The Plant Cell, Vol. 25: 3079–3088, August 2013, www.plantcell.org

Mutagenesis of Individual Pentatricopeptide Repeat Motifs Affects RNA Binding Activity and Reveals Functional Partitioning of *Arabidopsis* PROTON GRADIENT REGULATION3[C][W]

Sota Fujii,[a] Nozomi Sato,[a] and Toshiharu Shikanai[a,b,1]

[a] Department of Botany, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan
[b] CREST, Japan Science and Technology Agency, Chiyoda-ku, Tokyo 102-0076, Japan

Pentatricopeptide repeat (PPR) proteins bind RNA and act in multiple eukaryotic processes, including RNA editing, RNA stability, and translation. Here, we investigated the mechanism underlying the functional versatility of *Arabidopsis thaliana* PROTON GRADIENT REGULATION3 (PGR3), a chloroplast protein harboring 27 PPR motifs. Previous studies suggested that PGR3 acts in (1) stabilization of *photosynthetic electron transport L* (*petL*) operon RNA, (2) translation of *petL*, and (3) translation of *ndhA*. We showed here that replacement of the 4th amino acid of the 12th PPR with nonpolar or charged amino acids abolished functions (1) and (2) but not (3) of PGR3 by compromising the function of this specific PPR. This discovery enabled us to knock out the RNA binding ability of individual PPR motifs. Consequently, we showed that the 16 N-terminal PPRs were sufficient for function (1) via sequence-specific RNA binding, whereas the 11 C-terminal motifs were essential for functions (2) and (3) by activating translation. We also clarified that the 14th amino acid of the 12th PPR should be positively charged to make the PPR functionally active. Our finding opens up the possibility of selectively manipulating the functions of PPR proteins.

INTRODUCTION

Pentatricopeptide repeat (PPR) proteins form a large family of helical repeat proteins ubiquitously found in eukaryotes and are deeply involved in coevolution of the nucleus with the mitochondria and plastids (Small and Peeters, 2000; Schmitz-Linneweber and Small, 2008). Each PPR motif consists of 35–amino acid degenerate consensus related to tetratricopeptide motifs (Small and Peeters, 2000). A number of recent studies proved the direct RNA binding activity of the PPR motifs (Pfalz et al., 2009; Cai et al., 2011; Prikryl et al., 2011; Kobayashi et al., 2012; Okuda and Shikanai, 2012). PPR proteins harbor between two and 30 PPR motifs, and their tandem alignment allows the modular recognition of specific RNA sequences. Many studies have found that PPR mutants have compromised functions of specific genes in mitochondria or plastids; PPR proteins can promote versatile molecular processes involving RNA, such as editing, splicing, stabilization, and translation (reviewed in Schmitz-Linneweber and Small, 2008; Fujii and Small, 2011; Nakamura et al., 2012). As a consequence, the majority of PPR proteins are essential for important cellular processes, such as respiration and photosynthesis. The expansive evolution of PPR proteins in terrestrial plant species (ranging from 100 to over 1000 genes per genome) also suggests the plasticity of these proteins in adapting to a wide range of RNA targets (Delannoy et al., 2007; O'Toole et al., 2008; Fujii and Small, 2011). These findings suggest that it may be possible to redirect or even synthesize PPR proteins to bind to desired RNA targets in a manner similar to that of transcription activator-like effectors, a class of protein motif that binds to specific DNA sequences. Transcription activator-like effectors have already been applied to biotechnology via the rapid computational design of repeat arrays that bind to virtually any DNA sequence (Boch et al., 2009; Moscou and Bogdanove, 2009). With the unveiling of the primary version of the PPR code, manipulation of PPRs to bind to target RNAs is within reach (Barkan et al., 2012). Biochemical (Barkan et al., 2012; Kobayashi et al., 2012), statistical (Barkan et al., 2012; Yagi et al., 2013), and evolutionary constraint (Fujii et al., 2011) studies have postulated that the 1st, 3rd, and 6th amino acids within the PPR consensus govern RNA base recognition specificity. However, to put the final polish on the code for de novo prediction of PPR binding RNA sequences, we still need a better understanding of how PPR tracts recognize and bind to RNA.

PROTON GRADIENT REGULATION3 (PGR3) is a PPR protein that was identified through the screening of high-chlorophyll-fluorescence mutants at high light intensity in *Arabidopsis thaliana* (Shikanai et al., 1999). All the results obtained so far are consistent with a model in which PGR3 involves three distinct functions: (i) stabilization of *photosynthetic electron transport L* (*petL*) operon RNA, (ii) translational activation of *petL*, and (iii) translational activation of *ndhA* (Yamazaki et al., 2004; Cai et al., 2011; Figure 1). For the first two functions, PGR3 should bind to the 5′untranslated region (UTR) of *petL* operon RNA, as has been confirmed by in vitro RNA binding assay (Cai et al., 2011). As for the third function, although recombinant PGR3 also bound to the 5′UTR of *ndhA*, in a polysome analysis we were unable to detect the defect, probably because of technical limitations (Cai et al., 2011). We therefore still

[1] Address correspondence to shikanai@pmg.bot.kyoto-u.ac.jp.
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: Toshiharu Shikanai (shikanai@pmg.bot.kyoto-u.ac.jp).
[C] Some figures in this article are displayed in color online but in black and white in the print edition.
[W] Online version contains Web-only data.
www.plantcell.org/cgi/doi/10.1105/tpc.113.112193

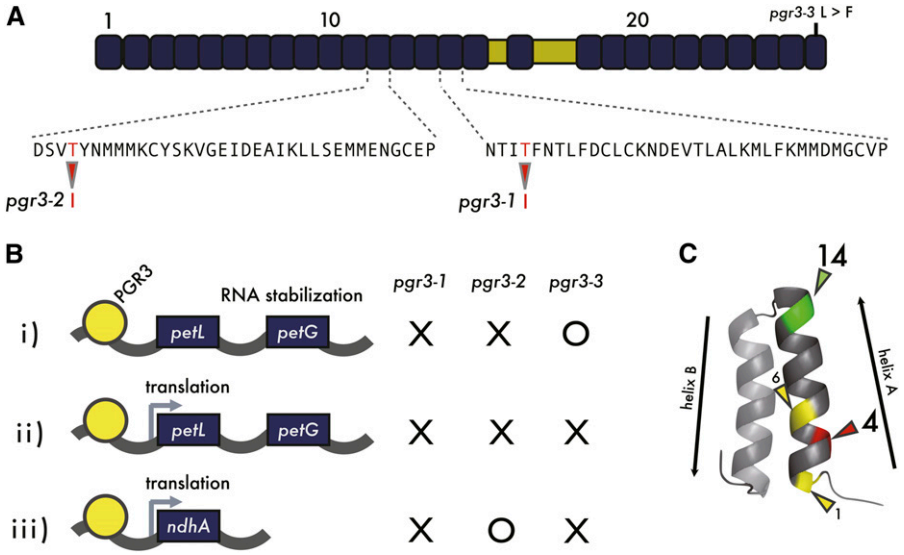

Figure 1. Summary of the Results of Previous Studies on PGR3.

(A) The three *pgr3* mutant alleles carry the amino acid point mutations indicated. *pgr3-1* and *pgr3-2* carry T-to-I replacements at the 4th amino acid position of the 15th and 12th PPR motifs, respectively. *pgr3-3* carries an L-to-F substitution in the last PPR motif.

