Curr Genet (2001) 39: 264–272
DOI 10.1007/s002940100204

ORIGINAL PAPER

Selvi Subramanian · Magid Fallahi · Linda Bonen

# Truncated and dispersed *rpl2* and *rps19* pseudogenes are co-transcribed with neighbouring downstream genes in wheat mitochondria

Received: 19 October 2000 / Accepted: 5 February 2001 / Published online: 9 May 2001


**Abstract** The wheat mitochondrial genome contains only partial coding sequences for the L2 and S19 ribosomal proteins, unlike in rice or liverwort mitochondria, where these genes are functional and have a bacterial-type linkage. A single-copy stretch corresponding to the extreme 3′ terminus of the wheat *rpl2* gene is co-transcribed with the *trans*-splicing *nad1* exon 4; and, at another unique location, the *rps19* segment lacking the 5′ coding region is co-transcribed with the downstream *nad4L* gene. In both cases, the 5′ termini of these transcripts map to promoter consensus motifs acquired through genomic reorganization, enabling continued expression of essential downstream genes. In both wheat and rice, the *rpl2* and *rps19* genomic regions differ in their RNA profiles between germinating embryos and seedlings. The absence of intact *rpl2* and *rps19* genes in wheat mitochondria is consistent with their inactivation through DNA rearrangement/deletion after the successful transfer of functional copies to the nucleus.

**Keywords** Mitochondria · Ribosomal protein · Rice · Wheat

## Introduction

The mitochondrial genomes of different eukaryotes vary with respect to the number of ribosomal proteins they encode. Those of fungi and animals have very few or no ribosomal protein genes, whereas certain protists have a relatively large number, notably 27 in *Reclinomonas americana* mitochondria (reviewed in Lang et al. 1999). Because mitochondrial ribosomes are expected to contain 80 or more proteins (reviewed in Graack and Wittmann-Liebold 1998), the rest are predicted to be nuclear-encoded and imported into the mitochondrion. In flowering plants so far, sequences homologous to 14 mitochondrial ribosomal protein genes have been identified in different species, namely ones encoding S1, S2, S3, S4, S7, S9, S10, S11, S12, S13, and S19 of the small subunit and L2, L5, and L16 of the large subunit (reviewed in Schuster and Brennicke 1994). These genes comprise a subset of the 16 ribosomal protein genes located in the mitochondrial genome of the liverwort, *Marchantia polymorpha* (Takemura et al. 1992). Furthermore, the number of mitochondrial-encoded ribosomal protein genes varies among flowering plants; and in some cases truncated rather than intact gene copies are present. For example, the completely sequenced mitochondrial genome of *Arabidopsis* (Unseld et al. 1997) contains only seven potentially functional ribosomal protein genes, namely *rps3*, *rps4*, *rps7*, *rps12*, *rpl2*, *rpl5*, and *rpl16* (although it should be noted that *rpl2* lacks the carboxy-terminal coding region) and there are truncated pseudogenes for *rps14* and *rps19*. Functional copies of these latter two genes have been identified in the *Arabidopsis* nucleus (Sanchez et al. 1996; Figueroa et al. 1999a), as has an *rps10* gene (Wischmann and Schuster 1995). Similarly, several other nuclear-located genes for mitochondrial ribosomal proteins have been characterized, including *rps12* in *Oenothera* (Grohmann et al. 1992), *rps11*, *rps14*, and *rps10* in rice (Kadowaki et al. 1996; Kubo et al. 1999, 2000), and *rps14* in maize (Figueroa et al. 1999b). Such movement of genes illustrates the dynamic nature of genetic information in plants, with the successful transfer of a functional copy to the nucleus usually being followed by inactivation or complete loss of the mitochondrial-encoded copy (reviewed in Blanchard and Lynch 2000; Palmer et al. 2000). Such events have been studied in detail for *cox2* (Adams et al. 1999) and a survey of other mitochondrial

Communicated by K. J. Newton

S. Subramanian · M. Fallahi · L. Bonen (✉)
Biology Department, University of Ottawa,
30 Marie Curie Street, Ottawa K1N 6N5, Canada
E-mail: lbonen@science.uottawa.ca

S. Subramanian
*Present address:* Centre for Plant Molecular Biology,
Tamil Nadu Agricultural University,
Coimbatore 641003, India

genes in a broad range of flowering plants suggests that ribosomal protein genes undergo loss from the mitochondrion much more frequently than do respiratory chain genes (Palmer et al. 2000).

In the mitochondrial genome of *Marchantia polymorpha*, ribosomal protein genes are for the most part clustered in the same (ancestral) gene order as in bacteria (Takemura et al. 1992), whereas in flowering plants they are often dispersed and differ in gene organization among species. Of particular focus in the present study are the *rpl2* and *rps19* genes. In *Marchantia* mitochondria, they are closely linked and followed by the *rps3* gene, whereas in rice mitochondria, although the *rpl2* and *rps19* genes are separated by 3 bp and are co-transcribed (Kubo et al. 1996), they are not physically close to *rps3*. No other monocot *rpl2* or *rps19* genes have been reported. In the dicots *Oenothera* and *Arabidopsis*, mitochondrial *rpl2* sequences (Sunkel et al. 1994; Unseld et al. 1997) lack the extreme 3′ terminal region and are physically distant from *rps19*. The latter is non-functional because of 5′ truncation in *Arabidopsis* (Unseld et al. 1997) and a nonsense mutation in *Oenothera* (Schuster and Brennicke 1991); and in both cases these pseudogenes are located immediately upstream of *rps3*. A similar arrangement is seen in petunia mitochondria, but in this case *rps19* is functional (Conklin and Hanson 1991). In *Arabidopsis*, the nuclear-located *rps19* copy contains a 5′ terminal extension with homology to RNA-binding domains and it has been proposed that this domain may serve as a functional substitute for the S13 ribosomal protein, because a corresponding *rps13* gene is absent from *Arabidopsis* mitochondrial DNA (Sanchez et al. 1996). Sequences homologous to parts of the *rpl2* gene are also present in mitochondrial DNA regions characterized in potato somatic hybrids (Loessl et al. 1999) and tobacco male sterile mutants (Vitart et al. 1992), but rice has the only documented, full-length mitochondrial *rpl2* gene, which is 3 kb and contains a single group II intron (Kubo et al. 1996). Both the L2 and S19 proteins perform important functions in the ribosome (reviewed in Uhlein et al. 1998; Culver and Noller 1999); and homologues of these proteins have been identified in eubacterial, archaeal, and eukaryotic cytosol ribosomes, so it is anticipated that plant mitochondrial ribosomes will also have functional counterparts.

