FEBS Letters 374 (1995) 152–156 FEBS 16172

# Transfer of *rps10* from the mitochondrion to the nucleus in *Arabidopsis thaliana*: evidence for RNA-mediated transfer and exon shuffling at the integration site

Christine Wischmann, Wolfgang Schuster*

*Institut für Genbiologische Forschung Berlin GmbH, Ihnestrasse 63, D-14195 Berlin, Germany*

Received 6 September 1995

**Abstract** ***Rps10*, a gene coding for ribosomal protein *S10* of *Arabidopsis* mitochondria has been transferred to the nuclear compartment, while in pea and potato the active *rps10* is mitochondrially located. The nuclear *rps10* gene contains an intron at the junction of the target signal sequence and the mitochondrial-derived sequence, indicating that exon shuffling may have been involved in the addition of the transit peptide signal. Sequence comparison of *Arabidopsis rps10* to the plant mitochondrial counterparts shows that the edited version is present in the nucleus of *Arabidopsis*. This finding corroborates RNA as an intermediate of a functional gene transfer between mitochondria and the nucleus. In vitro-translated *RPS10* protein is efficiently imported into potato mitochondria and a presequence of about 7 kDa is removed resulting in a mature protein that is larger compared to organellar and bacterial *RPS10* proteins.**

*Key words:* RNA editing; Plant mitochondria; Gene transfer; *rps10*; Protein import

## 1. Introduction

The gene content of mitochondrial genomes varies even between related species (for review see [1]). This is most obvious for ribosomal protein genes located on the mitochondrial DNA (mtDNA). In animal mtDNAs not a single gene coding for mitochondrial ribosomal proteins is present [2]. One ribosomal protein gene (*var1*) is present on the yeast mitochondrial genome [3], while the rest of the expected 60–80 proteins is encoded in the nucleus and imported from the cytosol into the organelle (for review see [4]). In *Marchantia polymorpha* sixteen genes coding for ribosomal proteins have been identified on the mitochondrial genome from homology to their bacterial counterparts [5]. The organization of ribosomal protein genes in the liverwort is similar to the situation found in *E. coli* [6] but differs in higher plants where the genes are scattered far apart in the genome [7]. In potato [8] and pea [9], a mitochondrial gene coding for ribosomal protein *S10* has been identified recently. In both species the gene seems to be functional, because a group II intron located in the *rps10* gene is spliced and the mRNA is edited [9,10]. In contrast, no *rps10* sequences have been found on the mtDNA of *Arabidopsis*, raising the question of whether this gene has been translocated to the nuclear compartment or has been lost completely in this plant. In *Oenothera*, a transcribed pseudogene of *rps12* is still present in the mitochondrial genome, while the functional gene is encoded in the nucleus [11]. The most extensively analysed example of a functional gene transfer was reported in legumes, where the *cox2* gene has been transferred to the nuclear compartment [12]. While in pea an active *cox2* gene is present in mitochondria, the mitochondrial copy in soybean is silent. In this plant the functional *COX2* is provided by a nuclear copy [13]. A similar situation was found in cowpea, however, the mitochondrial copy of *cox2* was removed completely from the mtDNA.

Here we report that the gene coding for the mitochondrial ribosomal protein *S10* of *Arabidopsis thaliana* has been transferred to the nucleus. We show that *rps10* is present as a single copy gene and describe by in vitro import studies that the protein is targeted to mitochondria.

