Editing in Plastids

All previously identified factors required for editing of specific sites are proteins of the E or DYW subclasses of the PPR family.

As the number of editing sites in chloroplasts is much lower than that in mitochondria and therefore easier to screen, we chose to focus on E and DYW members predicted to be targeted to plastids by TargetP (Emanuelsson et al., 2000) or Predotar (Small et al., 2004). Sixteen candidate genes were selected because of the availability of T-DNA insertion mutants for each gene (Table 1). For each mutant, the location of the T-DNA insertion was verified by PCR and sequencing, and homozygous mutant lines identified (see Methods for the insertion sites). To test if a T-DNA insertion in these genes leads to an RNA editing defect, the status of the 34 editing sites in *Arabidopsis* plastids was systematically examined in each mutant using a high-resolution melting screen (Chateigner-Boutin and Small, 2007).

Among these 16 genes, the disruption of six of them leads to a defect in RNA editing in plastids. We respectively named these genes *OTP80*, *OTP81*, *OTP82*, *OTP84*, *OTP85*, and *OTP86* (for *Organelle Transcript Processing*). *OTP80* encodes a PPR protein from the E subclass, whereas *OTP81*, *82*, *84*, *85*, and *86* encode members of the DYW subclass. *OTP82* was independently identified in another screen and will be described elsewhere. Disruption of the other 10 genes did not lead to an observable defect in RNA editing of any of the 34 sites we tested.

In the mutants *otp80*, *otp81*, *otp85*, and *otp86*, a single editing site was affected, but in *otp84*, three different editing sites appeared to remain unedited (Table 1; see Supplemental Figure 1 online). All of the editing defects were confirmed by a more sensitive and quantitative poisoned primer extension assay (Figure 1). In *otp81*, editing in the *rps12* intron (at genome position 69,553) is reduced to 2% (Figure 1B). We consider the residual editing to indicate that this is a weak allele, since the T-DNA is inserted into the 5'-untranslated region of *OTP81* (see Methods). The mutants *otp80*, *otp85*, and *otp86* are totally impaired in the editing of *rpl23* (86,056), *ndhD* (116494), and *rps14* (37,161), respectively (Figures 1A to 1D and 1E). In *otp84*, three defects in RNA editing were observed: *ndhB* (94,999), *ndhF* (112349), and *psbZ* (35,800) (Figure 1C). All three sites remain completely unedited in the mutant.

To investigate whether the editing defects we observed in these mutants were secondarily caused by altered RNA processing or modified RNA stability, the transcripts affected were analyzed by RNA gel blots (Figure 2). The hybridization patterns reveal no striking differences in the pattern and level of the transcripts whose editing is impaired in the mutants (*rpl23*, *rps12* intron, *ndhB*, *ndhF*, *psbZ*, *ndhD*, and *rps14*). Furthermore, none of the other editing sites in the same transcripts is affected in any of these mutants (see Supplemental Figure 1 online).

Complementation experiments were performed for all the *otp* mutants. The introduction of a wild-type copy of the corresponding *OTP* gene into each mutant restores the editing at the defective sites to wild-type levels (Figure 1). These results support a direct and specific role of these OTP factors in RNA editing of these sites.

OTP80, OTP81, OTP84, OTP85, and OTP86 Are Localized in Plastids

The chloroplast editing defects in these mutants suggest that the PPR proteins encoded by these genes reside in plastids. To

Table 1. Candidate and Known Genes Affecting Cchloroplast RNA Editing

<i>Arabidopsis</i> Locus	Subclass	Predotar	TargetP	Gene Name	Editing Defect (Genome Position)	AA Change	T-DNA Insertion Lines	T-DNA Location
At2g02980	DYW	pC	C	OTP85	<i>ndhD</i> (116,494)	S>L	SAIL_544_B03	CDS
At2g29760	DYW	C	C	OTP81	<i>rps12</i> intron (69,553)		SALK_092402	5' UTR
At3g03580	DYW	pC	C		None		GABI_895H11	CDS
At3g08820	DYW	C	N		None		SALK_023916	CDS
At3g57430	DYW	ER	C	OTP84	<i>ndhF</i> (112,349), <i>psbZ</i> (35,800)	S>L, S>L	SAIL_568_C04	CDS
					<i>ndhB</i> (94,999)	P>L	SALK_120902	CDS
At3g62890	DYW	N	C		None		SAIL_1249_D04	CDS
							SALK_044324	CDS
At3g63370	DYW	pC	C	OTP86	<i>rps14</i> (37161)	S>L	SALK_102445	CDS
At4g35130	DYW	C	C		None		SALK_118555	CDS
At1g74600	E	C	C		None		GABI_073C06	CDS
At1g77170	E	pM	C		None		SAIL_1291_C04	CDS
At3g22150	E	C	C		None		SALK_040629	CDS
At3g29230	E	pC	M		None		SAIL_205_G08	CDS
At4g04370	E	pC	N		None		SALK_025427	CDS
At4g25270	E	pC	C		None		SALK_090845	CDS
At5g59200	E	C	C	OTP80	<i>rpl23</i> (86,056)	S>L	SALK_060533	CDS
							SALK_111721	CDS
At1g08070	DYW	pC	C	OTP82	<i>ndhG</i> (118,858), <i>ndhB</i> (95,644)	S>F, S>L	SAIL_851_G04	CDS
							SALK_027812	CDS
At1g11290	DYW	N	C	<i>CRR22</i>	<i>ndhB</i> (96,419), <i>ndhD</i> (116,281) <i>rpoB</i> (25,779)		Okuda et al. (2009)	
At1g15510	DYW	C	C	<i>AtECB2</i>	<i>accD</i> (57,868)		Yu et al. (2009)	
At1g59720	DYW	N	C	<i>CRR28</i>	<i>ndhB</i> (96,698), <i>ndhD</i> (116,290)		Okuda et al. (2009)	
At3g22690	DYW	pC	C	<i>YS1</i>	<i>rpoB</i> (25,992)		Zhou et al. (2009)	
At5g13270	DYW	C	C	<i>RARE1</i>	<i>accD</i> (57,868)		Robbins et al. (2009)	
At5g48910	DYW	N	C	<i>LPA66</i>	<i>psbF</i> (63,985)		Cai et al. (2009)	
At2g45350	E	pC	N	<i>CRR4</i>	<i>ndhD</i> (117,166)		Kotera et al. (2005)	
At1g05750	E+	C	C	<i>CLB19</i>	<i>rpoA</i> (78,691), <i>clpP</i> (69,942)		Chateigner-Boutin et al. (2008)	
At5g55740	E+	C	C	<i>CRR21</i>	<i>ndhD</i> (116,785)		Okuda et al. (2007)	

pC, potentially plastid; C, plastid; pM, potentially mitochondrial; M, mitochondrial; ER, secretory pathway; N, none; AA, amino acid; UTR, untranslated region; CDS, coding sequence. The new editing factors revealed by this study are indicated in bold.

verify this, the N-terminal 100 amino acids of the proteins (72 amino acids in the case of OTP85) were fused with green fluorescent protein (GFP) driven by expression from the cauliflower mosaic virus 35S promoter. The plasmids containing the chimeric genes were transiently expressed in wild-type *Arabidopsis* cells by bombardment. Analysis of GFP fluorescence in transformed cells revealed that the fluorescence colocalized with a chloroplast marker: the fusion of red fluorescent protein (RFP) with the small subunit of *Arabidopsis* ribulose biphosphate carboxylase (Carrie et al., 2009) (see Supplemental Figure 2 online). Thus, OTP80, OTP81, OTP84, OTP85, and OTP86 are localized in plastids consistent with their role in editing plastid transcripts.

Requirement of RNA Editing at the Sites Affected by OTP Factors

The site that is not edited in *otp81* lies in the intron of *rps12*, and this altered sequence might affect splicing of *rps12* transcripts. However, RNA gel blot hybridization shows no difference in the pattern of *rps12* transcripts between *otp81* and the wild type (Figure 2B). Thus, editing of the *rps12* intron is not necessary for

the stabilization of the precursor transcripts, nor correct splicing, and there is no reason to think that the amount or function of Rps12 is deleteriously affected in the *otp81* mutant.

The conversion of cytidine to uridine at the sites affected in the other *otp* mutants is not silent and changes the nature of the amino acids in the proteins encoded by these transcripts (Table 1). The alteration of editing at these sites might be expected to have an impact on the function of these proteins. Nevertheless, *otp80*, *otp84*, *otp85*, and *otp86* show normal growth under standard conditions (see Supplemental Figure 3 online).

The mutants *otp80* and *otp86* are both affected in editing of transcripts encoding ribosomal proteins and thus could be expected to show translation defects. Translation in *Arabidopsis* plastids is thought to be essential (reviewed in Schmitz-Linneweber and Small, 2008). However, methylene blue staining of total RNA from *otp80* and *otp86* did not reveal any major defects in the integrity or quantity of the chloroplast rRNAs (indicated by arrows in Figures 2A and 2G). This suggests that these editing events are not essential for normal ribosome accumulation in plastids, in agreement with their normal growth phenotype, despite the presumed effects on the amino acid sequences of Rps14 and Rpl23 (Table 1).