We have examined the status of the *rpl2* and *rps19* genes in wheat mitochondria. Although sequences homologous to both genes are found as single genomic copies, neither comprises a fully intact gene. The *rpl2* segment is located upstream of the *trans*-splicing *nad1* exon 4, distant from the *rps19* fragment, which is upstream of the *nad4L* gene. However, both are expressed as co-transcripts with downstream genes and are preceded by promoter motifs, so that it appears that DNA rearrangements causing disruption of the structural genes also led to the recruitment of promoter sequences for the continued expression of the physically linked downstream genes.

## Materials and methods

### Isolation, cloning, and sequencing of plant mitochondrial DNA and RNA

Mitochondrial DNA and RNA were isolated from 24-h-old, germinating embryos and 6-day-old, etiolated seedlings of wheat (*Triticum aestivum* var. Frederick) and rice (*Oryza sativa* var. v20B) as previously described (Bonen 1987). After cloning regions of interest into pUC or M13 vectors, sequencing was carried out by the dideoxynucleotide chain termination method using Sequenase ver. 2.0 (US Biochemicals). RNA-editing sites were determined by direct sequencing of gel-purified reverse transcriptase (RT)-PCR products as previously described (Carrillo and Bonen 1997).

### RNA blot analysis

Northern hybridization experiments were performed using standard procedures (Sambrook et al. 1989). Gel loading was standardized, relative to equivalent amounts of mitochondrial ribosomal RNA from germinating embryos and etiolated seedlings. Hybridization probes were either DNA fragments which had been radiolabelled by random priming with Klenow polymerase and [$\alpha$-$^{32}$P] dATP or synthetic oligomers radiolabelled with [$\gamma$-32P] ATP and T4 polynucleotide kinase.

### Primer extension analysis

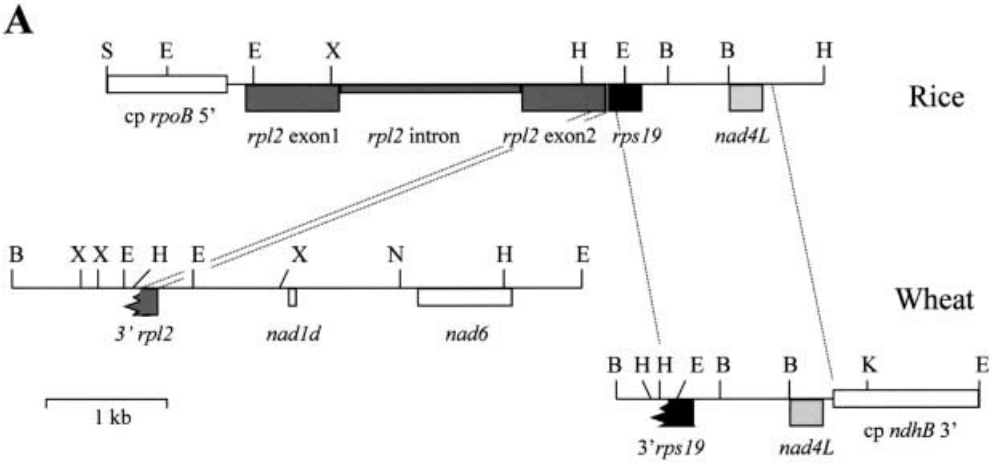
Synthetic oligomers which had been radiolabelled at the 5′ termini were annealed with approximately 25 μg mitochondrial RNA at 70 °C for 3 min. Extension was performed with Superscript II RT (Gibco BRL), accordingly to supplier's specifications; and products were resolved on denaturing sequencing gels. The sequences of the oligomers used were 5′ CACCATGAGGATGATCCACTGG 3′ for the wheat *rpl2* segment, 5′ AATTCCTCCCTGCGGATCTTC 3′ for rice *rpl2*, 5′ AGCAAACTCTCCAAATTTATGACC 3′ for the wheat *rps19* segment, and 5′ TGTATTTGGTAACTCTCCTG 3′ for rice and wheat *nad4L*. Size ladders were generated from sequencing reactions using the same oligomers with mitochondrial DNA as template.

## Results

### A single-copy *rpl2* segment in wheat mitochondria is co-transcribed with the fourth exon of the *trans*-splicing *nad1* gene

We determined that a short stretch homologous to the bacterial-type ribosomal protein L2 gene, is located approximately 1.1 kb upstream of the *trans*-splicing *nad1d* exon, which in turn is located upstream of the *nad6* gene (Fig. 1A). Both of the latter encode components of the NADH dehydrogenase complex and have previously been demonstrated to be co-transcribed in wheat mitochondria (Haouazine et al. 1993). The *rpl2*-related segment, which is present as a single genomic copy (based on Southern blot analysis), shares 94% nucleotide identity with the rice mitochondrial counterpart (Kubo et al. 1996), but it comprises only 165 bp at the extreme 3′ terminus of the gene (Fig. 1B), or about 10% of the entire coding sequence found in rice mitochondria.

The wheat *rpl2* segment does not have an initiation codon (or ACG codon which could be converted to




B

```
Wheat  GAATTCTTATCTATTTCTCGAGGCAATTATACTCATATCGACAATATGAACTGCATGAGAATTAGCATATAGAAAAGAAAGCTTTACCGAATACAGGTCTAGTAGGTATAG  110
       *    *    * ** *** * *    *   *          ** *   *  *  *    **    *  ** *         *  * *** *    *  *  *
Rice   GCCAGCCcCACAATGTCTTGTGCGGCTACCATCGGGTGTTGAAAAACTCATAGATTCCCGATGCCGAGCTACTATTGGTATAGTTTCCAATCCCAACCATGGTGCACGTA
         P  A  P>S  Q  C  L  V  R  L  P  S  G  V  E  K  L  I  D  S  R  C  R  A  T  I  G  I  V  S  N  P  N  H  G  A  R

                A  W  W  G  C  R  C  L  G  R  C  P  I  V  H  D  V  A  M  N  P  V  D  H  P  H  G  G  G  E  G  R  T  K
Wheat  AGCCGTAAGCGTGGTGGGGGTGCCGGTGCTTAGGCAGATGCCCCATTGTTCATGATGTTGCAATGAATCCAGTGGATCATCCTCATGGAGGAGGTGAGGGGCGCACGAAA  220
       ***    *      *     ******* * ******* ************ ** ************************************************
Rice   AGCTTAGAAAAGCAGGACAAAGCCGGTGGTCAGGCAGACGCCCCATTGTTCGTGGTGTTGCAATGAATCCAGTGGATCATCCTCATGGAGGAGGTGAGGGGCGCACGAAA
       K  L  R  K  A  G  Q  S  R  W  S  G  R  R  P  I  V  R  G  V  A  M  N  P  V  D  H  P  H  G  G  G  E  G  R  T  K

          G  G  R  P  S  V  S  P  W  G  K  L  T  K  A  G  F  R  A  V  V  G  V  V  K  L
Wheat  GGAGGTAGACCTTCGGTGTCACCTTGGGGAAAGCTCACCAAAGCAGGATTTCGGGCAGTAGTAGGGGTGGTGAAACTTTAGATGTCAAATGCCACCAATTTTTCTGCTCA  330
       ****************************** **** ********************** *** ******* *****     ** *       *      ***
Rice   GGAGGTAGACCTTCGGTGTCACCTTGGGGGAAGCCCACCAAAGCAGGATTTCGGGCAGGAGTGGGGGTGGGGAAACGCAGAATTTAGTTcATGCCACGACGATCTATATG
          G  G  R  P  S  V  S  P  W  G  K  P  T  K  A  G  F  R  A  G  V  G  V  G  K  R  R  I
```

