Theor Appl Genet (2000) 100:214–220

T. Kubo · M. P. Yamamoto · T. Mikami

The *nad4L-orf25* gene cluster is conserved and expressed in sugar beet mitochondria

Received: 12 April 1999 / Accepted: 22 June 1999

Abstract We have found that a gene coding for NADH dehydrogenase subunit 4L and a presumed gene, *orf25*, are linked and co-transcribed with each other in sugar beet mitochondria. Ten and twelve C-to-U editing events were observed in the mRNAs of *nad4L* and *orf25*, respectively; the amino-acid sequence specified after editing is better-conserved in comparison with the homologues of other organisms. It is interesting to note that the translation initiation codon of *nad4L* is created by editing. The conservation of the *nad4L-orf25* linkage was examined by PCR-amplification of the intergenic region. We obtained successful PCR products from five dicots (spinach, apple, snapdragon, petunia and tobacco) and two monocots (tulip and pineapple), but not in two poaceous plants, rice and maize. The intergenic region, when present, was found to be well-conserved in its sequence, suggesting a monophyletic origin of this linkage. Our result, together with previous reports of *Arabidopsis* and four poaceous species, favour the argument that the *nad4L-orf25* linkage is conserved throughout angiosperms except in the Poaceae.

Key words *Nad4L-orf25* linkage · RNA editing · Sugar beet · Mitochondria

Introduction

The mitochondrial DNAs (mtDNAs) of flowering plants contain a number of conserved but functionally unidentified ORFs. One such ORF is *orf25* which was originally described in the Texas male-sterile cytoplasm of maize (Dewey et al. 1986) and subsequently revealed to be conserved in the mitochondrial genomes of several other plant species (Stamper et al. 1987; Bonen et al. 1990; Brandt et al. 1992; Liu et al. 1992; Oda et al. 1992; Tang et al. 1996;) and of some protists (designated as *ymf39*) (Gray et al. 1998). Moreover, the translation product of *orf25* was found in a membrane fraction of maize, wheat and tobacco mitochondria (Prioli et al. 1993) suggesting that *orf25* is a functional gene.

In the Texas cytoplasm of maize, *orf25* is preceded by *urf13-T*, a gene responsible for cytoplasmic male sterility (Dewey et al. 1986). On the other hand, rice *orf25* is situated upstream of *cox3* and is co-transcribed with the latter (Liu et al. 1992). No accompanying genes are reported in the vicinity of the *orf25* locus of wheat (Bonen et al. 1990) and sorghum (Tang et al. 1996). Hence, poaceous plants show no evident common feature in the genomic organization of the *orf25* locus. In the present paper we report that this contrasts with the situation found in a variety of angiosperms, including *Arabidopsis* (Brandt et al. 1992) and sugar beet, where a highly conserved arrangement of *nad4L-orf25* genes is maintained with the *orf25* gene downstream from the *nad4L* gene.

It is well known that gene organization in flowering plant mitochondria is less conserved (Schuster and Brennicke 1994). Closer analysis of the *nad4L-orf25* gene cluster can thus be expected to provide valuable information about the evolution of plant mitochondrial genomes.

Communicated by K. Glimelius

T. Kubo · M. P. Yamamoto · T. Mikami (✉)
Laboratory of Genetic Engineering, Faculty of Agriculture, Hokkaido University, Sapporo 060–8589, Japan
e-mail: gelab@abs.agr.hokudai.ac.jp
Fax: +81 11 716 0879

Materials and methods

Sugar beet cv TK81-O (male-fertile cytoplasm) used in this study was obtained from the Hokkaido National Agricultural Experiment Station, Japan. Plant material of spinach, apple, petunia, maize, tulip and pineapple was purchased at a local market. Isolation of mtDNA and mtRNA from green leaves was described previously (Mikami et al. 1985; Senda et al. 1993). Total cellular DNA was isolated using an ISOPLANT kit (Nippon Gene, Tokyo, Japan). Rice and snapdragon DNAs were the gift of Dr. Y. Kishima, Hokkaido University, Japan. Northern-blot analysis and the reverse transcribed polymerase chain reaction (RT-PCR) were done as described in Kubo et al. (1993). Procedures for nucleic acid techniques were performed according to standard protocols (Sam-

brook et al. 1989). Sequencing of the plasmid DNA was done using Thermo Sequenase (Amersham, Little Chalfont, UK) and DNA sequencer Li-COR 4000L (Li-COR, Lincoln, Neb., USA) according to the instruction manual. PCR primers used in this study are as follows:

1, 5'-CTCTTACATTCTACGTTCCCG-3'
2, 5'-GGTAACTTTCTAGGAGTCATG-3'
3, 5'-GAAAGATAAAGATAAGCTTTC-3'
4, 5'-TCTTCTGTTAGCATGAACATG-3'
5, 5'-TATTACTTTCCGAGTCCGGGG-3'
6, 5'-TCTTCTTCGAACTTGATGCAC-3'.