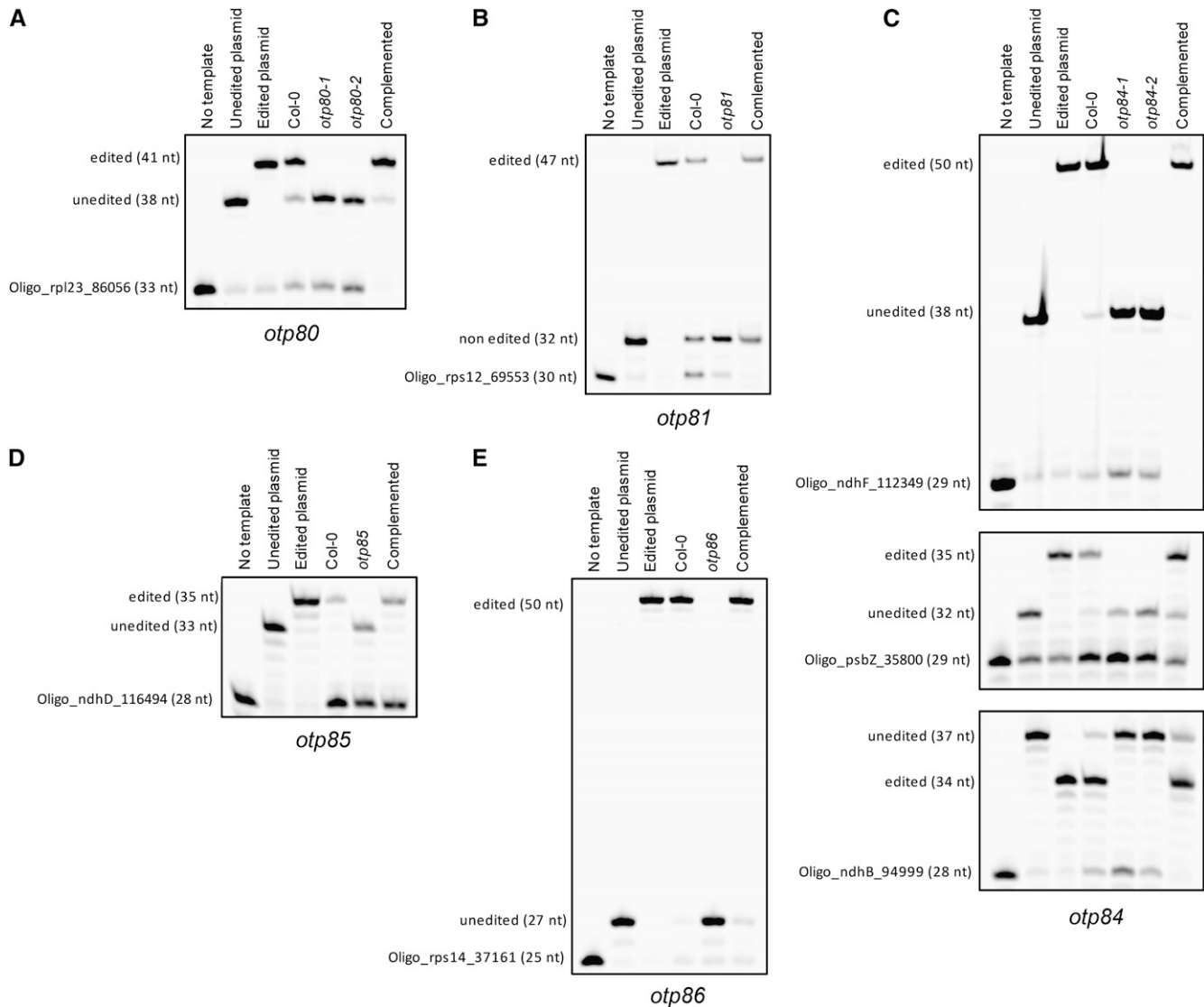


Figure 1. Editing Defects in *otp80*, *otp81*, *otp84*, *otp85*, and *otp86* Mutants.

Poisoned primer extension assays were conducted on the editing sites *rpl23* (nucleotide 86,056) for *otp80* (A), *rps12* intron (69,553) for *otp81* (B), *ndhF* (112,349), *psbZ* (35,800), and *ndhB* (94,999) for *otp84* (C), top to bottom; *ndhD* (116494) for *otp85* (D), and *rps14* (37,161) for *otp86* (E). The editing sites are specified relative to the nucleotide sequence of the complete *Arabidopsis* chloroplast genome.

The mutant *otp85* lacks editing of *ndhD* (nucleotide 116494). RNA editing at this site changes a Ser codon to a Leu codon. The mutant *otp84* is impaired in the editing of the three editing sites *ndhB* (94,999), *ndhF* (112,349), and *psbZ* (35,800). RNA editing at *ndhB* (94,999) converts a Pro codon to a Leu codon, while editing at *ndhF* (112349) and *psbZ* (35,800) convert Ser codons to Leu codons (Table 1). NdhB, NdhD, and NdhF are subunits of the chloroplast NAD(P)H dehydrogenase (NDH) complex involved in photosystem I cyclic electron flow (Shikanai, 2007). The NDH complex catalyzes electron donation to plastoquinone from the stromal electron pool. PsbZ is a core subunit of photosystem II (Swiatek et al., 2001). The editing defects in *ndhB*, *ndhF*, and *psbZ* could affect the function of the encoded proteins, which should alter photosynthetic parameters.

NDH complex activity can be monitored as a transient increase in chlorophyll fluorescence reflecting plastoquinone reduction after turning off actinic light (Shikanai et al., 1998). Figure 3 shows a typical chlorophyll fluorescence trace from wild-type *Arabidopsis* and in *crr21*, a mutant lacking NDH activity (Okuda et al., 2007). In *otp84*, the postillumination increase of chlorophyll fluorescence is modified but, unlike *crr21*, is not completely abolished. This indicates that NDH activity is diminished, but not absent, in the mutant. This also suggests that Leu residues at positions 494 in NdhB and/or 97 in NdhF are required for optimal NDH stability or activity. The postillumination increase of chlorophyll fluorescence in *otp85* is not modified (Figure 4), indicating that a Leu at position 225 in NdhD is not essential for NDH activity.

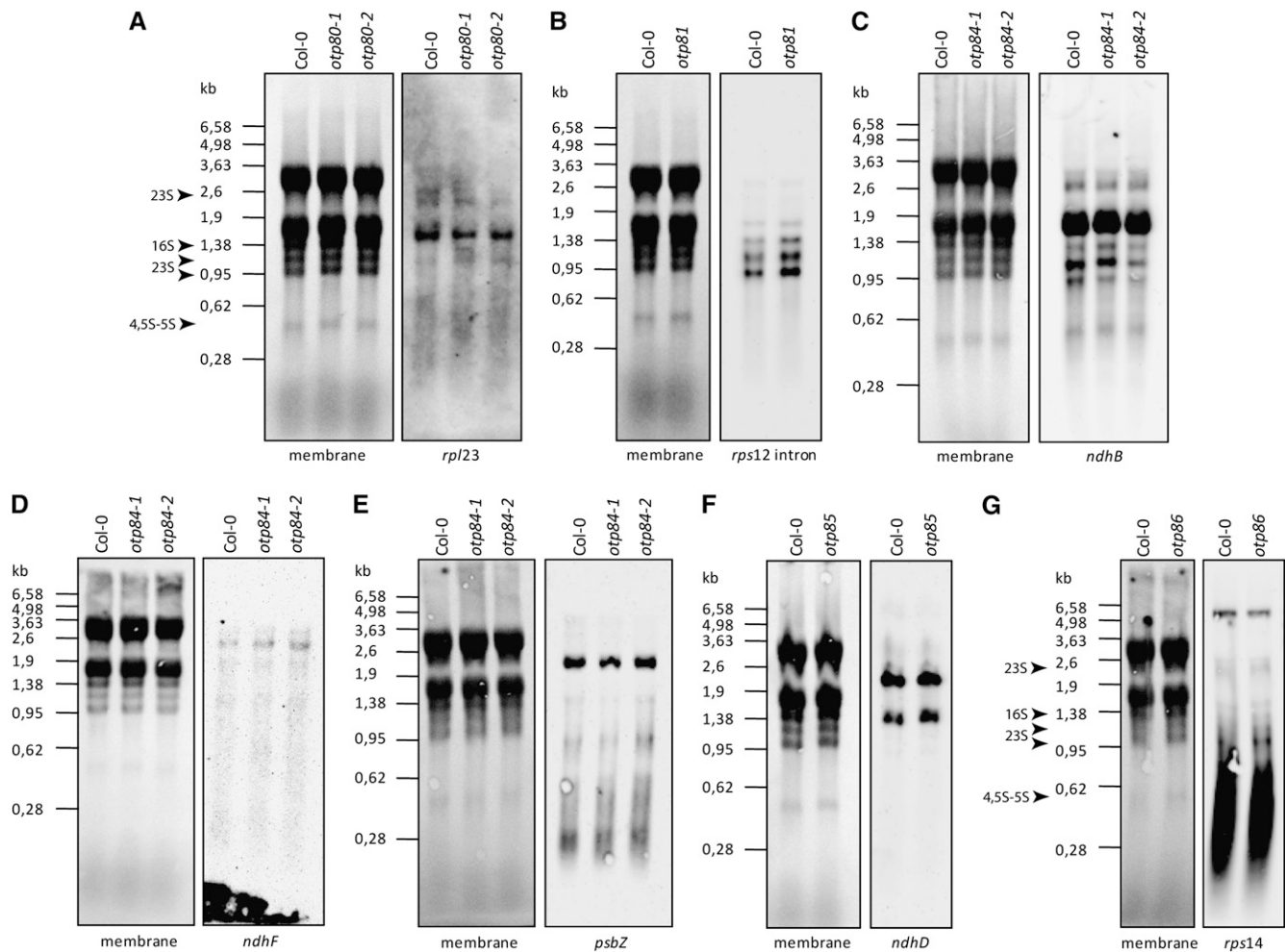


Figure 2. Transcript Profiles of Genes with the Editing Defects in *otp80*, *otp81*, *otp84*, *otp85*, and *otp86* Mutants.