(B) PGR3 is required for three functions in two distinct targets: stabilization of *petL* operon RNA i), and translation of ii) *petL* and iii) *ndhA*. *pgr3-1* is deficient in all three functions. *pgr3-2* is defective in functions i) and ii), and *pgr3-3* is defective in functions ii) and iii). It is still not certain that *ndhA* is the sole direct target of PGR3 among the 11 plastid-encoded *ndh* genes (Cai et al., 2011).

(C) Spatial demonstration of the 4th amino acid within 35 amino acids of a PPR motif. Known RNA recognition determination residues (1st and 6th amino acids) are indicated in yellow. X-ray structure of the PPR motif was retrieved from human mitochondrial RNA polymerase harboring two PPR motifs (PDB 3SPA) (Ringel et al., 2011).

cannot definitely conclude that the phenotype observed in the chloroplast NADH dehydrogenase-like (NDH) complex is due to a defect in *ndhA* translation (Cai et al., 2011). In *pgr3-1* and *pgr3-2* mutant alleles, lack of the *petL* operon RNA stabilization results in the absence of both PetL and PetG proteins (Yamazaki et al., 2004). *pgr3-1* is further defective in the *ndhA* translation, whereas *pgr3-2* retains function. Although *pgr3-3* mutants did not accumulate NDH, the level of cytochrome *f* was not affected; the phenotype of this mutant was explained by an inability to translate both *petL* and *ndhA*. In *pgr3-3*, PetG is likely translated (Yamazaki et al., 2004), resulting in a clear phenotypic difference from *pgr3-1* and supporting our hypothesis that PetG, rather than PetL, is required to stabilize cytochrome *f*.

pgr3-2 carries a Thr-to-Ile alteration at the 4th amino acid of the 12th PPR motif, resulting in specific loss of *petL* operon accumulation, but normal *ndhA* translation. Thus, the single amino acid alteration changes the RNA binding specificity. More specifically, the chemical properties of the amino acid at the 4th position of the PPR may be critical for PGR3 RNA binding affinity. We took advantage of the distinct phenotypes of these mutant alleles to evaluate the contribution of each PPR motif to PGR3 specificity and function.

RESULTS

Uncharged Polar Amino Acids Can Functionally Replace Thr at the 4th Amino Acid Position

Because the *pgr3-2* mutation did not affect *ndhA* translation (function iii) (Cai et al., 2011), we reasoned that this amino acid replacement (Thr to Ile) at the 4th amino acid position of the PPR motif consensus did not seriously affect the entire protein structure (which would have led to instability of PGR3), but instead specifically affected the RNA affinity of the motif in question. In addition, the 1st and 6th amino acids have been identified as RNA base recognition residues (Barkan et al., 2012). The residues proximal to these amino acids may also be required for RNA binding (Figure 1C); indeed, the 4th amino acid has also been proposed as an important site for RNA binding affinity (Kobayashi et al., 2012). This 4th amino acid is located on the solvent surface of the first α-helix in a PPR motif (Ringel et al., 2011; Kobayashi et al., 2012). However, these residues may not directly influence binding specificity (Fujii et al., 2011; Barkan et al., 2012; Yagi et al., 2013), since statistical test inferred that the 1st and the 6th amino acids serve as principle components for RNA recognition.

To gain an understanding of the chemical properties of the residues required to retain RNA affinity, we introduced a series of point mutations at the 4th amino acid of the 12th PPR of PGR3 (the *pgr3-2* mutation site). Mutant *PGR3* genes encoding nine different amino acid substitutions were introduced into the *pgr3-1* mutant, which lacks all three functions, under the control of the native *PGR3* promoter (Figure 2). Constructs carrying the amino acid replaced with Cys, Asn, or Ser complemented *petL* operon RNA accumulation and consequently cytochrome *f* accumulation (Figures 2A and 2B), although Cys was less preferential for full activity. Other substitutions at the site did not complement the function (Figure 2A), suggesting that the 4th position of the PPR motif should contain uncharged and polar

amino acids for binding to the 5′UTR of *petL* operon RNA. By contrast, all of the constructs complemented the accumulation of NdhK (Figure 2B). All attempts to directly detect PGR3 in vivo have failed, but our results suggest that substitution with other amino acids at this site did not seriously affect the stability of PGR3.

Scanning of PPR Motifs Conferring Versatility on PGR3 Function

Which specific PPR motifs of PGR3 contribute to each of the three hallmark functions (Figure 1)? To answer this question, we systematically substituted the 4th amino acid of each PPR motif (Thr) with Ile, a nonpolar amino acid; this abolished the function of this PPR motif but was unlikely to affect the stability of PGR3 (Figure 2). The series of *PGR3* variants was introduced into *pgr3-1* under the control of the *PGR3* promoter.

To monitor function (i), *petL* operon RNA was detected by RNA gel blots (Figure 3B). Because *petG* is also transcribed from the same promoter as *petL*, *petG* expression depends on function (i). The absence of PetG but not PetL destabilizes the cytochrome b_6f complex (Yamazaki et al., 2004), whose accumulation was monitored by measuring the cytochrome *f* level (Figure 3C). Because the reduced level of the cytochrome b_6f complex restricts linear electron transport, function (i) can be quantitatively evaluated as quantum yields of photosystem II (Figure 3D). Activity of translation was assessed using the antibody against PetL for function (ii) and NdhK and NdhL for function (iii) (Figure 3C). Since absence of NdhA totally destabilizes the chloroplast NDH complex (Peng et al., 2009), function (iii) can be monitored by the antibodies against NdhK and NdhL (Figure 3C).

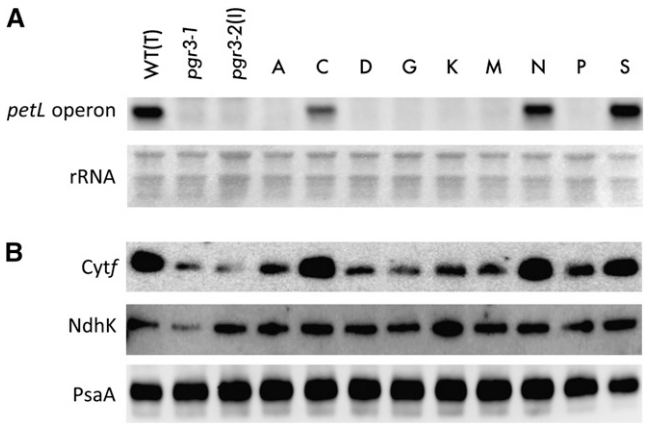

Figure 2. Investigation of Compensatory Amino Acids at the 4th Amino Acid.

(A) RNA gel blot analysis detecting *petL* operon RNA. The 4th amino acid of the 12th PPR motif was replaced by each amino acid indicated and then introduced into *pgr3-1*. The amino acid at this position was replaced by Ile in *pgr3-2*. WT, the wild type.

(B) Immunoblot detection of cytochrome *f*, a subunit of the cytochrome b_6f complex. Cytochrome *f* stability depends on PetG, rather than PetL, in *Arabidopsis* (Yamazaki et al., 2004); the analysis thus monitored function (i) rather than function (ii). NdhK is a subunit of chloroplast NDH and was detected instead of NdhA because the absence of NdhA destabilized the entire NDH complex (Peng et al., 2009). Note that all the PGR3 variants complemented the accumulation of NdhA, suggesting that any substitutions did not dramatically affect PGR3 stability. PsaA abundance is displayed as the loading control.

