

Robin K. Wilson · Maureen R. Hanson

## Preferential RNA editing at specific sites within transcripts of two plant mitochondrial genes does not depend on transcriptional context or nuclear genotype

Received: 16 November 1995 / 30 June 1996

**Abstract** Transcripts of most plant mitochondrial protein-coding genes exhibit C-to-U RNA editing events. In *Petunia*, two co-transcribed genes, *nad3* and *rps12*, exhibit transcripts which are not fully edited at all potential editing sites. We investigated the *nad3/rps12* transcript population in four different genotypes. In one pair of genotypes, the nuclear genome is identical but the *nad3/rps12* genes are in different transcriptional contexts. Both the *nad3/rps12* genes and the plant mitochondrial genomes are identical in a second pair of genotypes, but the nuclear background is derived from two different *Petunia* species. We found that the overall extent of editing varied greatly between genotypes and is affected by nuclear genotype but not by the global transcriptional context. Local sequence context around a particular site does affect editing frequency. In all genotypes, certain sites exhibit high editing frequency, but these sites do not share obvious primary sequence characteristics. In all genotypes examined, editing sites which do not affect the encoded amino acid are less frequently edited than sites which alter codons to non-synonymous forms. All these data indicate that an unidentified property of the sequences immediately surrounding a cytosine affect its selection as a target in the editing process.

**Key words** RNA editing · Mitochondria · Plant · NADH dehydrogenase

### Introduction

In higher plant mitochondria, C-to-U RNA editing occurs in transcripts from every examined open reading frame encoding a conserved functional polypeptide (Gray and Co-

vello 1993; Hanson et al. 1993; Araya et al. 1994). The extent to which all transcripts of a gene are edited varies greatly among genes. For example, when the cDNAs corresponding to mRNAs from wheat *cox3*, wheat *nad4*, and *Petunia atp9* were sequenced, all cDNAs displayed edits at each expected location (Gualberto et al. 1990; Lamatina and Grienberger 1991; Wintz and Hanson 1991). These transcripts can be considered fully edited; any translated proteins would include every expected amino acid change. The sequencing of cDNAs corresponding to other genes, such as wheat *cox2*, *Oenothera nad3*, *Petunia atp6*, and wheat *nad3* and *rps12*, has revealed a very different profile (Covello and Gray 1990; Schuster et al. 1990; Gualberto et al. 1991). Individual transcripts from these genes display various degrees of editing, ranging from unedited to edited at every identified editing site.

We previously used a reverse transcription/polymerase chain reaction (RT/PCR) technique to examine the editing extent in mitochondrial *nad3* transcript populations from cytoplasmic male-sterile (CMS) lines. We showed that the extent of editing at three assayed editing sites varies with nuclear background. After performing sexual crosses, we found that a single nuclear gene affects the extent of *nad3* transcript editing in *Petunia* (Lu and Hanson 1992). Thus, we identified nuclear genotypes which can be termed "high-editing extent" or "low-editing extent" depending on the presence or absence of this particular gene.

We were interested in further characterizing the partial editing of *nad3/rps12* transcripts by sequencing individual cDNA clones. In addition to examining editing in plants with different nuclear backgrounds, we examined the editing of transcripts from two *nad3/rps12* loci located in different transcriptional contexts. In lines containing the CMS-encoding mitochondrial genome, the *nad3/rps12* loci are co-transcribed with *pcf*, an abnormal CMS-associated gene. In contrast, wild-type *Petunia* lines do not contain the *pcf* gene and the *nad3/rps12* loci are transcribed with a 5' flanking region that differs from that found in CMS lines.

Analyzing individual cDNA clones allowed us to determine whether overall editing at all sites in *nad3/rps12* tran-

R. K. Wilson · M. R. Hanson (✉)  
Section of Genetics and Development, Cornell University,  
Biotechnology Building, Ithaca, NY 14853, USA

Communicated by K. J. Newton

scripts is high in plants identified by the RT/PCR assay as “high-editing extent” and low in those identified as “low-editing extent.” As shown in this paper, we find that the overall extent of editing is affected by the nuclear genotype but not by the transcriptional context. In addition, the sequenced cDNA clones allowed us to examine the frequencies with which individual sites are edited in partially edited transcripts. We hoped that an analysis of factors affecting the site-specific editing frequency would provide insights about the unknown determinants which specify Cs to be edited. By sequencing a number of *nad3/rps12* cDNA clones from four different genotypes, we were able to determine whether editing at the 27 sites is affected by nuclear genotype, transcriptional context, or both. We find that there are sites exhibiting high editing frequency in all lines, but these sites do not share obvious primary sequence characteristics. Additionally, we show that global transcriptional context (i.e., the other sequences or coding regions present on the transcript) does not affect the frequency of editing of individual sites, but that local sequence context around a particular site does affect editing frequency. In addition, editing sites which alter codons to non-synonymous forms (specifying a different amino acid) are more frequently edited than silent sites in all nuclear genotypes and transcriptional contexts examined.