C

```
                                                                  T  R  R  Q  Q  K  G  A  W  S  K  R  S  H  L  V  V  H  E  K  R  H  S  R  E  S  L
Wheat  AAGCTTGTTTCAATTGAACGAGGATGTGAACGAGAAGGCAACAAAAGGGAGCTTGGAGCAAGAGGAGTCATTTAGTAGTTCACGAAAAAAGGCATAGTAGAGAAAGTCTG  110
                                  *     * ** *     *     * * *       * **    *  *      ** ** * * * ***********
Rice                                ATGCCACGACGATCTATATGGAAGGGAAGTTTTGTTGATGCTTTCCTGTTTAGAATAAAGAAGAACAGAGAAAGTcTG
                                    M  P  R  R  S  I  W  K  G  S  F  V  D  A  F  L  F  R  I  K  K  N  R  E  S  L

       R  S  R  K  I  W  S  R  R  S  S  I  S>L  P  E  F  V  D  C  S  V  L  I  Y  N  G  K  T  P>L  V  R  C  K  I  T  E  -
Wheat  AGGAGCAGGAAAATTTGGTCACGTAGATCTTCTATTTcGCCGGAATTCGTTGATTGCTCCGTACTCATTTACAATGGAAAAACTCcTGTTCGTTGTAAGATCACTGAA--  218
       * ***********************************************************************************************************
Rice   ATGAGCAGGAAAATTTGGTCACGTAGATCTTCTATTTcGCCGGAATTcGTTGATTGCTCcGTACTCATTTACAATGGAAAAACTccTGTTCGTTGTAAGATCACTGAAGG
       M  S  R  K  I  W  S  R  R  S  S  I  S>L  P  E  F  V  D  C  S  V  L  I  Y  N  G  K  T  P>F  V  R  C  K  I  T  E  G

          -  -  G  H  K  F  G  E  F  A  F  T  R  R  R  R  P  Y  R  T  N  R  G  K  -  -  -  G  K  K  *
Wheat  -------GGTCATAAATTTGGAGAGTTTGCTTTTACACGGAGACGAAGACCCTATCGAACAAATAGAGGAAAG---------GGGAAAAAGTAAAGTCTAAGCGAC....  308
              ************************************************ ****************     *********************
Rice   AAAGGTTGGTCATAAATTTGGAGAGTTTGCTTTTACACGGAGACGAAGACCCTATCAAACAAATAGAGGAAAGGGCAGAAAGGGGAAAAAGTAAAGTCTAAGCGAC....
          K  V  G  H  K  F  G  E  F  A  F  T  R  R  R  R  P  Y  C  T  N  R  G  K  G  R  K  G  K  K  *
```

```
                                        T>M  D  P>L  I  K  Y  F  T  F  S  M  I  I  S>F  I  L  G  I  R>W  G  I  L  L  N  R
Wheat  ....TGAATTTCTCTGACATTCCATGTTTCCGAAAcGGATCcTATAAAATATTTCACTTTTTCTATGATCATCTcTATTTTAGGTATTcGGGGAATCCTCCTTAATAGAC  1270
           ***************************************************************************************************************
Rice   ....TGAATTTCTCTGACATTCCATGTTTCCGAAAcGGATCcTATAAAATATTTCACTTTTTCTATGATCATCTcTATTTTAGGTATTcGGGGAATCCTCCTTAATAGAC
                                        T>M  D  P>L  I  K  Y  F  T  F  S  M  I  I  S>F  I  L  G  I  R>W  G  I  L  L  N  R

       R  N  I  L  I  M  S>L  M  P>S  I  E  S>L  M  L  L  A  V  N  L  N  F  L  V  F  S  V  S  L  D  D  M  M  G  Q  S>L  F  A
Wheat  GAAATATTCTTATTATGTcAATGcCAATTGAATcAATGTTATTAGCTGTCAATTTGAACTTTTTGGTATTTTCCGTTTCTTTGGATGATATGATGGGTCAATcATTTGCT  1380
       ***************************************************************************************************************
Rice   GAAATATTCTTATTATGTcAATGcCAATTGAATcAATGTTATTAGCTGTCAATTTGAACTTTTTGGTATTTTCCGTTTCTTTGGATGATATGATGGGTCAATcATTTGCT
       R  N  I  L  I  M  S>L  M  P>S  I  E  S>L  M  L  L  A  V  N  L  N  F  L  V  F  S  V  S  L  D  D  M  M  G  Q  S>L  F  A

       S>L  L  V  P>L  T  V  A  A  A  E  S  A  I  G  L  A  I  F  V  I  T  F  R  V  R  G  T  I  A  V  E  S>F  I  N  C  I  Q
Wheat  TcATTAGTTCcAACAGTGGCAGCTGCGGAATCTGCTATTGGATTAGCCATTTTCGTTATTACTTTTCGAGTCCGAGGGACTATTGCTGTCGAATcTATAAATTGCATTCA  1490
       ***************************************************************************************************************
Rice   TcATTAGTTCcAACAGTGGCAGCTGCGGAATCTGCTATTGGATTAGCCATTTTCGTTATTACTTTTCGAGTCCGAGGGACTATTGCTGTCGAATcTATAAATTGCATTCA
       S>L  L  V  P>L  T  V  A  A  A  E  S  A  I  G  L  A  I  F  V  I  T  F  R  V  R  G  T  I  A  V  E  F  I  N  C  I  Q

          G  *
Wheat  AGGTTAAACATAACTACAGGAGAGTTACCAAATACAAAGTTCTGTTCTCCTTTCGTTCTCTTCTTTCTTTTctttttgcctgagtcagacatcaaatagcttcgatttgc  1600
       ********************* ************************** ************************   ***   *  **       ***       * *
Rice   AGGTTAAACATAACTACAGGGGAGTTACCAAATACAAAGTTCTGTTTTCCTTTCGTTCTCTTCTTTCTTTTTGATTTTTGTTTGTTGCGAGCAAAAATAAGCTCTATAGG
          G  *
```