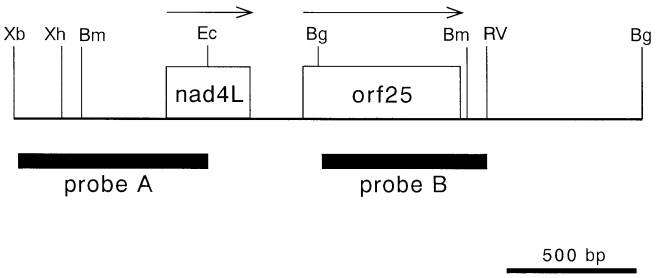

Fig. 1 Organization of the *nad4L-orf25* locus in sugar beet mitochondria. *Open boxes* represent ORFs for *nad4L* and *orf25*. The direction of transcription is indicated by *horizontal arrows*. The location and extent of probes used for Northern-blot analysis is shown by *bold lines*. Restriction sites are shown as: *Bg* for *Bgl*II, *Bm* for *Bam*HI, *Ec* for *Eco*RI, *RV* for *Eco*RV, *Xb* for *Xba*I, and *Xh* for *Xho*I

Fig. 2 Nucleotide sequence of the sugar beet *nad4L-orf25* locus with the corresponding amino acid-sequence. Numbering of nucleotides is from the beginning of the *nad4L* ORF. The cytidine residues altered by RNA editing are shown by *lower case letters* and numbered by *roman figures*. The amino-acid residues specified after editing are also shown. Positions of PCR primers are *underlined*. This sequence has been deposited in DDBJ/EMBL/GenBank under accession number AB020062

```
ttgaacgagattttggttttattataatataAATAACATTTGAATTTCTCTTACATTCTACGTTCCCGAA   -1
                                                1
nad4L
T>M  D   S   I   K   Y   F   T   F   S   M   I   I   F   I  S>L  G
AcG GAT TCT ATC AAA TAT TTT ACA TTT TCT ATG ATC ATC TTT ATT TcA GGT       51
 I                                                          II

 I  R>W  G   I   L   L   N   R   R   N   I   L   I   M  S>L  M   S
ATT cGG GGT ATT CTC CTT AAT AGA CGA AAT ATT CTT ATT ATG TcA ATG TCA      102
    III                                                  IV

 I   E  S>L  M   L   L   A   V   N  S>L  N   F   L   V   F   S   V
ATT GAA TcA ATG TTA TTA GCT GTG AAT TcG AAC TTT TTG GTA TTT TCT GTT      153
          V                        VI

 S  S>L  D   D   M   M   G   Q  S>L  F   A   L   L   V  P>L  T   V
TCT TcG GAT GAT ATG ATG GGT CAA TcA TTT GCT TTA TTG GTT CcA ACG GTG      204
     VII                       VIII                   IX

 A   A   A   E   S   A   I   G   L   A   I   F   V   I   T   F   R
GCA GCT GCG GAA TCT GCT ATT GGG TTA GCC ATT TTC GTT ATT ACT TTC CGA      255
                                                    5
 V   R   G   T   I   A   V   E  S>F  I   N   S   I   Q   G   *
GTC CGG GGT ACT ATT GCT GTA GAA TcT ATT AAT AGC ATT CAA GGT TAA ACATGA  309
                                   X

CTCCTAGAAAGTTACCAAAATACGAAGTTCTTTCGTTCTCTTCTTTCTTTTGACTTGGTTGGCAGGATCA  379
                2
GGGCCTTTCTCGCTGGGCGAGCGCATCCGATTCGAACTTTCCTAAAAAACTTCCCGTTCAGTTGCTGAAA  449
                                                                    3
                        orf25
                          M   R   K   S  S>F  T   K   M   Q   A   R   K
GATAAAGATAAGCTTTCTAA ATG AGA AAG AGT TcC ACG AAG ATG CAG GCT AGA AAG     505
                                       I
 M   L   F   A   A   I   L   S   I   C   A  S>L  S   S   K   K   I
ATG CTA TTT GCT GCT ATT CTA TCT ATT TGT GCA TcA AGT TCG AAG AAG ATC      556
                                           II                   6

S>L  I   Y   N   E   E   M   I   V   A  R>C  C   F   I   G   F   I
TcA ATC TAT AAT GAA GAA ATG ATA GTA GCT cGT TGT TTT ATA GGC TTT ATC      607
 III                                    IV

 I   F   S   R   K   S   L   G   N   T   F   K   V   T   L   D   E
ATA TTC AGT CGG AAG AGT TTA GGT AAT ACT TTC AAA GTG ACT CTC GAC GAG      658

 R   I   Q   A   I   Q   E   E  S>L  Q   Q   F  P>L  N   P   N   E
AGA ATC CAG GCT ATT CAG GAA GAA TcG CAG CAA TTC CcC AAT CCT AAC GAA      709
                                 V                VI

 V   V  P>L P>L  E   S   N   E   Q   Q   R   L   L   R   V   S   L
GTA GTT CcT ccG GAA TCC AAT GAA CAA CAA CGA TTA CTT AGG GTC AGC TTG      760
          VII   IX
            VIII

 R   I   C   G   T   V   V   E   S   L   P   M   A   R   C   A   P
CGA ATT TGT GGA ACC GTA GTA GAA TCA TTA CCA ATG GCA CGC TGT GCG CCT      811

 K   C   E   K   T   V   Q   A   L   L   C   R   N  L>L  N   V   K
AAG TGC GAA AAA ACA GTG CAA GCT TTG TTA TGT CGA AAC cTA AAT GTT AAG      862
                                             X

S>L  A   T   L  P>L  N   A   T   S   S   R   R   T   R   L   Q   D
TcA GCA ACA CTT CcA AAT GCC ACT TCT TCC CGT CGC ACC CGT CTT CAG GAT      913
XI              XII

 D   L   V   T   G   F   H   F   S   V   S   E   R   F   V   P   G
GAT CTA GTC ACA GGT TTT CAC TTC TCA GTG AGT GAA AGA TTT GTC CCC GGG      964

 S   T   L   K   A   S   I   V   E   L   I   R   E   G   L   A   V
TCT ACG TTG AAA GCT TCT ATA GTA GAA CTC ATT CGA GAA GGC TTG GCG GTC     1015

 L   R   M   V   R   V   G   G   S   L   K   N   K   E   D   K   *
TTA AGA ATG GTT CGG GTA GGA GGT TCT CTT AAG AAT AAA GAA GAC AAA TAG     1066
AATCTAATTCATGTTCATGCTAACAGAAGAGCGGATCCAATACCAAGACGACTTCTTTCTCAGAAAAGCA 1136
                                  4
```

