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

Tadamasa Sasaki, Yasushi Yukawa,* Tetsuya Miyamoto,† Junichi Obokata,† and Masahiro Sugiura**

*Graduate School of Natural Sciences, Nagoya City University, Mizuho, Nagoya, Japan; and †Center for Gene Research, Nagoya University, Nagoya, Japan

RNA editing alters genomic nucleotide sequences at the transcript level. In higher plant chloroplasts, C-to-U conversion is known to occur at around 30 specific sites. The tobacco cultivar *Nicotiana tabacum* is an amphidiploid derived from ancestors of *N. sylvestris* (maternal) and *N. tomentosiformis* (paternal). The chloroplast genome of *N. tabacum* is believed to originate from an ancestor of *N. sylvestris*. To study the evolution of RNA editing in higher plant chloroplasts, editing sites in the two likely progenitors have first been identified based on those found in *N. tabacum*. Altogether 34, 33, and 32 editing sites have been found in the chloroplast transcripts from *N. tabacum*, *N. sylvestris*, and *N. tomentosiformis*, respectively. Thirty-one sites are conserved among the three species, whereas remarkable differences are observed in the editing of *ndhB* and *ndhD* transcripts. Sites 7 and 8 in *ndhB* mRNAs are separated only by five nt, and both are edited in *N. tabacum* and *N. sylvestris*. However, site 8 is not edited in *N. tomentosiformis*, indicating that distinct *trans*-factors are involved in the two editing events. The first site in *ndhD* mRNAs is edited to produce an AUG start codon in *N. sylvestris* as well as in *N. tabacum* but not in *N. tomentosiformis*, suggesting that a distinct mechanism operates for the translational initiation of *N. tomentosiformis ndhD* mRNAs. Four to six sites are edited partially in green leaves. Some of these sites may represent evolutionary intermediates in the process of losing editing events.

Introduction

The chloroplast is known to possess its own genome and gene expression system (Sugiura 1992). Many chloroplast genes in land plants are transcribed as polycistronic RNAs, which are then processed into mature RNA species via complex pathways including RNA editing and splicing (Sugiura, Hirose, and Sugita 1998; Rochaix 2001). RNA editing in chloroplasts was first reported in the maize *rpl2* mRNA (Hoch et al. 1991), and a systematic search of chloroplast transcripts identified 27 editing sites in maize (Maier et al. 1995; Bock, Hermann, and Fuchs 1997), 26 editing sites in black pine (Wakasugi et al. 1996), 31 editing sites in tobacco (Hirose et al. 1999), and 21 editing sites in rice (Corneille, Lutz, and Maliga 2000). Editing in chloroplasts occurs generally in protein-coding regions and restored evolutionary conserved amino acid sequences (Maier et al. 1996). However, editing at the third position of a codon (silent editing) and editing in an untranslated region has also been reported (Hirose et al. 1996; Kudla and Bock 1999). In addition, extensive RNA editing, both C-to-U and U-to-C changes have been reported in the chloroplast of hornwort *Anthoceros formosae* (Yoshinaga et al. 1996). RNA editing has been found in chloroplast transcripts from all major lineages of land plants; however, neither frequency of editing nor the pattern of editing a specific transcript correlate with the phylogenic tree of the plant kingdom (Freyer, Kiefer-Meyer, and Kössel 1997).

A key question in chloroplast editing is how specific C residues are recognized precisely from all other C residues in transcripts. Using transgenic approaches in tobacco chloroplasts, *cis*-acting elements have been analyzed for *psbL* mRNAs (Chaudhuri, Carrer, and Maliga 1995; Chaudhuri and Maliga 1996), for *ndhB* mRNAs (sites 4 and 5) (Bock, Hermann, and Kössel. 1996; Bock, Hermann, and Fuchs 1997; Hermann and Bock 1999), and for *ndhF* and *rpoB* (site 2) mRNAs (Reed, Lyi, and Hanson 2001). These studies commonly showed that *cis*-acting elements reside in upstream regions of the editing sites. Furthermore, chloroplast transplastomic experiments suggested the involvement of *trans*-acting factors in editing (Chaudhuri, Carrer, and Maliga 1995; Chaudhuri and Maliga 1996; Bock and Koop 1997; Reed and Hanson 1997; Reed, Lyi, and Hanson 2001; Schmitz-Linneweber et al. 2001). These in vivo analyses show that at least some *trans* factors appear to be site specific and of extraplastidic origin. Recently, an in vitro RNA editing system from tobacco chloroplasts was developed in our laboratory to dissect biochemical processes of editing reactions in chloroplasts (Hirose and Sugiura 2001). Using this system, a tobacco chloroplast protein of 25 kd was found to bind specifically to the *cis*-acting element of *psbL* mRNA. This result provided the evidence that the protein, but not RNA, is the *trans*-acting factor that is likely to recognize the editing site of *psbL* mRNAs. An improved method was then reported for preparing chloroplast extracts supporting accurate RNA editing reactions in vitro not only from tobacco but also from pea (Miyamoto, Obokata, and Sugiura 2002). Using this improved system, we defined *cis* elements of *psbE* and *petB* mRNAs and detected *trans* factors that specifically bind to these elements, 56-kd and 70-kd proteins for *psbE* and *petB* mRNAs, respectively.

In the case of tobacco chloroplasts, the genome sequence has been completely determined (Shinozaki et al. 1986), the gene organization has been updated (Wakasugi et al. 1998; Wakasugi, Tsudzuki, and Sugiura 2001), and

Key words: chloroplast, RNA editing, *trans*-factor, transcript, tobacco.

E-mail: sugiura@nsc.nagoya-cu.ac.jp.