## 2. Materials and methods

### 2.1. Isolation of mitochondria and protein import into mitochondria

Mitochondria were isolated from potato tubers (*Solanum tuberosum* var. Bintje) and suspension cell cultures of *Arabidopsis thaliana* by four 5-s pulses on a high-speed blendor in extraction buffer containing 400 mM mannitol, 1 mM EGTA, 25 mM Tricine (pH 7.2), 10 mM $\beta$-mercaptoethanol, 0.2 mM phenylmethylsulfonyl fluoride (PMSF), 0.1% BSA and subsequent filtering through two layers of miracloth. The filtrate was spun at $2,000 \times g$ for 15 min at 4°C to remove cellular debris, then at $15,000 \times g$ for 20 min at 4°C to pellet mitochondria. After several washes in a buffer containing 400 mM mannitol, 10 mM Tricine (pH 7.2), 1 mM EGTA, 0.2 mM PMSF, mitochondria were layered onto a 14%–28%–45% Percoll step gradient and spun at $70,000 \times g$ for 45 min in a Beckman SW-28 rotor at 4°C. The interface band between 28% and 45% was collected and washed serveral times. For in vitro import studies the purified mitochondria were washed and resuspended in a buffer containing 400 mM mannitol, 10 mM $KH_2PO_4$ (pH 7.2), and 0.1% BSA. The precursor of *RPS10* was synthesized from a full-length *rps10* cDNA of *Arabidopsis* in the vector pBluescript by coupled transcription/translation in the presence of [$^{35}$S]methionine according to the supplier's instructions (Promega, Madison, WI, USA). Import assays contained 40 $\mu$l purified mitochondria (10 mg protein/ml), 160 $\mu$l import buffer and 10 $\mu$l of the reticulocyte lysate translation mix. Import buffer consisted of 250 mM mannitol, 20 mM HEPES (pH 7.5), 80 mM KCl, 1 mM $K_2HPO_4$, 1 mM ATP, 1 mM malate, 2 mM NADH, and 1 mM DTT. Import was allowed to proceed for 20 min at 10°C or 20°C. For protease treatment after the import reaction 200 $\mu$g/ml proteinase K was added and samples were incubated for 20 min at 20°C, then 1 mM PMSF was added and samples were incubated for a further 15 min. For detergent solubilization, 1% (w/v) Triton X-100 was added together with proteinase K (200 $\mu$g/ml). Inhibition of mitochondrial import was tested by the addition of 1 $\mu$M valinomycin. After the various import reactions were completed mitochondria were repurified by centrifugation through a 25% sucrose cushion. The repurified mitochondria were resuspended in loading buffer and heated for 2 min at 100°C prior to fractionation on SDS polyacrylamide gels. After separation the radiolabelled products were analyzed by fluorography.

### 2.2. Isolation of nucleic acids, DNA sequence analysis

Isolation of nucleic acids and Southern hybridization has been described previously [14]. Sequences were determined by the dideoxy

*Corresponding author. Fax: (49) (30) 8300-0736.
E-mail: schuster_w@mpimg-berlin-dahlem.mpg.de


*SSDI* 0014-5793(95)01100-5

chain termination method [15] using T7 DNA polymerase. Both strands of the DNA were sequenced. Computer analysis was performed on a VAX/VMS system using the GCG package, version 8 (Madison, USA). The sequence data of *rps10* have been deposited under accession number X80694/ATRIBS10 in the EMBL Nucleotide Sequence database.

*2.3. Screening for* rps10

Approximately $5 \times 10^5$ clones of a cDNA library (Uni-ZAP XR; Stratagene, La Jolla, CA, USA) made from poly(A)$^+$ mRNA of *Arabidopsis* were screened with the pea probe to isolate clones for *rps10*. The cloning of cDNA and excision of *rps10* cDNA clones from positive phages was done according to the manufacturer's conditions.

*2.4. PCR analysis of the genomic* rps10

The genomic *rps10* sequence was amplified by polymerase chain reaction (PCR) as already described [16] by using the following primers: (1) 5′-GTCATTTCTTCGTCTCACCG-3′; (2) 5′-GAATCGCTATACAAATC-3′; (3) 5′-CCAAGAATACGATAACG-3′; (4) 5′-AAGTTCGAAATTATTGC-3′ (for location see Fig. 2). PCR products obtained were cloned into the vector pCRII using a TA cloning kit (Invitrogen) and sequenced.