## Materials and methods

**Plant material.** Tissue was obtained from the *P. hybrida* isonuclear pair 3704/11127 and from the *P. parodii* isonuclear pair 3699/3988. Lines 3704 and 3699 are male fertile. Lines 3688 and 11127 contain the same male-sterility encoding mitochondrial genome. The mitochondrial genomes of lines 3688 and 11127 differ in organization and gene content from the mitochondrial genome of line 3704 (Folkerts and Hanson 1991). Accession numbers are as follows: M16770 (*pcf*), U30458 (3704 *nad3/rps12*), and U29764 (3699 *nad3/rps12*).

**Oligonucleotides.** The following oligonucleotides were used:

*Petunia nad3*:

(Male-sterile lines)

primer *ste5'* 5' GCCTTGACAAGTTAGTACGGGTACTG 3'

(Fertile lines)

primer *fer5'* 5' GAAGGAGCAGGCTCTCTTGGATATG 3'

(All lines)

primer *nad3'* 5' GGAATTCCTTTGTCCTTGTCCCTCCCCCT 3'

primer *rps3'* 5' GTCGACTTCGTACCTATCCTTACC 3'

**RNA isolation.** Total leaf RNA was prepared using standard procedures (Ausubel et al. 1989), scaled down for 3 g of frozen plant tissue. Total bud RNA was prepared by freezing 10–30 3–5-mm buds in liquid nitrogen, grinding in 5 ml of buffer (0.18 M Tris pH 8.2, 0.09 M LiCl, 4.5 mM EDTA, 1% SDS) adding 2 ml of phenol and grinding for another 2 min, adding 2 ml of chloroform, incubating at 65°C for 20 min, spinning at 5000 rpm for 5 min, extracting with phenol/chloroform three times, and precipitating with NaAcetate and ethanol.

**cDNA sequence analysis.** A few cDNA clones containing *nad3* and *rps12* were isolated from cDNA libraries prepared using random hexamers as primers for suspension cell total mitochondrial RNA of 3704 or 3688 (C. Sutton, personal communication). These cDNAs were cloned into the Lambda ZapII vector using the protocol suggested by Stratagene. RNA from leaves or buds was precipitated in

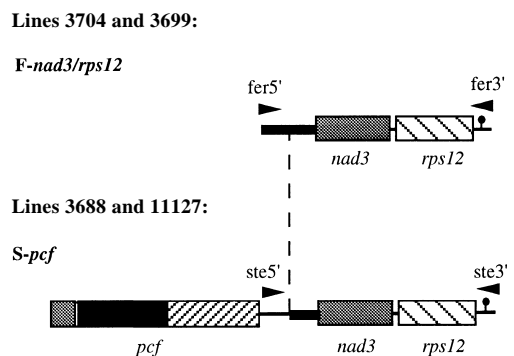
3 vol of ethanol. DNA was removed using RQ1 DNase (Promega). cDNAs were made by adding 5 mg of total RNA to 400 ng of primer *rps3'*, 15 U of RNasin (Promega), 0.5 mM of each deoxynucleotide, 10 mM DTT, 1× M-MLV-reverse transcriptase buffer (Life Technologies) in the presence of 200 units of M-MLV-reverse transcriptase (Life Technologies) at 37°C. After 1 h, these reactions were phenol/chloroform-extracted and ethanol-precipitated. PCR reactions were performed in a 100-ml vol by re-suspending each first-strand synthesis reaction in water and adding 0.2 mM of each deoxynucleotide, 600 ng of primer *ste5'* or 600 ng of primer *fer5'*, 600 ng of primer *rps3'*, 1× *Taq* DNA polymerase buffer (Cetus) in the presence of *Taq* DNA polymerase. The polymerase chain reaction was performed in a Hybaid cycler using the following conditions: cycle 1, 5 min 92°C, cycle 2, 30 s 92°C, 1 min 54°C, 1 min 72°C (repeat 30x, cycle 3, 10 min 72°C. To ensure that amplification was from RNA and not from DNA, control PCR reactions were minus the reverse transcription step. PCR products were chloroform-extracted, ethanol-precipitated, and cloned into the TA cloning vector (Invitrogen) using the recommended Invitrogen protocol. Primers used included M13 (–20), M13 (reverse), *ste5'*, *ste3'*, *rps3'*, and *nad3'*. Only one or two cDNAs from each PCR reaction were sequenced. Usually only one strand was sequenced and only C and T lanes were prepared to determine editing sites.

## Results

### Determination of RNA editing sites in *nad3* and *rps12* from different mitochondrial genomes

In CMS line 3688, *nad3* and *rps12* are co-transcribed with an abnormal open reading frame termed *pcf* (Hanson et al. 1989; Rasmussen and Hanson 1989; Pruitt and Hanson 1991). In wild-type lines 3704 and 3699, *nad3* and *rps12* are co-transcribed, but no *pcf* gene is present. The 5' transcribed regions upstream of *nad3* in wild-type and CMS lines are homologous only for 106 nt (Fig. 1). Primers 5' and 3' to the *nad3/rps12* regions found in the two different transcription units were used to produce PCR-amplified cDNAs from all four *Petunia* genotypes (Fig. 1).