Mol. Biol. Evol. 20(7):1028–1035. 2003
DOI: 10.1093/molbev/msg098

a systematic search for editing sites in the transcripts has been made (Hirose et al. 1999). In addition, both chloroplast transformation techniques (in vivo) (Svab and Maliga 1993) and chloroplast RNA editing system (in vitro) (Hirose and Sugiura 2001) are available only for tobacco. Therefore, tobacco is the organism of choice for analyzing detailed mechanisms of RNA editing in chloroplasts. The tobacco cultivar *Nicotiana tabacum* is a natural amphidiploid derived from two progenitors, and ancestors of *N. sylvestris* and *N. tomentosiformis* were the likely progenitors of *N. tabacum* (Smith 1974; Kenton et al. 1993). The chloroplast genome of *N. tabacum* is believed to have originated from *N. sylvestris* (Olmstead and Palmer 1991). Hence, these *Nicotiana* species offer a significant advantage for evolutional studies of RNA editing in higher plant chloroplasts. Recently, Schmitz-Linneweber et al. (2001) reported an interesting observation that *N. tabacum* lost an editing site but still possesses its *trans* factor, which probably originated from a progenitor of *N. tomentosiformis*.

Here we report the editing pattern of both *N. sylvestris* and *N. tomentosiformis* chloroplasts. Comparative analysis with editing sites in *N. tabacum* shows the opposite case as above, namely the presence of a C residue (to be edited in *N. tabacum* and *N. sylvestris*) but no corresponding editing activity in *N. tomentosiformis*. This analysis also indicates the involvement of distinct *trans* factors for two adjacent sites.

Materials and Methods

Nicotiana tabacum (var. Bright Yellow 4), *N. sylvestris*, and *N. tomentosiformis* leaves were harvested from 6-week-old plants grown in a growth chamber at 28°C under 16 h light/8 h dark conditions. Total cellular DNA and RNA were isolated from green leaves essentially as described by Wakasugi et al. (1994), with minor modifications. cDNA synthesis was carried out according to the instruction manual of the High Fidelity RNA PCR kit (TaKaRa) using 1 μg of total cellular RNA as template. Sequences encompassing editing sites were amplified by PCR from the cDNA and total cellular DNA using gene-specific primers listed in table 1, purified with GFX™ PCR DNA and Gel Band Purification kit (Amersham Pharmacia), and sequenced using the BigDye Terminator Cycle Sequencing FS Ready Reaction kit (ABI). Computer analysis of editing sites was carried out using GENETYX-MAC version 9.0 (Software Development Co.) and Sequencher version 3.0 (Gene Codes Co.).

Results and Discussion

RNA Editing Sites in Chloroplast Transcripts from *Nicotiana* Species

The *N. tabacum* chloroplast genome has so far been found to include 80 different protein-coding genes and 35 different genes encoding stable RNA species (Wakasugi et al. 2001) and 31 editing sites (all C-to-U conversion) which were observed from 16 gene transcripts (Hirose et al. 1999). Based on information from these sites, direct sequencing was performed for PCR-amplified cDNA fragments (145 to 1,869 bp) derived from *N. sylvestris* and *N. tomentosiformis* transcripts. During the course of this study, three additional C-to-U editing sites were found in *N. tabacum*, one from *rps2* mRNAs (site 1) and two in *ndhD* mRNAs (sites 3 and 4). Therefore, a total of 34 sites were surveyed. Table 2 compiles the editing sites identified in the three *Nicotiana* species. All the 34 sites are conserved at the DNA level (all C residues) among the three species, whereas no editing was observed at the transcript level for one site (*atpA* site 2) in *N. sylvestris* and three sites (*atpA* site 2, *ndhB* site 8, and *ndhD* site 1) in *N. tomentosiformis*. An additional editing site unique to *ndhA* mRNAs from *N. tomentosiformis* (already a T at this site in *N. tabacum* and *N. sylvestris*) was reported by Schmitz-Linneweber et al. (2001). Therefore, total numbers of editing sites so far identified are 33 in *N. sylvestris* and 32 in *N. tomentosiformis*.

Editing of *ndhB* Transcripts

The *ndhB* gene coding for NADH dehydrogenase subunit II is located within the large inverted repeat of the *N. tabacum* chloroplast genome and produces mature spliced mRNAs of ca. 1.5 kb (Matsubayashi et al. 1987). Nine editing sites have been identified in the transcript from the *N. tabacum ndhB* gene (Freyer, Kiefer-Meyer, and Kössel 1997; Hirose et al. 1999), the highest frequency in a single mRNA species. As five additional protein-coding genes are positioned within the repeat and no editing has been reported in their transcripts, high editing frequency of *ndhB* transcripts is not due to its gene location.

As shown in figure 1*A*, *ndhB* sites 7 and 8 in *N. tabacum* are separated only by 5 nt (one alanine codon, GCU) and both editing events cause an amino acid substitution from serine (UCA) to leucine (UUA). This is also the case for *N. sylvestris*, whereas editing was not observed in *N. tomentosiformis* for the C residue corresponding to site 8 of *N. tabacum* (fig. 1*B*). cDNA sequencing was repeated three times with different RNA preparations, and the same results were obtained. Therefore, we concluded that *N. tomentosiformis* lacks editing activity for site 8 despite conserved sequences around this position among the three *Nicotiana* species. On the other hand, another pair of *ndhB* sites 5 and 6, separated only by 8 nt, are edited in *N. tomentosiformis* as in *N. tabacum* and *N. sylvestris* (fig. 1*B*). These results indicate that site-recognition factors for sites 7 and 8 are different, because editing should occur in both sites if a single factor recognizes both C residues to be edited. This is consistent with the observation on transplastomic tobacco lines that the *ndhB* sites 7 and 8 are edited independently (Bock, Hermann, and Fuchs 1997).