A. NAD4L

```
     T                 S  R              S    S       S
     /                 /  /              /    /       /
Bv   MDSIKYFTFSMIIFILGIWGILLNRRNILIMLMSIELMLLAVNLNFLVFS    50
At   ..L.....␣.....␣.....␣.....␣.....S.....              50
Mp   ..LV..L.....L.L.....F...K.....                      50
Pw   ..LS..L.V...L.L.....F...K..IV.....L..               50
Nc         MNITL.L.LI..L.FV...K..IL..I...I....I..LI..S.  44

       S       S       P                          S
       /       /       /                          /
     VSLDDMMGQLFALLVLTVAAAESAIGLAIFVITFRVRGTIAVEFINSIQG   100
     .␣.....V.␣.␣.....                                   100
     .Y..........F.............L.....I.........CMK.      100
     .YI..CI.....I.............LL.VYY.I.........LMK.     100
     LNM..II..TY.IYIIV..G.........L.AFY.L..S.TI.YK        89
```

B. ORF25

```
         S                    S      S          R
         /                    /      /          /
Bv   MRKSFTKMQARKMLFAAILSICALSSKKILIYNEEMIVACCFIGFIIFSR   50
At   ..L.T.N.D.....................␣.....L.....            50
Zm   ..L.GMD.KK.....F.....␣...␣.....␣.....L....           50
Mp            M.EI.IF....FSV.............V...LS.VC.V...Q   41
Pw       MKPLFQPYI.LYIV.AF.VA...H.......TL.LI..F...FAIK   46

                              S   P       PP
                              /   /       //
     KSLGNTFKVTLDERIQAIQEELQQFLNPNEVVLLESNEQ-QRL----L--   93
     ....T.......GSL.....␣...P.....I␣.....-...----.--    93
     ....K...E...G..ES...S....C.....I.E..I..-.....----.NL 95
     .TF.E.I.AIF.A.SE.LLSD...WMSYQ.AM.S.LKK.-HE.RSIS.--   88
     H.F...ITESIQ..GDT.HD..K.SYVTQISQD.AIR..Y.QI----I--   90

                                            S   P
                                            /   /
     RVSLRICGT--VV-ESLP--MA-RCAPKCEKTVQALLCRNLNV-KLATLL   136
     .I.......--..-....--..-.....................-.␣...T   136
     QI.....S.VK..-.␣..--A.-.....................-.␣....  140
     .S.TQMI.E--SC-INDM--VT-.....KQ..KSV..QQIEQ-..K...    131
     PV.HVLDVLSTSLI.QSKATQKGAQSSLKQQIDFQIYKMYTHLY..QQSF   140

     NATSSRRTRLQDDLVTGFHFSVSERFVPGSTLKASIV----------      173
     ..L....I.F.....K.YTL.GKQ.AYSCIS..ER.----------      173
     ..L....I.....I.............S...TLVGAETVEQIREAFLLEP   190
     AIQEHS.IS..EKI..C.RET.CDE.RFSKLR.HQS----------       167
     PQKLQKKIHDNV-.LN-----------------------------       155

     -ELIREGLAVLRMVRVGGSLKNKEDK                           198
     -.F...S.V.....G.VFS                                  192
     RD.....V...K.....IPGKRSCGDGVGL                       221
     -K.VQQSM-..LKDGVPK                                   183
     ---.QQS..LSQKPIN.IKNAK.KL.AKNFA                      183
```