Comparisons of cDNA sequences with the *nad3/rps12* genomic sequences from the two different mitochondrial genomes allowed the identification of edited nucleotides. Nineteen editing sites were found in the *nad3* coding region and five editing sites were found in the *rps12* coding



**Fig. 1** Comparison of the region containing *pcf*, *nad3*, and *rps12* in lines 3688 and 11127 with the region containing *nad3* and *rps12* in lines 3704 and 3699. The vertical line represents a major point of divergence. The locations of PCR primers are shown by arrows

```

-100 TGTCGTAAGTTTCGTTCTTCGTTCCGTCGGTCGATCTTCCTATTTCTTC
          *                *                *
-51  CGGTATGCCCTCCGCCAACAGGAGCGAAGAACCAAGTTTCTGTGGTG

nad3  L                               L
1    M (S) E F A P I C I Y L V I S (P) L V
1    ATTCAGAAATTTGCACCTATTTGTATCTATTTAGTGTATCAGTCCGCTAGTT
          *                *

18   S L I (P) L G L P F L F S S N S S T
52   TCTTTGATCCACTCGGTCTTCCTTTCTATTTCTTCCAATAGTTCGACC
          **

35   Y P E K F S A Y E C G F D P (S) G D
106  TATCCAGAAAAATTTTCGGCCTACGAATGTGGTTTCGATCCTTCCGGTGAT
          **

52   A R S R F D I R F Y L V S I L F I
154  GCCAGAAGTCGTTTGTATATAAGATTTTATCTTGTTCCTATTTTATTTATT

          F L F S L
69   I (P) D (P) E V T F (S) F P W A V (P) (P) N
208  ATTCCGATCCGGAAGTAACCTTTTCCTTTCCTTGGCAGTACCTCCCAAC
          ** ** * * *

86   K I D (P) F G (S) W S M M A F L L I L
259  AAGATTGATCCGTTGGATCTTGGTCCATGATGGCCTTTTATTTGATTTTG
          * * *

103  T I G (S) L Y E W K R G A (S) D (R) E end
310  ACGATTGGATCTCTATGAATGAAAAGGGGTCTTCGGATCGGGAGTAA
          * * *

          rps12 M P
1    CCACTAGTGAGAGGGCAAAAATAGGGGAAGGACAAAGGAAAGACGATGCC

3    S L N Q L I R H G R E E K R R T D
412  ATCACTAAATCAATTGATTCGTTCATGGTAGAGAAGAAAACGGCGCACGGA

20   R T R A L D Q C P Q K Q G V C (P) R
463  CCGTACTCGAGCTTTGGATCAATGTCCCCAGAAGCAAGGAGTATGCCGCG
          *

37   V S T R T P K K P N S A (P) R K I A
514  TGTTTCAACGAGAACACCGAAAAACCTAATTTCAGTCCGCGTAAGATAGC
          *

54   K V R L S N R H D I F A (H) I P G E
565  CAAAGTACGGTTGAGCAATGCACATGATATATTTGCTCACATTCAGGCGCA
          *

71   G H N L Q E H S M V L I R G G R V
616  AGGTCATAATTTGCAGAACATTCATGCTTAAATAAGAGGAGGTAGAGT

          L F
88   K D (S) P G V K (S) H C I R G V K D L
667  GAAAGATTTCGCCAGGTGTGAAATCCCATTTGTATTCGAGGAGTCAAGGATTT
          * *

105  L G I P D R R R G R S K Y G A E K
718  GCTGGGAATTCGGATCGAAGAAGGCAGATCAAATATGGTCCGGAATA

122  P K S I end
769  ACCCAAATCGATATGAATGGAAGATGCCTCTGGAACCTGTTTTCTCGGTA

```

**Fig. 2** RNA editing sites in *Petunia* mitochondrial *nad3* and *rps12*. The deduced amino-acid sequences are shown directly above the genomic nucleotide sequence. Amino acids altered by editing are enclosed in parentheses, with the new amino acid written above. Asterisks mark the 27 edit sites. Some editing has been found in all lines at every site resulting in the creation of a non-synonymous codon. Sterile line 3688 and sterile line 11127 do not contain the underlined nucleotides upstream from the *nad3* coding region, so only fertile lines 3704 and 3699 contain the edit site in this region. The numbers adjacent to the nucleotide sequences refer to the distance from the *nad3* start codon. The numbers adjacent to the amino acid sequences refer to the distance from either the first *nad3* amino acid or the first *rps12* amino acid

```

          ↓
F-nad3: TTCTTCGTTCCGTCGTCGATCTTCCTATTTCTTCCGGTATG
          |||||
S-nad3: TTCTTCGTTCCG-----TCCTATTTCTTCCGGTATG
          ↑

```

**Fig. 3** Similar sequences with different editing cues. The sequences surrounding a frequently edited site C 55 bases upstream from the start of *nad3* in wild-type *Petunia* mitochondrial lines 3704 and 3699 compared to a similar never-edited site C 55 bases upstream from the start of *nad3* in CMS *Petunia* mitochondrial lines 11127 and 3688. The Cs marked with arrows are edited equally often in CMS and wild-type genotypes

region (Fig. 2). Every observed coding-region editing site was found edited in at least one transcript from each analyzed line. Non-coding-region editing sites were also found on the *nad3-rps12* transcript leader (Fig. 2).