Editing of *ndhD* (Site 1) Transcripts

The *ndhD* gene encodes NADH dehydrogenase subunit IV and is located in the middle of the small single-copy region of the *N. tabacum* chloroplast genome (Shinozaki et al. 1986). This gene is cotranscribed with the upstream *psaC*, and the resultant dicistronic pre-mRNA of approximately 2.5 kb is processed into monocistronic

Table 1
PCR Primers Used in This Work

Name[a]	Length[b]	Sequence (5′–3′)
atpA-A	481	AGACAGACCGGTAAACAGCAGTATCCACAGATA
atpA-B		CCAGAATTGAATAGGTCGGCGGATAAGAAGA
atpF-A	145	GATGTAGGTCCCCGTCGATGAAAAACGTAACCGATTCTT
atpF-B		GCTGTCGCCGTACTCGTACACTCCCTTTCCAAAAAAAATCAA
ndhA-A1	581	GATGTAGGTCCCCGTCGTATCTAACAGTTTAAGTACAGTTGATATAGTGG
ndhA-B1		GCTGTCGCCGTACTCGTTCTTGTTTGAGACAAGTCGTGAATC
ndhA-A2	554	GATGTAGGTCCCCGTCGAATGATAATTGATACAACAGAAATAGAAA
ndhA-B2		GCTGTCGCCGTACTCGTGTAGAGATATTGATAACACACATAGAGC
ndhA-A3	613	CTGAGGAGCCGTATGAGATGAAA
ndhA-B3		GAAAGAAGTTGGGACGAGGTTGT
ndhA-A4	1508	GTGGTCTACGGGCAGCTGCTCAATCAAT
ndhA-B4		TCCAACCTCCCAAATAAAGAACTGT
ndhB-A1	756	GATGTAGGTCCCCGTCGTCTCCCACTCCAGTCGTTGC
ndhB-B1		GCTGTCGCCGTACTCGTCTAAAAAAGGCTATCCTGAGCAAT
ndhB-A2	777	GATGTAGGTCCCCGTCGATGATCTGGCATGTACAGAATGAAA
ndhB-B2		GCTGTCGCCGTACTCGTTCCTTCGTATACGTCAGGAGTCCAT
ndhD-A1	731	CGAGCCTCGGGGACCGATCTATTAG
ndhD-B1		TCCTCCCCACATGCATACAACTACAT
ndhD-A2	383	CTATGCATGTGGGGAGGAAAGAAAC
ndhD-B2		CAAGGAGAAAATATAGAATGGGCATGAGGTAAT
ndhF-A	356	CCGTATGTGGGCTTTTCAGAGTGTTTTATTGTTA
ndhF-B		TGCCGCAACAGGTCGTG
petB-A1	1869	GCGCTGCAGCCCCCGTTTAGTTGACATGAATAGTTA
petB-B1		GCGGGGCCCAAGACGAACTGCGGTAGGAGAT
petB-A2	380	GCGCTGCAGCCCCCGTTTAGTTGACATGAATAGTTA
petB-B2		AAGTCGGTTATTGGGCAGTGAA
psbE-A	391	GATGTAGGTCCCCGTCGGTATACTCTAAAGACGCCTTCGGTA
psbE-B		GCTGTCGCCGTACTCGTGCCTCCTAAAACGATCTACTAAATTC
psbL-A	325	GATGTAGGTCCCCGTCGCCTAATGACTATAGATCGAACCTAT
psbL-B		GATGTAGGTCCCCGTCGAATTAAATTATGAGATCCTTCCGAAT
rpl20-A	464	TCGGGATAGGGATAGGAAAAAGAG
rpl20-B		AAAGCCCGAAATAAATAG
rpoA-A	238	CACACGGTTCCTTTATCCCCTTTTAC
rpoA-B		TGCCCAATATCCGTTTTACATCTTC
rpoB-A1	800	TCGGGGATGGAAATGAGG
rpoB-B1		TGGGTATATCAAGGTTCAGTCT
rpoB-A2	473	GATTCGGGGGCTCTTGCTAT
rpoB-B2		TTGCGGAGTAAATGGGCTTCTA
rpoC1-A	456	GATGTAGGTCCCCGTCGCAGTGTGATAAAAAAATCAATTAAAAA
rpoC1-B		GCTGTCGCCGTACTCGTAAAAAAAAGAGACGAGGAACCCTGA
rpoC2-A1	320	TAACTATAGCGCAAAGCCGAATCTCT
rpoC2-B1		TTCTCTTTCAGGCCTTTCAACCAAT
rpoC2-A2	1471	GATGTAGGTCCCCGTCGTTTTTGAAAAAGAATAAAAAAAAAAGG
rpoC2-B2		GCTGTCGCCGTACTCGTTTATTATCTTATTATTAATCAAGGATTTCTT
rps2-A1	496	TCTCGCAGGGCCGTATATTCTTCGTG
rps2-B1		TGGAGGCAGGAGTTCATTTTGGTCAT
rps2-A2	772	GATGTAGGTCCCCGTCGTTATTATGTTATTATTAATCAAGGATTTCTTA
rps2-B2		GCTGTCGCCGTACTCGTTTTTTTGAAAAAGAATAAAAAAAAAGG
rps14-A	468	GATGTAGGTCCCCGTCGTTCTATTTCTACATCTAGGATTCG
rps14-B		GCTGTCGCCGTACTCGTCGCCCCTCCACCCTATCTGTA

[a] A and B after gene names indicate DNA chains A and B, respectively. Sequences and chains can be positioned on the *N. tabacum* chloroplast genome (Wakasugi et al. 1998; accession number Z00044).