Fig. 3 A and B Alignment of the amino-acid sequence of *nad4L* and *orf25*. Gaps indicated by *dashes* are incorporated for maximum matching. Abbreviations are: *Bv* for sugar beet; At for *Arabidopsis* (Brandt et al. 1992); *Mp* for the liverwort *Marchantia polymorpha* (Oda et al. 1992); *Pw* for the chlorophyte alga *Prototheca wickerhamii* (Wolff et al. 1994); *Nc* for *Neurospora crassa* (Nelson and Macino 1987); *Zm* for *maize* (Ward and Levings 1991). The NAD4L (**A**) and ORF25 (**B**) polypeptides are deduced from cDNA (*Bv, At* and *Zm*) or genomic DNA (*Mp, Pw* and *Nc*) sequences. Amino-acid sequences encoded by genomic DNA in sugar beet mitochondria are shown above. The amino-acid residues altered by editing are *underlined* in At and Zm

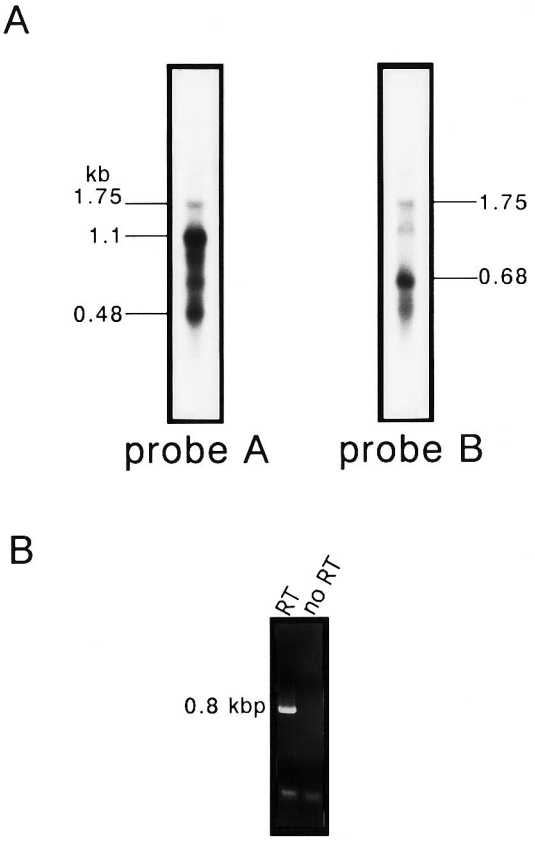

Fig. 4 A Northern-blot analysis of sugar beet mitochondrial RNA (mtRNA) with the probes shown in Fig. 1. Sizes of the transcripts are shown in kb. **B** PCR-amplification results demonstrating the co-transcription of *nad4L* and *orf25*. To test for DNA contamination in the RNA preparations, parallel PCR-amplifications were performed omitting the reverse transcriptase enzyme (*no RT*) in the reverse transcription

Results

Organization of the *nad4L-orf25* gene cluster in sugar beet

We previously observed that the gene probes for *nad4L* and *orf25* both hybridized to the same phage clone #336 of TK81-O mtDNA (Kubo et al. 1995). To confirm the linkage of *nad4L* and *orf25* in sugar beet mitochondria, two overlapping sub-clones (containing 1.7-kbp *Xba*I-*Eco*RV and 1.2-kbp *Bgl*II fragments, respectively) were isolated by colony hybridization and their nucleotide sequences determined (Fig. 1). Two ORFs having high homology to the *nad4L* and *orf25* sequences in other plant mitochondria were found to be adjacently encoded on the same DNA strand (Fig. 2). Unlike *Arabidopsis*, the third exon of *nad5* (*nad5/c*) is not coupled to the *nad4L-orf25* locus in sugar beet: it was mapped a minimum of 85 kbp away from the *nad4L-orf25* locus (Kubo et al. 1995; N. Itchoda et al., unpublished). The points of divergence between the sugar beet and *Arabidopsis* sequences are located 39 bp upstream of the *nad4L* ATG initiation codon and 118 bp downstream from the *orf25* termination codon. Our analysis failed to identify other coding sequences or the homologous sequences of *Arabidopsis* mitochondria in the region presented in Fig. 1.