## 3. Results and discussion

*3.1. Isolation of* rps10 *in* Arabidopsis

A probe of *rps10* (kindly provided by Dr. V. Knoop), covering the entire coding region of the mitochondrial gene of pea [9], was used in probing a Southern-blot to examine the presence of this gene in the mitochondrial genome of *Arabidopsis* (Fig. 1). Based on this hybridization it appears that *rps10* sequences have been completely lost from the mitochondrial genome in this plant and suggests that a gene transfer to the nuclear compartment has taken place. The possibility of a nuclear localization of *rps10* is supported by the hybridization results obtained with total DNA. While no signal is detectable in the mtDNA the presence of *rps10* in the nucleus is indicated by two hybridizing fragments. In order to examine the putative nuclear *rps10* gene, clones were isolated from an *Arabidopsis* cDNA library using the mitochondrial *rps10* of pea as a probe. Ten independant cDNA clones were found that were polyadenylated and showed high similarity to the mitochondrial *rps10* sequences of potato [10], pea [9], and *Marchantia* [5]. A putative start codon was identified near the 5′ end of the full-length cDNA consensus sequence, and the following open reading

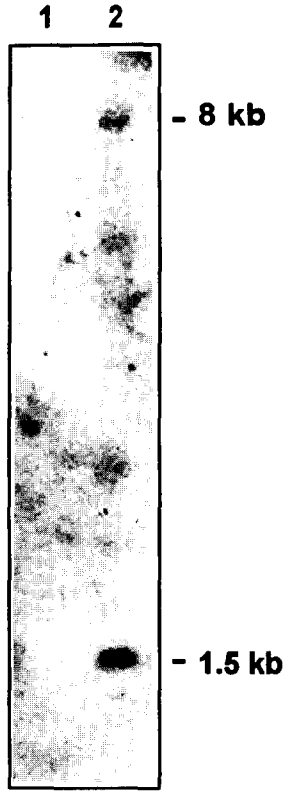


Fig. 1. *Rps10* hybridizes to nuclear DNA of *Arabidopsis thaliana*. Mitochondrial DNA (2 μg; lane 1) and total DNA (10 μg; lane 2) of *A. thaliana* were cut with *Hin*dIII, electrophoresed in a 0.8% agarose gel, transferred to a Biodyne A membrane (Pal), and hybridized with the pea *rps10* probe. Two signals of 1.5 kb and 8 kb are visible in the total DNA.

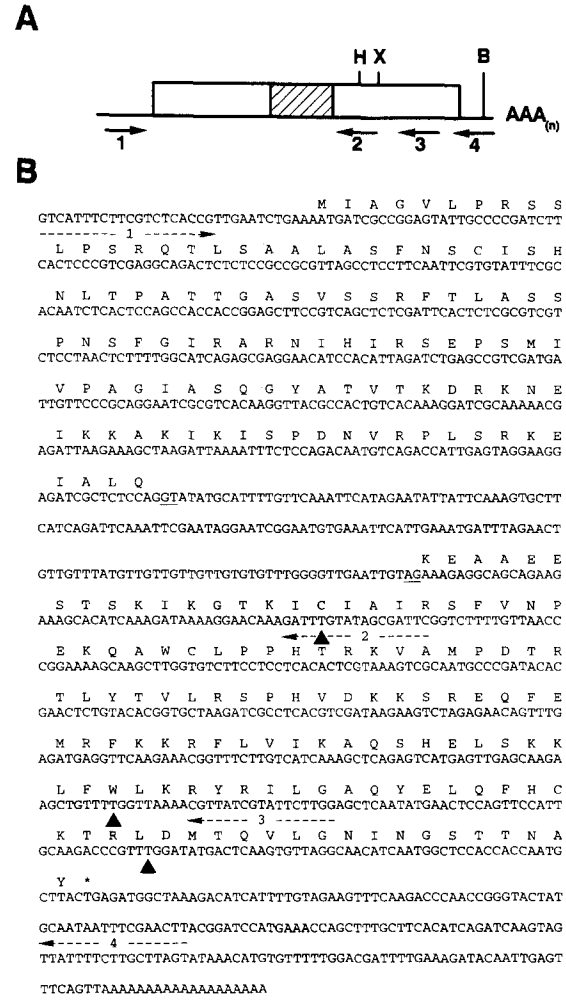


Fig. 2. Structure and nucleotide sequence of *rps10* in *Arabidopsis thaliana*. (A) Shows a scheme of *rps10* and the location of the primers used in PCR to determine the genomic structure and sequence. The location of the intron is given by a hatched box. Restriction recognition sites are indicated for *Bam*HI (B), *Hin*dIII (H), and *Xba*I (X). In (B) the nucleotide sequence of cDNA 8-1 is shown together with the intron sequence. The intron borders are underlined. The deduced amino acid sequence of *RPS10* is given above the nucleotide sequence. Arrowheads indicate nucleotide positions that are edited from C to U in mitochondrial transcripts of pea and/or potato. The location of the primers (1–4) used in PCR analysis of the genomic *rps10* structure are given below the sequence by arrows.