otp80 (A), *otp81* (B), *otp84* (C) to (E), *otp85* (F), and *otp86* (G) mutants. Fifteen micrograms of RNA from 15-d-old seedlings was loaded on formaldehyde agarose gels and transferred onto a membrane. Hybridizations were performed under high-stringency conditions using antisense RNA probes for the genes specified for each blot. The sizes of RNA markers are shown in kilobases. rRNA stained with methylene blue on the membranes is shown as a loading control (membrane). The arrows indicate the plastid rRNAs in *otp80* and *otp85* mutants.

RNA editing defects result in amino acid changes that may directly alter protein function or act by destabilizing the protein or by affecting its ability to form complexes with other proteins. To assess whether the subunits affected in *otp84* and *otp85* stably accumulate in vivo, protein blots were analyzed using antibodies against NdhH and PsbZ (Figure 4). The NDH complex is unstable without NdhH and NdhD (Peng et al., 2008), and an antibody against NdhH can be used to monitor accumulation of the complex. In *otp84*, the level of NdhH is reduced to 50 to 25% of the wild type. This result suggests that Leu residues at positions 494 in NdhB and/or 97 in NdhF are important for normal accumulation of the NDH complex. The decrease in the amount of the NDH complex largely explains the observed decrease in activity, although we cannot exclude that the specific activity of the complex is also decreased. In *otp85*, the accumulation of the NDH complex is not affected.

In addition to *ndhB* (94,999) and *ndhF* (112349), *otp84* is impaired in editing of *psbZ* (35,800), leading to retention of a Ser (rather than Leu) codon at position 17 in PsbZ, a core subunit of photosystem II (Swiatek et al., 2001). Immunodetection of PsbZ on blots of *otp84* extracts (Figure 4) reveals no defect in the accumulation of the protein in the mutant. Consistent with the normal growth of *otp84* (see Supplemental Figure 3 online), there were no alterations to the rate of electron transport through photosystem II nor to nonphotochemical quenching of chlorophyll fluorescence, which reflects ΔpH formation (see Supplemental Figure 4 online). These results suggest that a Leu at position 17 in PsbZ is not important for its function under standard growth conditions.

Multiple Editing Sites and Recognition Specificity

A major outstanding question is how editing factors recognize their target sites and of particular interest are those factors that

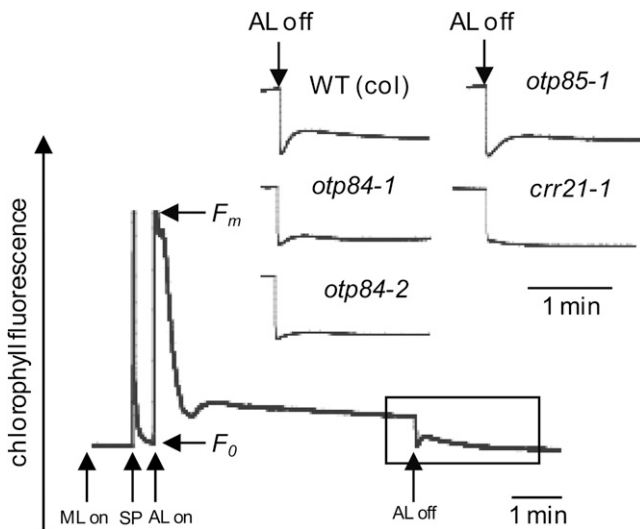


Figure 3. Monitoring of NDH Activity Using Chlorophyll Fluorescence Analysis for *otp84* and *otp85* Mutants.

The curve shows a typical trace of chlorophyll fluorescence in the wild type and a mutant totally impaired in NDH activity (*crr21*) (Okuda et al., 2007) compared with traces from *otp85* and two independent *otp84* mutants (*otp84-1* and *otp84-2*). Leaves were exposed to actinic light (AL) ($50 \mu\text{mol}$ of photons $\text{m}^{-2} \text{s}^{-1}$) for 5 min. AL was turned off, and the subsequent change in chlorophyll fluorescence level was monitored. The transient increase in chlorophyll fluorescence is due to the plastoquinone reduction based on NDH activity. Insets are magnified traces from the boxed area. The fluorescence levels were normalized by the maximum fluorescence at closed photosystem II centers in the dark (F_m) levels. ML, measuring light; SP, a saturating pulse of white light.

appear to recognize more than one site. Biochemical evidence suggests that most of the sequence recognition is concentrated within the 15 nucleotides immediately upstream of the editing site (Hirose and Sugiura, 2001; Miyamoto et al., 2002, 2004; Kobayashi et al., 2008). However, as remarked in previous work (Chateigner-Boutin et al., 2008; Okuda et al., 2009), editing sites requiring the same factor do not show extensive sequence identity in this crucial region, raising doubts as to how these putative specificity factors might operate. To investigate these questions, we undertook a systematic bioinformatic analysis of all plastid editing sites in *Arabidopsis* and the factors that recognize them.

Confirmation that editing sites requiring the same factor are not particularly similar to each other is shown in Figure 5. Only *ndhB* (94,999) and *ndhF* (112349) (requiring OTP84) pair together, and even this grouping is only weakly supported. The only strongly supported grouping in this tree is between *rpoB* (25,992) and *ndhB*(96,419), but these require different editing factors (YS1 and CRR22, respectively). At first sight, these data raise doubts that the identified editing factors can work as commonly assumed (i.e., by binding to the region immediately upstream of the target C). We reasoned that we had sufficient data to test this assumption and therefore examined whether the apparently dissimilar target sites requiring the same factors nevertheless contain sufficient conserved nucleotides to unambiguously

define a binding site consensus. If an unambiguous consensus could be found, then this would be consistent with the protein in question acting as a specificity factor by binding to this consensus sequence.

For each set of editing sites requiring a single factor, we calculated six possible consensus sequences using the following assumptions: (1) that PPR proteins can distinguish all four nucleotides; (2) that PPR proteins can distinguish purines (A and G) from pyrimidines (C and U); (3) that PPR proteins can distinguish the number of hydrogen bonding groups in each base (i.e., can distinguish A or U from G or C) or various combinations of these criteria (Table 2). Each consensus was then matched against both strands of the entire *Arabidopsis* chloroplast genome and the number of matches recorded. Finally, matches within duplicated or untranscribed sequences were discounted. The results indicate that with the exception of the sites recognized by CRR22, in all other cases there is sufficient conservation of sequence to unambiguously identify the edited sites from all other transcribed sequences within plastids, based on the assumptions (1) and (2) listed above.

DISCUSSION

A reverse genetic screen of T-DNA insertion mutants allowed us to identify six new PPR proteins required for RNA editing of nine sites in plastid RNAs. Unlike the majority of the previously characterized editing mutants (Okuda et al., 2006, 2007, 2009; Chateigner-Boutin et al., 2008; Kim et al., 2009; Yu et al., 2009; Zhou et al., 2009), which were identified via visible growth or fluorescence phenotypes, these mutants were identified directly via a screen for unedited RNAs. It is striking that of the eight mutants discovered in this way, including RARE1 (Robbins et al., 2009) and MEF1 (Zehrmann et al., 2009), none have growth or

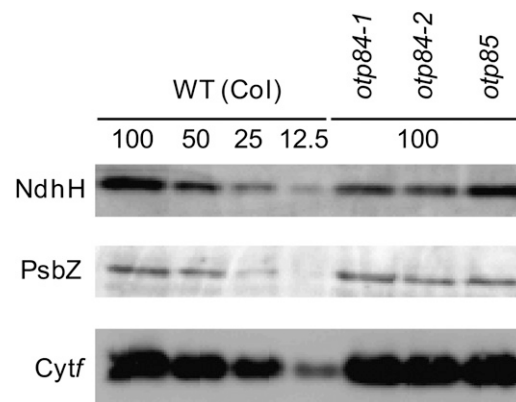


Figure 4. Protein Blot Analysis of the NDH Complex and the Major Photosynthetic Complexes.

Immunodetection of an NDH subunit, NdhH; a subunit of photosystem II, PsbZ; and a subunit of the *Cytb₆f* complex, cytochrome *f*. The proteins were extracted from thylakoid membrane fractions. Lanes were loaded with protein samples corresponding to 0.5 μg chlorophyll for *Cyt f*, 1 μg chlorophyll for PsbZ, and 5 μg chlorophyll for NdhH (100%) and the series of dilutions indicated as a percentage of the original sample.

