CHAPTER 14

Analysis of Chloroplast RNA Editing Sites in *Phalaenopsis aphrodite*

Ching-Chun Chang*,‡, Sin-Chung Liao† and Wun-Hong Zeng*

RNA editing sites were systematically examined for the transcripts of 74 known protein-coding genes in the chloroplasts of *P. aphrodite*. A total of 44 editing sites were identified in 24 transcripts, the highest number reported in seed plants to date. In addition, 21 editing sites were unique to *Phalaenopsis* as compared to other seed plants. All editing was C-to-U conversion, and 42 editing sites caused the change in amino acids. One of the remaining two editing sites occurred in transcripts of the *ndhB* pseudogene, and another in the 5′ untranslated region of *psbH* transcripts. However, RNA editing did not restore the continuous open reading frame in the frameshifted *ndh* genes, further confirming that they are pseudogenes.

14.1. Introduction

14.1.1. *RNA editing and regulation of gene function*

Chloroplast RNA editing, generally manifested as C-to-U conversion, but with U-to-C conversion also reported in a few cases, is one of the post-transcriptional regulation mechanisms of gene expression in land plants. Plastid RNA editing was first documented in maize *rpl2* transcripts for the creation of the initiation codon by ACG-to-AUG conversion, and since then instances of the generation and removal of the

*Institute of Biotechnology, National Cheng Kung University, Tainan 701, Taiwan.
†Department of Biological Science and Technology, Meiho Institute of Technology, Pingtung 912, Taiwan.
‡Corresponding author. E-mail: chingcc@mail.ncku.edu.tw

translational stop codon have also been identified (Bock, 2000; Fiebig *et al.*, 2004). However, editing sites generally reside at the internal positions in transcripts and most frequently affect the second codon positions. Sometimes they alter the first or third codon positions (Bock, 2000). Typically, the codon changes resulting from RNA editing restore the identity of conserved amino acids in plant phylogeny, suggesting that the conserved residues are important for the function of the protein. Previously, by assaying the activity of acetyl-CoA carboxylase expressed from unedited or edited *accD* mRNA in *E. coli*, it was suggested that editing is necessary for a functional enzyme (Sasaki *et al.*, 2001). In addition, transplastomic approaches using tobacco or *Chlamydomonas reinhardtii* have also demonstrated that the lack of specific RNA editing in *psbF* or *petB* transcripts, respectively, leads to a severe mutant phenotype (Bock *et al.*, 1994; Zito *et al.*, 1997). However, the editing frequency of chloroplast transcripts is dramatically affected by tissue types, developmental stages and environmental factors (Bock, 2000).

14.1.2. *Cis-elements surrounding the editing sites*

Chloroplast RNA editing has been systematically investigated for the protein-coding transcripts in the following plant species: dicot plants *Nicotiana tobacum* (Hirose *et al.*, 1999; Sasaki *et al.*, 2003, 2006), *Arabidopsis thaliana* (Lutz and Maliga, 2001; Tillich *et al.*, 2005), *Atropa belladonna* (Schmitz-Linneweber *et al.*, 2002), *Pisum sativum* (Inada *et al.*, 2004), *Solanum lycopersicum* (Kahlau *et al.*, 2006); monocot plants *Zea mays* (Maier *et al.*, 1995), *Oryza sativa* (Corneille *et al.*, 2000), *Saccharum officinarum* (Calsa Junior *et al.*, 2004); *Phalaenopsis aphrodite* (Zeng *et al.*, 2007); gymnosperm *Pinus thunbergii* (Wakasugi *et al.*, 1996); fern *Adiantum capillus-veneris* (Wolf *et al.*, 2004); and hornwort *Anthoceros formosae* (Kugita *et al.*, 2003). Genome-wide analysis of chloroplast transcripts has revealed that plants have undergone dramatic changes in both the levels and patterns of editing, from hornworts (1.2 percent conversions of all nucleotides examined) and ferns (0.38 percent) to seed plants (less than 0.05 percent) (Kugita *et al.*, 2003; Wolf *et al.*, 2004). In seed plants, a relatively constant number