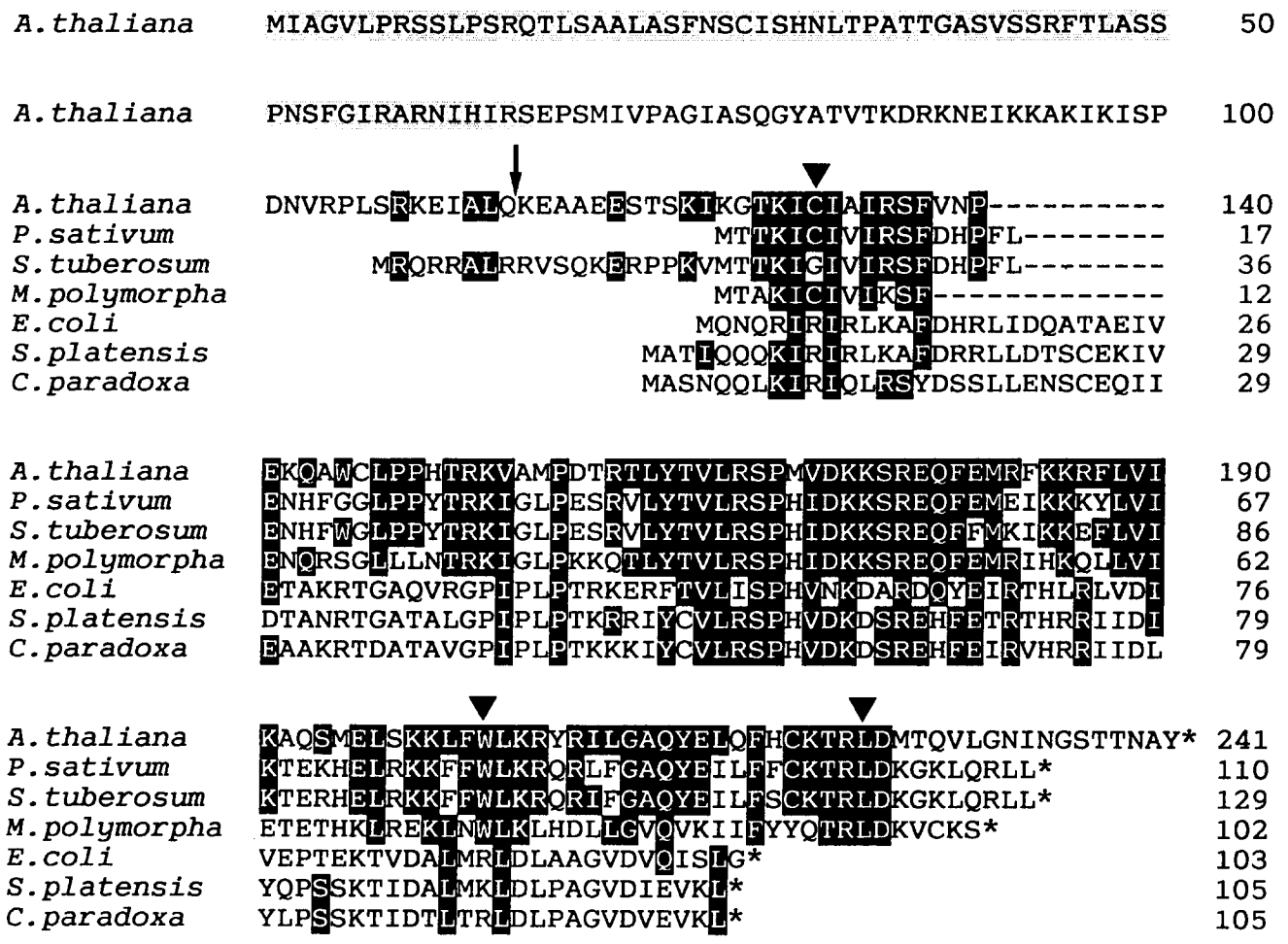


Fig. 3. Amino acid comparison of *RPS10* sequences. The cDNA-deduced amino acid sequence of *RPS10* from *Arabidopsis* is larger than the organellar/cyanellar and bacterial counterparts. Sequences of the mitochondrial *RPS10* of the liverwort *Marchantia polymorpha* (*M. polymorpha*) [5], *Pisum sativum* (*P. sativum*) [9], and *Solanum tuberosum* (*S. tubersosum*) [10], the cyanellar sequence of *Cyanophora paradoxa* (*C. paradoxa*) [24], the cyanobacterial sequence of *Spirulina platensis* (*S. platensis*) [25], and the eubacterial sequence of *Escherichia coli* (*E. coli*) [26] are shown. Homologous amino acids between the different *RPS10* polypeptides are highlighted by inverse contrast. The vertical arrow shows the intron location in the *rps10* gene of *Arabidopsis*. Amino acids changed by editing in plant mitochondrial transcripts of pea [9] and potato [10], that are likewise conserved in the *Arabidopsis* gene, are marked by arrowheads. The target signal sequence of the nuclear *rps10* is indicated by gray shading.

frame codes for a putative protein of 241 amino acids (Fig. 2). This reading frame is approximately 360 bp longer than all other known *rps10* genes. The cDNA clones varied at the 5′ end by 6 nucleotides due to incompletely synthesized cDNA. At the 3′ end poly(A) tails of different length and location were found (data not shown) indicating that different polyadenlyation sites are present. No further differences between the cDNA sequences were found suggesting a single copy gene for *RPS10* in *Arabidopsis*. This finding is supported by data from Southern blot hybridization (Fig. 1) where only two *Hin*dIII fragements of 1.5 and 8.0 kb were detected. Due to an internal *Hin*dIII restriction site in the *rps10* gene (in combination with the intensity of the hybridization signals) we conclude that the 5′-part is located on the 8.0 kb fragment whereas the 1.5 kb fragment encodes the 3′-part.

### 3.2. *The structure of the* rps10 *gene in* Arabidopsis

To examine the gene structure of *rps10*, we used different combinations of *rps10* primers to amplify the genomic region by PCR (Fig. 2). The PCR products obtained from cDNA and genomic DNA differed by 230 nucleotides in the 5′ terminal part of the gene (data not shown). The genomic PCR products were sequenced and revealed that a typical nuclear intron is located between the *rps10* homologous part and the upstream region (Fig. 2). If this upstream region is translated from the first ATG codon of the mRNA, the *rps10* gene of *Arabidopsis* would encode an N-terminal extension of approximately 120 amino acids. This putative target signal is required to direct the protein back to mitochondria, and resembles all features of a mitochondrial transit peptide in having many hydroxylated amino acids, and few negatively and many positively charged residues [17]. Separation of the transit peptide sequence by an intron from the evolutionary conserved, organellar-derived gene part was likewise observed in the transfers into the nucleus of the mitochondrial *cox2* gene in legumes [12] and of the plastid *rpl22* in all flowering plants [18]. The location of an intron at the insertion border suggests that a form of exon-shuffling [19] may be involved in a functional gene transfer. By this gene structure either a random nuclear sequence could be optimized for encoding a mitochondrial targeting peptide by