[b] Base pairs of amplified DNA fragments.

forms of approximately 2 kb (*ndhD*) and approximately 0.5 kb (*psaC*) in *N. tabacum* chloroplasts (Matsubayashi et al. 1987; Hayashida et al. 1987). The initiation codon of the *N. tabacum ndhD* gene was believed to be the ATG located 95 bp downstream from the *psaC* coding region, and a Shine-Dalgarno (SD)-like sequence (GAG) 12 bp upstream from the ATG was assigned as a potential ribosome binding site (fig. 2*A*; Matsubayashi et al. 1987). However, C-to-U editing has been shown to occur at ACG 25 nt downstream from the AUG in the *N. tabacum ndhD* transcript, which leads to the creation of an in-frame AUG codon (Neckermann et al. 1994). Our chloroplast in vitro translation assay has indicated that only the edited AUG codon acts as the actual initiation codon of *N. tabacum ndhD* mRNAs and that the upstream AUG has no function as an initiation codon (Hirose and Sugiura 1997). Recently, the leek *ndhD* transcript was reported to require editing to restore its start codon, which may be used as a marker for the processing of *psaC* and *ndhD* transcripts (Del Campo et al. 2002).

Our cDNA analysis showed that the ACG codon is also edited to produce AUG in *N. sylvestris*. However, editing of the ACG at the same position was not observed in *N. tomentosiformis* (fig. 2*B*). The experiment was

Table 2
RNA Editing Sites in *Nicotiana* Chloroplast Transcripts

Gene	Site	Position	Conversion[a]	*N. tabacum*[b]	*N. sylvestris*	*N. tomentosiformis*
atpA	1	264	P(cCc) → L(cUc)	+	+	Δ
	2	265	S(ucC) → S(ucU)	Δ	−	−
atpF		31	P(cCa) → L(cUa)	+	+	+
ndhA	1	114	S(uCa) → L(uUa)	+	+	+
	2[c]	189	S(uCa) → L(uUa)	(−)	(−)	+
	3	358	S(uCc) → F(uUc)	+	+	+
ndhB	1	50	S(uCa) → L(uUa)	+	+	+
	2	156	P(cCa) → L(cUa)	+	+	+
	3	196	H(Cau) → Y(Uau)	+	Δ	Δ
	4	204	S(uCa) → L(uUa)	+	Δ	Δ
	5	246	P(cCa) → L(cUa)	+	+	+
	6	249	S(uCu) → F(uUu)	+	+	+
	7	277	S(uCa) → L(uUa)	+	Δ	+
	8	279	S(uCa) → L(uUa)	+	+	−
	9	494	P(cCa) → L(cUa)	+	+	+
ndhD	1	1	T(aCg) → M(aUg)	Δ	Δ	−
	2	128	S(uCa) → L(uUa)	+	+	+
	3	200	S(uCa) → L(uUa)	Δ	Δ	Δ
	4	225	S(uCg) → L(uUg)	+	+	Δ
ndhF		97	S(uCa) → L(uUa)	+	+	+
petB		204	P(cCa) → L(cUa)	+	+	+
psbE		72	P(Cca) → S(Uca)	+	+	+
psbL		1	T(aCg) → M(aUg)	+	+	+
rpl20		103	S(uCa) → L(uUa)	+	+	+
rpoA		277	S(uCa) → L(uUa)	Δ	+	+
rpoC1		21	S(uCa) → L(uUa)	+	Δ	Δ
rpoC2		1248	S(uCa) → L(uUa)	+	+	+
rpoB	1	113	S(uCu) → F(uUu)	+	+	+
	2	158	S(uCa) → L(uUa)	+	+	+
	3	184	S(uCa) → L(uUa)	+	+	+
	4	667	S(uCu) → F(uUu)	+	+	+
rps2	1	45	T(aCa) → I(aUa)	+	+	+
	2	83	S(uCa) → L(uUa)	+	+	+
rps14	1	27	S(uCa) → L(uUa)	+	+	+
	2	50	P(cCa) → L(cUa)	+	+	+

[a] Position in nt is given with respect to the A of the initiation codon.

[b] *N. tabacum* editing sites *ndhD*-3, *ndhD*-4, and *rps2*-1 were identified by this work, and others are from Hirose et al. (1999). + indicates editing; (–) indicates no editing (T in the genome); – indicates no editing though C in the genome; Δ indicates partial editing.

[c] From Schmitz-Linneweber et al. (2001).

repeated two more times with different RNA preparations and the same negative results were obtained. Therefore, the mechanism of translational initiation for *ndhD* mRNAs is likely to be different between *N. tomentosiformis* and *N. tobacum*. *N. tomentosiformis* chloroplasts may possess a factor that can allow initiation from the upstream AUG codon, or its translation may start from the in-frame GUG codon 6 nt downstream from the AUG triplet (see fig. 2*A*). However, it cannot be ruled out that a functional *ndhD* gene is present in the nucleus of *N. tomentosiformis*.

Silent Editing in *atpA* Transcripts

The *atpA* gene encodes the α-subunit of ATP synthase complex and is located at the last gene in the *rps2-atpIHFA* operon (Wakasugi et al. 1998). Transcription of this operon in *N. tabacum* starts from at least four sites, and resulting polycistronic mRNAs are processed from at least four sites to produce a dicistromic *atpF-atpA* mRNA and other mRNAs (Miyagi et al. 1998). As shown in figure 3*A*, two editing sites were found in two successive codons of *N. tabacum atpA* transcripts (Hirose et al. 1996). The first C-to-U conversion caused proline (CCC) to leucine (CUC) substitution, whereas the second editing took place partially at the third position of the serine codon (UCC to UCU), leading to no amino acid change (silent editing). In the case of *N. sylvestris* and *N. tomentosiformis atpA* mRNAs, editing was detected in the first codon but not in the second position (fig. 3*B*), as was found in pea *atpA* mRNAs (Hirose et al. 1996). Therefore, silent editing at the serine codon seems to be unique to *N. tabacum*.