Three PGR3 variants (20, 21, and 27) stabilized the *petL* operon RNA but did not complement the translation of *petL* or *ndhA* (Figure 3). This result suggests that these PPR motifs are required for translation but are not required for binding to the 5′UTR of *petL* operon RNA. Another two lines, with *PGR3* variant 7 or 23, accumulated NdhK but not PetL (Figure 3). Although the stability of *petL* operon RNA was slightly complemented by introducing variants 7 and 23, the cytochrome *f* level was fully complemented (Figure 3), suggesting that low levels of RNA were sufficient for wild-type levels of *petG* translation. These results suggest that the 7th and 23th PPR motifs are not required *for ndhA* translation but are partially required for *petL* operon stabilization. Because *petL* translation requires its transcript accumulation, we were unable to conclude whether these PPR motifs could be specifically required for *petL* translation (see Discussion). The *petL* RNA level did not recover upon introduction of variant 9, resulting in a reduction in cytochrome *f* level (Figure 3). Puzzlingly, this line accumulated a level of PetL protein equivalent to that in the wild type (Figure 3); the reason for this is discussed later. The mutation at the 8th PPR motif resulted in overaccumulation of *petL* mRNA, implying that this PPR motif partly suppresses the activity to stabilize *petL* operon RNA in the wild type. Photosystem II (PSII) yields reflected the cytochrome *f* levels (Figures 3C and 3D) and were used to analyze multiple transgenic lines quantitatively for the evaluation of function (i) (see Methods for details).

Because we cannot evaluate the protein levels of PGR3 variants, it is possible that the observed differences in phenotypes depended on the level of PGR3 variants in vivo, rather than the function of each PPR motif. If this is the case, the mutant phenotypes depending on the defect in each function should be correlated because all the phenotypes depend on the protein levels. To study this possibility, *petL* RNA accumulation and PetL and NdhK protein levels were plotted against each other (see Supplemental Figure 1 online). We found no significant correlation between *petL* RNA and NdhK protein accumulation (see Supplemental Figure 1B online; $r = 0.28$, $P = 0.09$). This indicates that functional partitioning of PGR3 is not just visualizing the artifact resulting from the different stability among PGR3 variants. By contrast, there were correlations between *petL* RNA and PetL protein abundance, especially when plots were limited to N-terminal 16 PPR motifs (squared plots in Supplemental Figure 1A online). As discussed later, the N-terminal 16 PPR motifs are required for target recognition, whereas the C-terminal 11 PPR motifs likely function in translation. Meanwhile, correlation between PetL and NdhK protein abundances was also observed ($r = 0.60$, $P = 0.001$), and this was probably because the C-terminal 11 PPR motifs functions in a similar manner upon translation of both *petL* and *ndhA* (triangular plots in Supplemental Figure 1C online). We do not eliminate the effects of stability of PGR3 variants on the phenotypes completely, but the majority of observed phenotypes are explained by our model on the functional partitioning of PGR3.

Charges in the 14th Amino Acid Influence the Function of a PPR Motif

We established the method to evaluate the function of a PPR motif in question, leading to the conclusion that the 4th amino

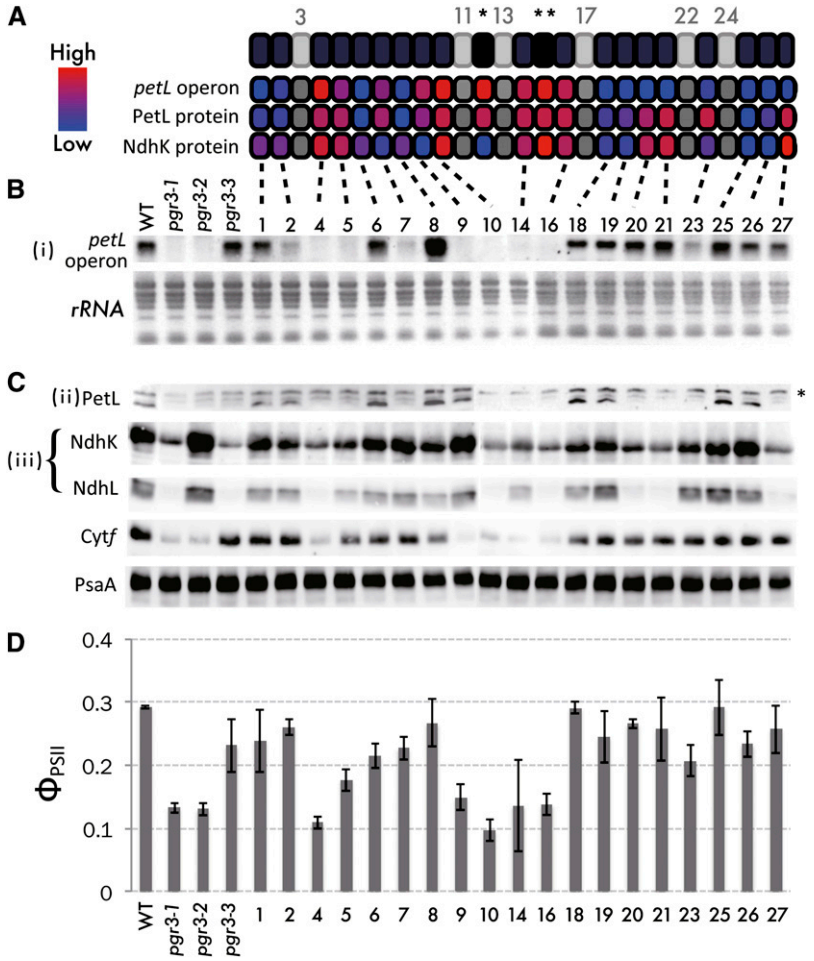

Figure 3. Systematic Replacement of the 4th Thr with Ile in Each PPR Motif.

(A) to **(C)** Protein structure of PGR3 **(A)** and summary of results (**[B]** and **[C]**).

(A) Light gray indicates PPR motifs that were not mutated, as they harbored amino acids other than Thr at the 4th position. Single and double asterisks indicate the PPR motifs mutated in *pgr3-2* and *pgr3-1*, respectively. The contribution of each PPR motif to the three functions was estimated from the RNA and protein abundance relative to that in the wild type, as digitized from the experiments below. When a point mutation severely affects PGR3 function, the color is shown as red (high contribution), but when a point mutation does not cause any defect the color is shown as blue (low contribution).

(B) RNA gel blot analysis of the *petL* operon RNA in transgenic *pgr3-1* lines expressing PGR3 variants carrying Thr-to-Ile replacements at the 4th position of each PPR motif indicated. WT, the wild type.

(C) Immunoblot detection of PetL, two subunits of the chloroplast NDH complex NdhK and NdhL, cytochrome *f*, and PsaA. Asterisk indicates non-specific signals.