of editing sites, 21 to 44, have been identified in plastids (Tsudzuki *et al.*, 2001; Sasaki *et al.*, 2003, 2006; Kahlau *et al.*, 2006; Zeng *et al.*, 2007). On comparing editing sites among dicots (tobacco) and monocots (maize and rice), 12 common editing sites between tobacco and monocot plants, and 20 common sites between the two monocots were identified (Tsudzuki *et al.*, 2001). However, when looking at more closely related taxa, the number of shared sites increases. For instance, maize, rice, and sugarcane in the Poaceae family share at least 23 editing sites (Calsa Junior *et al.*, 2004), while tobacco, tomato, and *Atropa* in the Solanaceae family share 30 (Schmitz-Linneweber *et al.*, 2002; Kahlau *et al.*, 2006). Furthermore, 31 conserved sites out of a total of 35 editing sites were identified among three species of tobacco, *N. tobacum*, *N. sylvestris*, and *N. tomentosiformis* (Sasaki *et al.*, 2003). Moreover, at a subspecies level, three different ecotypes of *Arabidopsis thaliana* have all 28 editing sites in common, though the consequences of RNA editing differ at one position between the ecotypes (Tillich *et al.*, 2005). These studies suggest that determining the distribution and pattern of editing sites across taxa and across the entire chloroplast genome is an important step in investigating the evolutionary process of RNA editing in angiosperms. However, these plants still represent a poor sample of major clades in the phylogeny of seed plants.

Although plastids of seed plants have a relatively constant number of editing sites (21 to 44), apparently conserved *cis*-elements surrounding plastid editing sites are scarce. The regions respectively about 30 and 10 nucleotides immediately upstream and downstream of the editing sites have been mapped for selection of the correct editing target and for editing efficiency by transplastomic and *in vitro* approaches (Bock, 2000; Chateigner-Boutin and Hanson, 2003).

14.1.3. *Protein factors involved in RNA editing*

The *cis*-elements are recognized by nuclear-encoded *trans*-acting factors that are believed to be either site-specific or bind to small clusters of related sites (Hirose and Sugiura, 2001; Chateigner-Boutin and Hanson, 2003; Miyamoto *et al.*, 2004). A biochemical approach to

UV cross-linking RNA template with chloroplast extracts *in vitro* has identified several proteins, e.g., 25-, 56- and 70-kD proteins, which are potentially involved in site-specific RNA editing and can bind to specific *cis*-elements (Hirose and Sugiura, 2001; Miyamoto *et al.*, 2002). Using a genetic approach focusing on NADH dehydrogenase activity, two pentatricopeptide repeat proteins, (PPR)-CRR4 and CRR21, acting as a required site-specific recognition factor for editing at the translational initiation codon and a second site in *ndhD* transcripts, respectively, were identified in *Arabidopsis* (Okuda *et al.*, 2006, 2007). Furthermore, using a reverse genetic approach to screen T-DNA insertion mutants, six PPR proteins that account for nine editing sites in the chloroplasts of *Arabidopsis* were identified (Hammani *et al.*, 2009). The PPR family, characterized by tandem arrays of the 35-amino-acid motif, with more than 450 members in *Arabidopsis*, is involved in post-transcriptional processes such as RNA splicing, stability, editing and translation in plastid and mitochondria (Schmitz-Linneweber and Small, 2008). Those PPR proteins involved in plastid RNA editing belong to the plant specific E and DYW subgroups in the PLS subfamily (Okuda *et al.*, 2007; Hammani *et al.*, 2009).

The Orchidaceae, with approximately 30,000 species, is one of the largest families in flowering plants, and *Phalaenopsis aphrodite* subsp.*formosana* is the first species in which the chloroplast genome has been completely determined (Chang *et al.*, 2006). Previously, the ACG rather than an ATG codon at the translation initiation sites was observed in the *rpl2* and *ndhD* genes of *P. aphrodite* (Chang *et al.*, 2006). To confirm the presence of the RNA editing system, we study the RNA editing patterns for the plastid transcripts of all known protein-coding genes in *P. aphrodite* and compare them with those of other seed plants.

14.2. Determination of Chloroplast RNA Editing Sites in *P. aphrodite*

14.2.1. *RNA extraction and RT-PCR*

Leaves of *P. aphrodite* subsp.*formosana* were obtained from seedlings at the two-leaf stage of development, from which total cellular RNA

was isolated according to the method described (Gehrig *et al.*, 2000). The RNA samples were then treated with DNase I (Promega, USA) for 30 min at 37°C to eliminate DNA contamination. To demonstrate the absence of DNA in the RNA preparation after DNase I digestion, the RNA quality was further checked by polymerase chain reaction (PCR) with at least three pairs of primers from plastid genes of the *Phalaenopsis* orchid. Primers were designed based on 74 known protein-coding genes encoded by the chloroplast genome of *P. aphrodite* (AY916449) using Vector NTI Suite software (InforMax, USA). Reverse transcriptase (Promega, USA) was used to synthesize cDNA from total RNA using a reverse primer for each gene at 42°C for 1 h. We multiplexed primers for up to 11 genes within a single reverse transcription (RT) reaction, and at least two independent RT reactions were performed. PCR was then applied to amplify cDNA with a primer pair of both forward and reverse primers for each gene separately. Genomic DNA was used as a template in a positive control to ensure the primer pairs were effective for PCR. The PCR reaction contained a final concentration of 200 nM gene-specific primers, 200 nM of each dNTP, three units of *Taq* DNA polymerase, and 5 μl of 10X*Taq* DNA polymerase buffer in a 50-μl reaction mixture. Amplification started with one 2-min cycle at 94°C, followed by 35–40 cycles of 1.5 min at 94°C, 2 min at 60°C, and 3 min at 72°C, and this was followed by one 5-min cycle at 72°C. Each PCR sample was electrophoresed on a 0.8 percent agarose gel and visualized by staining with ethidium bromide. PCR products were sliced from the gel and purified with a gel extraction kit (Viogene, Taiwan).