Partial Editing in Green Leaves

Most transcripts are essentially completely edited, whereas a limited number of editing sites are edited partially, suggesting that certain editing events play a regulatory role in gene expression (Bock et al. 1993; Hirose et al. 1996, 1999; Hirose and Sugiura 1997; Ruf and Kössel 1997; Nakajima and Mulligan 2001; Karcher and Bock 2002). Recently, a comprehensive analysis of the editing efficiency of each of the 27 known sites in maize chloroplasts was performed (Peeters and Hanson

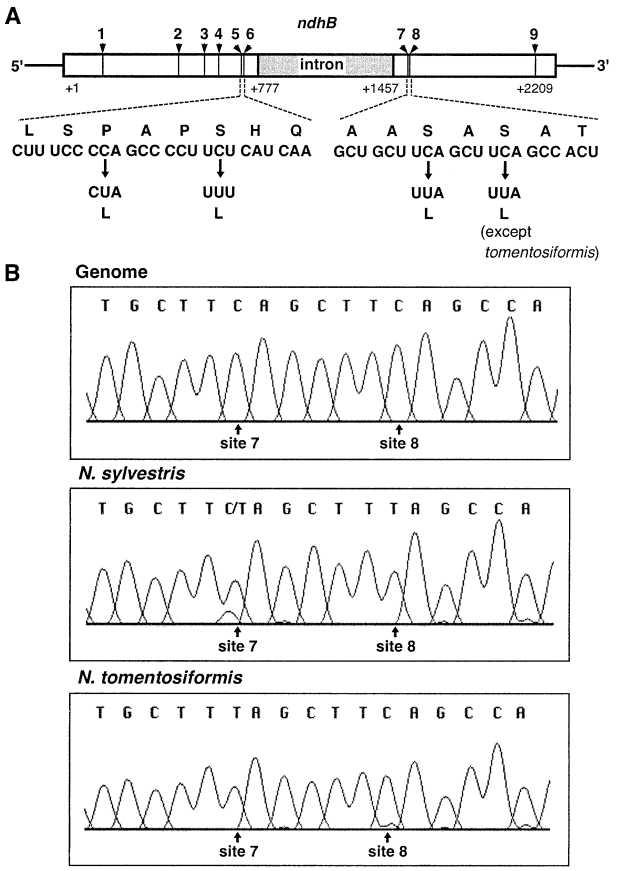

FIG. 1.—RNA editing of *ndhB* transcripts. (*A*) Schematic representation of the *N. tabacum ndhB* gene. Numbers at the top indicate editing sites. Nucleotide positions are from the A (+1) of its start codon. Partial sequences around sites 5 and 6 and sites 7 and 8 are shown below. Arrows indicate C-to-U conversion. (*B*) Sequence analysis of three cDNAs around sites 7 and 8 (indicated by arrows).

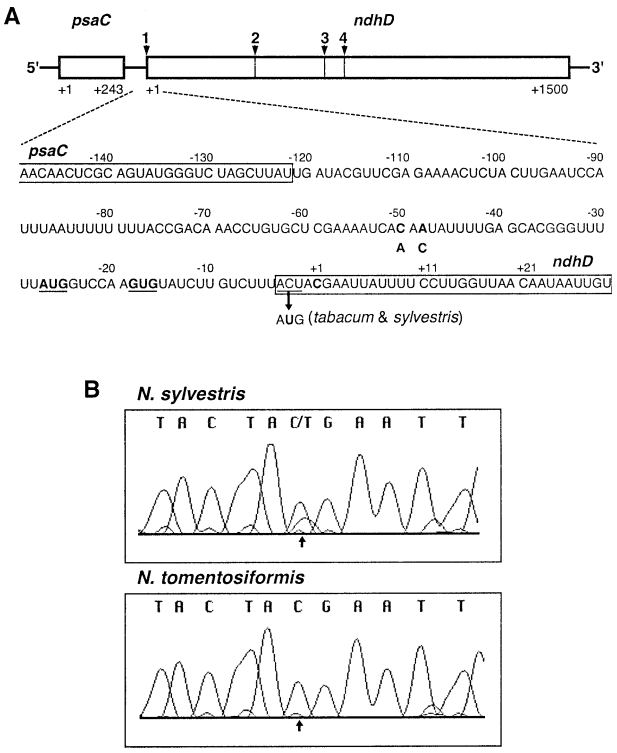

FIG. 2.—RNA editing of *ndhD* transcripts. (*A*) Schematic representation of the *N. tabacum psaC-ndhD* operon. Numbers at the top indicate editing sites. Nucleotide positions are from the A (+1) of its start codon. A partial sequence including the spacer is shown below. The sequence is identical in *N. tabacum* and *N. sylvestris* but differ in two positions –50 (C to A) and –49 (A to C) in *N. tomentosiformis*. Coding regions are boxed. Underlines show potential start codons. Arrow indicates C-to-U conversion. (*B*) Sequence analysis of two cDNA around site 1 (indicated by arrows).

2002). Editing efficiencies of some sites were found to be affected by the developmental stage, quite low in roots and calli, whereas editing of all 27 sites in young green leaves is close to 100% efficiency.

