RNA EDITING SITE OF *NDH* GENES IN *Borassus flabellifer*

Arpakorn Sakulsathaporn[1,2], Saowanee Suputtitada[3,4], Somsak Apisitwanich[1,2,3,4] and Supachai Vuttipongchikij[3,4]*

[1] Center for Agricultural Biotechnology, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand
[2] Center of Excellence on Agricultural Biotechnology: (AG-BIO/PERDO-CHE), Kasetsart University, Bangkok 10900, Thailand
[3] Department of Genetics Faculty of Science, Kasetsart University, Bangkok 10900, Thailand
[4] Center for Advanced Studies in tropical Natural Resource, National Research University, Kasetsart University, Bangkok 10900, Thailand

* e-mail: fsciscv@ku.ac.th

Abstract

RNA editing is a process that alters genetic information of the chloroplast genome by post transcriptional RNA modification. Commonly, specific cytosine residues in the RNA molecules are substituted to the uridine residues in order to provide coding sequences for functional proteins. However, the mechanism of this RNA editing process is still unclear. The highest RNA editing event has been found to occur in *ndh* genes which encode NDH complexes that are important component for electron transport chain of photosynthesis and therefore, *ndh* genes represent an interesting target for the RNA editing study. This study, RNA editing sites in the chloroplast genome of *Borassus flabellifer* were investigated using RNA editing prediction program to compare the *ndhA*, *ndhB*, *ndhD*, *ndhF* and *ndhG* obtained sequences . The results showed that of all RNA editing sites, C bases were substituted with U bases and these mostly occurred at the second base of each codon, and sometimes it occurred at the first base position. However, substitution of U could not observe at the third base position of the codon. In all observed transcripts, we found 11, 4, 2 and 1 RNA editing sites in *ndh B, ndh D*, *ndh A* and *ndh G* transcripts, respectively. In contrast, no RNA editing was observed in *ndhF* transcript. Consequently, these U substitutions caused amino acid change from original transcripts in all edited mRNA. These amino acids were changed from neutral molecules to be hydrophobic molecules. This made NDH complex be stable and attached to the thylakoid fastenly.

Keywords: RNA editing, *ndh* gene, *Borassus flabellifer*

Introduction

RNA editing is one of the post-transcriptional processes, which modifies the genetic information of RNA molecules. In higher plant, RNA editing was the first evident in *rpl2* transcript of maize chloroplast genome (Hoch et al. 1991). Most of RNA editing events occur through a specific alteration in coding sequences of mRNA that change C to U bases (Shikanai 2006). In chloroplast genome, RNA editing is essential for expressing functional proteins by modifying the coding sequence and amino acid sequence is changed. Non-edited transcripts generally are resulted in un-functional proteins. For example, *Arabidopsis atecb2* mutant line, that was defected in the editing of *AccD* transcript, displayed albino cotyledons

and seedling lethality (Yu et al. 2009). Trans proteins that play roles in RNA editing process were identified in the group of pentatricopeptide repeat protein (PPR), which recognizes as short sequence surrounding the target C residue. All PPRs involved in RNA editing process belong to E and DYW subclasses. DYW subclass contains invariant cysteine and histidine residues matching the active site of cytidine deaminases. Okuda at el. (2009) identified *crr22* and *crr28* that were *DYW* genes requiring for RNA editing controlling NDH activity in *Arabidopsis*.

Previous reports had found the highest editing frequency in *ndh* genes of chloroplasts such as *Cycas taitungensis, Arabidopsis thaliana* and *Gossypium hirsutum*. (Chen et al. 2011, Tillich et al. 2005 and Jian et al. 2012) Most plastid genomes of higher plant contain 11 *ndh* genes producing NDH polypeptides form a thylakoid NDH complex to catalyze electron transfer from NADH to plastoquinone in the cyclic photosystem. (Martin and Sabater, 2010). Under environmental stress and leaf senescence, NDH complex will be increased to reduce H_2O_2 and other reactive oxygen species. (Casano et al. 2001 and Lascano et al. 2003). Therefore the target for this investigated RNA editing of *Borassus flabellifer* was the *ndh* genes. *Borassus flabellifer*, an Arecaceae family member, is widespread in South-East Asia. It lives in Coryphoideae subfamily same *Phoenix dactylifera* (date palm).