(D) Quantum yields of PSII determined at a light intensity of 500 μmol photons m^{-2} s^{-1}. Data represent means of three biological replicates, with standard deviations. The cytochrome *f* level and PSII yield are linked to the PetG level, which could not be tested directly (Yamazaki et al., 2004; Cai et al., 2011); the analysis thus is equivalent to monitoring *petL* operon accumulation.

acid should be uncharged polar to make the motif functionally active (Figure 2). This strategy may be also efficient to determine other amino acid residue(s) or amino acid chemical properties that determine the contribution of each motif to RNA binding or translation. First, we surveyed if there is any correlation between amino acid charges and the accumulation levels of RNA or protein in PGR3 variants. We calculated the average isoelectric points for every two pair of residues in a PPR motif (595 amino acid combinations), and their correlation coefficient was calculated with the levels of *petL* RNA, PetL protein, and NdhK protein and also with the average functional importance estimated (see Methods for details) from all the three functions. As a result, the average isoelectric points of combinations that included the 14th amino acid were prone to have positive correlation with the average functional importance of each motif (Figure 4) and negative correlation with the individual RNA or protein accumulation level (see Supplemental Figure 2 online), suggesting that the residue should be positively charged to make the PPR motif functionally active. Precisely, the functional contributions

of the PPR motifs were higher when a positively charged amino acid (e.g., Lys or Arg) was located at the 14th position in PGR3 (Figure 4B). Functional importance of each PPR motif did not have a significant correlation with the bit scores from the HMM search ($r = -0.21$, $P = 0.18$) or with that from the PROSITE motif scanning ($r = -0.14$, $P = 0.27$) (Figure 4B). This indicates that soundness of the PPR motif, or closeness to the PPR consensus, did not have great influence on its functional importance in PGR3. We also used the molecular weights and the Kyte-Dolittle hydrophobicity scale (Kyte and Doolittle, 1982) of the amino acids as the parameters to correlate with average functional importance (see Supplemental Figure 3 online). The correlation patterns calculated from these parameters were more obscure than that with the pI of the amino acids (Figure 4), showing a mosaic coloring structure that indicates the absence of a particular amino acid position with strong effect.

We attempted to verify the influence experimentally by introducing amino acids with different charges into the 14th amino acid position of the 12th PPR, which is required for the stabilization of *petL* operon RNA (Figure 2). As a result, *petL* operon RNA was equally stabilized by the substitution of this site (Lys in the wild type) with positively charged Arg, slightly less effectively when replaced with His, and much less effectively when non-positively charged amino acids Ile and Asp (Figure 5). Also importantly, differences of the molecular weights or the hydrophobicities of the replaced amino acids do not adequately explain the effectiveness of the *petL* RNA stabilization function of the PGR3 variants (Figure 5C).

Truncation of the Last 11 PPR Motifs Does Not Substantially Affect the *petL* Operon Stabilization Function of PGR3

We noticed that none of the 4th amino acid point mutations within the last 10 PPR motifs (18 to 27) affected the *petL* operon accumulation, although the 23rd PPR was required for full activity (Figure 3). This result suggests that the N-terminal 16 PPR motifs are sufficient for stabilizing *petL* operon RNA. To test this possibility, several versions of truncated PGR3 were expressed in *pgr3-1*. Consistent with our hypothesis, P16, the PGR3 variant carrying only the first 16 N-terminal PPR motifs, showed recovery of *petL* operon RNA accumulation to the same extent as P18 and P26 (Figure 6). However, further truncation of two PPR motifs (P14) impaired this function to the level of *pgr3-1* (Figure 6), suggesting that the first 16 PPR motifs of PGR3 are required, and are sufficient, for *petL* operon stabilization.

To test our hypothesis further, we expressed the truncated versions (P14 and P16) of recombinant PGR3 and tested whether they possess the affinity against the *petL* RNA ligand (see Supplemental Figure 4 online). As expected from the in vivo phenotype, P16 retained the ability to bind to the *petL* RNA with only a slight compromise compared with the full-length PGR3, but P14 did not, suggesting that the functional partitioning structure was reproducible in vitro.

DISCUSSION

By introducing a charged or nonpolar amino acid into the 4th position, we succeeded in specifically impairing the RNA binding affinity of a particular motif without affecting the stability of the whole protein for most of the PPR motifs in PGR3. This approach is supported by the fact that the 4th amino acid is exposed to the putative RNA binding surface of a PPR, as predicted by the PPR-RNA docking model (Fujii et al., 2011) and by in vitro experimental evidence (Kobayashi et al., 2012). A mutation in such a residue is unlikely to affect the protein structure drastically. Furthermore, if amino acid substitutions at the site affect PGR3 function via a reduction in protein stability, we should observe high correlations among the extents of functional disruption in all the three PGR3 functions. This was not the case (see Supplemental Figure 1 online), indicating that the functional partitioning of PGR3 cannot be explained by the different stability of PGR3 variants.

We introduced systematic mutations into the 4th site of a PPR motif, which revealed that some PPR motifs contribute to only one or two of the three proposed functions of PGR3 (Yamazaki et al., 2004; Cai et al., 2011). Not all of the PPR motifs participate in the same function. For instance, the 7th PPR motif was required for *petL* RNA accumulation, or *petL* translation, but the motif was not essential for *ndhA* translation (Figure 3). For several point mutations that severely impaired all three functions of PGR3 (PPR motifs 4, 10, and 16), we cannot exclude the possibility that these mutations destabilized PGR3. More probably, however, these PPR motifs are required for all three functions.

We found that PPR motifs carrying positively charged amino acids at position 14 (Figure 1C) contributed more to PGR3 function (Figure 4; see Supplemental Figure 2 online). This amino acid is located at the bottom of the central groove of the two α-helices of a PPR motif that may interact with the phosphate skeleton of RNA (Small and Peeters, 2000). Yeast PET309 with point mutations in basic amino acids in the proximity of this site was only partially functional; more importantly, double point mutations in basic residues in this region gave stronger mutant phenotype (Tavares-Carreón et al., 2008). Our study suggests, in agreement with the previous studies, that surface charge of the amino acid at the bottom of the central groove provides quantitative information on RNA binding strength. Notably, the statistical analysis also highlighted some other amino acids that may contribute to PPR function (Figure 4). Although not as strong as the 14th amino acid, positive charges on the 19th amino acid might also contribute to RNA binding by a similar mechanism, as they are located nearby. The PPR motifs with negatively charged amino acids at the 29th amino acid were prone to have a stronger influence on PGR3 function (Figure 4). It will be worth further investigating if the charges on these amino acids might also influence PGR3 function, as do the amino acids at position 14.