**Fig. 1 A–C** Genomic organization and nucleotide sequence of the wheat mitochondrial *rpl2* (3′ segment), *rps19* (3′ segment), and *nad4L* regions, and their comparison to rice mitochondrial homologues. **A** Restriction maps of the rice and wheat mitochondrial 3′ *rpl2* (grey), 3′ *rps19* (black), and *nad4L* (light grey) regions, with other genes and chloroplast (*cp*)-derived sequences shown in white. Selected restriction sites are shown as: *B Bam*HI, *E Eco*RI, *H Hin*dIII, *K Kpn*I, *N Nar*I, *S Sal*I, *X Xba*I. **B, C** Nucleotide and amino acid sequences of wheat 3′ *rpl2* (**B**) and wheat 3′ *rps19–nad4L* (**C**) aligned with rice counterparts. Editing sites are shown by *lower case letters* and include the rice *rpl2–rps19* sites previously determined by Kubo et al. (1996). Amino acid substitutions are *blocked*, breakpoints in homology are indicated by *vertical broken lines*, and the positions of 5′ termini of transcripts in wheat are marked by *open triangles*. The highly-conserved amino-terminal S19 residues missing from wheat (see Results) are in *bold italics* in the rice sequence. Positions of oligomers used for primer extension analysis are *underlined*. In **B**, the 5′ end of the rice *rps19* gene is shown in *uppercase italics* and in [C] the chloroplast-derived *ndhB* 3′ flanking sequence downstream of wheat *nad4L* is shown in *lower case*. The wheat mitochondrial 3′ *rpl2* region and 3′ *rps19-nad4L* sequences are deposited under accession numbers AJ295995 and AJ295996, respectively. The rice mitochondrial *nad4L* gene sequence (AJ295997) is adjacent to that previously determined for *rpl2* and *rps19* (Kubo et al. 1996; D78336)

AUG by RNA editing) and there is an in-frame termination codon 13 bp upstream of the homologous stretch (Fig. 1B, bold). Of the ten nucleotide substitutions between the wheat and rice *rpl2* sequences, eight are non-synonymous and their nature is consistent with a lack of functional constraint in the former. Sequences corresponding to the rest of the *rpl2* gene were not detected elsewhere in the wheat mitochondrial genome, using various rice mitochondrial *rpl2* probes in Southern blot analysis (data not shown), although it cannot be completely excluded that either very divergent or very short sequences might escape detection. However, sequence analysis of approximately 4 kb upstream of the *rpl2* segment did not reveal any homology to *rpl2*-coding sequences or to structures/sequences diagnostic of group II introns, except for a 26 bp out of 27 bp match with the 5′ part of domain 5 of *nad1* intron 1 (Chapdelaine and Bonen 1991). However, this directly precedes the *rpl2* segment (Fig. 1B, positions 104–130) rather than being associated with the other half of domain 5 and domain 6, as in group II introns. Furthermore, RT-PCR analysis did not provide any evidence for *rpl2* mRNA generated by splicing *in cis* or *in trans*; and PCR experiments using oligomers designed from conserved *rpl2* regions did not reveal the presence of any low-level forms of mitochondrial DNA containing a complete gene.

The *rpl2*-homologous region of the wheat mitochondrial genome gives rise to low levels of large heterogeneous transcripts, as judged by RNA blot analysis (Fig. 2A, lanes 3, 4). They range in size from about 1.5 kb to 4 kb and are much more abundant in germinating (24 h) wheat embryos than in etiolated (6 day) seedlings, when standardized for mitochondrial ribosomal RNA levels. Similarly, mitochondrial RNA isolated from etiolated wheat seedlings show only very low levels of transcripts hybridizing to *nad1d* probes (Haouazine et al. 1993), in contrast to RNA from callus tissue where transcripts of 2.1, 3.4, and 5.6 kb are detected with *nad1d* and *nad6* probes. The *nad6* gene also shows an abundant mature mRNA of approximately 1.2 kb in wheat embryos (Fig. 2A, lane 5), seedlings (Fig. 2A lane 6; Haouzaine et al. 1992), and callus tissue (Haouzaine et al. 1993). In contrast, using the *rpl2* upstream region as probe, only very faint signals at the positions of the cellular ribosomal RNAs are observed (Fig. 2, lanes 1, 2), suggesting that *rpl2*-containing transcripts do not have long 5′ extensions. The regions further upstream for at least 4 kb also lack abundant stable transcripts (data not shown).

Primer extension analysis using an *rpl2*-specific oligomer (Fig. 1B, underlined) identifies the 5′ terminus of the major transcript in wheat embryos as a doublet positioned about 60 nt upstream of the *rpl2* homologous sequences (Fig. 2C, open triangle). It maps to the motif GAGAATTAG*CATA*TAGAAAGAA (position of 5′ termini underlined; Figs. 1B, 2C, open triangles) and this conforms to the monocot mitochondrial promoter consensus motif, in having a CRTA element preceded by a purine rich stretch (reviewed in Fey and Marechal-Drouard 1999), as shown blocked in Fig. 2A. This is consistent with transcripts of 1.5 kb or larger containing both the *rpl2* and *nad1d* regions, and those 3.2 kb or larger containing the whole *rpl2–nad1d–nad6* region. This extends the observations of Haouazine et al. (1993), where co-transcripts of approximately 3.5–5.6 kb were detected in callus tissue, using *nad1d* or *nad6* probes. Taking these observations together, we conclude that this wheat mitochondrial *rpl2* segment is a truncated pseudogene and that there is no intact *rpl2* gene in wheat mitochondria. Nevertheless, transcription of this region enables expression of the downstream *nad1d*-coding and intronic sequences, which must undergo *trans*-splicing to generate the mature *nad1* mRNA (Chapdelaine and Bonen 1991).