Methodology

DNA isolation

Total genomic DNA of *B. flabellifer* was isolated from a young leaf using Dneasy Plant Mini kit (QIAgen, USA).

RNA extraction and cDNA synthesis

Total RNA of *B. flabellifer* was isolated from a young leaf using Spectrum™ Plant Total RNA kit (Sigma-aldrich, USA), then cDNA was synthesized using Protoscrip M-MuLV First Strand cDNA kit (NEB, UK).

PCR amplification and sequencing

The specific primers were designed according to *Elaeis guineensis* (NC_017602.1) and *Phoenix dactylifera* (NC_013991.2) palm chloroplast genome data. The *ndh* genes were amplified from *B. flabellifer* DNA and cDNA with PCR profiles: initial denaturation for 5 min at 95˚C, followed by 35 cycles of 30 s at 95˚C, 30 s at 55-60˚C depending on melting temperatures of primers, 30 s at 72˚C and final extension for 10 min at 72˚C. Amplification products were separated using 1% agarose gel electrophoresis and purified using gel extraction kit (Flavogen, Taiwan), then they were sequenced and analyzed.

Prediction of RNA editing site

All *ndhA, ndhB, ndhD, ndhF,* and *ndhG* sequences obtained from *B. flabellifer* chloroplast genome were analyzed by PREP-Cp (http://prep.unl.edu/) for RNA editing prediction.

Identification of RNA editing site

RNA editing sites of *B. flabellifer* were identified by sequence alignment of DNA and cDNA sequences of each gene using the clustalW program.
(http://www.ebi.ac.uk/Tools/msa/clustalw2/)

Results and discussion

Prediction and identification of RNA editing of *ndh* genes in *Borassus flabellifer*

Coding sequences of *ndhA*, *ndhB*, *ndhD*, *ndhF* and *ndhG* were amplified from both DNA and cDNA of the same plant sample and analyzed the RNA editing sites. Analysis of five coding sequences obtained from DNA sample using PREP-cp online service program indicated totally 27 predicted RNA editing sites. Sequence alignments between DNA and cDNA with clustalW (Table 1) founding 11, 4, 2, 1 and 0 RNA edited site in *ndh B, ndh D*, *ndh A, ndh G* and *ndh F* transcripts showed the 16 actually edited sites at the five *ndh* genes. This indicated the 55.17% accuracy value of PREP-Cp program because the PREP-Cp program contains only data of 7 plant species. They were *Arabidopsis thaliana, Nicotiana tabacum, Atropa belladonna, Zea mays, Oryza sativa, Phalaenopsis aphrodite* and *Pisum sativum.* So the data were analyzed by RNA-tagged sequences of PREP-Aln to get the higher accuracy. Nevertheless, PREP is a useful tool for RNA editing site prediction of chloroplast and mitochondrial genome. PREP-Mt is excellent for prediction the mitochondrial genome of *Magnolia tripetala*, *Mimulus guttatu, Citrallu lanatus* and *Cucurbita pepo* (Alverson et al. 2010; Muwer et al. 2012 and Hepburn et al. 2012).

Among the 18 editing sites, the *ndhB* gene was the most frequent RNA editing sites up to 11 positions while that of the *ndhF* gene has not any editing. The *ndh* plays role on stress response and leaf senescence in the form of NDH complex. (Bieker et al. 2011 and Zepata et al. 2005) This could be that *ndhF* mRNA from young leaf does not have any post-transcription process and the amount of ndhF protein at this stage is enough. This phenomenon, *ndhF* might be a key gene controlling activity of NDH complex. It supports the report of Favory et al. (2005). Total NDH complexes could be determined by the quantity of each subunit, such as. *ndhF*-deficient tobacco extended the duration time of leaf senescence (Zepata et al. 2005) and the phosphorylation of ndhF affected to activity of chloroplast NDH complex in the barley leaf (Lascano et al. 2003). Further study, the RNA editing of *ndhF* gene of old leaf and its senescence will be investigated.