14.2.2. *Determination of RNA editing sites*

The cDNA fragments were either directly sequenced or first cloned into pGEM-T vector (Promega, USA), and then the propagated plasmid DNA was sequenced. The sequencing reaction was performed using the BigDye terminator cycle sequencing kit (Applied Biosystems, USA), according to the protocol recommended by the manufacturer. The DNA sequencer was an Applied Biosystems ABI 3700. To determine RNA editing sites, the cDNA sequences were then aligned to that of genomic DNA.

14.3. Chloroplast RNA Editing Patterns in *P. aphrodite*

14.3.1. *Overall properties of chloroplast RNA editing in P. aphrodite*

The chloroplast genome of *P. aphrodite* subsp.*formosana* contains 110 genes including 34 RNA genes, 74 known protein-coding genes and 2 conserved reading frames with unknown function (Chang *et al.* 2006). To extensively study the pattern of RNA editing, we sequenced 60,651 bp of cDNA representing the 74 known chloroplast protein-coding transcripts of *P. aphrodite* (Zeng *et al.*, 2007). A total of 44 editing sites were identified in the 24 transcripts of *P. aphrodite* chloroplast genes, which represented an average of 0.07 percent of the nucleotides examined (Table 14.1). This is the highest number of RNA editing sites reported in seed plants to date. All the RNA editing sites were of the C-to-U conversion type. Of the 42 sites that involved codons, four (9.5 percent) were in the first position, and 38 (90.5 percent) were in the second position; all resulted in the substitution of one amino acid for another (Table 14.1). This result is consistent with previous reports regarding the patterns of RNA editing across widely divergent taxa, which show a bias in favor of second codon position edits (Bock, 2000). The consequence of RNA editing in codon position for *Phalaenopsis* is to mostly restore the conservation of amino acids with other seed plants. The most frequently edited codon was Ser converted to Leu, followed by Ser to Phe, and Pro to Leu (Table 14.1). One of the remaining two editing sites occurred in transcripts of the *ndhB* pseudogene, and another in the 5′ untranslated region of *psbH* transcripts. Among 44 sites, seven partially edited sites were detected in the transcripts of *atpA* (site 2), *clpP* (site 2), *ndhB*, *psbF*, *rpoA* (sites 1, 3) and *rps8* genes in the *Phalaenopsis* orchid (Table 14.1). This is not surprising, since RNA editing efficiency has been reported to vary in different organs, developmental stages, and environmental conditions (Ruf and Kossel, 1997; Chateigner-Boutin and Hanson, 2003). In addition, there are no apparently conserved *cis*-elements surrounding the 44 editing sites in *Phalaenopsis*; however, a U_A context bias immediately before and after the editing site were observed (Fig. 14.1).

Table 14.1 RNA Editing Sites in Chloroplast Transcripts of the *Phalaenopsis* Orchid

Gene	Site	Nucleotide Position	Codon Position	Edited Codon	Amino Acid Change
accD	1^{φ}	1184	395	uCa	S → L
	2^{φ}	1412	471	cCa	P → L
	3^{φ}	1430	477	cCu	P → L
atpA	1	773	258	uCa	S → L
	2^{*}	1148	383	uCa	S → L
atpB	1^{φ}	1184	395	uCa	S → L
atpF	1	92	31	cCa	P → L
atpI	1^{φ}	428	143	cCu	P → L
	2^{φ}	629	210	uCa	S → L
clpP	1^{φ}	82	28	Cau	H → Y
	2^{*}	559	187	Cau	H → Y
matK	1^{φ}	533	178	uCu	S → F
	2^{φ}	718	240	Cau	H → Y
	3	1066	356	Cac	H → Y
ndhB$^{\#}$	1^{*}	1977			C → U
petB	1	611	204	cCa	P → L
petL	1	5	2	cCu	P → L
psaI	1^{φ}	80	27	uCu	S → F
psbF	1^{*}	77	26	uCu	S → F
psbH	1^{φ}	−30			C → U
rpl2	1	2	1	aCg	T → M
rpl23	1^{φ}	71	24	uCu	S → F
rpoA	$1^{*\varphi}$	200	67	uCu	S → F
	2^{φ}	368	123	uCa	S → L
	3^{*}	830	277	uCa	S → L
rpoB	1^{φ}	92	31	uCc	S → F
	2	401	134	uCu	S → F
	3	536	179	uCg	S → L
	4	614	205	uCa	S → L
	5	629	210	uCg	S → L
	6	686	229	cCg	P → L
	7	2489	830	uCa	S → L
rpoC1	1	62	21	uCa	S → L
	2^{φ}	203	68	uCu	S → F
	3^{φ}	509	170	uCa	S → L
	4^{φ}	638	213	uCg	S → L
rpoC2	1^{φ}	2846	949	uCu	S → F
rps2	1	134	45	aCa	T → I