## The single-copy, truncated *rps19*-homologous sequences are co-transcribed with *nad4L* in wheat mitochondria

The region immediately downstream of the wheat mitochondrial *rpl2* segment does not contain the *rps19* gene, as in rice where there is an intergenic spacer of only 3 bp (Kubo et al. 1996). However, *rps19* sequences have been identified by hybridization analysis to be located at a single site in the wheat mitochondrial genome, about 1 kb upstream of the *nad4L* gene (Fig. 1A). Sequence analysis has established that a stretch of 213 bp shares 98% nucleotide identity with the rice *rps19* gene of 279 bp. However, the 5′ terminal section with the initiation codon is missing (Fig. 1C) . The open reading frame extends for 69 bp upstream of the *rps19* segment, but within that region there is no in-frame potential initiation codon; and the single ACG which could be

edited to AUG lies upstream of the 5′ terminus of the major transcript (see below). The missing part corresponds to a conservative region of the S19 protein, with seven residues being invariant among rice, petunia, and liverwort mitochondrial, and *Escherichia coli* S19 proteins (Fig. 1C, bold italics for residues in rice). The wheat DNA sequence has two deletions of 9 bp and two non-synonymous nucleotide substitutions, compared to its rice counterpart. In addition, differences in RNA editing alter the specificity of a third codon (see below).

The homology between wheat and rice sequences extends about 1,260 bp downstream of the *rps19* termination codon and includes the *nad4L* gene (Fig. 1A, C). Beyond the *nad4L* gene, conservation between the two monocot sequences continues for only 64 bp, at which point there are chloroplast-derived *ndhB* 3′ flanking sequences present in the wheat but not in the rice mitochondrial genome (Fig. 1C, nucleotides in lower case). Various hybridization and PCR-based experiments (similar to those described above for *rpl2*) were conducted to search for the missing *rps19* piece, but we found no evidence for its presence elsewhere in the genome; and we conclude that there is unlikely to be a functional *rps19* gene in wheat mitochondria.

The wheat *rps19*-homologous region gives rise to an abundant transcript of approximately 1.5 kb, detected with probes specific to *rps19* or *nad4L* in germinating embryos (Fig. 2B, lanes 1, 3). This RNA is not abundant in wheat seedlings (Fig. 2B, lanes 2, 4), whereas short *nad4L*-specific RNAs are prominent in both stages of development (Fig. 2B, lanes 3, 4), although with somewhat differing profiles, ranging in size from about 0.25 kb to 0.65 kb. The 5′ terminus of the 1.5 kb cotranscript has been mapped by primer extension analysis to a position 3 nt downstream of the breakpoint in homology between wheat and rice, i.e., within the *rps19*-homologous region (Fig. 2C). It is preceded by the sequence GAAAAAAGG*CATA*GTAG<u>A</u>GAA

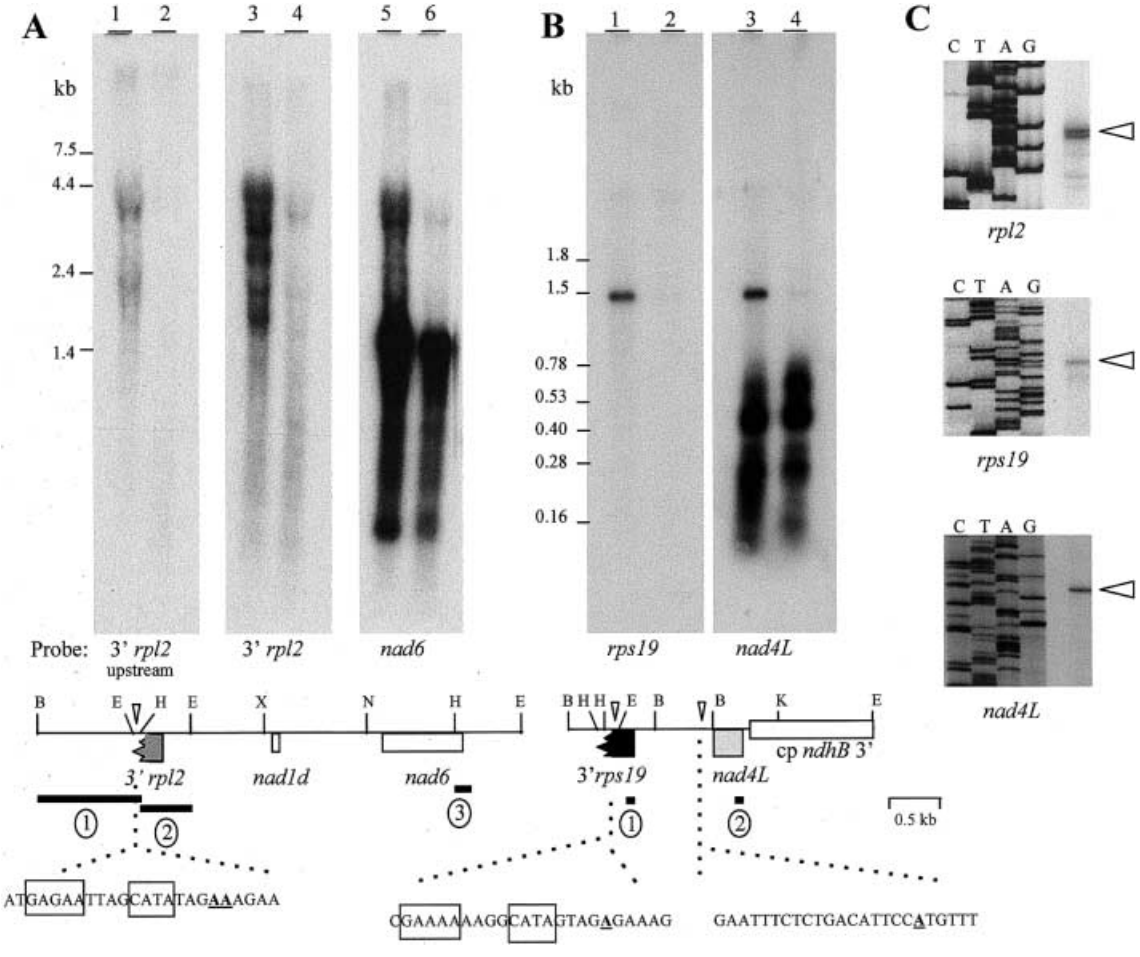



**Fig. 2 A–C** Transcript analysis of the mitochondrial 3′ *rpl2* and 3′ *rps19–nad4L* regions in wheat embryos and seedlings. **A, B** Northern blot analysis of 3′ *rpl2* region (**A**) and 3′ *rps19* region (**B**), with schematics showing probe positions. RNA from germinating wheat embryos (*lanes 1, 3, 5*) or etiolated seedlings (*lanes 2, 4, 6*) was hybridized with probes as shown. *Probes 1 and 2* in panel **B** are synthetic oligomers. *Lanes* were standardized for equivalent amounts of mitochondrial ribosomal RNA. Restriction sites are as in Fig. 1. **C** Primer extension data mapping the 5′ termini of transcripts for the *rpl2* segment, *rps19* segment, and *nad4L* from wheat embryos. Primer positions are shown by *open triangles* on the schematics in **A–C** and are underlined in Fig. 1B, C. Sequencing ladders were generated using the same oligomers. *Open triangles* show the positions of the transcript termini and these sequences are shown on the schematic, with consensus promoter motifs *blocked*

(position of 5′ terminus underlined), which fits the consensus promoter motif. It therefore appears that DNA rearrangements have resulted in the acquisition of regulatory sequences which allow co-transcription of the *rps19* segment and *nad4L*.