RNA editing sites of *ndh* coding sequences

Transcript sequence analysis of *ndhA*, *hdhB*, *ndhD*, *ndhF* and *ndhG* in *B. flabellifer* revealed 18 edited base positions. All of them were the same pattern of which cytocine base was substituted with uridine base. Total 17 of 18 edited sites, uridine substitution occurred at the second base position in the codon, the only one occurred at the first position of *ndhB* gene. Predominance of base substitution at the second base in the codon of this study is harmoniously found in other plants, including *Arabidopsis thaliana*, *Nicotiana tabacum*, *Atropa belladonna*, *Zea mays*, *Oryza zativa* and *Lolium perenne* (table 2). Consequently, base substitution caused the amino acid change of translated protein products. Most *ndh* RNA editing of *B. flabellifer* was position and codon bias of U_A (55%), C_A (22.22%) and U_C, U_G, C_U and A_G (5.56%), respectively. Edited 18 codon sites made amino acid change which is effected on the polarity of the molecule. They are changed from being neutral and hydrophilic molecule to be hydrophobic molecule. The serine was substituted with leucine at 11 codon sites (66.11%), and with phenylalanine (6.25%). Proline was substituted with leucine at 4 codon sites (22.22%). Threonine was substituted with methionine and histidine was substituted with tyrosine at 1 each codon site with the percentage of 6.25. The change of second base in a codon from C to U commonly occurs in chloroplast genome (Cuenca et al., 2010), especially for UU_ which is TT_ in DNA. The thimine dimer often arised by UV and

reactive oxygen species and these thimine dimer are easily mutated. This transcript editing mechanism from U to C at the second position of a codon helps plants to avoid mutation and it is an evolution mechanism of terrestrial plants which evolved from aquatic plants (Yura and Go, 2008). Another reason, RNA editing is often found in the serine codon which is hydrophilic and changed to be leucine which is hydrophobic. This makes NDH complex be stable and attached to the thylakoid fastenly. (Jobson and Qiu, 2008) These results have been reported in *Cycas taitungensis, Gossypium hirsutum, Phalaenopsis aphrodite* and *Elaeis guineensi* (Chen et al., 2011; Jiang et al., 2012; Chang et al., 2011 and Uthaipaisanwong et al., 2012). For *B. flabellifer ndh* transcripts, codons were edited to leucine 83.33%.

Table 1 Prediction and experimentally identified RNA editing site of *B. flabellifer ndh* genes.

gene	nucleotide position	amino acid position	edited codon position	base	edited codon	amino acid change	predicted PREP	Experimental identification
ndhA	476	159	2	c-U	uca-uUa	S-L	+	+
	566	189	2	c-U	uca-uUa	S-L	+	+
	1073	358	2	c-U	ucc-uUc	S-F	+	-
ndhB	149	50	2	c-U	uca-uUa	S-L	+	+
	467	156	2	c-U	cca-cUa	P-L	+	+
	542	181	2	c-U	acg-aUg	T-M	+	+
	586	196	1	c-U	cau-Uau	H-Y	+	+
	704	235	2	c-U	ucc-uUc	S-F	+	+
	737	246	2	c-U	cca-cUa	P-L	+	+
	830	277	2	c-U	uca-uUa	S-L	+	+
	836	279	2	c-U	uca-uUa	S-L	+	+
	1112	371	2	c-U	uca-uUa	S-L	+	+
	1193	398	2	c-U	uca-uUa	S-L	+	+
	1255	419	1	c-U	cau-Uau	H-Y	+	-
	1481	494	2	c-U	cca-cUa	P-L	+	+
ndhD	5	2	2	c-U	acg-aUg	T-M	+	-
	62	21	2	c-U	uca-uUa	S-L	+	-
	386	129	2	c-U	uca-uUa	S-L	+	+
	677	226	2	c-U	ucg-uUg	S-L	+	-
	687	326	2	c-U	ucg-uUg	S-L	-	+
	1196	399	2	c-U	uca-uUa	S-L	+	+
	1313	438	2	c-U	uca-uUa	S-L	+	+
ndhF	62	21	2	c-U	uca-uUa	S-L	+	-
	290	97	2	c-U	uca-uUa	S-L	+	-
	442	148	1	c-U	cau-Uau	H-Y	+	-
	586	196	1	c-U	cuu-Uuu	L-F	+	-
	1393	465	1	c-U	cac-Uac	H-Y	+	-
ndhG	314	105	2	c-U	aca-aUa	T-I	+	-
	347	116	2	c-U	cca-cUa	P-L	-	+

