THE MITOCHONDRIAL S7 RIBOSOMAL PROTEIN GENE: EVIDENCE OF GENE TRANSFER FROM THE MITOCHONDRION TO THE NUCLEUS

by

DEGEN ZHUO, M.Sc.

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfilment of the requirements for the
PhD degree in the

Ottawa-Carleton Institute of Biology
University of Ottawa
Ottawa, Ontario

1994

ISBN 0-612-00561-5

ACKNOWLEDGEMENTS

First of all, I would like to thank Dr. Linda Bonen for her guidance in this project, for her all support, constant encouragements and patience in completing this project, and for her great help in my writing this thesis and her critical comments about this thesis.

I would also like to thank Dr Douglas Johnson, Dr. Brian Miki and Dr. Rob Charlebois for serving on my research advisory committee and for their helpful suggestions and comments.

I would like to express my deep appreciation to Karen Williams, Pascale Garber and Caroline Wood for their excellent technical help.

Many thanks to Shaojiu Wang and Yangshen Liu for their help and useful comments about screening the pea and soybean nuclear genomic libraries.

I would like to thank Professor Donal A. Hickey for allowing me to draw several figures on his computer.

Many thanks to Mrs. Bea Valentine and Chen Xia for their help in preparing and printing photography.

I would like to express my thanks to Mr. Jacques Hélie for his excellent graphic service.

Special thanks to my fellow graduate students, faculty and staff members in Biology Department for their help, encouragements and service.

I would like to thank my wife Shoukang, who gave birth of our lovely son, Dawei,

five years ago, for her all support and constant encouragements and her sharing responsibility in taking care of our son, who deserve my thanks for his company, sweet hugs and nice behaviour.

Finally, I would like to give my most sincere thanks to my grandfather who has raised me to a grown-up man after my mother passed away when I was only five months and six day. His humanity, kindness, generosity and sacrifice have encouraged me to overcome tremendous difficulties and hardships.

ABSTRACT

This work was first directed at studying plant mitochondrial gene expression during different early stages of plant development. Because of the approaches used, it also led to the isolation and characterization of new plant mitochondrial genes and in turn that led to the study of gene transfer from the mitochondrion to the nucleus during evolution.

Mitochondrial cDNA clone libraries were constructed from mitoplast RNAs of 24 hour germinating wheat embryos and six-day etiolated seedlings. One such cDNA clone was found to share similarity with bacterial and chloroplast S7 ribosomal proteins. This protein is important for ribosome assembly and translation initiation. This single-copy gene (*rps7*) was found to be actively transcribed and the 5' termini of its major transcripts mapped immediately upstream of a repeat element, which also precedes the wheat mitochondrial genes for *coxII*, *orf25* and *atp6*. The precursor *rps7* RNA undergoes C-to-U type of RNA editing at three positions within the coding region and one site unexpectedly decreases amino acid similarity with homologues. This site and the second non-silent one show full editing in RNA from germinating embryos whereas some cDNA clones show incomplete editing for these or the silent site in RNA from etiolated seedlings. The patterns indicate no polarity to the RNA editing process.

Sequences homologous to the wheat mitochondrial *rps7* gene are present in the rice and pea mitochondrial genomes, but absent from that of soybean. Interestingly, the pea homologue lacks 119 bp of the 5' *rps7* coding region and it is not present anywhere else in

the pea mitochondrial genome nor is the 3' coding region transcribed. Therefore, it is a truncated pseudogene. Immediately upstream, there is an open reading frame of 248 amino acids, which is actively transcribed and homologous to the *orf228* gene in liverwort mitochondria. The derived protein of this ORF shares similarity with those of the *helC* genes in *Rhodobacter capsulatus* and *orf263* gene in *Bradyrhizobium japonicum*. This gene product is believed to be involved in the translocation of heme from the matrix to the intermembrane space in cytochrome *c* biogenesis.

Southern blot analysis showed that there are *rps7* homologous sequences present in pea and soybean nuclear DNAs. In order to study them further, pea nuclear DNA, pea cDNA and soybean nuclear DNA libraries obtained from Clontech were screened with the wheat mitochondrial *rps7* probes. The potentially positive clones turned out to be yeast-bacterial shuttle vectors (in the case of the pea nuclear DNA library) and mitochondrial DNA contamination (in the pea cDNA library). The potentially positive clones from a soybean nuclear DNA library turned out to contain mitochondrial DNA of the *orf228* gene plus a 92 bp DNA sequence homologous to the wheat mitochondrial *rps7* gene, although the latter was undetected by Southern hybridization analysis.