Mutagenesis within the C-terminal 11 PPR motifs, or even complete depletion of these motifs, did not impair the PGR3 function to stabilize the *petL* operon RNA (Figures 3 and 6). At least PPR motifs 20, 21, and 27 were definitely required for the other two functions, *petL* and *ndhA* translation (Figure 3). A number of PPR proteins have been reported to target multiple RNA sequences (Schmitz-Linneweber et al., 2005; Chateigner-Boutin et al., 2008; Hammani et al., 2009; Pfalz et al., 2009; Zehrmann et al., 2009; Bentolila et al., 2010; Okuda et al., 2010); here, we report of the partitioning of a single PPR protein into regions harboring different functions. A straightforward question arises: What is the function of each region? The first 16 PPR

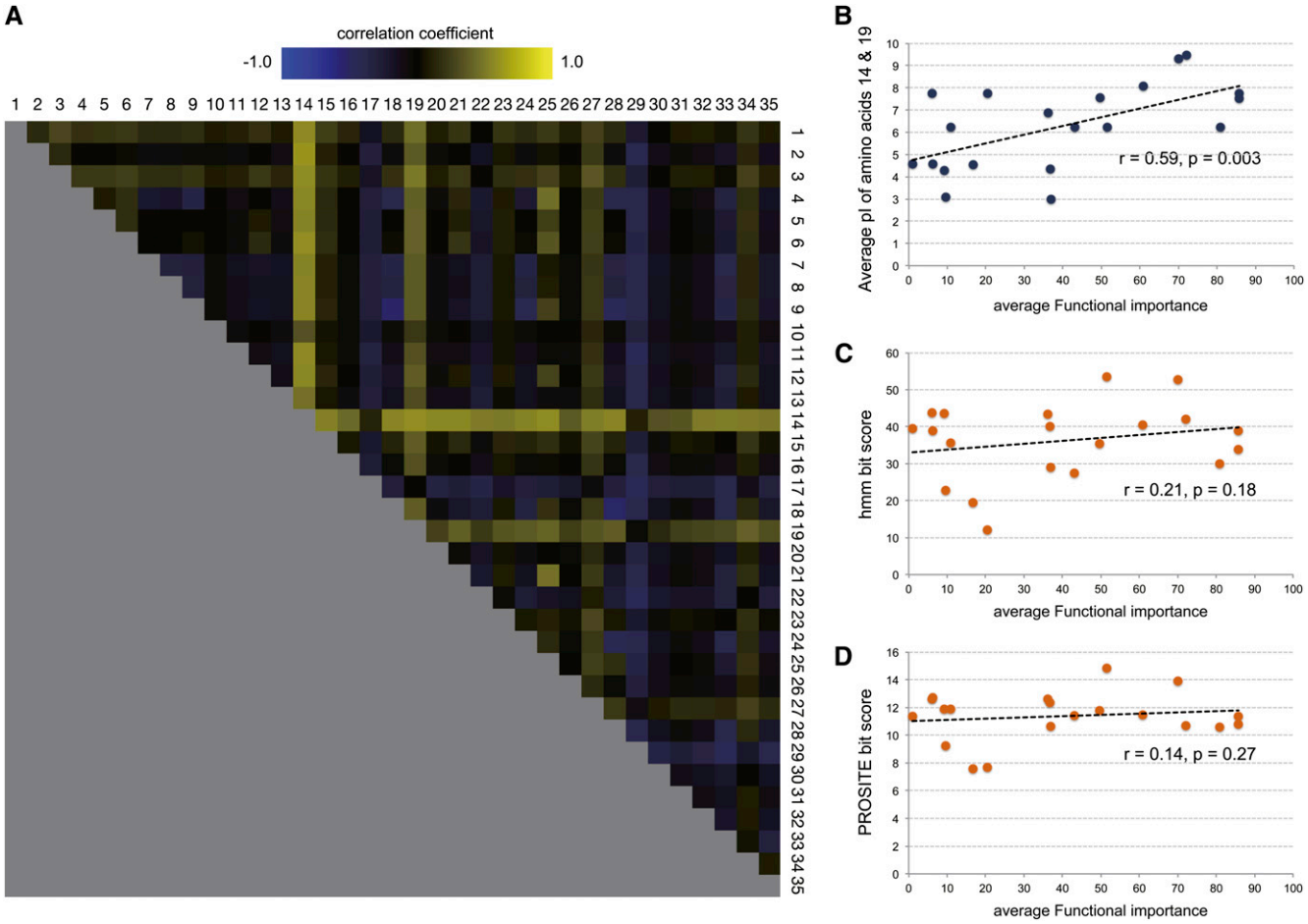

Figure 4. Calculation of Correlation between the pI of Amino Acids and the Functional Importance of Each PPR Motif.

(A) Pearson correlation coefficient calculated for every possible combination of two amino acid positions (595 pairs) and average functional importance equivalent to the average of contribution value of three functions from Figure 3A.
(B) Plots between average pI of amino acid positions, 14 and 19, and the average functional importance. Shown as the example of significant positive correlation between pI of amino acid 14 and average functional importance.
(C) Plots between hmmsearch bit score and the average functional importance.
(D) Plots between PROSITE scan bit score and the average functional importance.
For panels **(B)** to **(D)**, data points from 18 out of 19 PGR3 point mutation transgenic lines generated in this study, *pgr3-1* and *pgr3-2* were used. Data points for PGR3 variants in the last 27th PPR motif (including *pgr3-3*) were not used because the motif is truncated and amino acids 14 and 19 are absent.

motifs are likely essential for RNA binding; this hypothesis was supported by the results of our in vitro RNA binding assay of PGR3-2 (Cai et al., 2011). Recent transcriptomic studies reported the presence of PPR footprints in chloroplast RNA: binding of PPR proteins protects short RNA sequences from exonucleases (Ruwe and Schmitz-Linneweber, 2012; Zhelyazkova et al., 2012). The footprints included the predicted binding sites of PGR3 present in *petL* and *ndhA* 5′UTRs. N-terminal 16 PPR motifs most likely bind to the leader sequence of the *petL* operon and protect the entire transcript from a 5′→3′ exonuclease. This was also supported by this study, demonstrating the correlation between the levels of *petL* RNA and PetL protein when the analyzed region was restricted to the N-terminal 16 PPR motifs rather than the entire PGR3 (see Supplemental Figure 1 online). The N-terminal PPRs likely contribute to *petL* translation via the stabilization of *petL* operon RNA. Because *petG* translation also depends on the stabilization of *petL* operon RNA, the defect in function (i) results in a reduction in cytochrome *f* levels and consequently reduced PSII yields (Figures 3C and 3D). By contrast, the C-terminal 11 PPR motifs have functions specifically required for *petL* and *ndhA* translation (see Supplemental Figure 1C online). How could the 11 C-terminal PPR motifs be involved in translation? A likely scenario is that the C terminus of PGR3 binds the target RNA with less sequence specificity compared with the N terminus, thereby promoting translation via a mechanism such as conformational alteration of RNA secondary structure. For example, PPR10 can release the RNA secondary structure flanking the Shine-Dalgarno sequence of the plastid *atpH* gene, likely

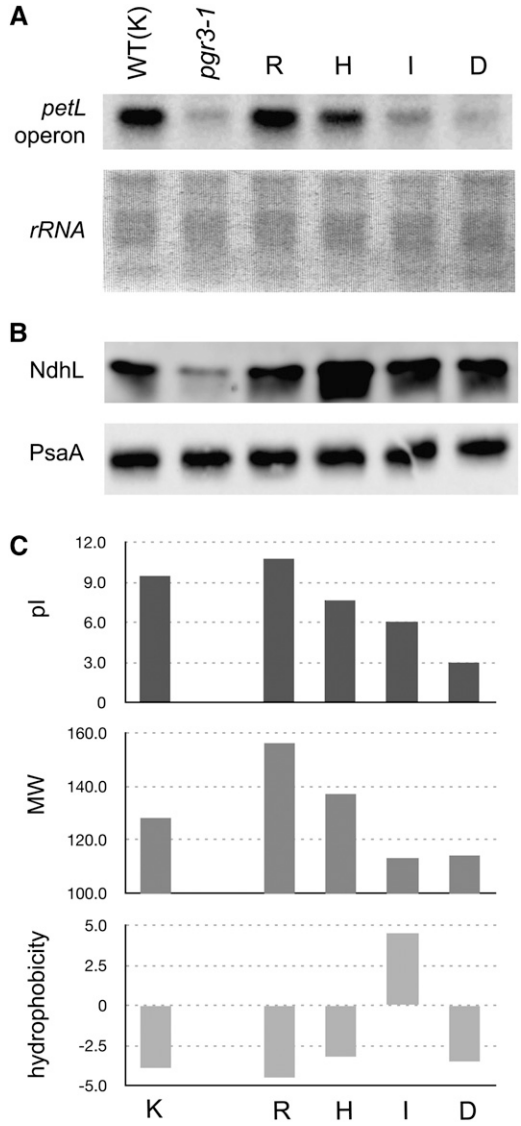

Figure 5. Influence of pI at the 14th Amino Acid of the 12th PPR.