A comparison of the *orf25* amino-acid sequences from a variety of flowering plants indicates variable and conserved regions of the gene (Bonen et al. 1990). In sugar beet *orf25*, we found that the conserved region is confined to the 5' half of the ORF (Fig. 3 B). The inser-

Table 1 Editing patterns in sugar beet *nad41* and *orf25* mRNAs

nad4L

Site of editing										Number of cDNAs analyzed
I	II	III	IV	V	VI	VII	VIII	XI	X	
+	+	+	+	+	+	+	+	+	+	6
–	+	+	+	+	+	+	+	+	+	1
+	+	+	+	+	+	+	+	+	–	1
–	+	–	+	–	+	–	–	+	+	1
–	–	–	–	–	–	–	–	–	–	1

orf25

Site of editing												Number of cDNAs analyzed
I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	
+	+	+	+	+	+	+	+	+	+	+	+	1
+	+	+	+	+	+	+	–	+	+	+	+	1
+	+	+	+	+	+	+	+	+	–	+	+	1
–	+	+	+	+	+	+	–	+	+	+	+	1
+	+	+	+	+	+	+	–	+	–	+	+	4

tion of 14 amino-acid residues in the 3' region of maize and sorghum *orf25* is missing in the sugar beet *orf25*, as was reported in *Arabidopsis* and tobacco (Fig. 3 B and data not shown).

Transcription analysis

To examine the expression of sugar beet *nad4L* and *orf25* genes, we tried to identify their transcripts by Northern-blot analysis using total mtRNA. As shown in Fig. 4, the *nad4L*- and *orf25*-specific probes (see Fig. 1) detected several transcripts ranging in size from 1.75 to 0.48 kb. The presence of a 1.75-kb low-abundant transcript common to the two RNA blots suggests the co-transcription of *nad4L* and *orf25*. This was confirmed by RT-PCR experiments using a set of primers of 5+4 (Fig. 2); a PCR product of the size expected (0.8 kbp) was obtained as shown in Fig. 4 and the nature of the PCR product was checked by nucleotide sequencing. The smaller and predominant transcripts specifically hybridized to either of the two probes: the *nad4L* probe hybridized to the 1.1- and 0.48-kb bands and the *orf25* probe to the 0.68-kb band. These transcripts are large enough to contain the entire coding region for each gene.

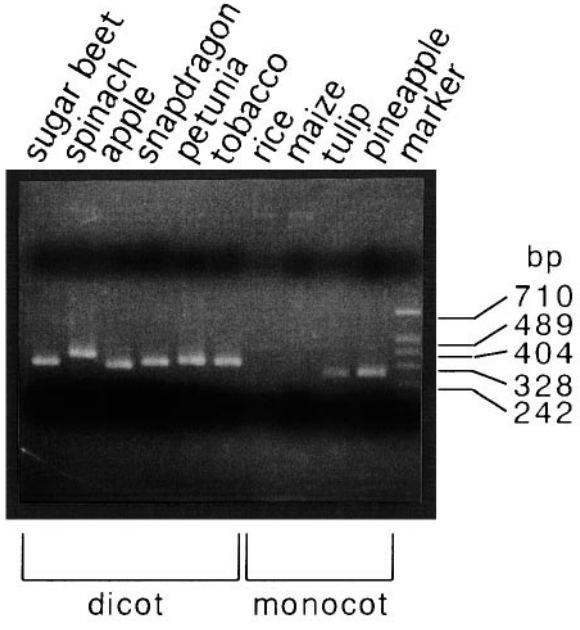

Fig. 5 Ethidium bromide staining of the 2% agarose gel after electrophoresis of the PCR-amplification of the *nad4L-orf25* intergenic region. Total cellular DNA was isolated from sugar beet, spinach, apple, snapdragon, petunia, tobacco, rice, maize, tulip and pineapple. pBluescript SK (+) DNA digested with *Hpa*II is included as a size marker