We used transcripts from young green leaves to identify editing sites in the *Nicotiana* species. Most identified sites are fully edited, whereas a limited number of identified sites are edited partially (table 2). Partial editing was confirmed by three independent cDNA sequencing from different RNA preparations. The extent of editing is dependent on the site. In *N. tabacum*, partial editing was reported in *atpA* site 2 (Hirose et al. 1996), *ndhD* site 1 (Hirose and Sugiura 1997) and *rpoA* (Hirose et al. 1999). An additional site, *ndhD* site 3, was found from our own analysis to be partially edited (fig. 4). Altogether four sites are partially edited in *N. tabacum* green leaves.

In *N. sylvestris* and *N. tomentosiformis* green leaves, six sites were found to be edited partially, of which four sites are common (*ndhB* sites 3 and 4, *ndhD* site 3, and *rpoC1*) between them (fig. 4). Editing at *ndhD* site 1 was partial in *N. sylvestris* (see fig. 4), similar to that in *N. tabacum* (Hirose and Sugiura, 1997). As this editing creates AUG from ACG, approximately half of the *ndhD* mRNA seems to be nonfunctional in *N. sylvestris* green leaves, as in the case of *N. tabacum*. The *ndhD* site 4 from *N. tomentosiformis* was edited slightly, but the site was edited fully in the other two species (fig. 4; Hirose et al. 1999).

Editing of *ndhB* sites 3 and 4 and *rpoC1* occurs fully in *N. tabacum* (Hirose et al. 1999) but not in the other species. This phenomenon could be explained by doubling nuclear genes encoding editing factors. Recently, it was reported that editing of the two *ndhB* sites in *N. tabacum* is temperature sensitive (Karcher and Bock 2002). Therefore, interaction of *cis* elements with editing machineries for these editing sites may be more fragile when compared with other sites. The corresponding sites in maize plastids were poorly edited in roots and calli (Peeters and Hanson 2002), suggesting low expression of the *trans* factors in nonphotosynthetic cells. However, the opposite case was observed for *rpoA* editing; partial in the tetraploid but full in the two diploid species. This may be due to interference in *N. tabacum* between the expression of nuclear genes encoding the editing factors or between these factors derived from the two progenitors. Partial editing implies the presence of two or more different mRNA species from single genes, which potentially leads to the microheterogeneity of relevant protein products. If this is not the

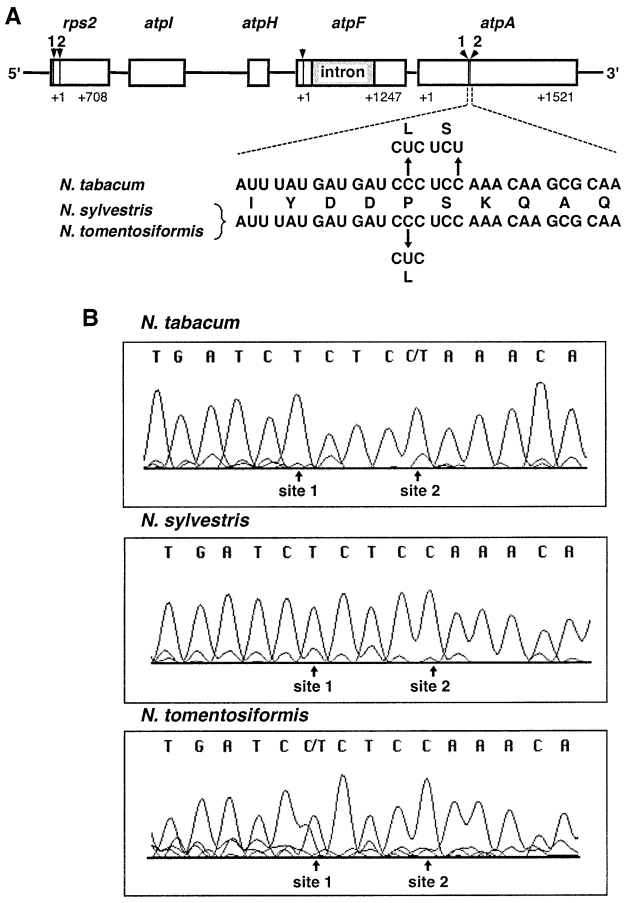

FIG. 3.—RNA editing of *atpA* transcripts. (*A*) Schematic representation of the *N. tabacum rps2-atpIHFA* operon. Numbers at the top indicate editing sites. A partial sequence around sites 1 and 2 is shown below. Arrows indicate C-to-U conversion. (*B*) Sequence analysis of two cDNAs around sites 1 and 2 (indicated by arrows).