(*Continued*)

Table 14.1 (*Continued*)

Gene	Site	Nucleotide Position	Codon Position	Edited Codon	Amino Acid Change
rps8	1*	182	61	uCa	S → L
rps14	1	149	50	uCa	S → L
rps16	1♀	143	48	uCa	S → L
ycf3	1	44	15	uCu	S → F
	2	185	62	aCg	T → M
	3♀	191	64	cCa	P → L
Total	44				42

#Pseudogene.
*Partial editing.
♀Unique to *Phalaenopsis*.

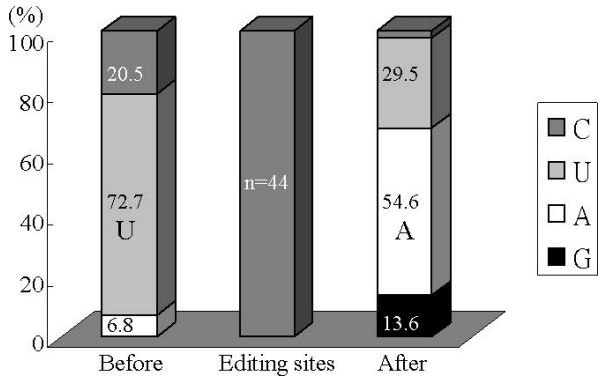

Fig. 14.1. Nearest-neighbor bias towards a U_A context immediately before and after RNA editing sites in the chloroplast transcripts of *P. aphrodite*. The numbers indicate the frequency (percent) of specific nucleotides immediately before or after editing sites. The total number of RNA editing sites is indicated by *n*.

14.3.2. *RNA editing in transcripts encoding for genetic apparatuses*

Transcripts of the plastid *rpoA, rpoB, rpoC1* and *rpoC2* genes are the most extensively edited (15 sites) among the functional gene groups in *P. aphrodite*. Seven editing sites are unique to *Phalaenopsis* orchids.

Two unique sites at Ser^{67} (UCU) and Ser^{123} (UCA) were identified in the *rpoA* transcripts, with conversion to Phe^{67} (UUU) and Leu^{123} (UUA), respectively. Both editing events result in the restoration of codon conservation in plants. There are seven editing sites in *rpoB* transcripts, the most of any chloroplast transcripts in the *Phalaenopsis* orchid or in the *rpoB* transcripts of the other angiosperm chloroplasts studied to date. Interestingly, the first editing site with conversion of Ser^{31} (UCC) to Phe^{31} (UUC) is the only site unique to the *Phalaenopsis* orchid and results in a codon diversification from other angiosperms, but it is conserved with the gymnosperm black pine. In the *Phalaenopsis* orchid, four of the seven editing sites (sites 3 to 6) in *rpoB* transcripts are the same as in maize, clustered in the region corresponding to the dispensable domain I of *E. coli* RNA polymerase β subunit (Corneille *et al.*, 2000). Three unique RNA editing sites were identified in the *rpoC1* transcripts, converting Ser^{68} (UCU) to Phe^{68} (UUU), Ser^{170} (UCA) to Leu^{170} (UUA), and Ser^{213} (UCG) to Leu^{213} (UUG), respectively. One unique RNA editing site was identified in *rpoC2* transcripts, converting Ser^{949} (UCU) to Phe^{949} (UUU). Conversion of the above codon position in the *rpoC1* and *rpoC2* transcripts generally leads to codon conservation among seed plants. The transcripts of six ribosomal protein-coding genes, *rpl2, rpl23, rps2, rps8*, *rps14* and *rps16*, are converted by RNA editing with one codon for each transcript in *Phalaenopsis*. However, there are two unique editing sites in this orchid. One occurs at the Ser^{24} (UCU) codon in *rpl23* transcripts, and the other occurs at the Ser^{48} (UCA) codon in *rps16* transcripts, and they converted to Phe^{24} (UUU) and Leu^{48} (UUA), respectively. C-to-U conversion at these two sites restores codon conservation among seed plants. Interestingly, compared with previously analyzed seed plants, the 103rd codon of *rpl20* transcripts is Leu (UUA), either with or without a C-to-U conversion from Ser (UCA), but the corresponding codon of *rpl20* transcripts in *Phalaenopsis* is Ser, and no apparent RNA editing was observed by directly sequencing the RT-PCR products (Zeng *et al.*, 2007). This suggests that independent loss of this editing site might have occurred in *Phalaenopsis* during the evolutionary process.