(A) RNA gel blot analysis detecting *petL* operon RNA. The 14th amino acid of the 12th PPR motif was replaced by each amino acid indicated and then introduced into *pgr3-1*. WT, the wild type.
(B) Immunoblot detection of NdhL and PsaA.
(C) pI, molecular weight (MW), and hydrophobicity of each amino acid replaced. pI at pH 7.0 is indicated. Molecular weight is in daltons, hydrophobicity according to the Kyte-Doolittle scale (Kyte and Doolittle, 1982).

facilitating the access of ribosomes to initiate translation (Prikryl et al., 2011). It is also possible that the C terminus sequence recruits cofactors required for translational activation. The Pumilio and FBF homology repeat proteins contain a class of RNA binding motifs, each of which binds a single base in modular fashion, as do PPRs (Wang et al., 2002; Wickens et al., 2002; Quenault et al., 2011). Pumilio and FBF homology proteins can recruit other proteins to destabilize or repress the translation of target RNAs (Goldstrohm et al., 2006). From an evolutionary perspective, these C-terminal sequences could be modules of PPR motifs that are redirected to acquire novel functions (e.g., protein translation) other than specific RNA recognition. This is reminiscent of the function of the E domain, which consists of C-terminal peptides present in a subgroup of PPR genes known as the PLS subfamily (Lurin et al., 2004; Shikanai, 2006; Schmitz-Linneweber and Small, 2008; Fujii and Small, 2011). The sequence structure of the E domain is similar to that of the PPR motif because they conceivably share the same origin, except that the E domain is considered to participate in the core RNA editing mechanism rather than in RNA recognition (Shikanai, 2006; Schmitz-Linneweber and Small, 2008; Fujii and Small, 2011). The correlation between PetL and NdhK protein abundance within the PGR3 variants was stronger in the C-terminal 11 PPR motifs than in the entire PGR3 (see Supplemental Figure 1 online), suggesting that the region is required for the translational mechanism shared by *petL* and *ndhA*. Introduction of the C-terminal 16 PPR motifs of PGR3 fused to its chloroplast targeting peptides into the *pgr3-1* phenotype did not result in complementation of any of the three phenotypes (data not

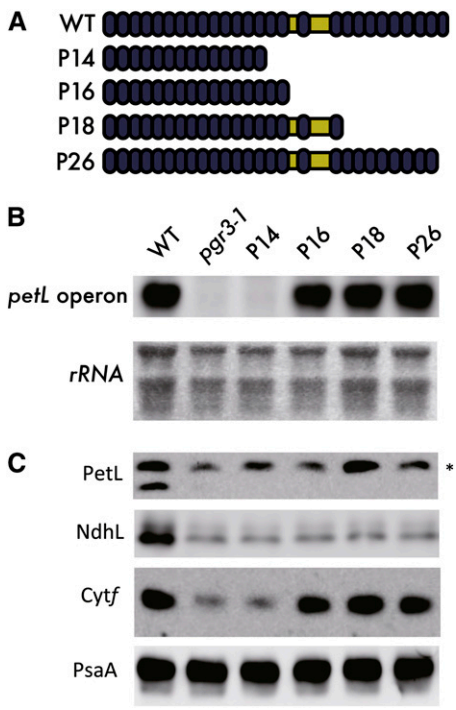

Figure 6. Investigation of the PGR3 Region Essential for Stabilizing the *petL* Operon RNA.

(A) Introduction of truncated PGR3 series into the *pgr3-1* mutant (e.g., P14 indicates the line expressing the 14 N-terminal PPR motifs). WT, the wild type.
(B) and **(C)** Levels of *petL* operon RNA **(B)** and also cytochrome *f* protein **(C)** were determined to monitor the complementation of *petL* transcript stabilization. None of the truncated series recovered the PetL or NdhL accumulation. PsaA level was used as the internal control. Asterisk indicates nonspecific signals.
[See online article for color version of this figure.]

shown). This suggests that the C terminus is unlikely to be able to function on its own without the N terminus.

Our model cannot explain the phenotype of the line with PGR3 variant 9; in this line, PetL accumulated normally without recovery of the *petL* operon RNA level (Figures 3B and 3C). The defect in function (i) resulted in reduced cytochrome *f* level and lower PSII yield, probably via the reduced PetG level (Figures 3C and 3D). Comparison with the results for the line harboring variant 7, which accumulated slightly more *petL* operon RNA but less PetL protein than did the line with variant 9 (Figures 3B and 3C), reveals that the RNA level does not simply determine the protein level. The 9th PPR motif is essential for *petL* RNA stabilization, but its mutation may activate *petL* translation: This PPR motif suppresses *petL* translation in the wild-type protein. Further analysis of polysome-associated RNA in these variant transgenic lines should address this question in the future. Also puzzlingly, *petL* RNA abundance was only partially recovered in the line with variant 23, although truncation of the last 11 PPR motifs did not affect the level of *petL* operon RNA (Figures 3B and 4). Other than the possibility that the mutation in the 23th PPR partially destabilizes PGR3, the mutation may cause the aberrant interference of the motif with the 16 N-terminal PPR motifs upon PGR3–*petL* RNA interaction. We propose that the N-terminal (PPR1 to PPR16) and C-terminal (PPR17 to PPR27) halves of PGR3 have distinct functions, but to fully explain these results we must speculate that there is yet identified intramolecular communication between the two regions.

METHODS

Plant Materials and Growth Conditions

All *pgr3* alleles were obtained by ethyl methanesulfonate mutagenesis, as described in a previous study (Yamazaki et al., 2004). Plants were grown in Metromix potting soil under controlled conditions (light intensity of 40 μmol photons m^{-2} s^{-1}, 16-h-light/8-h-dark cycle at 23°C).