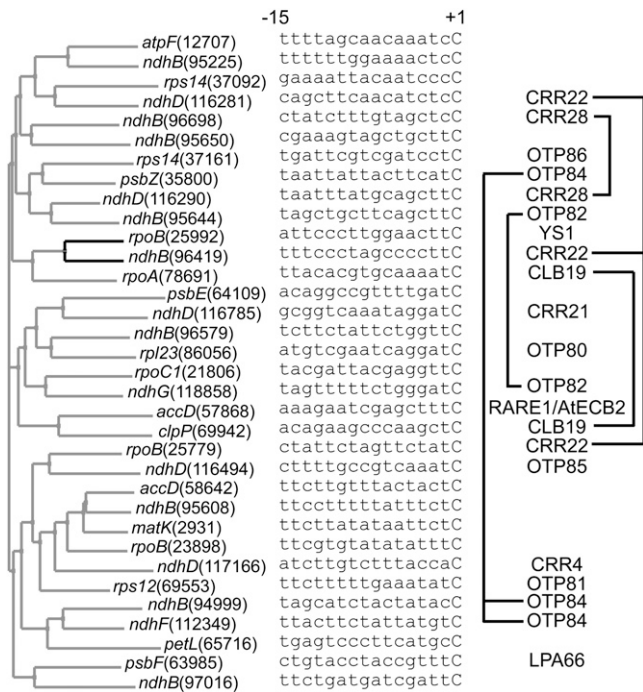


Figure 5. Assessment of Sequence Similarities among Editing Sites.

For the 34 known editing sites in the *Arabidopsis* chloroplast genome (Chateigner-Boutin and Small, 2007), the 15 nucleotides upstream of the editing site were aligned with respect to the editing site (the final C in each sequence) and a distance tree was constructed. Only the grouping of *rpoB* (25,992) with *ndhB* (96,419) has bootstrap support of >50%, and this is indicated by the black lines in the tree. The right panel indicates the PPR proteins involved in the recognition of each editing site.

physiological defects, apart from the partial decrease in NDH activity observed for *otp84*. This rather suggests that many more RNA editing events may be phenotypically silent than previously expected, at least under optimal growth conditions. This observation has implications for the evolutionary stability of editing events (Tillich et al., 2009), as it implies that selection pressure on both the editing sites and the factors that recognize them may be weaker than often thought. It should be noted, however, that as editing almost always restores codons for conserved amino acids (Figure 6), it is likely that under some conditions it would be possible to observe deleterious effects of a lack of editing of these apparently silent sites.

More than half of the candidate genes we examined could not be linked to RNA editing in plastids. This might be for a number of different reasons, including genetic redundancy, errors in targeting predictions (such that the protein functions in mitochondria, not plastids), incomplete surveying of editing sites (it is still possible that there are more than the 34 known editing sites in *Arabidopsis* plastid RNAs), or because the protein has functions unrelated to editing. As an example of the last case, the DYW protein CRR2 is required for RNA processing (accumulation of monocistronic *ndhB* mRNA), not RNA editing (Hashimoto et al., 2003).

Although prediction of editing factors is not yet infallible, nevertheless 22 of the 34 known editing sites in plastids have

been assigned corresponding PPR *trans*-factors, and we expect that the remaining editing events are also likely to require a PPR protein from the plant-specific PLS subfamily. A total of 44 of these proteins are predicted to be targeted to plastids (Lurin et al., 2004), which is more than sufficient to account for all the known editing sites in plastids. In mitochondria, the number of editing sites is considerably higher than the number of PLS subfamily PPR proteins predicted to be targeted to mitochondria: 82 PPR proteins (Lurin et al., 2004) for >500 editing sites (Giegé and Brennicke, 1999; Bentolila et al., 2008; Zehrmann et al., 2008). This suggests that in mitochondria, a single PPR editing factor should, on average, cover more sites than in plastids. Indeed, in plastids, the majority of editing factors are required for a single site, with a maximum of three in the cases of CRR22 (Okuda et al., 2009) and OTP84. The first mitochondrial editing factors to be found, OGR1 and MEF-1, are required for at least seven and three editing sites in rice (*Oryza sativa*) and *Arabidopsis* mitochondria, respectively (Kim et al., 2009; Zehrmann et al., 2009).

The sequences surrounding the editing sites identified in plant organelles do not show any unequivocal consensus, apart from the one or two bases immediately surrounding the site (reviewed in Mulligan et al., 1999 and reexamined in Cummings and Myers, 2004). This probably represents preferences of the enzyme catalyzing the editing reaction rather than a target for postulated specificity factor(s). The key question of how specific C residues are recognized for editing from all other C residues in organelle transcripts remained unanswered for many years. Experiments with transgenic plastids and in vitro RNA editing assays using organelle extracts delimited the primary region recognized by putative *trans*-factors to the 20 or so nucleotides immediately upstream of the editing site, with, in most cases, the most important recognition elements situated from -5 to -15 with respect to the edited C (reviewed in Shikanai, 2006). Binding to this region was confirmed for CRR4, the first editing factor to be genetically identified (Okuda et al., 2006), and it has been assumed that the other PPR proteins identified as editing factors function in a similar way (Okuda et al., 2007, 2009; Chateigner-Boutin et al., 2008; Cai et al., 2009; Kim et al., 2009; Robbins et al., 2009; Yu et al., 2009; Zehrmann et al., 2009; Zhou et al., 2009). However, no proof of specific binding has been shown for any of these subsequent editing factors, including the six new ones identified in the screen described here.

Some confirmation of specific binding comes from comparing the factors genetically identified in *Arabidopsis* to those biochemically identified in tobacco. Cross-linking experiments trapping putative protein *trans*-factors bound to the *cis*-acting elements of editing sites (Kobayashi et al., 2008) identified a 91-kD protein binding the *rpoB*-1 site which, based on size, could correspond to the tobacco ortholog of YS1, a PPR protein required for the editing of the same site in *Arabidopsis* plastids (Zhou et al., 2009) and whose predicted molecular mass (discounting the plastid targeting sequence) is 90.3 kD. An even more compelling match is seen between the 95-kD protein binding the two RNA editing sites *ndhB*-9 and *ndhF*-1 in tobacco (Kobayashi et al., 2008) and OTP84, which we show to recognize the equivalent *Arabidopsis* editing sites *ndhB* (94,999) and *ndhF* (112349). The predicted molecular mass of OTP84 (discounting

Table 2. Analyses of the Editing Sites Requiring the Same PPR Proteins

PPR		Consensus	Hits	Filtered Hits	Location
OTP82	a	UAG-U-U-G-UC	15		
	b	YRRYY-YYYYRR-YC	27		
	c	UAGYU-YUYRGR-UC	3	2	<i>ndhB, ndhG</i>
	d	WWS-W-W-SSWWC	260		
	e	UAG-U-U-GSWUC	4		
	f	UAGYU-YUYRGSWUC	3	2	<i>ndhB, ndhG</i>
CRR22	a	---A---C	17,078		
	b	Y-YYYRRY-YY-YC	59		
	c	Y-YYYARY-YY-YC	35	12	<i>psbK, ycf2, ycf1, matK, 3'UTR atpH, rpoB, 3'UTR rps4, ndhB, ndhD, ndhG, ndhA, ycf1</i>
	d	-W--W--W-C	15,439		
	e	-W--A--W-C	7,646		
	f	YW-YYYARY-YYWYC	23	8	<i>psbK, ycf1, 3'UTR atpH, rpoB, ndhB, ndhD, ndhG, ndhA</i>
OTP84	a	U--UA-U--C	444		
	b	Y-RY-YYRY-YRYC	28		
	c	U-RY-YUAYU-YRYC	7	3	<i>psbZ, ndhB, ndhF</i>
	d	WW-WW-WW-WW-C	1,732		
	e	UW-WW-UA-UW-C	88		
	f	UWRYWWYUAYUWYRYC	5	3	<i>psbZ, ndhB, ndhF</i>
CLB19	a	-A-A-G-CAA-UC	16		
	b	-YR-R-RY-YRRR-YC	74		
	c	-YA-A-GY-CAAR-UC	4	2	<i>clpP, rpoA</i>
	d	W-WSW-S-SSWW-WC	39		
	e	W-ASA-G-SCAA-UC	2		
	f	WYASA-GYSCAAR-UC	2	2	<i>clpP, rpoA</i>
CRR28	a	-AU-U-UG-AGCU-C	16		
	b	Y-RYY-YRYRRYYC	6		
	c	Y-AUYU-UG-AGCU-C	3	2	<i>ndhB, ndhD</i>
	d	-WWW-WWWS-WSSW-C	79		
	e	-WAU-UWUG-AGCU-C	3		
	f	YWAUYUWUG-AGCU-C	3	2	<i>ndhB, ndhD</i>

Consensus sequences were derived as follows from the 15 nucleotides immediately upstream of the edited C (and including the editing site itself): (a) full conservation of nucleotides (A, U, G, and C); (b) conservation of purines (A or G = R) or pyrimidines (U or C = Y); (c) combination of a and b; (d) conservation of number of hydrogen bonding groups (A or U = W, G or C = S); (e) combination of a and d; (f) combination of a, b, and d. The "Hits" column indicates the number of times each consensus is found within both strands of the *Arabidopsis* plastid genome. "Filtered Hits" removes duplicate matches within the inverted repeats, matches to noncoding strands, and matches to intergenic regions that do not form stably accumulated transcripts. The editing sites known to be recognized by each factor are highlighted in bold. UTR, untranslated region.

the plastid targeting sequence) is 94.8 kD. All the observations so far are entirely consistent with editing site specificity being determined by binding of the relevant PPR protein to the sequence upstream of the target C, in accordance with the general view of how PPR proteins function in RNA metabolism (Delannoy et al., 2007). It is worth noting that there is no significant difference in length, number of motifs, or any other aspect of the protein sequence or structure that we can find to distinguish PPR proteins recognizing multiple target sites from those that recognize single sites (Figure 7).