14.3.3. *RNA editing in transcripts encoding for photosynthetic apparatuses*

In *P. aphrodite*, concerning the genes encoding subunits in the complexes (or assembly the complexes) of the photosynthetic electron transport chain, five of them, *psaI, psbF*, *psbH*, *petB* and *petL*, have transcripts with one RNA editing site, and *ycf3* transcripts have three sites. Three editing sites were unique to the *Phalaenopsis* orchid. The first was at the Ser^{27} (UCU) codon in *psaI* transcripts, and the second was at the Pro^{64} (CCA) codon in the *ycf3* transcripts. They changed to Phe^{27} (UUU) and Leu^{64} (CUA) in their respective transcripts, leading to amino acid conservation in plants. The third unique editing site resided in the -30 nucleotide (C) position in the upstream untranslated region (5′ UTR) of *psbH* transcripts, and this nucleotide position is located at the variable region when the 5′ UTR nucleotide sequences of *psbH* transcripts from eight species of seed plants were aligned (Fig. 14.2a). Previously, a C-to-U conversion identified at nucleotide position -10 of the *ndhG* 5′ UTR in monocot plants was predicted to modify the RNA secondary (stem/loop) structure (Drescher *et al.*, 2002). To find out if C-to-U conversion in the 5′ UTR of *psbH* transcripts in *Phalaenopsis* also affected the RNA secondary structure, the sequence extending 32 nt to each side of the editing site (-62 nt to $+3$ nt) was examined in its edited and unedited form with the RNA shapes prediction program (Giegerich *et al.*, 2004). The result suggests that the edited RNA sequence can form an energetically less stable secondary structure (Fig. 14.2b). Therefore, it is possible that editing in the *psbH* 5′ UTR might influence *psbH* expression. C-to-U conversions at the Ser^{26} (UCU) to Phe^{26} (UUU) and Pro^{204} (CCA) to Leu^{204} (CUA) codons were identified in the *psbF* and *petB* transcripts of *Phalaenopsis*, respectively. The functional importance of the corresponding editing events has been previously demonstrated in transplastomic tobacco and *Chlamydomonas reinhardtii*, respectively, in which the lack of RNA editing causes a severe mutant phenotype (Bock, 2000). Concerning the six plastid genes involved in the ATP synthase complex, four of them, *atpA*, *atpB*, *atpF* and *atpI*, have transcripts that are edited at six sites in *Phalaenopsis*. Three sites are unique to the orchid. One occurs at the Ser^{395}