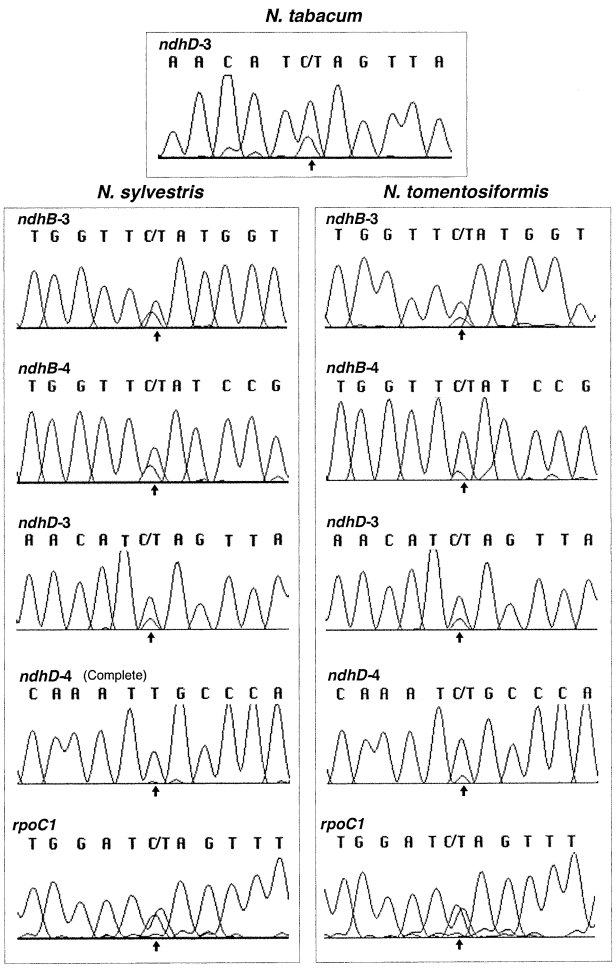

FIG. 4.—Partial RNA editing of *ndhB*, *ndhD* and *rpoC1* transcripts. Sequence analysis of cDNAs around editing sites where editing is partial. Arrows indicate editing sites.

case, mRNA surveillance mechanisms should operate in chloroplasts. In the case of an editing event creating AUG start codons, unedited mRNAs are most likely nonfunctional.

Conclusion

Recent findings show that an editing activity can be present despite the absence of the target site; *N. tabacum* is capable of editing the exogenous *ndhA* site 2 even though its plastid genome lacks this site (Schmitz-Linneweber et al. 2001). The nucleus of *N. tomentosiformis* is suggested to be the donor of the corresponding *trans* factor. Here we present evidence to the contrary; *N. tabacum* and *N. sylvestris* have *ndhB* site 8 and *ndhD* site 1 and their cognate editing activities, whereas no editing activity for these sites in *N. tomentosiformis* was detected even though these positions hold C residues in their genome. Therefore, these editing factors are thought to originate from *N. sylvestris*. As for the *ndhB* site 8, the *ndhB* protein may be functional with either serine or leucine at this position. If this is the case, editing of this site is dispensable and *N. tomentosiformis* lost its editing activity. This suggests that conserved amino acid residues are not always essential for protein function. Comparison of editing patterns among *N. tabacum* (amphidiploid) and its progeny representatives, *N. sylvestris* and *N. tomentosiformis*, will provide clues for better understanding of the evolution of editing events in plastids.

Supplementary Material

The sequences reported here have been deposited in the DNA Databank of Japan (DDBJ, accession numbers AB098210 to AB098251).

Acknowledgments

We thank T. Tsudzuki, T. Wakasugi, T. Nakamura, and K. Hasegawa for continuous discussions and suggestions. This work was supported in part by a Grant-in-Aid for Scientific Research in Priority Areas C (No. 13201001) from the Ministry of Education, Science, Sports and Culture.

Literature Cited

Bock, R., R. Hagemann, H. Kössel, and J. Kudla. 1993. Tissue- and stage-specific modulation of RNA editing of the *psbF* and

psbL transcript from spinach plastids—a new regulatory mechanism? Mol. Gen. Genet. **240**:238–244.

Bock, R., M. Hermann, and M. Fuchs. 1997. Identification of critical nucleotide positions for plastid RNA editing site recognition. RNA **3**:1194–1200.

Bock, R., M. Hermann, and H. Kössel. 1996. In vivo dissection of *cis*-acting determinants for plastid RNA editing. EMBO J. **15**:5052–5059.

Bock, R., and H. U. Koop. 1997. Extraplastidic site-specific factors mediate RNA editing in chloroplasts. EMBO J. **16**: 3282–3288.

Chaudhuri, S., H. Carrer, and P. Maliga. 1995. Site-specific factor involved in the editing of the *psbL* mRNA in tobacco plastids. EMBO J. **14**:2951–2957.

Chaudhuri, S., and P. Maliga. 1996. Sequences directing C to U editing of the plastid *psbL* mRNA are located within a 22 nucleotide segment spanning the editing site. EMBO J. **15**:5958–5964.

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

Del Campo, E., M. B. Sabater, and M. Martin. 2002. Post-transcriptional control of chloroplast gene expression. J. Biol. Chem. **27**:36457–36464.

Freyer, R., M.-C. Kiefer-Meyer, and H. Kössel. 1997. Occurrence of plastid RNA editing in all major lineages of land plants. Proc. Natl. Acad. Sci. USA **94**:6285–6290.

Hayashida, N., T. Matsubayashi, K. Shinozaki, M. Sugiura, K. Inoue, and T. Hiyama. 1987. The gene for the 9 kd polypeptide, a possible apoprotein for the iron-sulfur centers A and B of the phytosystem I complex, in the tobacco chloroplast DNA. Curr. Genet. **12**:247–250.

Hermann, M., and R. Bock. 1999. Transfer of plastid RNA-editing activity to novel sites suggests a critical role for spacing in editing-site recognition. Proc. Natl. Acad. Sci. USA **96**:4856–4861.

Hirose, T., H. Fan, J. Y. Suzuki, T. Wakasugi, T. Tsudzuki, H. Kössel, and M. Sugiura. 1996. Occurrence of silent RNA editing in chloroplasts: its species specificity and the influence of environmental and developmental conditions. Plant Mol. Biol. **30**:667–672.

Hirose, T., T. Kusumegi, T. Tsudzuki, and M. Sugiura. 1999. RNA editing sites in tobacco chloroplast transcripts: editing as a possible regulator of chloroplast RNA polymerase activity. Mol. Gen. Genet. **262**:462–467.