Only a few PPR proteins have been shown unambiguously to bind to multiple target sites, of which the best-studied are the maize (*Zea mays*) proteins CRP1 (Schmitz-Linneweber et al., 2005) and PPR10 (Pfalz et al., 2009). In these two cases the

multiple target sites have almost identical sequences. This is not true of the multiple editing sites targeted by single factors. In our study, a distance tree relating the *Arabidopsis* plastid editing sites (Figure 5) does not group the editing sites known to be recognized by a single PPR factor except for the *ndhB* (94,999) and *ndhF* (112349) sites mentioned above. Previously proposed clusters based on apparent sequence similarity (Chateigner-Boutin and Hanson, 2002; Chateigner-Boutin and Hanson, 2003) do not exactly correspond to groups of sites recognized by single factors. This prompted us to test whether the commonly held assumption that these editing factors bind the biochemically defined *cis*-elements is consistent with the data.

We reasoned that an unambiguous consensus between sites requiring the same editing factor would be consistent with the

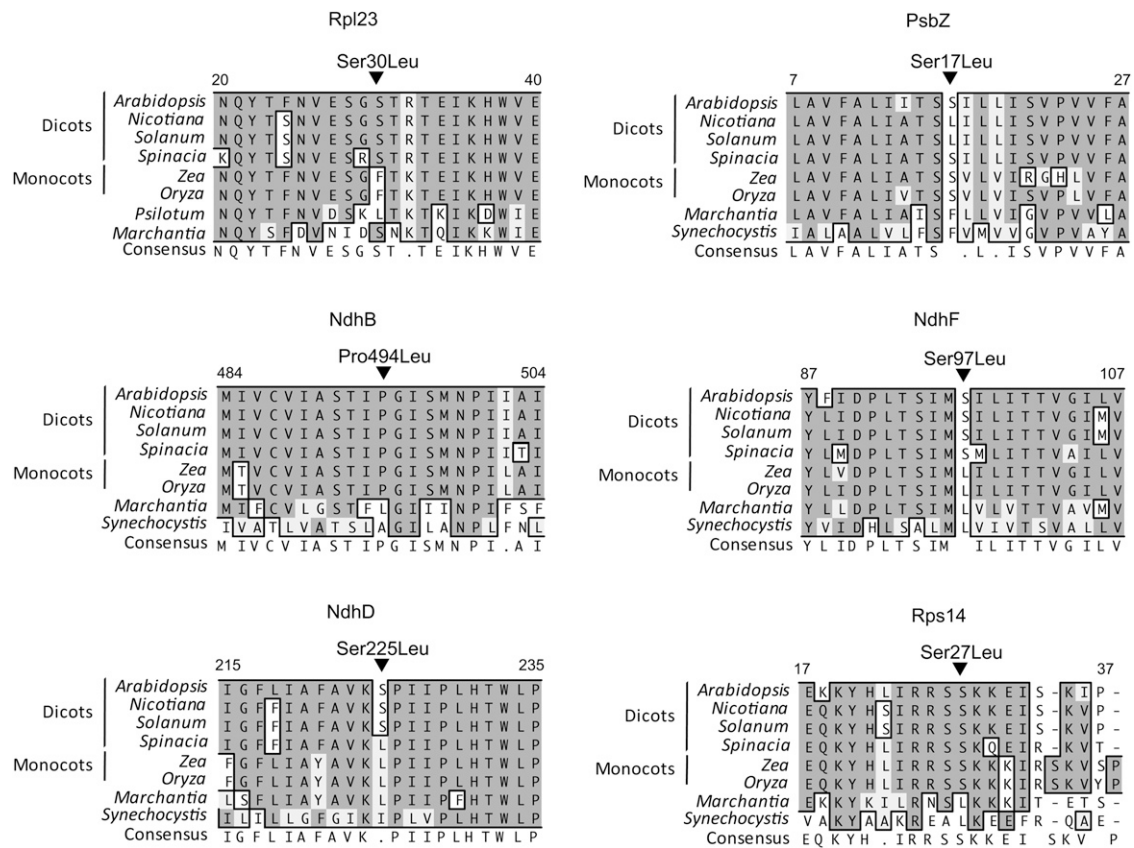


Figure 6. Partial Sequence Alignments of Rpl23, PsbZ, NdhB, NdhF, Rps14, and NdhD around the Amino Acids Affected by RNA Editing.

Arabidopsis Rpl23, PsbZ, NdhB, NdhF, Rps14, and NdhD proteins were aligned with their homologs from other species. The alignment was performed using ClustalW (Thompson et al., 1994). Amino acids that are fully or semiconserved are shaded black or gray, respectively. Numbers indicate amino acid positions in the protein. The arrows above the sequences indicate the positions of edited codons.

hypothesis that the factor is involved in target site recognition. We addressed this by generating various consensus sequences (Table 2) from the multiple target sites, using the results from published in vitro experiments (Hirose and Sugiura, 2001; Miyamoto et al., 2002, 2004; Kobayashi et al., 2008) to delimit the region of interest from -15 to the editing site. We found that it was possible to define an unambiguous 15-nucleotide consensus for each of the editing factors required for editing of multiple sites, except in the case of CRR22. We conclude that the data are consistent with the hypothesis that these PPR editing factors bind to multiple target sites and consistent with the biochemical evidence that recognition involves the nucleotides upstream of the editing site. The one exception is CRR22, which will be discussed in more detail later.

In the cases where a consensus can be found, the pattern of nucleotide similarity is informative regarding what sequence features might be being recognized. RNA nucleotides tend to interact base-specifically with proteins in one of two ways: via stacking interactions with the aromatic rings or via hydrogen bonding. Sequence specificity can be achieved by positive base-specific bonding interactions or by negative steric hindrance (Auweter et al., 2006). Based on these considerations, we

considered three (mutually compatible) hypotheses: that editing factors can distinguish all four RNA nucleotides, that they can distinguish purines (A and G) from pyrimidines (C and U), or that they can distinguish bases by Watson-Crick hydrogen bonding patterns (G and C versus A and U). None of these consensus alone provided sufficient specificity to unambiguously define the edited sites within the plastid transcriptome, but combining conserved nucleotide positions with conserved purines or pyrimidines is sufficient (in the cases where a consensus could be achieved). We conclude that to achieve the required specificity, editing factors must be able to distinguish purine nucleotides from pyrimidine nucleotides, and at least at some positions, be able to uniquely distinguish one of the four nucleotides. The relatively poor performance of the consensus based on A/U versus C/G suggests that this distinction is not important in determining binding specificity.

This analysis is consistent with previous biochemical data. Kobayashi et al. (2008) conducted experiments using RNAs competing with the factor required for editing *ndhB-9* and *ndhF-1* (as explained above, this is almost certainly the tobacco ortholog of OTP84). Competitor RNAs were constructed by scanning mutagenesis in which each successive five-nucleotide