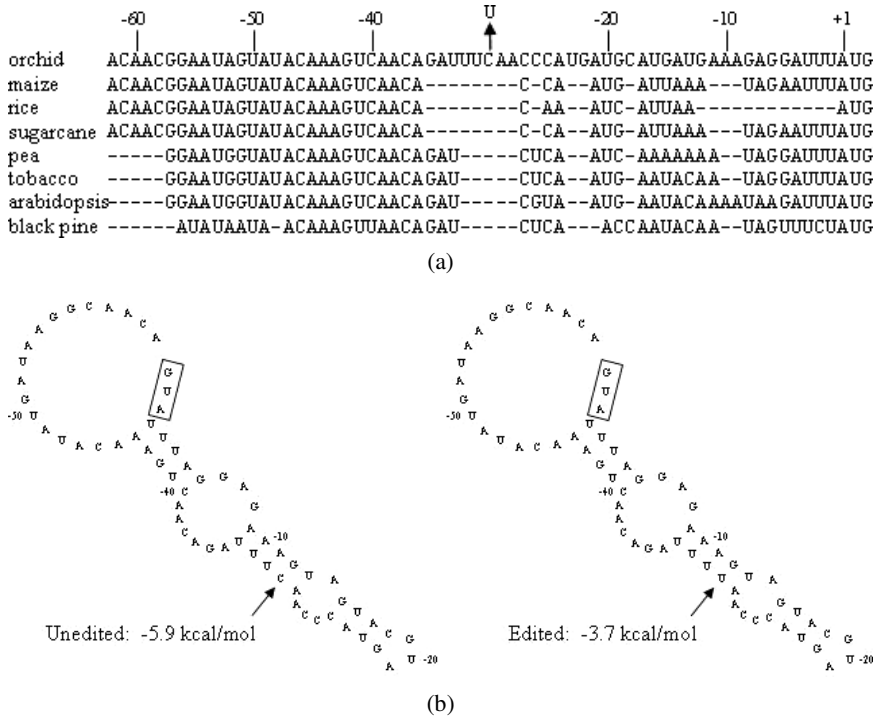

Fig. 14.2. RNA editing site of *psbH* transcripts in the chloroplasts of *P. aphrodite*. (a) The nucleotide sequences in the 5′ untranslated region of *psbH* transcripts from eight species of seed plants were aligned by GeneDoc. Position from translational start site A (+1) is indicated. RNA editing at position −30 is converted from C to U in the orchid as indicated by the arrow. Orchid: *P. aphrodite*; maize: *Zea mays*; rice: *Oryza sativa*; sugarcane: *Saccharum officinarum*; pea: *Pisum sativum*; tobacco: *Nicotiana tabacum*; arabidopsis: *Arabidopsis thaliana*; black pine: *Pinus thunbergii*. (b) Predicted RNA secondary structures formed by the unedited and edited *psbH* 5′UTR using the RNAshapes algorithm. The edited and unedited nucleotides are indicated by arrows. The *psbH* translational start codon is boxed.

(UCA) codon in *atpB* transcripts, while the other two occur at the Pro^{143} (CCU) and Ser^{210} (UCA) codons in *atpI* transcripts, and they convert to Leu^{395} (UUA), Leu^{143} (CUU), and Leu^{210} (UUA), respectively. Conversion of C to U at those sites leads to the restoration of codon conservation among seed plants (Zeng *et al.*, 2007).

14.3.4. *RNA editing in transcripts encoding for miscellaneous proteins*

Conversion of Ser^{267} or Pro^{267} to Leu^{267} by RNA editing was identified in *accD* transcripts of pea, soybean, canola, *Arabidopsis* and black pine, and resulted in amino acid conservation among seed plants (Sasaki *et al.*, 2001; Zeng *et al.*, 2007). RNA editing at this codon position was previously demonstrated to be required for functional acetyl-CoA carboxylase *in vitro* (Sasaki *et al.*, 2001). However, in the *Phalaenopsis* orchid, the corresponding codon is Phe^{266}, and RNA editing is not apparent (Zeng *et al.*, 2007). Considering the similarities in the hydrophobic properties of Phe and Leu, the *accD* gene is probably still functional in *Phalaenopsis* chloroplasts. On the other hand, three editing sites, which converted Ser^{395} (UCA) to Leu^{395} (UUA), Pro^{471} (CCA) to Leu^{471} (CUA), and Pro^{477} (CCU) to Leu^{477} (CUU), were uniquely observed in the *accD* transcripts of *Phalaenopsis*. Furthermore, three other unique editing sites were present in *clpP* and *matK* transcripts of *Phalaenopsis*. The former occurred at His^{28} (CAU) with conversion to Tyr^{28} (UAU), and the latter occurred at Ser^{178} (UCU) with conversion to Phe^{178} (UUU), and His^{240} (CAU) changed to Tyr^{240} (UAU). All of the above editing events tended to restore the codon conservation in seed plants (Zeng *et al.*, 2007).

Although *Ndh* genes encoding the subunits of the NADH dehydrogenase complex are involved in the cyclic electron flow of photosystem I and chlororespiration in tobacco, they are not essential for a plant's growth under normal conditions (Burrows *et al.*, 1998). All 11 subunits of *ndh* genes are present in the chloroplast genomes of photosynthetic vascular plants so far sequenced, with the exception of black pine, *Phalaenopsis* and *Oncidium* (Wakasugi *et al.*, 1994; Chang *et al.*, 2006; Wu *et al.*, 2010). RNA editing occurs frequently in the *ndh* transcripts, which account for more than 40 percent of the editing sites in the chloroplast transcripts of higher plants, except in *Phalaenopsis* and black pine (Zeng *et al.*, 2007). In particular, the *ndhB* transcripts of barley, tobacco, and *Arabidopsis* differ from the corresponding genomic sequence at nine sites, the highest number of editing events for a single chloroplast mRNA reported to date