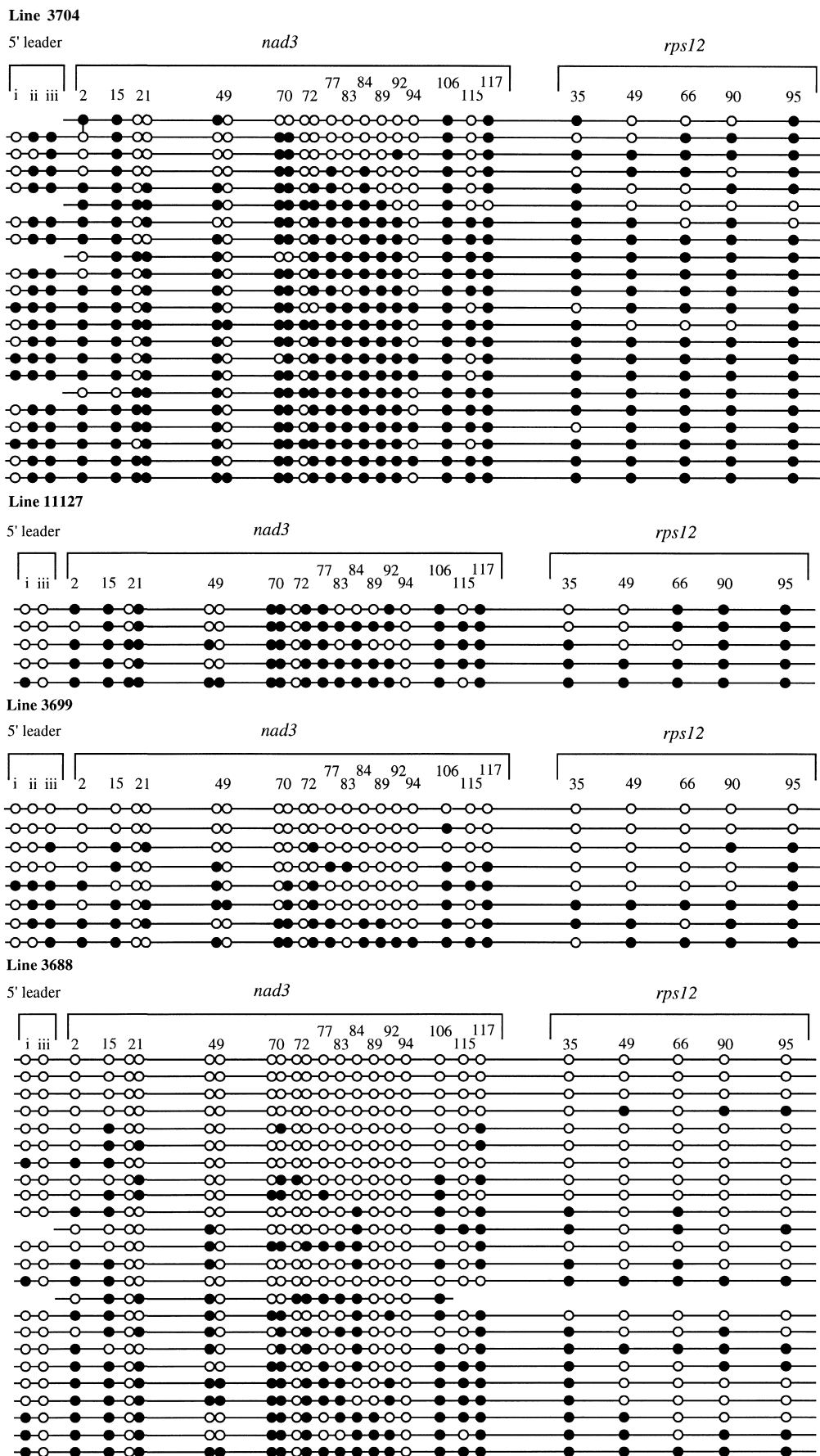
An interesting difference in editing was detected between transcripts from the two different *nad3/rps12* loci in the two *Petunia* mitochondrial genomes. There are three editing sites upstream from the *nad3* gene present in the mitochondrial genome carried by wild-type lines 3699 and 3704. A 10-nt segment encompassing the second edit (underlined in Fig. 2) is missing from male-sterile lines 3688 and 11127. The wild-type lines contain a third edited C upstream of *nad3*, but in the CMS lines the corresponding C was never found to be edited. The sequence surrounding the unedited C is identical 8 nt 5' (Fig. 3) and at least 900 nt 3' to the sequence surrounding the edited C in the wild-type lines (Fig. 1). It is therefore possible that the small 5' sequence alteration in the CMS line has resulted in the loss of an editing site in this line.