This work, in addition to identifying and characterizing new plant mitochondrial genes, provides evidence for the recent transfer of functional genetic information from the organelle to the nucleus.

RÉSUMÉ

Le but initial de ce travail a été d'étudier l'expression des gènes mitochondriaux des plantes durant leurs premières étapes de développement. Les approches utilisées ont rendu possible l' isolation et la caractérisation de nouveaux gènes mitochondriaux. Ceci nous a permis d' étudier le processus du transfert des gènes des mitochondries au noyau au courss de l'évolution dans les plantes.

Des librairies d'ADN complémentaires (ADN-c) ont été construites à partir d'ARN de mitoplaste d'embryons de germe de blé (24 heures) et à partir de plantules étiolées de six jours. Un clone d'ADN complémentaire ayant de la similarité avec les protéines ribosomales S7 chloroplastiques et bactériennes a été isolé. La protéine ribosomale S7 (*rps7*) est importante pour l'assemblage des ribosomes ainsi que pour l'initiation de la traduction. Ce gène à copie unique (*rps7*) est transcrit activement. Les extrémités 5' des produits de transcription majeurs se situent juste avant un élément répétitif qui précède aussi les gènes mitochondriaux *coxII*, *orf25*, et *atp6* du blé. L'ARN precurseur du *rps7* subit des modifications du type C-à-U (édition d'ARN) à trois positions différentes dans la région traduite. Deux de ces positions sont non-silencieuses. La modification qui se passe à l' une des positions a comme effet de diminuer la similarité (au niveau de la séquence d'acides aminés) envers d'autres gènes homologues. Cette position ainsi que l'autre position non-silencieuse sont modifiées au complet dans l'ARN des embryons en phase de germination, tandis que quelques clones d' ADN complémentaire montrent des modifications incomplètes aux trois positions dans l'ARN

des plantules etiolées. Ces caractéristiques semblent indiquer le manque de polarité dans le processus de modification de ces ARN.

Des séquences homologues au gène mitochondrial *rps7* du blé sont aussi présentes dans les génomes mitochondriaux du riz et du pois, mais sont absentes dans le soya. Il est intéressant de noter que la séquence homologue au *rps7* du pois manque 119 bp de l'extremité 5' de la région traduite du *rps7* et n'est pas présente ailleurs dans le génome mitochondrial du pois. Le fait que la région codante de l'extrémité 3' ne soit pas transcrite suggère que celui-ci est un pseudogène tronqué. Il y a un cadre de lecture ouvert (ORF) de 248 acides aminés qui précède l'extrémité 5' de cet homologue du gène *rps7* et qui semble être activement transcrit. Celui-ci est homologue au gène *orf228* dans le génome mitochondrial de l' hépatique *Marchantia*. La protéine derivée de cette région codante a de la similarité avec les gènes *helC* de *Rhodobacter capsulatus* et *orf263* de *Bradyrhizobium japonicum*. Cette protéine semble être impliquée dans la translocation de l'hème de la matrice à l'espace intermembranaire pendant la biogenèse du cytochrome *c*.

Une analyse de Southern démontre qu'il y a des séquences homologues au gène *rps7* dans l'ADN nucléaire du pois et du soya. Afin de mieux étudier ces séquences, des librairies faites à partir d'ADN nucléaire de pois, d'ADN complémentaire de pois, ainsi que d'ADN nucléaire de soya obtenues de Clontech ont été sondées en utilisant le gène *rps7* mitochondrial de blé. Des clones positifs potentiels se sont révélés soit des séquences de vecteur-navette levure/bactérie (dans le cas de la librairie d'ADN nucléaire de pois), soit de l'ADN mitochondrial contaminant (dans le cas de la librarie d'ADN-c de pois). Les clones potentiellement positifs de la librairie d'ADN nucléaire de soya contiennent de l'ADN mitochondriel du gène *orf228* plus une séquence de 92 paires de bases de séquences ayant de l'homologie avec le gène mitochondrial *rps7* du blé, biên que celur-ci n'ait pas été détecté par l'analyse de Southern.

Ce travail, en plus d'identifier et de caractériser des nouveaux gènes mitochondriaux de plantes, démontre qu'il y a eu un transfert récent de l'information génétique fonctionelle du génome mitochondrial au génome nucléaire.