They were 66.67% UUA, 26.67% CUA and 6.67% UUG. The PPR protein recognizes and edits the U_A which has 15 upstream and 4 downstream nucleotides (totally 22 nucleotides). (Castandet and Araya, 2011) So that 22 nucleotides sequence alignment of 18 sites were analyzed. The result did not occur conserved sequence. Shikanai (2012) suggested that RNA editing could occur at any site. In *Arabidopsis thaliana*, PPR, OTP84, having *DYW* domain edited the *ndhB* at the 279th amino acid and also did the *ndhF* at the 97th amino acid. For *B. flabellifer* sample, the editing was the same in *ndhB* (279) but it did not change the 97th amino acid of *ndhF*. Further research, the *ndhF* would be revealed in its RNA editing process.

Table 2 Summary data of RNA editing in 6 plant species.

species	*Arabidopsis thaliana*[a]	*Nicotiana tabacum*[a]	*Atropa belladonna*[a]	*Zea mays*[a]	*Oryza zativa*[b]	*Lolium perenne*[b]	*Borassus flabellier*
Number of editing site	17	15	15	12	12	15	18
C-U	100%	100%	100%	100%	100%	100%	100%
1st codon editing	11.76%	6.67%	6.67%	8.33%	8.33%	6.67%	5.56%
2nd codon editing	88.24%	93.33%	93.33%	91.67%	91.67%	93.33%	94.44%
3rd codon editing	0%	0%	0%	0%	0%	0%	0%

[a] from RNA editing database (http://biologia.unical.it/py_script/search.html)
[b] from Diekmann et al. (2009)

Acknowledgments

This research is partially supported by the Center of Excellence on Agricultural Biotechnology, Science and Technology Postgraduate Education and Research Development Office, Office of Higher Education Commission, Ministry of Education. (AG-BIO/PERDO-CHE)

References

1. Hoch B, Maier R.M., Appel K, Igloi G.L, Kössel H. (1991) Internal editing of the maize chloroplast ndhA transcript restores codons for conserved amino acids. Nature. 353: 178–180.
2. Shikanai T. (2006) RNA editing in plant organelles: machinery, physiological function and evolution. Cell Mol Life Sci. 63:698–708.
3. Yu Q-B, Jiang Y, Chong K, Yang Z-N. (2009) AtECB2, a pentatricopeptide repeat protein, is required for chloroplast transcript *accD* RNA editing and early chloroplast biogenesis in Arabidopsis thaliana. The Plant J. 59:1011–1023.
4. Okuda K, Chateigner Boutin A, 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.
5. Chen H, Deng L, Jiang Y, Lu P, Yu J. (2011) RNA editing sites exist in protein-coding genes in the chloroplast genome of *Cycas taitungensis*. J Integr Plant Biol. 5312: 961–970.
6. Tillich M, Funk HT, Schmitz-Linneweber C, Poltnigg P, Sabater B, Martin M, Maier RM. (2005) Plastid RNA editing in *Arabidopsis thaliana* ecotypes. Plant J., 43:708–715.