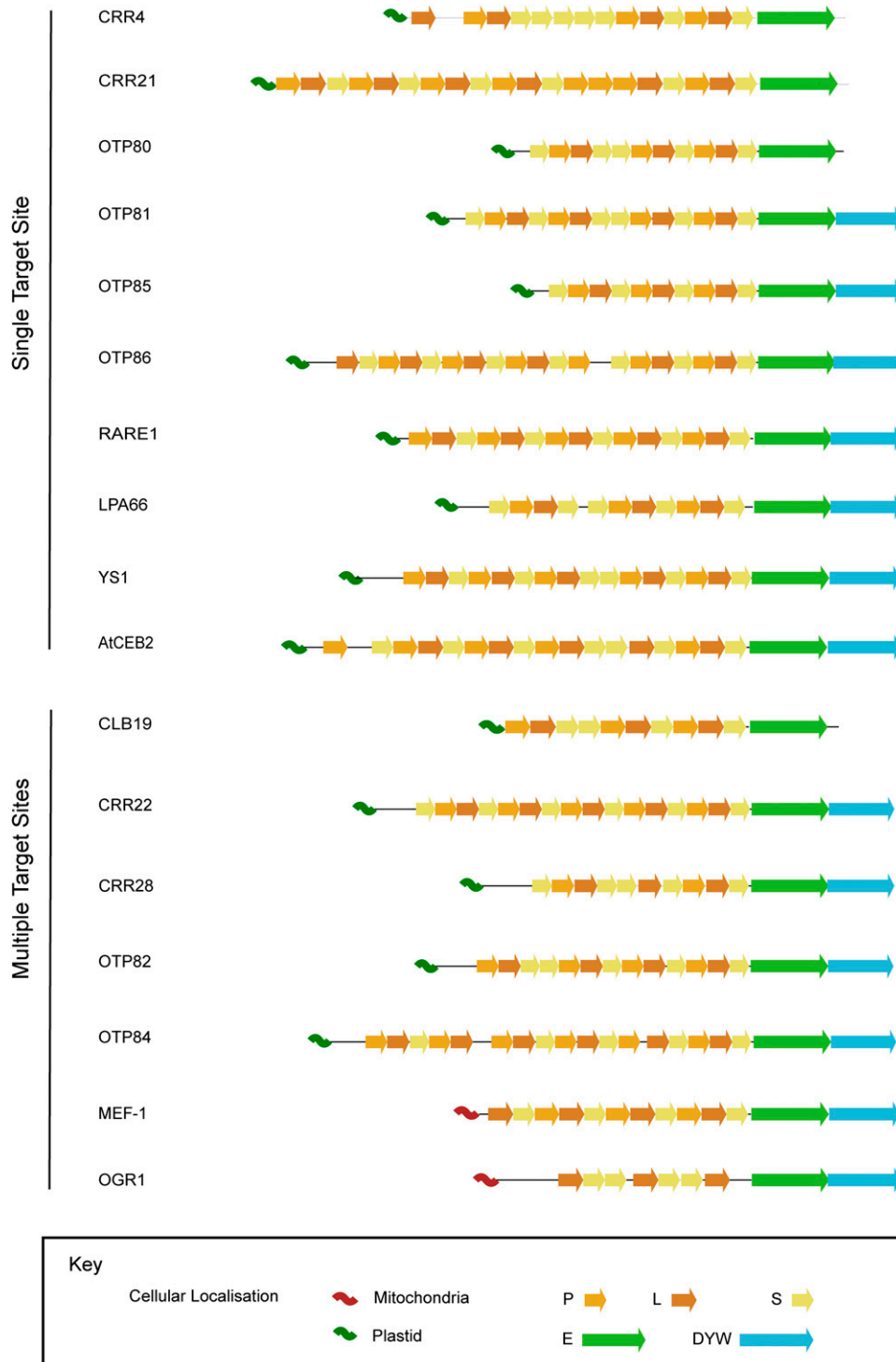


Figure 7. Schematic Structure of Plant PPR Proteins Involved in Editing of Plastid or Mitochondrial Transcripts.

PPR proteins are grouped according to their number of target sites (single or multiple). All the identified PPR motifs are indicated (as defined in Lurin et al., 2004), as well as the C-terminal E and DYW domains and the N-terminal organelle targeting sequences. P represents canonical 35-amino acid PPR motifs; L (long) and S (short) represent variant PPR motifs with generally 36 and 31 amino acids, respectively.

block in the region from -15 to -1 was substituted with its complementary nucleotide sequence. These mutations therefore switched the purine/pyrimidine nature of the residues in the upstream *cis*-elements. For both sites, mutations introduced into the -15 to -6 region abolished competition for the binding of the editing factor, whereas mutations in the -5 to -1 region had a weaker effect. Consistent with this, mutagenesis of the *ndhF* RNA substrate showed that purine/pyrimidine exchanges in the -15 to -1 upstream region abolish *in vitro* editing of the *ndhF*-1 site (Sasaki et al., 2006).

Throughout the analysis discussed above, the exception is CRR22. This PPR editing factor is required for editing the three sites *ndhB* (96,419), *ndhD* (116,281), and *rpoB* (25,779) (Okuda et al., 2009), but the sequence conservation between these three sites is not sufficient to explain how these three sites can be distinguished from other equally similar sequences scattered throughout the plastid transcriptome (Table 2). One possibility might be that CRR22 recognizes sequences outside the region studied here, but the sequence similarity around these sites is no greater over a more extended region. Another possibility is that different arrays of PPR motifs within CRR22 recognize different target sites, thus allowing a single protein to bind to several unrelated targets. However, CRR22 is not longer than other editing factors (Figure 7), which might be expected if it contained two or more independent binding domains. We therefore need to consider whether CRR22 truly acts as a specificity factor at all three sites it is required for or whether there might be an alternative explanation for its involvement in editing at one or more of these sites. Recently, two different PPR editing factors, RARE1 and ECB2, were demonstrated to be required for editing of a single site, *accD* (57,868), in *Arabidopsis* chloroplasts (Robbins et al., 2009; Yu et al., 2009). As all the other evidence suggests that a single PPR protein is sufficient to define editing specificity, the implication of two proteins in the editing of *accD* (57,868) implies that one of them may have a different role, as might be the case for CRR22. The most attractive hypothesis would be that this different role is to recruit the editing enzyme or even to catalyze the editing reaction itself. Like CRR22, both RARE1 and AtECB2 belong to the DYW subfamily, and one model for RNA editing proposes that the C-terminal DYW domain carries the cytidine deaminase editing activity (Salone et al., 2007).

METHODS

All the primers used in this study are listed in Supplemental Table 1 online.

Plant Material

Arabidopsis thaliana ecotype Columbia (Col-0) was used in this study. The T-DNA insertion mutant lines were obtained from the ABRC Stock Center. Accession numbers are provided in Table 1.

Genetic Analysis

Total cellular DNA was isolated as described by Edwards et al. (1991). Plants were genotyped for homozygous lines by PCR, and the insertion position was confirmed by sequencing with a T-DNA left border primer.

otp80-1 (SALK_111721, Col-0, insertion site +141), *otp80-2* (SALK_060533, Col-0, insertion site +221), *otp81* (SALK_092402, Col-0, insertion site -15), *otp84-1* (SAIL_568_C04, Col-0, insertion site +932), *otp84-2* (SALK_120902, Col-0, insertion site +642), *otp85* (SAIL_544_B03, Col-0, insertion site +1328), and *otp86* (SALK_102445, Col-0, insertion site +698).

Analysis of Targeting via GFP Fusions

The first 300 bp of the coding sequences of the PPR genes were amplified using Phusion DNA polymerase (Finnzymes) with primers listed in Supplemental Table 1 containing the *attB* sites for Gateway cloning according to the manufacturer's instructions (Invitrogen). The GFP vector used and the chloroplast targeting marker encoding RFP fused to the small subunit of *Arabidopsis* ribulose biphosphate carboxylase were kindly provided by James Whelan (University of Western Australia) (Carrie et al., 2009). Biolistic transformations of GFP and RFP constructs were performed on *Arabidopsis* cell culture (Carrie et al., 2007). The GFP construct and the chloroplast RFP marker (5 μ g each) were coprecipitated onto gold particles and transformed using the biolistic PDS-1000/He system (Bio-Rad). Particles were bombarded onto 2 mL of *Arabidopsis* cell suspension resting on filter paper on osmoticum plates (2.17 g/L Murashige and Skoog modified basal salt mixture, 30 g/L sucrose, 0.5 mg/L naphthalene acetic acid, 0.05 mg/L kinetin, and 36.44 g/L mannitol). After bombardment, the cells were placed in the dark at 22°C for 24 h. Observation of transient GFP and RFP expression was performed using an Olympus BX61 fluorescence microscope with excitation wavelengths of 460/480 nm (GFP) and 535/555 nm (RFP) and emission wavelengths of 495 to 540 nm (GFP) and 570 to 625 nm (RFP). Subsequent images were captured using Cell imaging software as previously described (Carrie et al., 2007; Murcha et al., 2007).

Genetic Complementation

The 1870-, 3138-, 2673-, 2278-, and 2883-bp fragments containing the respective coding sequence of *OTP80*, *OTP81*, *OTP84*, *OTP85*, and *OTP86* were amplified by PCR on total cellular DNA. These constructs were cloned into pGWB1 (*OTP80*, *OTP81*, and *OTP85*) or pGWB2 (*OTP84* and *OTP86*) binary vectors and introduced into *otp* mutants via *Agrobacterium tumefaciens* GV3101. Transformants were obtained by selection on Murashige and Skoog agar plates containing 25 μ g/mL hygromycin and confirmed by PCR.

Analysis of RNA Editing

High-resolution melting analysis of amplicons was performed as previously described (Chateigner-Boutin and Small, 2007) using the primers listed by Okuda et al. (2009). Poisoned primer extension of RT-PCR products was performed as described by Chateigner-Boutin and Small (2007). RT-PCR products were obtained with primers surrounding the editing sites and serve as templates for the extension reaction from a labeled 6-carboxyfluorescein primer that anneals next to the target editing site. The extension is stopped by the incorporation of ddGTP or ddCTP at the location of the editing site for unedited molecules producing a short unedited product. The extension is stopped at the next G/C for the edited molecules producing a longer edited product. For the *ndhB* (94,999) site, the extension is stopped by the incorporation of ddATP leading to a short edited product and a long unedited product.

RNA Preparation and Analysis

Total RNA from leaves of 15-d-old plantlets was isolated using TRIzol reagent (Invitrogen) as recommended by the manufacturer. Fifteen micrograms of RNA was fractionated on 1.2% (w/v) formaldehyde agarose gels and transferred onto Hybond N⁺ nylon membranes

(GE Healthcare). RNA integrity, loading, and transfer were checked by staining the membrane with methylene blue.