RNA editing

RNA editing of the sugar beet *nad4L* and *orf25* transcripts was analyzed by cDNA sequencing. Two sets of primers were used to generate cDNAs of the *nad4L* (primers 1+2) and *orf25* (3+4) loci. We identified ten C-to-U editing sites for *nad4L* and 12 C-to-U editing sites for *orf25*, respectively (Fig. 2). Most of the editing events improve the similarity between sugar beet NAD4L and ORF25 proteins and their counterparts in other organisms (Fig. 3). It should also be noted that the initiation codon of *nad4L* is created by editing. The distribution of editing sites varies between *nad4L* and *orf25*

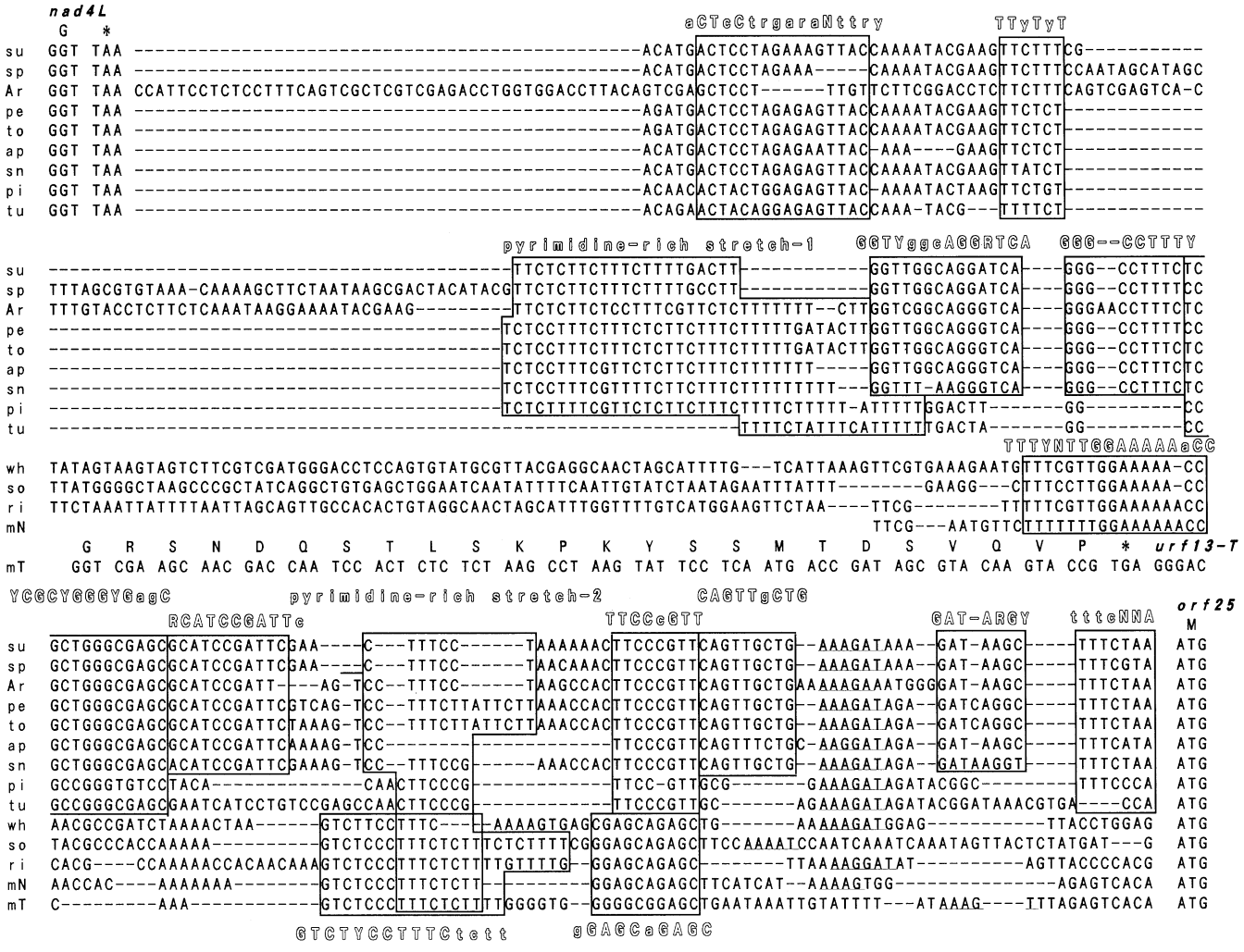

Fig. 6 Nucleotide sequence of the *nad4L-orf25* intergenic region of sugar beet (*su*), spinach (*sp*), *Arabidopsis* (*Ar*) (Brandt et al. 1992), petunia (*pe*), tobacco (*to*), apple (*ap*), snapdragon (*sn*), pineapple (*pi*) and tulip (*tu*). The upstream sequence of the *orf25* of wheat (*wh*) (Bonen et al. 1990), sorghum (*so*) (Tang et al. 1996), rice (*ri*) (Liu et al. 1992), and maize with normal cytoplasm (*mN*) (Stamper et al. 1987) and with Texas-cytoplasm (*mT*) (Dewey et al. 1986) are also shown. Gaps indicated by *dashes* are incorporated for maximum matching. The short stretches of conserved sequence and the pyrimidine-rich stretches are *boxed*. The nucleotide sequence beyond the 91 bp upstream of mN *orf25* is not registered in the data base. These sequence have been deposited in DDBJ/EMBL/GenBank under accession numbers AB026712 (tulip), AB026713 (spinach), AB026714 (snapdragon), AB026715 (pineapple), AB026716 (petunia), AB026717 (tobacco) and AB026718 (apple)

transcripts. The editing sites were found to be scattered throughout the entire *nad4L* ORF, while those in *orf25* were confined to the conserved region (Figs. 2 and 3).

The pattern of editing observed is summarized in Table 1. Six out of twelve cDNA clones for *nad4L* were modified at all observed editing sites. On the other hand, analysis of the *orf25* cDNA clones indicates the completely edited molecules to be less abundant; seven out of the eight clones lack the editing events at either one or both of the two sites VIII and X.

The *nad4L-orf25* gene cluster is conserved among angiosperms

The finding of the *nad4L-orf25* linkage in sugar beet as well as in *Arabidopsis* (Brandt et al. 1992) prompted us to investigate whether this linkage is phylogenetically conserved within angiosperms. We designed a set of primers (primers 5, 6) based on the conserved sequences of the *nad4L* and *orf25* coding region (Fig. 2). The *nad4L/orf25* locus in angiosperms, including five dicots and four monocots, was subjected to PCR analysis using these primers. As shown in Fig. 5, seven plant species other than rice and maize gave PCR products ranging in size from 300 to 380 bp, which suggests the conservation of the *nad4L-orf25* linkage in the seven species.