Contents

List of Figures

List of Tables

List of Abbreviations

A	adenosine
AMV	avian myeloblastosis virus
atp	genes for ATP synthase subunits
bp	base pairs
BSA	bovine serum albumin
C	cytidine
cDNA	complementary DNA
CMS	cytoplasmic male-sterility
cob	gene for cytochrome$_{b-c1}$ subunit
cox	genes for cytochrome oxidase subunits
ddNTP	dideoxynucleotide triphosphate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EGTA	ethyleneglycol-bis-tetraacetic acid
G	guanosine
I	inosine
kb	kilo base pair
mp	mitoplast
mRNA	messenger RNA
mt	mitochondrial
mtDNA	mitochondrial DNA
mtRNA	mitochondrial RNA
nad	genes for NADH dehydrogenase subunits
nt	nucleotide
ORF	open reading frame

PCR	**polymerase chain reaction**
RNA	**ribonucleic acid**
RNase	**ribonuclease**
rpl	**genes for proteins of the ribosomal large subunits**
rpm	**revolutions per minute**
rps	**genes for proteins of the ribosomal small subunits**
rRNA	**ribosomal RNA**
RT-PCR	**reverse-transcriptase PCR**
SDS	**sodium dodecyl sulphate**
SSC	**0.15 mM NaCl, 0.015 mM sodium citrate (pH7)**
T	**thymidine**
TE	**10 mM Tris-HCl (pH7.5), 1 mM EDTA**
Tris	**tris-(hydroxymethyl)aminomethane**
tRNA	**transfer RNA**
U	**uridine**

CHAPTER 1 LITERATURE REVIEW

1.1 The eukaryotic cell and endosymbiotic origin of organelles.

Living organisms are divided into three fundamentally different groups: eubacteria, archaebacteria, and eukaryotes (Woese et al. 1990). Unlike the first two groups, eukaryotic cells compartmentalize the nuclear DNA away from the cytoplasm and have discrete subcellular organelles including the nucleus, mitochondrion, and in the case of plants and algae, plastids. The majority of traits are controlled by genes on the nuclear chromosomes. Eukaryotic genome sizes are extremely large compared to those in prokaryotic cells, ranging from 1×10^7 bp in fungi to 1×10^{13} bp in some plants (reviewed in Cavalier-Smith 1993). In addition, DNAs are present in chloroplasts and mitochondria. Mitochondrial genomes range in size from about 16 kb in mammals to 2400 kb in some plants; the majority of the plastid genomes are between 120 and 160 kb in size (reviewed in Gray 1992). The main function of mitochondria is oxidative phosphorylation and the genes they contain contribute to this function of respiration. They also contribute to the biosynthesis of pyrimidines, amino acids, phospholipids, nucleotides, folate coenzymes, heme, urea, and many other metabolites (Douce and Neuburger 1989). The main function of chloroplasts is photosynthesis, but they also carry out the storage and synthesis of pigments and other metabolites (Mullet 1988).

In eukaryotic cells, some aspects of gene expression are different from that of prokaryotes. Their transcription takes place in the nucleus and is controlled by various regulatory elements, including promoters, enhancers, and silencers (reviewed in Maniatis et al. 1987; Goldberg et al. 1989). After introns are removed from transcripts and other modifications are made, they are exported into the cytoplasm where the mRNAs are translated into proteins. In contrast, the regulation of eubacterial and archaebacterial genes is simpler. For example, only a few introns have been found and transcription and translation are coupled in the single cellular compartment (Cavalier-Smith 1993).

In eukaryotic cells, the nuclear genes, to which both parents contribute equally except for the sex chromosomes in animals, segregate in offspring according to Mendelian rules. On the other hand, inheritance of mitochondrial and chloroplast genes is uniparental, usually from the maternal parent. However, in some rare cases, chloroplast and mitochondrial genes are paternally transmitted (Neale et al. 1989).

Mitochondria and chloroplasts have their own unique genomes. Their genes are transcribed and translated by their own machinery. There have been two theories proposed to explain the origin of organelles: the autogenous origin theory (origin from within) (Raff and Mahler 1972) and the endosymbiosis theory (xenogenous origin, or origin from outside) (Margulis 1981). The autogenous origin theory suggested that the nuclear, mitochondrial and plastid genomes came from a single ancestral genome by a process of intracellular fragmentation, compartmentalization and functional specialization within the same cell. If so, the nuclear, mitochondrial and chloroplast genomes would share a more recent common ancestor compared to other organisms.