Hirose, T., and M. Sugiura. 1997. Both RNA editing and RNA cleavage are required for translation of tobacco chloroplast *ndhD* mRNA: a possible regulatory mechanism for the expression of a chloroplast operon consisting of functionally unrelated genes. EMBO J. **16**:6804–6811.

———. 2001. Involvement of a site-specific *trans*-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: development of a chloroplast in vitro RNA editing system. EMBO J. **20**:1144–1152.

Hoch, B., R. M. Maier, K. Appel, G. L. Igloi, and H. Kössel. 1991. Editing of a chloroplast mRNA by creation of an initiation codon. Nature **353**:178–180.

Karcher, D., and R. Bock. 2002. Temperature sensitivity of RNA editing and intron splicing reactions in the plastid *ndhB* transcript. Curr. Genet. **41**:48–52.

Kenton, A., A. S. Parokonny, Y. Y. Gleba, and M. D. Bennett. 1993. Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. Mol. Gen. Genet. **240**:159–169.

Kudla, J., and R. Bock. 1999. RNA editing in an untranslated region of the ginkgo chloroplast genome. Gene **234**:81–86.

Maier, R. M., K. Neckermann, G. L. Igloi, and H. Kössel. 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.

Maier, R. M., P. Zeltz, H. Kössel, G. Bonnard, J. M. Gualberto, and J. M. Grienenberger. 1996. RNA editing in plant mitochondria and chloroplasts. Plant Mol. Biol. **32**:343–365.

Matsubayashi, T., T. Wakasugi, K. Shinozaki et al. (12 co-authors). 1987. Six chloroplast genes (*ndhA-F*) homologous to human mitochondrial genes encoding components of the respiratory chain NADH dehydrogenase are actively expressed: determination of the splice sites in *ndhA* and *ndhB* pre-mRNAs. Mol. Gen. Genet. **210**:385–393.

Miyagi, T., S. Kapoor, M. Sugita, and M. Sugiura. 1998. Transcript analysis of the tobacco plastid operon *rps2/atpI/H/F/A* reveals the existence of a non-consensus type II (NCII) promoter upstream of the *atpI* coding sequence. Mol. Gen. Genet. **257**:299–307.

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

Nakajima, Y., and R. M. Mulligan. 2001. Heat stress results in incomplete C-to-U editing of maize chloroplast mRNAs and correlates with changes in chloroplast transcription level. Curr. Genet. **40**:209–213.

Neckermann, K., P. Zeltz, G. L. Igloi, H. Kössel, and R. M. Maier. 1994. The role of RNA editing in conservation of start codons in chloroplast genomes. Gene **146**:177–182.

Olmstead, R., and J. D. Palmer. 1991. Chloroplast DNA and systematics in the *solanaceae*. Pp.301–320 *in* J. G. Hawkes, R. N. Lester, M. Nee and N. Estrada, eds. *Solanaceae* III. Taxonomy, chemistry and evolution. Kew, Royal Botanic Gardens, Linnean Society of London, London.

Peeters, N. M., and M. R. Hanson. 2002. Transcript abundance supercedes editing efficiency as a factor in developmental variation of chloroplast gene expression. RNA **8**: 497–511.

Reed, M. L., and M. R. Hanson. 1997. A heterologous maize *rpoB* editing site is recognized by transgenic tobacco chloroplasts. Mol. Cell. Biol. **17**:6948–6952.

Reed, M. L., S. M. Lyi, and R. M. Hanson. 2001. Edited transcripts compete with unedited mRNAs for *trans*-acting editing factors in higher plant chloroplasts. Gene **272**:165–171.

Rochaix, J. D. 2001. Posttranscriptional control of chloroplast gene expression: from RNA to photosynthetic complex. Plant Physiol. **125**:142–144.

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

Schmitz-Linneweber, C., M. Tillicch, R. G. Herrmann, and R. M. Maier. 2001. Heterologous, splicing-dependent RNA editing in chloroplasts: allotetraploidy provides *trans*-factors. EMBO J. **20**:4874–4883.

Shinozaki, K., M. Ohme, M. Tanaka et al. (19 co-authors). 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J. **5**: 2043–2049.

Smith, H. H. 1974. *Nicotiana*. Pp.281–314 *in* R. C. King, ed. Handbook of genetics 2. Plenum Press, New York.

Sugiura, M. 1992. The chloroplast genome. Plant Mol. Biol. **18**: 149–168.

Sugiura, M., T. Hirose, and M. Sugita. 1998. Evolution and mechanism of translation in chloroplasts. Annu. Rev. Genet. **32**:437–459.

Svab, Z., and P. Maliga. 1993. High frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. Proc. Natl. Acad. Sci. USA **90**:913–917.

Wakasugi, T., T. Hirose, M. Horihata, T. Tsudzuki, H. Kössel, and M. Sugiura. 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**: 8766–8770.

Wakasugi, T., M. Sugita, T. Tsudzuki, and M. Sugiura. 1998. Updated gene map of tobacco chloroplast DNA. Plant Mol. Biol. Reptr. **16**:231–241.

Wakasugi, T., J. Tsudzuki, S. Ito, K. Nakashima, T. Tsudzuki, and M. Sugiura. 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., T. Tsudzuki, and M. Sugiura. 2001. The genomics of land plant chloroplasts: gene content and alteration of genomic information by RNA editing. Photosynthesis Res. **70**:107–118.

Yoshinaga, K., H. Iinuma, T. Masuzawa, and K. Ueda. 1996. Extensive RNA editing of U to C in addition to C to U substitution in the *rbcL* transcripts of hornwort chloroplasts and the origin of RNA editing in green plants. Nucleic Acids Res. **24**:1008–1014.

Kenneth Wolfe, Associate Editor

Accepted January 14, 2003