7. Jiang Y, Fan SL, Song MZ, Yu JN, Yu SX. (2012) Identification of RNA editing sites in cotton *(Gossypium hirsutum)* chloroplasts and editing events that affect secondary and three-dimensional protein structures. Genet Mol Res. 4112:987-1001.
8. Martin M, Sabater B. (2010) Plastid ndh genes in plant evolution Plant Physiol Biochem. 48:636–645.
9. Casano LM, Martin M, Sabater B. (2001) Hydrogen peroxide mediates the induction of chloroplastic Ndh complex under photooxidative stress in barley. Plant Physiol. 125:1450–1458.
10. Lascano HR, Casano LM, Martin M., Sabater B. (2003) The activity of the chloroplastic Ndh Complex is regulated by phosphorylation of the NDH-F Subunit. Plant Physiol. 132: 256–262.
11. Alverson AJ, Wei X, Rice DW, Stern DB, Barry K, Palmer JD. (2010) Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). Mol Biol Evol. 276:1436-48.
12. Mower JP, Case AL, Floro ER and Willis JH. (2012) Evidence against equimolarity of large repeat arrangements and a predominant master circle structure of the mitochondrial genome from a monkeyflower (Mimulus guttatus) lineage with cryptic CMS. Genome Biol Evol., 0145:670-86.
13. Hepburn N J, Schmidt D W, Mower J P. (2012) Loss of two introns from the *Magnolia tripetala* mitochondrial *cox2* gene implicates horizontal gene transfer and gene conversion as a novel mechanism of intron loss Mol Biol Evol. 29:3111-20.
14. Bieker S, Riester L, Stahl M, Franzaring J, Zentgraf U.(2012) Senescence-specific Alteration of Hydrogen Peroxide Levels in Arabidopsis thaliana and Oilseed Rape Spring Variety Brassica napus L. cv. Mozart(F). J Integr Plant Biol. 54(8):540-54.
15. *Zapata* JM, Guéra A, Esteban-Carrasco A, Martín M, Sabater B. (2005) Chloroplasts regulate leaf senescence*:* delayed senescence in transgenic ndhF-defective tobacco. Cell Death Differ. 1210:1277-84.
16. Favory JJ, Kobayshi M, Tanaka K, Peltier G, Kreis M, Valay JG, Lerbs-Mache S. (2005) Specific function of a plastid sigma factor for *ndhF* gene transcription. Nucleic Acids Res. 33*:*5991-5999.
17. Lascano HR, Casano LM, Martin M, Sabater B. (2003). The activity of the chloroplastic Ndh complex is regulated by phosphorylation of the NDH-F subunit. Plant Physiol. 132:256–262.
18. Cuenca A, Petersen G, Seberg O, Davis JI, Stevenson DW. (2010) Are substitution rates and RNA editing correlated? BMC Evol Biol. 10**:**349.
19. Yura K, Go M. (2008) Correlation between amino acid residues converted by RNA editing and functional residues in protein three-dimensional structures in plant organelles. BMC Plant Biol. 8:79.
20. Jobson RW, Qiu YL. (2008) Did RNA editing in plant organellar genomes originate under natural selection or through genetic drift? Biol Direct. 3:43.
21. Chang CC, Lin HC, Lin IP, Chow TY, Chen HH, Chen WH, Cheng CH, Lin CY, Liu SM, Chang CC, Chaw SM. (2006) The chloroplast genome of *Phalaenopsis aphrodite* (Orchidaceae): Comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. Mol Biol Evol. 23:279–291.
22. Uthaipaisanwong P, Chanprasert J, Shearman JR, Sangsrakru D, Yoocha T, Jomchai N, Jantasuriyarat C, Tragoonrung S, Tangphatsornruang S. (2012) Characterization of the chloroplast genome sequence of oil palm (*Elaeis guineensis Jacq*). Gene. 500:172-180.

23. Castandet B, Araya, A. (2011) RNA editing in plant organelles. Why make it easy? Biochemistry (Mosc) 76:924-931.
24. Shikanai T. (2012) Why Do Plants Edit RNA in Plant Organelles? Organelle Genetics. pp 381-397.
25. Diekmann K, Hodkinson TR, Wolfe KH, Van Den Bekerom R, Dix PJ, Barth S, (2009) Complete chloroplast genome sequence of a major allogamous forage species, perennial ryegrass *(Lolium perenne L.).* DNA Res. 16:165-176.