Transgenic *Arabidopsis thaliana* Plants

All site-directed mutagenesis was performed by fusion PCR. Specifically, two DNA fragments carrying the desired point mutations at the overlapping position were amplified with PCR using KOD Plus Neo DNA polymerase (Toyobo). Two fragments were gel extracted, mixed, and subjected to a second PCR amplification to obtain the mutagenized single fragment. DNA fragments harboring the *PGR3* (At4g31850) promoter and the mutagenized *PGR3* coding region sequences were cloned into pDONR-zeo (Life Technologies) or pENTER/D-TOPO (Life Technologies). The *PGR3* promoter corresponds to the 689-bp DNA fragment immediately upstream of the *PGR3* start codon, which reaches to the last exon of the adjacent gene (At4g31840). All of the primers used in this study are listed in Supplemental Table 1 online. These gene fragments were further introduced into pGWB-NB1 (Nakagawa et al., 2007) by homologous recombination using LR clonase II (Life Technologies). Mutations were only introduced into the 19 PPR motifs that carried Thr in the 4th amino acid. Transgenic plants were obtained using an *Agrobacterium tumefaciens* transfection method with Basta resistance selection on Murashige and Skoog medium (Clough and Bent, 1998). We confirmed that the presence of the original mutation of *pgr3-1* and the additional copy of the mutant PGR3 in the transgenic lines by PCR. For all of the constructs used in this study, at least three independent transgenic lines were used for chlorophyll fluorescence analysis that monitors PSII yield and NDH activity. As observed in small standard deviations, PSII yield was not divergent among lines (Figure 3D), indicating that the gene expression driven by the *PGR3* native promoter was sufficient and stable enough. For each construct, representative lines with intermediate values of PSII yield were selected to further quantify protein or RNA. Chlorophyll fluorescence analysis of NDH activity (Okegawa et al., 2008) did not provide quantitative information, but no lines showed a phenotype inconsistent with the others. In summary, a single line represented each construct for the molecular phenotypes (protein and RNA), but the electron transport phenotypes were confirmed in multiple lines.

RNA Preparation and RNA Gel Blot Analysis

RNA was isolated from the leaves of 3-week-old plants using RNAiso (TaKaRa-Bio). Three micrograms of total RNA was electrophoresed and then transferred onto a nylon membrane. The membrane was hybridized with a digoxigenin-labeled DNA probe corresponding to the *petL-petG* genic region (primers used are listed in Supplemental Table 1 online), and the signals were detected with a Gene Image CDP-Star detection kit (GE Healthcare). *rRNA* was visualized by staining with either ethidium bromide or methylene blue. Band intensities were quantified and digitized using ImageJ (Schneider et al., 2012).

Immunoblot Analysis

Immunoblot analysis was performed as described previously (Yamazaki et al., 2004). Protein corresponding to 2.0 μg of chlorophyll was loaded to detect NdhK and NdhL; 0.4 μg was loaded to detect cytochrome *f*, 0.75 μg to detect PetL, and 0.5 μg to detect PsaA. The chlorophyll contents were used to normalize signal intensities. Band intensities of each PGR3 variant line were quantified using relative protein abundance in *pgr3-1*, *pgr3-2*, and *pgr3-3*, which was measured in the previous study using a dilution series (Yamazaki et al., 2004). Signals were digitized using ImageJ (Schneider et al., 2012). Signal saturation was avoided using the saturated-coloring function implemented in LAS-3000 (Fuji-Film).

Chlorophyll Fluorescence Analysis

Quantum yields of PSII were determined with a Mini-PAM fluorometer (Walz) as described previously (Shikanai et al., 1999).

Statistical Analysis to Correlate Amino Acid Chemical Properties with PPR Functions

All of the bioinformatic analysis was done using custom PERL scripts. PPR motifs were identified by hmmsearch (Finn et al., 2011) or by the stand-alone version of PROSITE (Gattiker et al., 2002). In order to evaluate the possible combinatorial effects of independent positions, we calculated the average isoelectric points of two amino acids. Average functional importance of individual PPR motifs was calculated based on the quantification of *petL* RNA, PetL protein, and the NdhK protein abundance of each transgenic line (Figure 3). Specifically, reduction rate of RNA or protein was subtracted from 100%, and the remaining value was considered as the functional contribution rate, as indicated in Figure 3A. The functional contribution rates of all three functions were averaged to obtain the average functional importance. Pearson correlation coefficients between the average functional importance and the amino acid chemical property parameters were calculated using R (R Development Core Team, 2013).

Expression and Purification of Recombinant PGR3 Proteins

The sequences encoding mature PGR3, or the two truncated forms of PGR3 (P14 and P16), were cloned into the *Not*I-*Eco*RI sites of pMAL-c5x (New England Biolabs). Primers used are listed in Supplemental Table 1

online. Constructs were transformed into Rosetta (DE3) pLysS (Novagen). Expression of the maltose binding protein–fused PGR3 fragments was induced by the addition of 0.3 mM isopropyl thio-β-galactoside and incubation at 16°C for 16 h. The cells were disrupted in column buffer (20 mM Tris-HCl, pH 7.4, 100 mM NaCl, and 10 mM β-mercaptoethanol) by sonication. The lysate was centrifuged for 30 min at 18,000*g*, and the supernatant was filtered through a 0.45-μm filter (Millipore). The flow-through was incubated with Amylose Resin (New England Biolabs) for 2 h at 4°C, and the resin was washed with wash buffer (20 mM Tris-HCl, pH 7.4, 1 M NaCl, and 10 mM β-mercaptoethanol) three times. The fused protein was eluted with the wash buffer containing 10 mM maltose.

RNA Gel Mobility Shift Assays

The purified recombinant PGR3 protein fragments were dialyzed against buffer containing 100 mM NaCl, 40 mM Tris-HCl, pH 7.5, 4 mM DTT, and 10% glycerol. RNA binding reaction with designated concentrations of the recombinant proteins were performed under 25°C for 20 min, with the 25-μL scale reaction mixture containing 100 mM NaCl, 40 mM Tris-HCl, pH 7.5, 4 mM DTT, 10% glycerol, 0.1 mg/mL BSA, 0.5 mg/mL heparin, 10 units of RNase inhibitors (Life Technologies), and 1.0 nM *petL* RNA probe labeled with digoxigenin at both 5′ and 3′ ends. RNA sequence of the *petL* RNA probe was 5′- UUAGGGAAGUACUUUAAGAAACAUAUGUAUAA-3′, which harbors the entire RNA footprint observed in the *petL* 5′UTR (Ruwe and Schmitz-Linneweber, 2012) and thus is most likely the minimal PGR3 binding site. Gel electrophoresis under native conditions was performed as described previously (Cai et al., 2011), except that 10% acrylamide gels were used. RNA was electroblotted onto a nylon membrane in 4°C for 1 h. Chemical luminescence was detected using the equivalent method to that used for the RNA gel blot analysis described above.

Accession Number

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under accession number At4g31840 (*PGR3*).

Supplemental Data

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

Supplemental Figure 1. Linear Plots of *petL* RNA Abundance (Function i), PetL Protein Abundance (Function ii), and NdhK Protein Abundance (Function iii) in the Transgenic Lines with T-to-I Point Mutations at Each PPR Motif.

Supplemental Figure 2. Correlation between the Isoelectric Points of Amino Acids and the Accumulation Levels of RNA or Proteins in Each T-to-I Mutation Transgenic Line.

Supplemental Figure 3. Calculation of Correlation between the Molecular Weights and Hydrophobicities of Amino Acids and the Functional Importance of Each PPR Motif.

Supplemental Figure 4. In Vitro RNA Gel-Shift Assay Using the Full-Length Recombinant PGR3 (MBP-FL) and the Truncated Forms (MBP-P16 and MBP-P14).

Supplemental Table 1. List of Primers Used in This Study.