The 5′ terminus of the most abundant RNA species containing the complete *nad4L* coding sequence is located 11 nt upstream of the *nad4L* initiation codon in both wheat embryos and seedlings (Fig. 2C). The non-translated leader of this mRNA is therefore exceptionally short, and its 5′ terminus maps to the sequence TCTCTGACATTCCATGTTT (5′ terminal nucleotide underlined). A consensus promoter motif is lacking, so these transcripts may arise through processing events; however, it is worth noting that not all monocot mitochondrial promoters conform to the consensus sequence (Fey and Marechal-Drouard 1999). Within the wheat *nad4L*-coding region, there are 11 C-to-U RNA editing sites (Fig. 1C), including one which creates the initiation codon. All of these editing events lead to changes in the amino acid sequence and increase its similarity with those from other organisms. We observed a similar editing pattern in rice *nad4L*, but with one less edit where there is a genomically-encoded T (Fig. 1C). Thus the wheat and rice *nad4L* sequences are identical at the RNA level. Within the wheat *rps19*-homologous region of the 1.5-kb transcript, there was one edit observed in both embryos and seedlings (Fig. 1C; P > L, blocked) and another observed only in seedlings (Fig. 1C; S > L, blocked). These comprise two of the three predicted editing sites, based on those observed in the functional *rps19*-coding sequences of rice (Kubo et al. 1996) and petunia (Conklin and Hanson 1991). Their presence also indicates that although the truncated *rps19* is non-functional, the RNA editing machinery is still present and can recognize these sites in their new genomic context.

### Developmentally-specific differences in RNA profiles of the *rpl2–rps19–nad4L* region of rice mitochondria

In rice mitochondria, the *nad4L* gene is located downstream of the *rpl2–rps19* genes, both of which are intact; and the RNA profile of this genomic region shows a complex set of transcripts, which differ between the embryo and seedling stages of development, relative to constant ribosomal RNA levels (Fig. 3A). It has previously been shown that in young rice seedlings, the *rpl2* and *rps19* genes are co-transcribed as RNA species of approximately 2.5, 2.9, and 3.8 kb (Kubo et al. 1996). We observe comparable species (designated as 2.6, 3.0, and 3.5 kb) in seedlings, although their relative abundance differs in embryos. For example, with a *rpl2* cDNA probe, the 3.5-kb species is more prominent in embryos than in seedlings, whereas it is the reverse for the 2.6-kb species (Fig. 3A, lanes 7 vs 8). In addition, a 5.0-kb species, which is abundant in embryos, is detected with *rpl2*, *rps19*, and *nad4L* probes (Fig. 3A, lanes 3, 7, 9, 11), consistent with co-transcription of this entire region. The *rpl2* intron probe detects a 1.5-kb species, which corresponds to the excised group II intron and is present at much higher levels in embryos than in seedlings (Fig. 3A, lanes 5 vs 6, asterisk). This is similar to observations made for certain wheat mitochondrial introns (cf. Carrillo and Bonen 1997).The *rpl2* intron probe also detects a prominent 3-kb species in both developmental stages and its size is consistent with a splicing intermediate in which only the first transesterification step has occurred, so that the intron is still attached to the downstream exon. In keeping with this interpretation, the *rpl2* exon 1 probe does not give a strong 3-kb signal in embryos (Fig. 3A, lane 3). However, because there is a co-migrating, 3-kb hybridizing species detectable in seedlings (Fig. 3A, lane 4), it can be deduced that the putative 2.9-kb mRNA is present at much reduced levels in embryos compared to seedlings. Consequently, the predominant mRNA for *rpl2* and *rps19* in rice embryos is predicted to be the 3.5-kb species.

The 5′ termini of transcripts extending upstream from the initiation codons for *rpl2* and *nad4L* were determined by primer extension experiments using oligomers mapping within their coding sequences (Fig. 3B). The former maps to a position 290 nt upstream of the *rpl2* initiation codon, located within the chloroplast-derived *rpoB* 5′ flanking sequences (Kubo et al. 1996). That sequence (AAAATTCCATTATCAG; with the 5′ terminal nucleotide underlined) does not conform to the consensus promoter motif. The major 5′ terminus for *nad4L* maps to the same position as in wheat (Fig. 3B), so that it too has a 5′ leader of only 11 nt. It can also be seen that the processing pathways differ between the two stages of development with respect to *nad4L*-containing transcripts (Fig. 3A, lanes 11 vs 12). Several large species are more abundant in embryos (5.0 kb and 3.5 kb), whereas other, smaller ones (2.6 kb and 1.5 kb) are more abundant in seedlings. The profiles are consistent with co-transcription of the *rpl2–rps19–nad4L* region as 5-kb and larger transcripts, followed by complex processing events which generate smaller, discrete transcripts, including ones which are developmentally-specific. In both stages, the most abundant *nad4L* transcripts are the monocistronic ones of approximately 0.45 kb (Fig. 3A, lanes 11, 12).

## Discussion

The mitochondrial *rpl2* and *rps19* genes have a markedly different status in rice and wheat. In rice, they are intact and closely linked, whereas in wheat both genes are truncated and the residual segments are far apart in the mitochondrial genome. The degeneration of these two wheat ribosomal protein genes indicates an absence of functional constraint, but because L2- and S19-type proteins perform essential roles in all other ribosomes (in organelles, eubacteria, eukaryotes, and archaea), it is anticipated that plant mitochondrial ribosomes also have functional homologues. The simplest explanation is