RNA probes were internally labeled with biotinylated cytidine by transcription of PCR products cloned in pGEM-T Easy vector (Promega). The primers used for the PCR are listed in Supplemental Table 1 online. Clones with inserts in antisense orientation were amplified by PCR using the forward primer and M13/pUC reverse. The PCR products served as a template for *in vitro* transcription with SP6 polymerase following the manufacturer's instructions (Maxiscript Ambion).

To prepare the *ndhF* RNA probe, the *ndhF* sequence was amplified using an antisense primer linked to the T7 promoter. The PCR product served as template for *in vitro* transcription with T7 polymerase (Maxiscript Ambion).

Prehybridization of the membrane was performed for 1 h in hybridization buffer (5× SSC, 50% [v/v] formamide, 0.5% SDS, and 100 μg/mL heparin) at 68°C. Hybridization with RNA probes was performed in the same buffer overnight at 68°C, followed by three 15-min washes at room temperature in 1× SSC/0.5% SDS and two washes at 60°C in 0.1× SSC/0.1% SDS for 20 min and 1 h, respectively. Signal detection was performed using the Chemiluminescent Nucleic Acid Detection Module (Pierce) and read in an ImageQuant-RT ECL (Amersham).

Chlorophyll Fluorescence Analysis

Chlorophyll fluorescence was measured using a MINI-PAM portable chlorophyll fluorometer (Waltz). The transient increase in chlorophyll fluorescence after turning off actinic light was monitored as previously described (Shikanai et al., 1998).

Immunoblot Analysis

Chloroplasts were isolated from the leaves of 4-week-old plants as previously described (Okuda et al., 2007). Samples were normalized by measuring chlorophyll concentration. The protein samples were separated by 12.5% SDS-PAGE. After electrophoresis, the proteins were transferred onto a Hybond-P membrane (GE Healthcare) and incubated with specific antibodies. The signals were detected using an ECL Advance Western Blotting Detection Kit (for NdhH; GE Healthcare) or an ECL Plus Western Blotting Detection Kit (for the others; GE Healthcare). The signals were visualized by a LAS1000 chemiluminescence analyzer (Fuji Film).

Bioinformatic Analysis

The 15 nucleotides sequences upstream of editing sites were aligned and clustered with ClustalW 1.83 (Thompson et al., 1994) using the default parameters. Consensuses for editing sites recognized by the same factor were calculated by hand and searched against the *Arabidopsis* plastid genome sequence using fuzznuc from the EMBOSS package (Rice et al., 2000).

Accession Numbers

The genes described in this article correspond to the following Arabidopsis Genome Initiative codes: At5g59200 (*OTP80*), At2g29760 (*OTP81*), At3g57430 (*OTP84*), At2g02980 (*OTP85*), and At3g63370 (*OTP86*). Accession information for T-DNA insertion lines is provided in Table 1. Editing sites are specified relative to the nucleotide sequence of the complete *Arabidopsis* chloroplast genome (GenBank accession number AP000423).

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. High-Resolution Melting Screen of *otp80*, *otp81*, *otp84*, *otp85*, and *otp86* Mutants.

Supplemental Figure 2. Analysis of Subcellular Localization of the *OTP80*, *OTP81*, *OTP84*, *OTP85*, and *OTP86* Proteins.

Supplemental Figure 3. Phenotype of *otp80*, *otp81*, *otp84*, *otp85*, *otp86*, and Wild-Type *Arabidopsis* Plants.

Supplemental Figure 4. In Vivo Analysis of Electron Transport Activity Using Light Intensity Dependence of Electron Transport Rate.

Supplemental Table 1. Oligonucleotides Used in This Study.

Supplemental Data Set 1. Text File of the Alignment Used for the Analysis Shown in Figure 5.

ACKNOWLEDGMENTS

This work is funded through the Australian Research Council Centres of Excellence scheme (CE0561495), an International Science Linkages grant (CG120098) from the Australian Government Department of Innovation, Industry, Science, and Research, and scholarships from the Region of Alsace and the University of Western Australia. K.H. is the holder of a Lavoisier Fellowship, and I.S. is a West Australian State Premier's Fellow. T.S. was supported by Grants-in-Aid (17GS0316 and 16085206) and by Genomics for Agricultural Innovation (GPN0008).

Received September 17, 2009; revised October 9, 2009; accepted October 30, 2009; published November 29, 2009.

REFERENCES

- Auweter, S.D., Oberstrass, F.C., and Allain, F.H. (2006). Sequence-specific binding of single-stranded RNA: Is there a code for recognition? *Nucleic Acids Res.* **34**: 4943–4959.
- Bentolila, S., Elliott, L.E., and Hanson, M.R. (2008). Genetic architecture of mitochondrial editing in *Arabidopsis thaliana*. *Genetics* **178**: 1693–1708.
- Bock, R., Hermann, M., and Fuchs, M. (1997). Identification of critical nucleotide positions for plastid RNA editing site recognition. *RNA* **3**: 1194–1200.
- Bock, R., Hermann, M., and Kössel, H. (1996). *In vivo* dissection of *cis*-acting determinants for plastid RNA editing. *EMBO J.* **15**: 5052–5059.
- Bock, R., and Koop, H.U. (1997). Extraplasmidic site-specific factors mediate RNA editing in chloroplasts. *EMBO J.* **16**: 3282–3288.
- Cai, W., Ji, D., Peng, L., Guo, J., Ma, J., Zou, M., Lu, C., and Zhang, L. (2009). LPA66 is required for editing *psbF* chloroplast transcripts in *Arabidopsis*. *Plant Physiol.* **150**: 1260–1271.
- Carrie, C., Kühn, K., Murcha, M.W., Duncan, O., Small, I.D., O'Toole, N., and Whelan, J. (2009). Approaches to defining dual-targeted proteins in *Arabidopsis*. *Plant J.* **57**: 1128–1139.
- Carrie, C., Murcha, M.W., Millar, A.H., Smith, S.M., and Whelan, J. (2007). Nine 3-ketoacyl-CoA thiolases (KATs) and acetoacetyl-CoA thiolases (ACATs) encoded by five genes in *Arabidopsis thaliana* are targeted either to peroxisomes or cytosol but not to mitochondria. *Plant Mol. Biol.* **63**: 97–108.
- Chateigner-Boutin, A.L., and Hanson, M.R. (2002). Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared *cis* elements. *Mol. Cell. Biol.* **22**: 8448–8456.
- Chateigner-Boutin, A.L., and Hanson, M.R. (2003). Developmental co-variation of RNA editing extent of plastid editing sites exhibiting similar *cis*-elements. *Nucleic Acids Res.* **31**: 2586–2594.
- Chateigner-Boutin, A.L., Ramos-Vega, M., Guevara-García, A., Andrés, C., de la Luz Gutiérrez-Nava, M., Cantero, A., Delannoy,