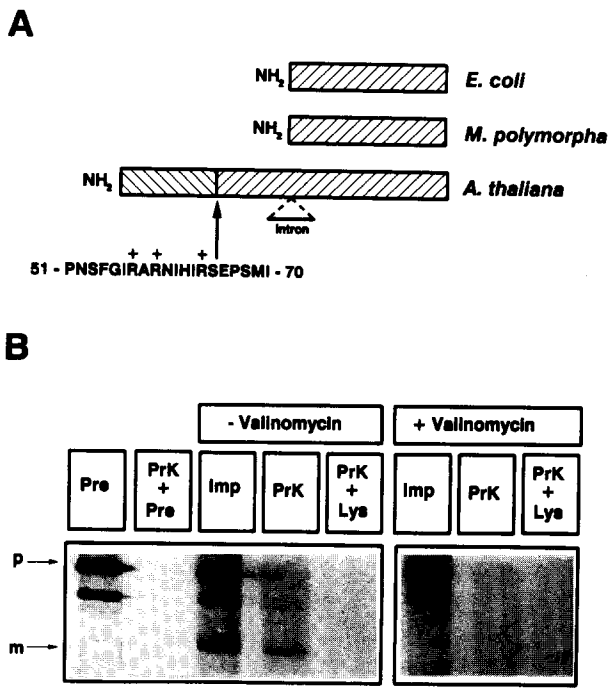


Fig. 4. Import of *RPS10* into isolated mitochondria from potato. In (A), the size differences between *RPS10* of different species are shown schematically. The intron location in the *Arabidopsis* gene differs from the estimated processing site (indicated by a vertical arrow above the amino acid sequence) by 49 residues. Positively charged amino acids around the processing site of the *RPS10* precursor are indicated. (B) shows the import of *RPS10* into isolated potato mitochondria. In vitro-translated *RPS10* precursor (Pre, p) of *Arabidopsis* is imported (Imp) into purified potato mitochondria. A polypeptide of 7 kDa is removed during the import process resulting in a mature (m) *RPS10* of about 21 kDa that is resistant to proteinase K treatment (PrK). When mitochondria are lysed by the addition of Triton X-100 (Lys) the protected protein is digested by proteinase K. Import of *RPS10* is completely abolished in the presence of valinomycin (+ Valinomycin).

using alternative splicing sites, or, in a different scenario, the transferred gene could be put into different nuclear locations after the integration into the nuclear genome. This shuffling process would also allow to find optimal conditions for activating the gene in the nuclear compartment by adding signals for transcription and regulation.

### 3.3. *The* Arabidopsis *gene corresponds to the spliced and edited* rps10 *sequence*

A typical organellar group II intron [20] is located in the *rps10* gene of both, potato and pea. This group II intron is absent from the transferred *rps10* sequence present in *Arabidopsis*. Most likely this intron has been removed before gene transfer started, either on the RNA level by splicing or was already absent (or never present in this gene) at the DNA level. Amino acid comparison of *RPS10* from *Arabidopsis* with its mitochondrial counterparts revealed that the edited version is present in the nucleus (Figs. 2B,3). In the *rps10* of pea three RNA editing sites are observed in the coding region [9], each of which changes the polypeptide sequence to a more conserved mitochondrial ribosomal protein. All these codon positions are found as edited sequence in the *Arabidopsis* gene. Surprisingly, these edited codons are highly characteristic of mitochondrial *RPS10* polypeptides. The unedited codons, however, would specify the amino acid conserved in the bacterial proteins. This finding strongly suggests the idea that RNA may act as intermediate in the translocation of organellar sequences into the nucleus.

### 3.4. *In vitro import of RPS10 into potato mitochondria*

In order to analyse the import of *RPS10* into the organelle, the precursor protein was incubated with isolated potato mitochondria. Fig. 4 shows, that after incubation of *RPS10* with isolated mitochondria, the precursor is converted to a smaller protein of about 21 kDa which may represent the mature polypeptide. This protein is resistent to added protease and represents the in vitro-imported form of the *RPS10* polypeptide. The protection was abolished when the potato mitochondria were solubilized with Trition X-100 prior to the protease treatment (Fig. 4B). Import into mitochondria requires a membrane potential, therefore addition of valinomycin inhibited the translocation of *RPS10* through the mitochondrial membranes. From comparison of the amino acid sequences of *RPS10* and the location of the intron in the *Arabidopsis* gene, a target signal sequence of about 14 kDa was expected, however, during the import process a target peptide of only 7 kDa was cleaved from the precursor protein. Therefore, the mature *RPS10* of *Arabidopsis* is about 60 amino acids larger compared to the mitochondrial and bacterial counterparts. This could mean that additional functions for this ribosomal protein may be encoded by the N-terminal extension. In yeast it has been shown that specific N- and C-terminal extensions of the ribosomal protein *MrpS28* are required for respiratory growth [21]. These extensions are not encoded in the *E. coli* counterpart of *Mrps28*, namely *rps15*, and may encode specific functions required in yeast [22]. As *RPS10* belongs to the tertiary binding proteins in ribosome assembly [23] it is more likely that the extended N-terminus is tolerated due to the location on the surface of the ribosome.