The PCR products were further sequenced either directly or after cloning. Figure 6 shows an alignment of the nucleotide sequences determined, as well as the *Arabidopsis* sequence (Brandt et al. 1992). The upstream sequences of *orf25* in wheat (Bonen et al. 1990), sorghum (Tang et al. 1996), rice (Liu et al. 1992) and the Texas male-sterile and normal fertile cytoplasms of maize (Dewey et al. 1986; Stamper et al. 1987) are also

included in the comparison. The intergenic spacers between *nad4L* and *orf25* are similar in the seven dicots and two monocots (tulip and pineapple) compared here, though the degree of conservation between these sequences is not so high as in coding regions. The spacer region can be subdivided into several blocks of homology, almost all of which are absent from the poaceous plants shown in Fig. 6. This suggests a common phylogenetic origin of the *nad4L-orf25* linkage. Sequence divergencies include single base substitutions and insertions/deletions; the intergenic spacers in *Arabidopsis* and spinach are characterized by somewhat large insertions in the 5' half of the spacer.

Discussion

Most plant mitochondrial genes are scattered singly in the genome and gene order is highly variable as a result of frequent inversions and other rearrangements (Schuster and Brennicke 1994). Notable exceptions are the intimate associations *rrn18-rrn5*, *orfB* (recently proposed to be *atp8*; see Gray et al. 1998)-*cox3*, *nad1-rps13*, *rps3-rpl16* and *nad3-rps12* that are commonly found in the mtDNAs of different plant species (Schuster and Brennicke 1994).

The present paper shows that the clustered organization of the *nad4L* and *orf25* genes can be added to the list. Our analysis leads to three observations. (1) Sugar beet mitochondrial *nad4L* and *orf25* genes are 166-bp apart and are co-transcribed. (2) The *nad4L* and *orf25* transcripts are edited in several positions, one of these resulting in the creation of an initiation codon of *nad4L*, indicating that the two genes are functional. (3) Our data also demonstrate that the *nad4L-orf25* gene cluster is conserved in five dicot plants (spinach, apple, snapdragon, petunia and tobacco) and two monocot plants (tulip and pineapple), but not in two poaceous plants, rice and maize. These genes were previously reported to be physically and transcriptionally unlinked in four poaceous plants, maize, wheat, sorghum and rice, thus leaving one with the inference that the disconnection of the two genes is limited in its taxonomic distribution to the Poaceae.

Conservation of a gene cluster against the extensive genome rearrangements in higher-plant mitochondria (Schuster and Brennicke 1994) raises the possibility of some structural and/or functional constraint. One can imagine that the *nad4L-orf25* intergenic sequence (some 150–230 bp) might be implicated in reducing the chance of recombination with other regions. Otherwise, the *nad4L-orf25* linkage might be selectively preserved probably because of the mode of expression in angiosperms other than the Poaceae. Suggestive evidence for the functional significance of this linkage also includes an observation that the primary sequence of the *nad4L-orf25* intergenic region shows a relatively high degree of conservation among the nine plant species compared here (see Fig. 6).