(Tsudzuki *et al.*, 2001; Tillich *et al.*, 2005; Kahlau *et al.*, 2006; Zeng *et al.*, 2007). The "relative neutrality hypothesis" explains well why editing sites evolve more readily in those genes in which a transitory loss of function can be tolerated (Fiebig *et al.*, 2004). In *Phalaenopsis*, the *ndhA, ndhF* and *ndhH* genes are completely absent from the chloroplast genome. The other eight *ndh* genes (*ndhB, ndhC, ndhJ, ndhD, ndhE, ndhG, ndhI* and *ndhK*) have various degrees of nucleotide insertion/deletion as compared to tobacco, and they are all frame-shifted (Chang *et al.*, 2006). In addition, the plastid *ndhD* genes of *Phalaenopsis* have an ACG rather than an ATG codon at their translation initiation sites (Chang *et al.*, 2006). Previously, C-to-U conversion at the initiation codon of *ndhD* transcripts were reported in dicot plants (Tsudzuki *et al.*, 2001). Therefore, RNA editing was hypothesized to be required in repairing internal stop codons and/or initiation sites and thus restore the normal function of *ndh* genes in *Phalaenopsis*. However, from RT-PCR assays, no apparent RNA editing sites for *ndhC, ndhD, ndhE, ndhG, ndhI, ndhJ* or *ndhK* transcripts were detected. Only one partial editing site, corresponding to the maize codon 494, was identified in *ndhB* transcripts of the *Phalaenopsis* orchid, but it did not repair the internal stop codon of frame-shifted *ndhB* transcripts. Therefore, all the *ndh* genes are pseudogenes in the plastid genome of *Phalaenopsis*. The single remaining partial editing site is likely an evolutionary remnant from before the complete loss of plastid RNA editing sites for non-functional *ndh* genes.

14.4. Future Prospects

Determining the distribution and pattern of editing sites across taxa and across the entire chloroplast genome is an important step in investigating the evolutionary process of RNA editing in angiosperms. With the available information on the RNA editing patterns in *P. aphrodite*, the next step is to identify the protein factors as well as the editing enzymes involved in the RNA editing process in the chloroplasts of *P. aphrodite*. RNA interference (RNAi) will be a powerful tool for knocking out/down the potential PPR proteins involved in site-specific recognition of RNA editing in *P. aphrodite*.

Acknowledgments

This work was financially supported in part by grants from the National Science Council of Taiwan.

References

Bock R, Kossel H, Maliga P. (1994) Introduction of a heterologous editing site into the tobacco plastid genome the lack of RNA editing leads to a mutant phenotype *EMBOJ* **13**:4623–4628.

Bock R. (2000) Sense from nonsense: How the genetic information of chloroplasts is altered by RNA editing. *Biochimie* **82**(6–7): 549–557.

Burrows PA, Sazanov LA, Svab Z *et al.* (1998) Identification of a functional respiratory complex in chloroplasts through analysis of tobacco mutants containing disrupted plastid *ndh* genes. *EMBO J* **17**: 868–876.

Calsa Junior T, Carraro DM, Benatti MR *et al.* (2004) Structural features and transcript-editing analysis of sugarcane (*Saccharum officinarum* L.) chloroplast genome. *Curr Genet* **46**: 366–373.

Chang CC, Lin HC, Lin IP *et al.* (2006) The chloroplast genome of *Phalaenopsis aphrodite* (Orchidaceae): Comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. *Mol Biol Evol* **23**(2): 279–291.

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

Corneille S, Lutz K, Maliga P. (2000) Conservation of RNA editing between rice and maize plastids: Are most editing events dispensable? *Mol Gen Genet* **264**(4): 419–424.

Drescher A, Hupfer H, Nickel C *et al.* (2002) C-to-U conversion in the intercistronic *ndhI/ndhG* RNA of plastids from monocot plants: Conventional editing in an unconventional small reading frame? *Mol Genet Genom* **267**(2): 262–269.

Fiebig A, Stegemann S, Bock R. (2004) Rapid evolution of RNA editing sites in a small non-essential plastid gene. *Nucleic Acids Res* **32**(12): 3615–3622.

Gehrig HH, Winter K, Cushman J *et al.* (2000) An improved RNA isolation method for succulent plant species rich in polyphenols and polysaccharides. *Plant Mol Biol Rep* **18**: 369–376.

Giegerich R, Voss B, Rehmsmeier M. (2004) Abstract shapes of RNA. *Nucleic Acids Res* **32**(16): 4843–4851.

Hammani K, Okuda K, Tanz SK *et al.* (2009) A study of new *Arabidopsis* chloroplast RNA editing mutants reveals general features of editing factors and their target sites. *Plant Cell* **21**(11): 3686–3699.

Hirose T, Kusumegi T, Tsudzuki T, Sugiura M. (1999) RNA editing sites in tobacco chloroplast transcripts: Editing as a possible regulator of chloroplast RNA polymerase activity. *Mol Gen Genet* **262**(3): 46–27.

Hirose T, 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**(5): 1144–1152.

Inada M, Sasaki T, Yukawa M *et al.* (2004) A systematic search for RNA editing sites in pea chloroplasts: An editing event causes diversification from the evolutionarily conserved amino acid sequence. *Plant Cell Physiol* **45**(11): 1615–1622.

Kahlau S, Aspinall S, Gray JC, Bock R. (2006) Sequence of the tomato chloroplast DNA and evolutionary comparison of solanaceous plastid genomes. *J Mol Evol* **63**(2): 194–207.