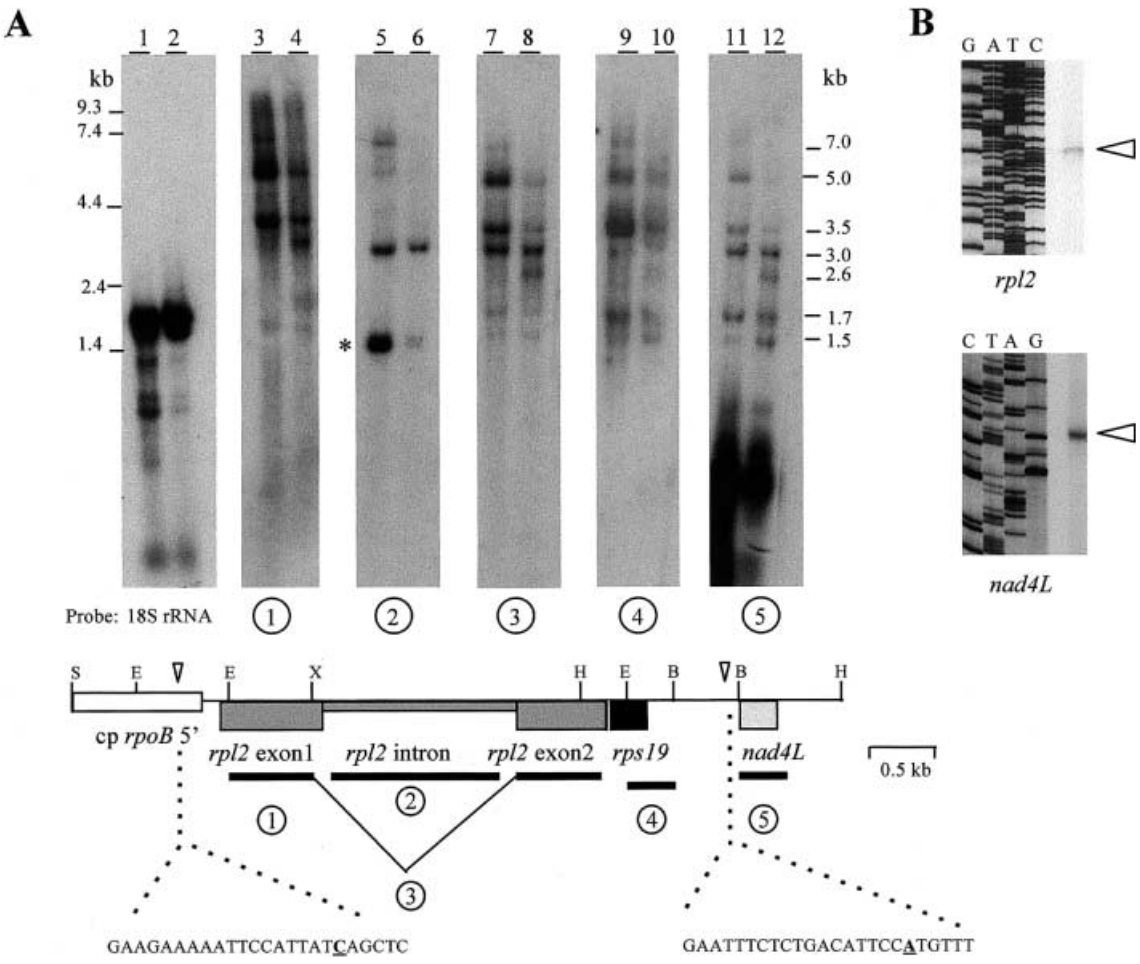



**Fig. 3 A,B** Transcript analysis of the mitochondrial *rpl2–rps19–nad4L* region in rice embryos and seedlings. **A** Northern blot analysis of RNAs isolated from germinating rice embryos (*lanes 1, 3, 5, 7, 9, 11*) and etiolated seedlings (*lanes 2, 4, 6, 8, 10, 12*) and hybridized with probes as indicated on the schematic (*numbers inside circles*). *Probe 1* consists of *rpl2* exon 1 and *probe 3* is a *rpl2* cDNA containing both exons. The *asterisk* shows the position of the *rpl2* excised intron. Restriction sites are as in Fig. 1. **B** Primer extension data mapping the 5′ termini of *rpl2* and *nad4L* transcripts from rice embryos. Sequencing ladders were generated using the same oligomers. *Open triangles* show the positions of the transcript termini and these sequences are shown on the schematic in **A**

that active copies of these two ribosomal protein genes have been translocated to the nucleus in the lineage leading to wheat. Assuming that the nuclear-located copies are appropriately expressed and the proteins are targeted back to the mitochondrion, then the mitochondrial copies are redundant and could be lost. Indeed in the case of *rpl2*, we have identified actively transcribed *rpl2* homologues in the wheat nuclear genome (S. Subramanian and L. Bonen, unpublished observations). Recent large-scale studies indicate that functional gene transfer to the nucleus within the flowering plant lineage occurs at a remarkably high frequency (Palmer et al. 2000) and it might be expected that events influencing *rpl2* and *rps19* in monocot lineages are independent from those in dicots. Indeed comparisons of the nuclear-located mitochondrial *rps19* sequences from two dicots, *Arabidopsis* and soybean, has led to the suggestion of separate activation events and probably also independent transfer events in these two lineages (Palmer et al. 2000). Interestingly in *Arabidopsis*, an expressed *rpl2*-homologous sequence corresponding to the 3′ terminal region (which is missing from the mitochondrial-located copy) has been identified in the nuclear genome (W. Schuster, X82556).

The wheat mitochondrial *rpl2* and *rps19* genes have both been inactivated through DNA deletion; however the truncated segments are still co-transcribed with downstream genes, *nad1d* and *nad4L* respectively. Their 5′ termini map to monocot mitochondrial promoter motifs, so it appears that rearrangements have brought new promoter elements into context, to ensure continued expression of downstream functional genes. In fact, without the recruitment of new promoters, such recombination events would presumably have been lethal. In both plants, the RNA profiles differ between embryos and seedlings; and in rice, two of the three co-transcripts (2.9 kb and 2.5 kb) identified as potential mRNAs for *rpl2* and *rps19* (Kubo et al. 1996) are present at very low levels in rice embryos. The rice expression signals are provided by chloroplast-derived *rpoB* 5′ flanking

sequences, which again illustrates the dynamic and chimeric nature of plant mitochondrial genomes.

Wheat and rice are estimated to have shared a common ancestor about 60 million years ago but, because of the low rate of nucleotide substitution in plant mitochondrial DNA (Wolfe et al. 1987) and the added conservative effect of RNA editing, their protein-coding sequences typically show very high similarity. This is illustrated by the rice and wheat *nad4L* genes, which share 100% nucleotide identity at the RNA level. In contrast, the remnants of the wheat *rpl2* and *rps19* genes show 94% and 98% nucleotide identity, respectively, with their rice counterparts. Because the wheat mitochondrial *rpl2* pseudogene has accumulated more mutations than *rps19*, it might be argued that its inactivation preceded that of *rps19*. Whether this might also reflect independent transfer to the nucleus is unknown, because the length of time of co-existence of functional copies in both compartments may vary. Even if transfer occurred at the same time (through an edited co-transcript), separate activation events would be necessary for both genes to acquire appropriate expression signals. This study illustrates the plasticity of flowering plant mitochondrial genome organization and gene location; and it will be of interest to learn more about the timing, frequency, and mode of these events during the evolutionary history of the cereals.