ACKNOWLEDGMENTS

We thank Asako Tahara for her excellent technical assistance. We thank Yuichiro Takahashi and Amane Makino for giving us antibodies. T.S. was supported by Grants 22114509 and 22247005 from the Ministry of Education, Culture, Sports, Science, and Technology of Japan and a grant from the Ministry of Agriculture, Forestry, and Fisheries of Japan (Genomics for Agricultural Innovation; GPN0008). S.F. was supported by the Grant-in-Aid for Japan Society for the Promotion of Science fellows.

AUTHOR CONTRIBUTIONS

S.F. designed research, performed research, analyzed data, and wrote the article. N.S. performed research. T.S. designed research, analyzed data, and wrote the article.

Received March 31, 2013; revised July 16, 2013; accepted August 2, 2013; published August 23, 2013.

REFERENCES

Barkan, A., Rojas, M., Fujii, S., Yap, A., Chong, Y.S., Bond, C.S., and Small, I. (2012). A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. PLoS Genet. **8:** e1002910.

Bentolila, S., Knight, W., and Hanson, M. (2010). Natural variation in Arabidopsis leads to the identification of REME1, a pentatricopeptide repeat-DYW protein controlling the editing of mitochondrial transcripts. Plant Physiol. **154:** 1966–1982.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. Science **326:** 1509–1512.

Cai, W., Okuda, K., Peng, L., and Shikanai, T. (2011). PROTON GRADIENT REGULATION 3 recognizes multiple targets with limited similarity and mediates translation and RNA stabilization in plastids. Plant J. **67:** 318–327.

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.

Clough, S.J., and Bent, A.F. (1998). Floral dip: A simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:** 735–743.

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.

Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: Interactive sequence similarity searching. Nucleic Acids Res. **39** (Web Server issue)**:** W29–W37.

Fujii, S., Bond, C.S., and Small, I.D. (2011). Selection patterns on restorer-like genes reveal a conflict between nuclear and mitochondrial genomes throughout angiosperm evolution. Proc. Natl. Acad. Sci. USA **108:** 1723–1728.

Fujii, S., and Small, I. (2011). The evolution of RNA editing and pentatricopeptide repeat genes. New Phytol. **191:** 37–47.

Gattiker, A., Gasteiger, E., and Bairoch, A. (2002). ScanProsite: A reference implementation of a PROSITE scanning tool. Appl. Bioinformatics **1:** 107–108.

Goldstrohm, A.C., Hook, B.A., Seay, D.J., and Wickens, M. (2006). PUF proteins bind Pop2p to regulate messenger RNAs. Nat. Struct. Mol. Biol. **13:** 533–539.

Hammani, K., Okuda, K., Tanz, S.K., Chateigner-Boutin, A.L., Shikanai, T., and Small, I. (2009). A study of new *Arabidopsis*

chloroplast RNA editing mutants reveals general features of editing factors and their target sites. Plant Cell **21:** 3686–3699.

Kobayashi, K., Kawabata, M., Hisano, K., Kazama, T., Matsuoka, K., Sugita, M., and Nakamura, T. (2012). Identification and characterization of the RNA binding surface of the pentatricopeptide repeat protein. Nucleic Acids Res. **40:** 2712–2723.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:** 105–132.

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.

Moscou, M.J., and Bogdanove, A.J. (2009). A simple cipher governs DNA recognition by TAL effectors. Science **326:** 1501.

Nakagawa, T., et al. (2007). Improved Gateway binary vectors: High-performance vectors for creation of fusion constructs in transgenic analysis of plants. Biosci. Biotechnol. Biochem. **71:** 2095–2100.

Nakamura, T., Yagi, Y., and Kobayashi, K. (2012). Mechanistic insight into pentatricopeptide repeat proteins as sequence-specific RNA-binding proteins for organellar RNAs in plants. Plant Cell Physiol. **53:** 1171–1179.

Okegawa, Y., Kagawa, Y., Kobayashi, Y., and Shikanai, T. (2008). Characterization of factors affecting the activity of photosystem I cyclic electron transport in chloroplasts. Plant Cell Physiol. **49:** 825–834.

Okuda, K., Hammani, K., Tanz, S.K., Peng, L., Fukao, Y., Myouga, F., Motohashi, R., Shinozaki, K., Small, I., and Shikanai, T. (2010). The pentatricopeptide repeat protein OTP82 is required for RNA editing of plastid ndhB and ndhG transcripts. Plant J. **61:** 339–349.

Okuda, K., and Shikanai, T. (2012). A pentatricopeptide repeat protein acts as a site-specificity factor at multiple RNA editing sites with unrelated cis-acting elements in plastids. Nucleic Acids Res. **40:** 5052–5064.

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., Fukao, Y., Fujiwara, M., Takami, T., and Shikanai, T. (2009). Efficient operation of NAD(P)H dehydrogenase requires supercomplex formation with photosystem I via minor LHCI in *Arabidopsis*. Plant Cell **21:** 3623–3640.

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.

Prikryl, J., Rojas, M., Schuster, G., and Barkan, A. (2011). Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. Proc. Natl. Acad. Sci. USA **108:** 415–420.

Quenault, T., Lithgow, T., and Traven, A. (2011). PUF proteins: Repression, activation and mRNA localization. Trends Cell Biol. **21:** 104–112.

R Development Core Team (2013). R: A Language and Environment for Statistical Computing. (Vienna, Austria: R Foundation for Statistical Computing).

Ringel, R., Sologub, M., Morozov, Y.I., Litonin, D., Cramer, P., and Temiakov, D. (2011). Structure of human mitochondrial RNA polymerase. Nature **478:** 269–273.

Ruwe, H., and Schmitz-Linneweber, C. (2012). Short non-coding RNA fragments accumulating in chloroplasts: Footprints of RNA binding proteins? Nucleic Acids Res. **40:** 3106–3116.

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.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods **9:** 671–675.

Shikanai, T. (2006). RNA editing in plant organelles: Machinery, physiological function and evolution. Cell. Mol. Life Sci. **63:** 698–708.

Shikanai, T., Munekage, Y., Shimizu, K., Endo, T., and Hashimoto, T. (1999). Identification and characterization of *Arabidopsis* mutants with reduced quenching of chlorophyll fluorescence. Plant Cell Physiol. **40:** 1134–1142.

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.

Tavares-Carreón, F., Camacho-Villasana, Y., Zamudio-Ochoa, A., Shingú-Vázquez, M., Torres-Larios, A., and Pérez-Martínez, X. (2008). The pentatricopeptide repeats present in Pet309 are necessary for translation but not for stability of the mitochondrial COX1 mRNA in yeast. J. Biol. Chem. **283:** 1472–1479.

Wang, X., McLachlan, J., Zamore, P.D., and Hall, T.M. (2002). Modular recognition of RNA by a human pumilio-homology domain. Cell **110:** 501–512.

Wickens, M., Bernstein, D.S., Kimble, J., and Parker, R. (2002). A PUF family portrait: 3'UTR regulation as a way of life. Trends Genet. **18:** 150–157.

Yagi, Y., Hayashi, S., Kobayashi, K., Hirayama, T., and Nakamura, T. (2013). Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. PLoS ONE **8:** e57286.

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

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.

Zhelyazkova, P., Hammani, K., Rojas, M., Voelker, R., Vargas-Suárez, M., Börner, T., and Barkan, A. (2012). Protein-mediated protection as the predominant mechanism for defining processed mRNA termini in land plant chloroplasts. Nucleic Acids Res. **40:** 3092–3105.