- E., Jiménez, L.F., Lurin, C., Small, I., and León, P.** (2008). CLB19, a pentatricopeptide repeat protein required for editing of *rpoA* and *clpP* chloroplast transcripts. *Plant J.* **56**: 590–602.
- Chateigner-Boutin, A.L., and Small, I.** (2007). A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons. *Nucleic Acids Res.* **35**: e114.
- Chaudhuri, S., Carrer, H., and Maliga, P.** (1995). Site-specific factor involved in the editing of the *psbL* mRNA in tobacco plastids. *EMBO J.* **14**: 2951–2957.
- Chaudhuri, S., and Maliga, P.** (1996). Sequences directing C to U editing of the plastid *psbL* mRNA are located within a 22 nucleotide segment spanning the editing site. *EMBO J.* **15**: 5958–5964.
- Cummings, M.P., and Myers, D.S.** (2004). Simple statistical models predict C-to-U editing sites in plant mitochondrial RNA. *BMC Bioinformatics* **5**: 132.
- Delannoy, E., Stanley, W.A., Bond, C.S., and Small, I.D.** (2007). Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. *Biochem. Soc. Trans.* **35**: 1643–1647.
- Edwards, K., Johnstone, C., and Thompson, C.** (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* **19**: 1349.
- Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G.** (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **300**: 1005–1016.
- Giegé, P., and Brennicke, A.** (1999). RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. *Proc. Natl. Acad. Sci. USA* **96**: 15324–15329.
- Hashimoto, M., Endo, T., Peltier, G., Tasaka, M., and Shikanai, T.** (2003). A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast *ndhB* in *Arabidopsis*. *Plant J.* **36**: 541–549.
- Hayes, M.L., Reed, M.L., Hegeman, C.E., and Hanson, M.R.** (2006). Sequence elements critical for efficient RNA editing of a tobacco chloroplast transcript *in vivo* and *in vitro*. *Nucleic Acids Res.* **34**: 3742–3754.
- Hegeman, C.E., Hayes, M.L., and Hanson, M.R.** (2005). Substrate and cofactor requirements for RNA editing of chloroplast transcripts in *Arabidopsis in vitro*. *Plant J.* **42**: 124–132.
- Hirose, T., and Sugiura, M.** (2001). Involvement of a site-specific trans-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: Development of a chloroplast *in vitro* RNA editing system. *EMBO J.* **20**: 1144–1152.
- Kim, S.R., Yang, J.I., Moon, S., Ryu, C.H., An, K., Kim, K.M., Yim, J., and An, G.** (2009). Rice OGR1 encodes a pentatricopeptide repeat-DYW protein and is essential for RNA editing in mitochondria. *Plant J.* **59**: 738–749.
- Kobayashi, Y., Matsuo, M., Sakamoto, K., Wakasugi, T., Yamada, K., and Obokata, J.** (2008). Two RNA editing sites with *cis*-acting elements of moderate sequence identity are recognized by an identical site-recognition protein in tobacco chloroplasts. *Nucleic Acids Res.* **36**: 311–318.
- Kotera, E., Tasaka, M., and Shikanai, T.** (2005). A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature* **433**: 326–330.
- Lurin, C., et al.** (2004). Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* **16**: 2089–2103.
- Miyamoto, T., Obokata, J., and Sugiura, M.** (2002). Recognition of RNA editing sites is directed by unique proteins in chloroplasts: Biochemical identification of *cis*-acting elements and *trans*-acting factors involved in RNA editing in tobacco and pea chloroplasts. *Mol. Cell. Biol.* **22**: 6726–6734.
- Miyamoto, T., Obokata, J., and Sugiura, M.** (2004). A site-specific factor interacts directly with its cognate RNA editing site in chloroplast transcripts. *Proc. Natl. Acad. Sci. USA* **101**: 48–52.
- Mulligan, R.M., Williams, M.A., and Shanahan, M.T.** (1999). RNA editing site recognition in higher plant mitochondria. *J. Hered.* **90**: 338–344.
- Murcha, M.W., Elhafez, D., Lister, R., Tonti-Filippini, J., Baumgartner, M., Philippar, K., Carrie, C., Mokranjac, D., Soll, J., and Whelan, J.** (2007). Characterization of the preprotein and amino acid transporter gene family in *Arabidopsis*. *Plant Physiol.* **143**: 199–212.
- Neuwirt, J., Takenaka, M., van der Merwe, J.A., and Brennicke, A.** (2005). An *in vitro* RNA editing system from cauliflower mitochondria: editing site recognition parameters can vary in different plant species. *RNA* **11**: 1563–1570.
- Okuda, K., Chateigner-Boutin, A.L., Nakamura, T., Delannoy, E., Sugita, M., Myouga, F., Motohashi, R., Shinozaki, K., Small, I., and Shikanai, T.** (2009). Pentatricopeptide repeat proteins with the DYW motif have distinct molecular functions in RNA editing and RNA cleavage in *Arabidopsis* chloroplasts. *Plant Cell* **21**: 146–156.
- Okuda, K., Myouga, F., Motohashi, R., Shinozaki, K., and Shikanai, T.** (2007). Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. *Proc. Natl. Acad. Sci. USA* **104**: 8178–8183.
- Okuda, K., Nakamura, T., Sugita, M., Shimizu, T., and Shikanai, T.** (2006). A pentatricopeptide repeat protein is a site recognition factor in chloroplast RNA editing. *J. Biol. Chem.* **281**: 37661–37667.
- O'Toole, N., Hattori, M., Andres, C., Iida, K., Lurin, C., Schmitz-Linneweber, C., Sugita, M., and Small, I.** (2008). On the expansion of the pentatricopeptide repeat gene family in plants. *Mol. Biol. Evol.* **25**: 1120–1128.
- Peng, L., Shimizu, H., and Shikanai, T.** (2008). The chloroplast NAD(P) H dehydrogenase complex interacts with photosystem I in *Arabidopsis*. *J. Biol. Chem.* **283**: 34873–34879.
- 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.
- Reed, M.L., Peeters, N.M., and Hanson, M.R.** (2001). A single alteration 20 nt 5' to an editing target inhibits chloroplast RNA editing *in vivo*. *Nucleic Acids Res.* **29**: 1507–1513.
- Rice, P., Longden, I., and Bleasby, A.** (2000). EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **16**: 276–277.
- Robbins, J.C., Heller, W.P., and Hanson, M.R.** (2009). A comparative genomics approach identifies a PPR-DYW protein that is essential for C-to-U editing of the *Arabidopsis* chloroplast *accD* transcript. *RNA* **15**: 1142–1153.
- Rudinger, M., Polsakiewicz, M., and Knoop, V.** (2008). Organellar RNA editing and plant-specific extensions of pentatricopeptide repeat proteins in jungermanniid but not in marchantiid liverworts. *Mol. Biol. Evol.* **25**: 1405–1414.
- Salone, V., Rudinger, M., Polsakiewicz, M., Hoffmann, B., Groth-Malonek, M., Szurek, B., Small, I., Knoop, V., and Lurin, C.** (2007). A hypothesis on the identification of the editing enzyme in plant organelles. *FEBS Lett.* **581**: 4132–4138.
- Sasaki, T., Yukawa, Y., Wakasugi, T., Yamada, K., and Sugiura, M.** (2006). A simple *in vitro* RNA editing assay for chloroplast transcripts using fluorescent dideoxynucleotides: Distinct types of sequence elements required for editing of *ndh* transcripts. *Plant J.* **47**: 802–810.
- Schmitz-Linneweber, C., and Small, I.** (2008). Pentatricopeptide repeat proteins: A socket set for organelle gene expression. *Trends Plant Sci.* **13**: 663–670.
- Schmitz-Linneweber, C., Williams-Carrier, R., and Barkan, A.** (2005). RNA immunoprecipitation and microarray analysis show a chloroplast pentatricopeptide repeat protein to be associated with the 5' region of mRNAs whose translation it activates. *Plant Cell* **17**: 2791–2804.

- Shikanai, T.** (2006). RNA editing in plant organelles: Machinery, physiological function and evolution. *Cell. Mol. Life Sci.* **63**: 698–708.
- Shikanai, T.** (2007). Cyclic electron transport around photosystem I: Genetic approaches. *Annu. Rev. Plant Biol.* **58**: 199–217.
- Shikanai, T., Endo, T., Hashimoto, T., Yamada, Y., Asada, K., and Yokota, A.** (1998). Directed disruption of the tobacco *ndhB* gene impairs cyclic electron flow around photosystem I. *Proc. Natl. Acad. Sci. USA* **95**: 9705–9709.
- Simpson, L., and Emeson, R.B.** (1996). RNA editing. *Annu. Rev. Neurosci.* **19**: 27–52.
- Small, I., Peeters, N., Legeai, F., and Lurin, C.** (2004). Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* **4**: 1581–1590.
- Small, I.D., and Peeters, N.** (2000). The PPR motif - A TPR-related motif prevalent in plant organellar proteins. *Trends Biochem. Sci.* **25**: 46–47.
- Swiatek, M., Kuras, R., Sokolenko, A., Higgs, D., Olive, J., Cinque, G., Müller, B., Eichacker, L.A., Stern, D.B., Bassi, R., Herrmann, R. G., and Wollman, F.A.** (2001). The chloroplast gene *ycf9* encodes a photosystem II (PSII) core subunit, PsbZ, that participates in PSII supramolecular architecture. *Plant Cell* **13**: 1347–1367.
- Takenaka, M., Neuwirt, J., and Brennicke, A.** (2004). Complex *cis*-elements determine an RNA editing site in pea mitochondria. *Nucleic Acids Res.* **32**: 4137–4144.
- Takenaka, M., Verbitskiy, D., van der Merwe, J.A., Zehrmann, A., and Brennicke, A.** (2008). The process of RNA editing in plant mitochondria. *Mitochondrion* **8**: 35–46.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J.** (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680.
- Tillich, M., Le Sy, V., Schulerowitz, K., von Haeseler, A., Maier, U.G., and Schmitz-Linneweber, C.** (2009). Loss of matK RNA editing in seed plant chloroplasts. *BMC Evol. Biol.* **9**: 201.
- Verbitskiy, D., van der Merwe, J.A., Zehrmann, A., Brennicke, A., and Takenaka, M.** (2008). Multiple specificity recognition motifs enhance plant mitochondrial RNA editing in vitro. *J. Biol. Chem.* **283**: 24374–24381.
- Yu, Q.B., Jiang, Y., Chong, K., and Yang, Z.N.** (2009). AtECB2, a pentatricopeptide repeat protein, is required for chloroplast transcript *accD* RNA editing and early chloroplast biogenesis in *Arabidopsis thaliana*. *Plant J.* **59**: 1011–1023.
- Zehrmann, A., van der Merwe, J.A., Verbitskiy, D., Brennicke, A., and Takenaka, M.** (2008). Seven large variations in the extent of RNA editing in plant mitochondria between three ecotypes of *Arabidopsis thaliana*. *Mitochondrion* **8**: 319–327.
- Zehrmann, A., Verbitskiy, D., van der Merwe, J.A., Brennicke, A., and Takenaka, M.** (2009). A DYW domain-containing pentatricopeptide repeat protein is required for RNA editing at multiple sites in mitochondria of *Arabidopsis thaliana*. *Plant Cell* **21**: 558–567.
- Zhou, W., Cheng, Y., Yap, A., Chateigner-Boutin, A.L., Delannoy, E., Hammani, K., Small, I., and Huang, J.** (2009). The *Arabidopsis* gene *YS1* encoding a DYW protein is required for editing of *rpoB* transcripts and the rapid development of chloroplasts during early growth. *Plant J.* **58**: 82–96.