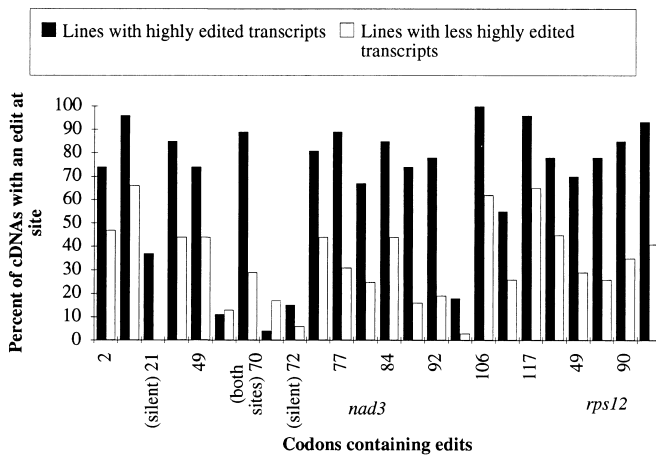
#### Analysis of partial editing of *nad3* and *rps12* in four different genotypes

Individual cDNAs were sequenced from four lines; two contain the same *pcf/nad3/rps12* transcriptional unit but differ in nuclear genotype, while another two lines carry wild-type *nad3/rps12* loci but also differ in nuclear genotype. Lines 3688 and 3699 carry the nuclear allele previously found to encode low-editing extent at sites in codons 49, 72, 83 in *nad3*, while the nuclear genome of lines 11127 and 3704 encode high-editing extent at these three sites. As shown in the diagram of individual cDNAs (Fig. 4), *nad3/rps12* transcripts from lines carrying the low-editing extent allele are overall less edited than transcripts from lines carrying the high-editing extent allele. These data show that the effect of the allele is not limited to the three sites previously assayed. Inspection of these individual cDNAs also reveals no clear 5' to 3' or 3' to 5' orderly progression of editing; edited sites and unedited sites are interspersed.

Some sites are edited more frequently in the population of transcripts than other sites. We grouped together the cDNAs from the isonuclear lines with high-editing extent (3704 and 11127) and the cDNAs from lines with low-ed-

**Fig. 4** Edited sites in cDNAs prepared from RNA isolated from iso-nuclear lines 3704 and 11127 and iso-nuclear lines 3699 and 3688. *Filled circles* represent edited residues, *unfilled circles* represent unedited residues. *Numbers* indicate codons in *nad3* and *rps12*. cDNAs were derived from either suspension cultures, leaves, or buds. No differences in the degree of *nad3/rps12* transcript editing have been detected between different tissues (Lu and Hanson 1992 and unpublished)





**Fig. 5** Comparison of the frequency of editing at individual sites. The bars indicate the percentage of edited residues at each editing site in cDNAs from highly edited isonuclear lines 3704 and 11127 and from less highly edited isonuclear lines 3699 and 3688. Numbers refer to codons in *nad3* and *rps12*

iting extent (3688 and 3699). Then we compared the frequency of editing at each editing site in the population of transcripts from the two different nuclear backgrounds (Fig. 5). We found that the sites most likely to be edited in the high-editing-extent lines were also the sites most likely to be edited in the low-editing-extent lines. When the primary sequence of the most highly edited sites was inspected carefully, we did not detect any motifs or common features. When we examined the sites that were least likely to be edited, we observed that sites in which editing resulted in silent codon alterations – in which the encoded amino acid was not affected – were less likely to be edited in partially edited transcripts (Fig. 5). Except for the general low level editing of silent sites, we found no obvious features of primary sequence which could predict whether a site would be highly edited or rarely edited.

#### Effect of editing on the predicted amino acid sequence

The translation of *Petunia* transcripts edited at all non-silent sites would result in NAD3 and RPS12 proteins very

similar to the proteins predicted from other higher-plant *nad3* and *rps12* edited transcripts (Fig. 6). Remarkably, the *Petunia* and wheat RPS12 proteins predicted from fully edited transcripts are identical, despite the evolutionary distance between this dicot and monocot. In addition, the NAD3 protein predicted from the edited *Petunia* transcript would be very similar to the protein predicted from the genomic *nad3* sequence of the lower plant *Marchantia polymorpha*, in which RNA editing has never been observed (Fig. 6). The high degree of amino acid conservation suggests that most codon alterations resulting from editing could be important for the production of a functional protein. However, transcripts that are fully edited at all non-silent sites are rather low in abundance. Only 6 out of 22 transcripts represented by cDNAs from line 3704, and only 1 out of 24 cDNAs from line 3688 (Fig. 4), could encode the NAD3 protein shown in Fig. 6. Thirteen cDNAs out of twenty-two from wild-type line 3704 but only 3 out of 24 cDNAs from line 3688 (Fig. 4) represent transcripts that could encode the RPS12 protein shown in Fig. 6.

Many of the same editing events resulting in changes to non-synonymous codons occur in *nad3* transcripts of *Petunia*, *Oenothera*, and wheat (Figs. 6, 7). In contrast, silent editing positions are less conserved (Fig. 7). The two dicots share two silent sites and each contains an additional unshared site. The monocot does not share any of these four silent sites with the dicots, but contains four other unshared silent sites. One other silent site is shared by *Petunia* and wheat but not *Oenothera*.