*Acknowledgements:* We thank Dr. Hugo Sanchez for critical reading of this manuscript and Iris Gruska and Waltraut Jekabsons for technical assistance. We are grateful to Dr. Volker Knoop for the pea *rps10* probe. This work was supported by grants from the Deutsche Forschungsgemeinschaft and the Bundesministerium für Forschung und Technologie.

## References

[1] Schuster, W. and Brennicke, A. (1994) Annu. Rev. Plant Physiol. Mol. Biol. 45, 61–78.
[2] Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. (1981) Nature 290, 457–465.
[3] Terpstra, P., Zanders, E. and Butow, R.A. (1979) J. Biol. Chem. 254, 12653–12661.
[4] Curgy, J.J. (1985) Biol. Cell. 54, 1–38.
[5] Takemura, M., Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Nozato, N., Akashi, K. and Ohyama, K. (1992) Nucleic Acids Res. 20, 3199–3205.
[6] Zurawski, G.and Zurawski, S.M. (1985) Nucleic Acids Res. 13, 4521–4526.
[7] Grohmann, L., Brennicke, A. and Schuster, W. (1993) in: The Translational Apparatus (K.H. Nierhaus, F. Franceschi, A.R. Subramanian, Erdmann V.A. and B. Wittmann-Liebold eds.) pp. 599–607, Plenum Press, New York.

[8] Zanlungo, S., Quiñones, V., Moenne, A., Holuigue, L. and Jordana, X. (1994) Plant Mol. Biol. 25, 743–749.
[9] Knoop, V., Ehrhardt, T., Lättig, K. and Brennicke, A. (1995) Curr. Genet. 27, 559–564.
[10] Zanlungo, S., Quiñones, V., Moenne, A., Holuigue, L. and Jordana, X. (1995) Curr. Genet. 27, 565–571.
[11] Grohmann, L., Brennicke, A. and Schuster, W. (1992) Nucleic Acids Res. 20, 5641–5646.
[12] Nugent, J.M. and Palmer, J.D. (1991) Cell 66, 473–481.
[13] Covello, P.S. and Gray, M.W. (1992) EMBO J. 11, 3815–3820.
[14] Schuster, W., Hiesel, R., Wissinger, B., Schobel, W. and Brennicke, A. (1988) in: Plant Molecular Biology (C.H. Shaw ed.) pp. 79–102, IRL Press, Oxford.
[15] Sanger, F., Nicklen, S. and Coulson, A. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[16] Yu, W. and Schuster, W. (1995) J. Biol. Chem. 270, 18227–18233.
[17] von Heijne, G., Steppuhn, G. and Herrmann, G. (1989) Eur. J. Biochem. 180, 535–545.
[18] Gantt, J.S., Baldauf, S.L., Calie, P.J., Weeden, N.F. and Palmer, J.D. (1991) EMBO J. 10, 3073–3078.
[19] Gilbert, W. (1978) Nature 271, 501.
[20] Michel, F. and Dujon, B. (1983) EMBO J. 2, 33–38.
[21] Huff, M.O., Hanic-Joyce, P.J., Dang, H., Rodrigues, L.-A. and Ellis S.R. (1993) J. Mol. Biol. 233, 597–605.
[22] Li, Y., Huff, M.O., Hanic-Joyce, P.J. and Ellis S.R. (1993) J. Mol. Biol. 233, 606–614.
[23] Tate, W.P., McCaughan, K.K., Kastner, B., Trotman, C.N.A., Stöffler-Meilicke, M. and Stöffler, G. (1988) Biochem. Int. 17, 179–186.
[24] Neumann-Spallart, C., Jakowitsch, J., Kraus, M., Brandtner, M., Bohnert, H.J. and Löffelhardt, W. (1991) Curr. Genet. 19, 313–315.
[25] Sanangelantoni, A.M. and Tiboni, O. (1993) J. Gen. Microbiol. 139, 2579–2584.
[26] Yaguchi, M., Roy, C. and Wittmann, H.-G. (1980) FEBS Lett. 121, 113–116.