**Acknowledgement** The financial support of the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged.

## References


Adams KL, Song K, Roessler PG, Nugent JM, Doyle JL, Doyle JJ, Palmer JD (1999) Intracellular gene transfer in action: dual transcription and multiple silencings of nuclear and mitochondrial *cox2* genes in legumes. Proc Natl Acad Sci USA 96:13863–13868

Blanchard JL, Lynch M (2000) Organellar genes: why do they end up in the nucleus? Trends Genet 16:315–320

Bonen L (1987) The mitochondrial S13 ribosomal protein gene is silent in wheat embryos and seedlings. Nucleic Acids Res 15:10393–10404

Carrillo C, Bonen L (1997) RNA editing status of *nad7* intron domains in wheat mitochondria. Nucleic Acids Res 25:403–409

Chapdelaine Y, Bonen L (1991) The wheat mitochondrial gene for subunit I of the NADH dehydrogenase complex: a *trans*-splicing model for this gene-in-pieces. Cell 65:465–472

Conklin PL, Hanson MR (1991) Ribosomal protein S19 is encoded by the mitochondrial genome in *Petunia hybrida*. Nucleic Acids Res 19:2701–2705

Culver GM, Noller HF (1999) Efficient reconstitution of functional *Escherichia coli* 30S ribosomal subunits from a complete set of recombinant small subunit ribosomal proteins. RNA 5:832–843

Fey J, Marechal-Drouard L (1999) Compilation and analysis of plant mitochondrial promoter sequences: an illustration of a divergent evolution between monocot and dicot mitochondria. Biochem Biophys Res Comm 256:409–414

Figueroa P, Gomez I, Carmona R, Holuigue L, Araya A, Jordana X (1999a) The gene for mitochondrial ribosomal protein S14 has been transferred to the nucleus in *Arabidopsis thaliana*. Mol Gen Genet 262:139–144

Figueroa P, Gomez I, Holuigue L, Araya A, Jordana X (1999b) Transfer of *rps14* from the mitochondrion to the nucleus in maize implied integration within a gene encoding the iron-sulphur subunit of succcinate dehydrogenase and expression by alternative splicing. Plant J 18:601–609

Graack HR, Wittmann-Liebold B (1998) Mitochondrial ribosomal proteins (MRPs) of yeast. Biochem J 329:433–448

Grohmann L, Brennicke A, Schuster W (1992) The mitochondrial gene encoding ribosomal protein S12 has been translocated to the nuclear genome in *Oenothera*. Nucleic Acids Res 20:5641–5646

Haouazine N, Pereira de Souza A, Jubier MF, Lancelin D, Delcher E, Lejeune B (1992) The wheat mitochondrial genome contains an ORF showing sequence homology to the gene encoding the subunit 6 of the NADH-ubiquinone oxidoreductase. Plant Mol Biol 20:395–404

Haouazine N, Takvorian A, Jubier MF, Michel F, Lejeune B (1993) The *nad6* gene and the exon d of *nad1* are co-transcribed in wheat mitochondria. Curr Genet 24:533–538

Kadowaki K, Kubo N, Ozawa K, Hirai A (1996) Targeting presequence acquisition after mitochondrial gene transfer to the nucleus occurs by duplication of existing targeting signals. EMBO J 15:6652–6661

Kubo N, Ozawa K, Hino T, Kadowaki K (1996) A ribosomal protein L2 gene is transcribed, spliced, and edited at one site in rice mitochondria. Plant Mol Biol 31:853–862

Kubo N, Harada K, Hirai A, Kadowaki K (1999) A single nuclear transcript encoding mitochondrial RPS14 and SDHB of rice is processed by alternative splicing: common use of the same mitochondrial targeting signal for different proteins. Proc Natl Acad Sci USA 96:9207–9211

Kubo N, Jordana X, Ozawa K, Zanlungo S, Harada K, Sasaki T, Kadowaki K (2000) Transfer of the mitochondrial *rps10* gene to the nucleus in rice: acquisition of the 5′ untranslated region followed by gene duplication. Mol Gen Genet 263: 733–739

Lang BF, Gray MW, Burger G (1999) Mitochondrial genome evolution and the origin of eukaryotes. Annu Rev Genet 33:351–397

Loessl A, Adler N, Horn R, Frei U, Wenzel G (1999) Chondriome-type characterization of potato: mt alpha, beta, gamma, delta, epsilon and novel plastid-mitochondrial configurations in somatic hybrids. Theor Appl Genet 99:1–10

Palmer JD, Adams KL, Cho Y, Parkinson CL, Qiu YL, Song K (2000) Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. Proc Natl Acad Sci USA 97:6960–6966

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sánchez H, Fester T, Kloska S, Schröder W, Schuster W (1996) Transfer of *rps19* to the nucleus involves the gain of an RNP-binding motif which may functionally replace RPS13 in *Arabidopsis* mitochondria. EMBO J 15:2138–2149

Schuster W, Brennicke A (1991) RNA editing makes mistakes in plant mitochondria: editing loses sense in transcripts of a *rps19* pseudogene and in creating stop codons in *cox1* and *rps3* mRNAs of *Oenothera*. Nucleic Acids Res 19:6923–6928

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

Sunkel S, Brennicke A, Knoop V (1994) RNA editing of a conserved reading frame in plant mitochondria increases its similarity to two overlapping reading frames in *Escherichia coli*. Mol Gen Genet 242:65–72

Takemura M, Oda K, Yamato K, Ohta E, Nakamura Y, Nozato N, Akashi K, Ohyama K (1992) Gene clusters for ribosomal proteins in the mitochondrial genome of a liverwort, *Marchantia polymorpha*. Nucleic Acids Res 20:3199–3205

Uhlein M, Weglohner W, Urlaub H, Wittmann-Liebold M (1998) Functional implications of ribosomal protein L2 in protein biosynthesis as shown by in vivo replacement studies. Biochem J 331:423–430

Unseld M, Marienfeld JR, Brandt P, Brennicke A (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. Nat Genet 15:57–61

Vitart V, De Paepe R, Mathieu C, Chetrit P, Vedel F (1992) Amplification of substoichiometric recombinant mitochondrial DNA sequences in a nuclear, male sterile mutant regenerated from protoplast culture in *Nicotiana sylverstris*. Mol Gen Genet 233:193–200

Wischmann C, Schuster W (1995) Transfer of *rps10* from the mitochondrion to the nucleus in *Arabidopsis thaliana*: evidence for RNA-mediated transfer and exon shuffling at the integration site. FEBS Lett 374:152–156

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058