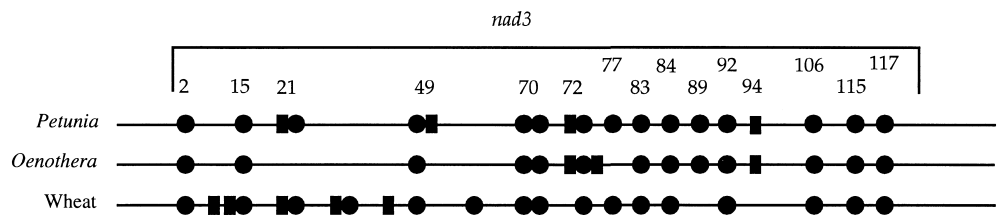
#### Discussion

Despite the identification of hundreds of editing sites in the transcripts arising from dozens of genes in many plant species, the actual mechanism for editing-site recognition is not known. Our study of *nad3/rps12* transcript editing in isonuclear lines containing divergent mitochondrial genomes supports the hypothesis that the sequence immediately surrounding an edited C is more important than distant features. Though the 5' portion of the *nad3/rps12* transcript in lines containing the CMS-encoding mitochondrial genome differs greatly from that found in wild-type lines,

**Fig. 6** Predicted mitochondrial NAD3 and RPS12 amino-acid sequences in *Petunia*, *Oenothera* (Schuster et al. 1990), wheat (Gualberto et al. 1988) and *Marchantia* (Yamato et al. 1993). *Oenothera* does not contain a complete mitochondrial *rps12* gene. The sequences presented for *Petunia*, *Oenothera*, and wheat would result from the translation of transcripts edited at all identified editing sites. Amino acids altered by editing are underlined. RNA editing has not been detected in *Marchantia*

|                         |    |               |               |               |               |               |               |               |       |        |               |       |        |       |       |    |
|-------------------------|----|---------------|---------------|---------------|---------------|---------------|---------------|---------------|-------|--------|---------------|-------|--------|-------|-------|----|
| <i>Petunia</i> NAD3     | 1  | <u>M</u> LEFA | PIC           | IY            | <u>L</u> VISL | <u>L</u> VS   | <u>L</u> LSLI | <u>L</u> LVVP | FLFSS | NSSTY  | PEKFS         | AYECG | FDPEG  | DARSR | FDIRF | 60 |
| <i>Oenothera</i> NAD3   | 1  | <u>M</u> LEFA | PICIS         |               | <u>L</u> VISL | <u>L</u> LSLI | <u>L</u> LVVP | FLFSS         | NSSRY | PEKLS  | AYECG         | FDPEG | DARSR  | FDIRF | 60    |    |
| Wheat NAD3              | 1  | <u>M</u> LEFA | PIC           | IY            | <u>L</u> VISL | <u>L</u> BSLI | <u>L</u> LVGP | FLFAS         | NSSTY | PEKLS  | AYECG         | FDPEG | DARSR  | FDIRF | 60    |    |
| <i>Marchantia</i> NAD3  | 1  | MEFA          | PIFVY         |               | <u>L</u> VISL | <u>L</u> LSLI | LIGVS         | FLPASSSSSLAY  |       | PEKLS  | AYECG         | FDPPD | DARSR  | FDIRF | 60    |    |
| <i>Petunia</i> NAD3     | 61 | YLVSI         | <u>L</u> FIIE | <u>D</u> LEVT | FFFPW         | AVSLN         | KIDLF         | GEWSM         | MAFLL | ILITIG | <u>E</u> LYEW | KRGAL | DWE    | 118   |       |    |
| <i>Oenothera</i> NAD3   | 61 | YLVSI         | <u>L</u> FIIE | <u>D</u> LEVT | FFFPW         | AVSEN         | KIDLF         | GEWSM         | MAFLL | ILITIG | <u>E</u> LYEW | KRGAL | DWE    | 118   |       |    |
| Wheat NAD3              | 61 | YLVSI         | <u>L</u> FIIE | <u>D</u> LEVT | FFFPW         | AVSLN         | KIDLF         | GEWSM         | MAFLL | ILITIG | <u>E</u> LYEW | KRGAL | DWE    | 118   |       |    |
| <i>Marchantia</i> NAD3  | 61 | YLVSI         | <u>L</u> FIIF | DPEVT         | FLFPW         | AVSLN         | KIGLF         | GFWSM         | MVFLF | ILITIG | FVYEW         | KRGAL | DWE    | 118   |       |    |
| <i>Petunia</i> RPS12    | 1  | MPSLN         | QLIRH         | GREEK         | RRTDR         | TRALD         | QCPQK         | QGVCL         | RVSTR | TPKKP  | NSALR         | KIAKV | RLSNR  | HDIFA | 65    |    |
| Wheat RPS12             | 1  | MPSLN         | QLIRH         | GREEK         | RRTDR         | TRALD         | QCPQK         | QGVCL         | RVSTR | TPKKP  | NSALR         | KIAKV | RLSNR  | HDIFA | 65    |    |
| <i>Marchantia</i> RPS12 | 1  | MPTMN         | QLVRK         | GRESK         | RRTKR         | TRALN         | KCPQK         | QGVCL         | RVSTR | SPKKP  | NSALR         | KIAKV | RLTNR  | NEIIA | 65    |    |
| <i>Petunia</i> RPS12    | 66 | <u>Y</u> IPGE | GHNLO         | EHSMV         | LIRGG         | RVKDL         | PGVKE         | HCIRG         | VKDLL | GIPDR  | RRGRS         | KYGAE | KPKSI  | 125   |       |    |
| Wheat RPS12             | 66 | <u>Y</u> IPGE | GHNLO         | EHSMV         | LIRGG         | RVKDL         | PGVKE         | HCIRG         | VKDLL | GIPDR  | RRGRS         | KYGAE | KPKSI  | 125   |       |    |
| <i>Marchantia</i> RPS12 | 66 | <u>Y</u> IPGE | GHNLO         | EHSVY         | MVRGG         | RVQDL         | PGVKY         | HCIRG         | VKDLO | GIPGR  | RRGRS         | KYGTK | KPKDYI | 126   |       |    |