Kugita M, Yamamoto Y, Fujikawa T *et al.* (2003) RNA editing in hornwort chloroplasts makes more than half the genes functional. *Nucleic Acids Res* **31**: 2417–2423.

Lutz KA, Maliga P. (2001) Lack of conservation of editing sites in mRNAs that encode subunits of the NAD(P)H dehydrogenase complex in plastids and mitochondria of *Arabidopsis thaliana. Curr Genet* **40**(3): 241–249.

Maier RM, Neckermann K, Igloi GL, Kossel H. (1995) Complete sequence of the maize chloroplast genome: Gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *J Mol Biol* **251**: 614–628.

Miyamoto T, Obokata J, Sugiura M. (2002) Recognition of RNA editing sites is directed by unique proteins in choroplasts: biochemical identification of cis-acting elements and trans-acting factors involved in RNA editing in tobacco and pea chloroplasts. *Mol Ceu Biol* **22**(19):6726–6734.

Miyamoto T, Obokata J, Sugiura M. (2004) A site-specific factor interacts directly with its cognate RNA editing site in chloroplast transcripts. *Proc Natl Acad Sci USA* **101**(1): 48–52.

Okuda K, Nakamura T, Sugita M *et al.* (2006) A pentatricopeptide repeat protein is a site-recognition factor in chloroplast RNA editing. *J Biol Chem* **281**(49): 37661–37667.

Okuda K, Myouga F, Motohashi R *et al.* (2007) Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. *Proc Natl Acad Sci USA* **104**(19): 8178–8183.

Ruf S, Kossel H. (1997) Tissue-specific and differential editing of the two *ycf3* editing sites in maize plastids. *Curr Genet* **32**(1): 19–23.

Sasaki Y, Kozaki A, Ohmori A *et al.* (2001) Chloroplast RNA editing required for functional acetyl-CoA carboxylase in plants. *J Biol Chem* **276**(6): 3937–3940.

Sasaki T, Yukawa Y, Miyamoto T *et al.* (2003) Identification of RNA editing sites in chloroplast transcripts from the maternal and paternal progenitors of tobacco (*Nicotiana tabacum*): Comparative analysis shows the involvement of distinct trans-factors for *ndhB* editing. *Mol Biol Evol* **20**(7): 1028–1035.

Sasaki T, Yukawa Y, Wakasugi T *et al.* (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, Regel R, Du TG *et al.* (2002) The plastid chromosome of *Atropa belladonna* and its comparison with that of *Nicotiana tabacum*: The role of RNA editing in generating divergence in the process of plant speciation. *Mol Biol Evol* **19**: 1602–1612.

Schmitz-Linneweber C, Small I. (2008) Pentatricopeptide repeat proteins: A socket set for organelle gene expression. *Trends Plant Sci* **13**(12): 663–670.

Tillich M, Funk HT, Schmitz-Linneweber C *et al.* (2005) Editing of plastid RNA in *Arabidopsis thaliana* ecotypes. *Plant J* **43**(5): 708–715.

Tsudzuki T, Wakasugi T, Sugiura M. (2001) Comparative analysis of RNA editing sites in higher plant chloroplasts. *J Mol Evol* **53**: 327–332.

Wakasugi T, Tsudzuki J, Ito S *et al.* (1994) Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. *Proc Natl Acad Sci USA* **91**: 9794–9798.

Wakasugi T, Hirose T, Horihata M *et al.* (1996) Creation of a novel protein-coding region at the RNA level in black pine chloroplasts: The pattern of RNA editing in the gymnosperm chloroplast is different from that in angiosperms. *Proc Natl Acad Sci USA* **93**(16): 8766–8770.

Wolf PG, Rowe CA, Hasebe M. (2004) High levels of RNA editing in a vascular plant chloroplast genome: Analysis of transcripts from the fern *Adiantum capillus-veneris*. *Gene* **339**: 89–97.

Wu FH, Chan MT, Liao DC *et al.* (2010) Complete chloroplast genome of *Oncidium* Gower Ramsey and evaluation of molecular markers for identification and breeding in Oncidiinae. *BMC Plant Biol* **10**(1): 68.

Zeng WH, Liao SC, Chang CC. (2007) Identification of RNA editing sites in chloroplast transcripts of *Phalaenopsis aphrodite* and comparative analysis with those of other seed plants. *Plant Cell Physiol* **48**(2): 362–368.

Zito F, Kuras R, Choquet Y, Kossel H, Wollman FA. (1997) Mutations of cytochrome b_6 in *Chlamydomonas reinhardtii* disclose the functional significance for a proline to leucine conversion by *per B* editing in maize and tobacco. *Plant Mol Biol* **33**(1): 79–86.