The question inevitably arises when this gene cluster emerged. The *nad4L* and *orf25* genes (or their homologues) are found in some protists as well as *Marchantia* (Oda et al. 1992; Gray et al. 1998). However, neither protists nor *Marchantia* retain the *nad4L-orf25* gene cluster. Thus, it seems likely that this linkage arose at some time after the divergence of the lineage leading to *Marchantia* but before the species divergence in angiosperms. Further analysis of a wider range of plant species is necessary to allow more definite conclusions about the evolution of the *nad4L-orf25* gene cluster.

Acknowledgements We gratefully acknowledge Dr. Y. Kishima for providing DNA samples and valuable suggestions. This work was done in part at the Research Center for Molecular Genetics, Hokkaido University, and was supported in part by Grants in Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan, and a Grant from the Program for Promotion of Basic Research Activities for Innovative Biosciences, Japan. M.P.Y. is the recipient of the JSPS Fellowship for Japanese Junior Scientists.

References

Bonen L, Bird S, Belanger L (1990) Characterization of the wheat mitochondrial *orf25* gene. Plant Mol Biol 15:793–795

Brandt P, Sunkel S, Unseld M, Brennicke A, Knoop V (1992) The *nad4L* gene is encoded between exon c of *nad5* and *orf25* in the *Arabidopsis* mitochondrial genome. Mol Gen Genet 236:33–38

Dewey RE, Levings CS III, Timothy DH (1986) Novel recombinations in the maize mitochondrial genome produce a unique transcriptional unit in the Texas male-sterile cytoplasm. Cell 44:439–449

Gray MW, Lang BF, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Brossard N, Delage E, Littlejohn TG, Plante I, Rioux P, Saint-Louis D, Zhu Y, Burger G (1998) Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res 26:865–878

Kubo T, Mikami T, Kinoshita T (1993) The sugar beet mitochondrial genome contains an ORF sharing sequence homology with the gene for the 30-kDa subunit of bovine mitochondrial complex I. Mol Gen Genet 241:479–482

Kubo T, Satoh Y, Muro T, Kinoshita T, Mikami T (1995) Physical and gene organization of mitochondrial DNA from the fertile cytoplasm of sugarbeet (*Beta vulgaris* L.). Curr Genet 28:235–241

Liu AW, Narayanan KK, Andre CP, Kaleikau EK, Walbot V (1992) Co-transcription of *orf25* and *coxIII* in rice mitochondria. Curr Genet 21:507–513

Mikami T, Kishima Y, Sugiura M, Kinoshita T (1985) Organelle genome diversity in sugar beet with normal and different sources of male-sterile cytoplasms. Theor Appl Genet 71:166–171

Nelson MA, Macino G (1987) Structure and expression of the overlapping ND4L and ND5 genes of *Neurospora crassa* mitochondria. Mol Gen Genet 206:307–317

Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T, Ohyama K (1992) Gene organization deduced from the complete sequence of the liverwort *Marchantia polymorpha* mitochondrial DNA. J Mol Biol 223:1–7

Prioli LM, Huang J, Levings CS III (1993) The plant mitochondrial open reading frame *orf221* encodes a membrane-bound protein. Plant Mol Biol 23:287–295

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual (2nd edn), Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Schuster W, Brennicke A (1994) The plant mitochondrial genome: physical structure, information content, RNA editing, and gene migration to the nucleus. Annu Rev Plant Physiol Plant Mol Biol 45:61–78

Senda M, Mikami T, Kinoshita T (1993) The sugar beet mitochondrial gene for the ATPase alpha-subunit: sequence, transcription and rearrangements in cytoplasmic male-sterile plants. Curr Genet 24:164–170

Stamper SE, Dewey RE, Bland MM, Levings CS III (1987) Characterization of the gene *urf13-T* and an unidentified reading frame, ORF25, in maize and tobacco mitochondria. Curr Genet 12:457–463

Tang HV, Pring DR, Muza FR, Yan B (1996) Sorghum mitochondrial *orf25* and a related chimeric configuration of a male-sterile cytoplasm. Curr Genet 29:265–274

Ward GC, Levings CS III (1991) The protein-coding gene T-*urf13* is not edited in maize mitochondria. Plant Mol Biol 17:1083–1088

Wolff G, Plante I, Lang BF, Kuck U, Burger G (1994) Complete sequence of the mitochondrial DNA of the chlorophyte alga *Prototheca wickerhamii*. J Mol Biol 237:75—86