**Fig. 7** *Nad3* editing sites in *Petunia*, *Oenothera*, and wheat. Edits which change a codon to a non-synonymous form are represented by a *circle*. Silent edits are represented by a *square*. Some codons have both a silent edit site and an editing site that changes the codon to a non-synonymous form



there is remarkable consistency in the Cs selected for editing and the frequency of editing at each particular site in the transcript population. Most partially edited transcripts from all lines are edited at certain editing sites – for example, codons 15, 49, 106, and 117 in *nad3* are always highly edited. These sites are distributed throughout the transcript; there does not seem to be a localized “hot spot” for editing.

The editing sites identified when transcripts of similar sequences within divergent loci are compared also support the hypothesis that surrounding sequences affect the likelihood that a C will be edited. Similar sequences present in both *coxII* and *pcf* transcripts in *Petunia* exhibit identical editing except at one site where the homologous C is edited in *coxII* but not in the aberrant *coxII*-homologous portion of *pcf* transcripts (Nivison et al. 1994). Inspection of the sequence around the unedited C in *pcf* reveals altered nucleotides at –19 and +18 from the C edited in *coxII*. In Fig. 3, we show that a C upstream of the *nad3* coding region in two different mitochondrial genomes is edited when a 10-nt sequence is present at –9 but has never been observed to be edited when this sequence is absent, even though the sequences 3' are identical for at least 900 nt. This example suggests that the 5' sequence may be more critical than the 3' sequence in editing-site recognition.

While nuclear genotype does affect the overall editing extent of *nad3/rps12* transcripts, the nuclear genotype does not affect the relative editing frequency of one site versus another site in the same transcript. Highly edited sites in a genotype displaying a highly edited transcript population were also the sites most likely to be edited in a genotype with a less-edited transcript population. We have previously determined that a single nuclear allele is correlated with both the editing extent and the transcript abundance of *nad3*, but is not correlated with the editing extent of several other mitochondrial genes, *coxII*, *nad1*, *atp6*, or *atp9* (Lu and Hanson 1992). When plants segregating for this allele were examined, those containing more abundant *nad3* transcripts exhibited transcripts with a higher editing extent than plants with less abundant transcripts. In Fig. 5 we show that the presence of this allele does not affect which sites are most likely to be edited. One interpretation of the correlation between abundance and editing extent is that most transcripts are turned over before editing is completed in the genotypes exhibiting low transcript abundance, resulting in a transcript population exhibiting a low editing extent. In this scenario, in genotypes with abundant

transcripts, the half-life of a transcript is longer, resulting in a greater degree of completion of editing. The editing mechanism would have a longer period of time during which to encounter marginally or transiently vulnerable editing sites. An alternative hypothesis is that less-efficient editing actually causes greater RNA turnover; transcripts are less abundant because the lack of editing de-stabilizes them.

As shown in Fig. 6, editing restores codons for conserved amino acids. If editing is required to produce a functional protein, then there should be selection for conservation of those features of a transcript which make a site subject to the editing machinery. Because far more unedited Cs occur in plant mitochondrial transcripts than edited Cs, the sequence requirements for editing of a site must be rather exacting. In the absence of selection, mutations occurring at locations reducing the likelihood of editing at a site can accumulate. These mutations could either be proximal to the site itself, or in genes for unidentified trans-acting factors that might participate in site recognition. There should be no selection against mutations that destroy the efficiency of editing at a site where the alteration of C to U has no effect on the encoded protein. Due to this, the location of silent editing sites should be less conserved than the location of editing sites which alter a critical amino acid, as we have observed in *nad3* and *rps12* (Fig. 7). Furthermore, in genes that exhibit partially edited transcripts, lack of selection for editing efficiency would be expected to result in lower editing frequency at silent sites. We observed that silent sites in *nad3* and *rps12* were relatively less edited in all genotypes (Fig. 5). Most silent editing sites were also less frequently edited than non-silent sites in 51 *atp6* cDNAs from four sorghum lines, except that one unusual site was found to be edited in 25 of the cDNAs. Kempken et al. (1995) noted that this unusual silent editing site had a high sequence similarity to another site in the *atp6* gene where editing increases protein conservation. Perhaps a trans-acting factor can recognize both the non-silent and the silent editing site, resulting in the maintenance of a high level of editing of the silent site.

Our data is consistent with the unordered selection of editing sites by the editing machinery. Interspersion of completed editing events with unedited sites on the same transcript of *nad3/rps12* (Fig. 4) and other genes (Conklin et al. 1991; Sutton et al. 1991; Yang and Mulligan 1991) indicates that there is no obligatory 5' to 3' or 3' to 5' order of editing. That certain sites are more frequently edited

than others in all examined transcriptional contexts and nuclear backgrounds indicates that the editing activity is more effective at these positions than at others within the sequence. Perhaps some feature of these sites results in a greater affinity for one or more components of the editing machinery, causing these sites to be edited first in a new transcript. Other editing sites – especially silent sites – are rarely edited, and appear less vulnerable to the editing mechanism. Testing this hypothesis will require dissecting the editing complex in vitro and determining how efficiently deliberately altered editing substrates are edited.

**Acknowledgements** This work was supported by a DOE Energy Biosciences and an NIH GMS grant to M.R.H.; R.K.W. was the recipient of an NSF predoctoral fellowship. We thank C.A. Sutton for providing cDNA libraries and helpful discussion.

## References

- Araya A, Begu D, Litvak S (1994) RNA editing in plants. *Physiol Plant* 91:543–550
- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1989) Phenol/SDS method for plant RNA preparation. In: Short protocols in molecular biology. John Wiley and Sons, New York, pp 144–145
- Conklin PL, Wilson RW, Hanson MR (1991) Multiple *trans*-splicing events are required to produce a mature *nad1* transcript in a plant mitochondrion. *Genes Dev* 5:1407–1415
- Covello PS, Gray MW (1990) Differences in editing at homologous sites in messenger RNAs from angiosperm mitochondria. *Nucleic Acids Res* 18:5189–5196
- Folkerts O, Hanson MR (1991) The male sterility associated *pcf* gene and the normal *atp9-1* gene in *Petunia* are located on different mitochondrial DNA molecules. *Genetics* 129:885–895
- Gray MW, Covello PS (1993) RNA editing in plant mitochondria and chloroplasts. *FASEB J* 7:64–71
- Gualberto JM, Wintz H, Weil JH, Grienenberger JM (1988) The genes coding for subunit 3 of NADH dehydrogenase and for ribosomal protein S12 are present in the wheat and maize mitochondrial genomes and are co-transcribed. *Mol Gen Genet* 215:118–127
- Gualberto JM, Weil JH, Grienenberger JM (1990) Editing of the wheat *coxIII* transcript: evidence for twelve C to U and one U to C conversions and for sequence similarities around editing sites. *Nucleic Acids Res* 18:3771–3776
- Gualberto JM, Bonnard G, Lamattina L, Grienenberger JM (1991) Expression of the wheat mitochondrial *nad3-rps12* transcription unit: correlation between editing and mRNA maturation. *Plant Cell* 3:1109–1120
- Hanson MR, Pruitt KD, Nivison HT (1989) Male sterility loci in plant mitochondrial genomes. *Oxford Surveys Plant Mol Cell Biol* 6:61–85
- Hanson MR, Sutton CA, Lu B, Conklin PL, Wintz H, Wilson RK, Pruitt KD (1993) RNA editing in *Petunia* mitochondria. In: Brennicke A, Kuck U (eds) *Plant mitochondria*. VCH, Weinheim, pp 71–81
- Kempken F, Höfken G, Pring DR (1995) Analysis of silent RNA editing sites in *atp6* transcripts of *Sorghum bicolor*. *Curr Genet* 27:555–558
- Lamattina L, Grienenberger JM (1991) RNA editing of the transcript coding for subunit 4 of NADH dehydrogenase in wheat mitochondria: uneven distribution of the editing sites among the four exons. *Nucleic Acids Res* 19:3275–3282
- Lu B, Hanson MR (1992) A single nuclear gene specifies the abundance and extent of RNA editing of a plant mitochondrial transcript. *Nucleic Acids Res* 20:5699–5703
- Nivison HT, Sutton CA, Wilson RK, Hanson MR (1994) Sequencing, processing, and localization of the *Petunia* CMS-associated mitochondrial protein. *Plant J* 5:613–623
- Pruitt KD, Hanson MR (1991) Transcription of the *Petunia* mitochondrial CMS-associated *pcf* locus in male-sterile and fertility restored lines. *Mol Gen Genet* 227:348–355
- Rasmussen J, Hanson MR (1989) A NADH dehydrogenase subunit gene is co-transcribed with the abnormal *Petunia* mitochondrial CMS-associated *pcf* locus in male-sterile and fertility restored lines. *Mol Gen Genet* 215:332–336
- Schuster W, Wissinger B, Unseld M, Brennicke A (1990) Transcripts of the NADH-dehydrogenase subunit 3 gene are differentially edited in *Oenothera* mitochondria. *EMBO J* 9:263–269
- Sutton CA, Conklin PL, Pruitt KD, Hanson MR (1991) Editing of pre-mRNAs can occur before *cis*- and *trans*-splicing in *Petunia* mitochondria. *Mol Cell Biol* 11:4274–4277
- Wintz H, Hanson MR (1991) A termination codon is created by RNA editing in the *Petunia* mitochondrial *atp9* gene transcript. *Curr Genet* 19:61–64
- Yamato K, Nozato N, Oda K, Ohta E, Takemura M, Akashi K, Ohyama K (1993) Occurrence and transcription of genes of *nad1*, *nad3*, *nad4l*, and *nad6* coding for NADH dehydrogenase subunits 1, 3 4L, and 6 in liverwort mitochondria. *Curr Genet* 23:526–531
- Yang AJ, Mulligan RM (1991) RNA editing intermediates of *coxII* transcripts in maize mitochondria. *Mol Cell Biol* 11:4278–4281