The endosymbiotic hypothesis (Margulis 1981) proposed that the nuclear, mitochondrial and plastid genomes once existed independently in different cells. The mitochondria and plastids originated from bacteria, which were engulfed by proto-eukaryotes.

They may have been acquired at the same time or sequentially with mitochondria acquired first (reviewed in Gray 1992). By providing efficient energy-transducing systems, respiring and photosynthetic bacteria would confer selective advantage to respiration-inefficient host proto-eukaryotic cells facing an increasingly aerobic environment. They established an internal symbiotic relationship. During that period, most respiring and photosynthetic bacterial genes were transferred into the nucleus of the host and the organelles became indispensable parts of the resulting eukaryotic cells (reviewed in Gray 1989; 1992). According to the endosymbiotic theory, a number of assumptions have been made. The major ones are: (i) that nuclear, plastid and mitochondrial genomes originated from distantly separated lineages that had a long time of independent existence (Gray 1989); ii) that most genes from the endosymbiont were transferred into the nucleus of the host cell and these genes were integrated into the nuclear genome during the early stages of endosymbiosis (Harington and Thornley 1982); iii) that some genes of the endosymbiont were degraded during evolution since they were no longer needed or their functions were provided by the nuclear ones (Harington and Thornley, 1982); and iv) a remaining subset of genes have been maintained in mitochondria and chloroplasts (Obar and Green 1985).

There are many lines of evidence to support the endosymbiotic theory (reviewed in Gray 1992). Some of the strongest are: (1) the mitochondrial and chloroplast ribosomal RNA genes are more similar to those found in prokaryotes than those in the nucleus; (2) the organellar ribosome features and resistance to antibiotics are very similar to eubacterial ones; (3) some membrane components of endosymbionts are similar to those of photosynthetic and respiring bacteria; (4) the chloroplast genome organization is more similar to the prokaryotic one than the nuclear counterpart. It is generally believed that the nuclear genomes originated from the lineage leading to archaebacteria (reviewed in Gray 1992). It has been shown by comparison of elongation factor-1α protein sequences that extremely thermophilic, sulphur-metabolizing bacteria (eocytes) appear to be the closest surviving relatives of the eukaryotes (Rivera and Lake 1992). Phylogenetic studies also suggest that mitochondria and chloroplasts originated about 1 to 1.5 billion years ago from purple nonsulphur bacteria (alpha group) and blue green algae (photosynthetic cyanobacteria), respectively (Gray et al. 1989).

Since gene expression in prokaryotes is different from that in eukaryotes, the genes transferred during early endosymbiosis had to acquire DNA sequences that control eukaryotic gene expression. Since the gene products are imported from the cytoplasm to mitochondria or plastids (Pfanner and Neupert 1990), the genes also had to acquire DNA sequences that code for signal peptides to be specifically targeted into the mitochondrion or plastid. Such signal peptides contain about 15-70 amino acids (usually located at the amino acid terminus). Although sequence comparisons reveal no obvious homology, signal peptide sequences are generally rich in positively charged and hydroxylated amino acids and have the capability of forming amphiphilic α-helices (or possibly ß-sheets) when in contact with the lipid bilayer (Glover and Lindsay 1992). They must then be removed to generate mature proteins after transfer into the mitochondrion (Pfanner and Neupert 1990). Many such examples of transferred genes, consistent with the endosymbiosis theory, have been found (reviewed in Glover and Lindsay 1992).

1.2. Mitochondrial genome organization and gene content.

Mitochondrial genomes are very divergent in organization, size and mode of gene expression among different lineages of eukaryotic organisms (reviewed in Bonen 1991). It is a question of interest why they have taken different evolutionary paths. In almost all cases they encode at least two ribosomal RNAs, a set of transfer RNAs and 8-12 components of the respiratory chain (reviewed in Gray 1992).

1.2.1 Animal, fungal and protist mitochondria.

Animal mitochondrial genomes have been extensively characterized, in part because they are small circular molecules with no recombinationally-active repeats (reviewed in Clayton 1991, 1992). They vary in size from only 14 kb to 20 kb (Table 1.1) and are extremely economical in sequence utilization. Each genome has two rRNA genes, a full set of 22 tRNA genes for translation and genes coding for 12 polypeptides (Table 1.1). In most animals, the mitochondrial tRNA genes are dispersed throughout the genome and act as RNA processing signals (reviewed in Clayton 1991, 1992).

In contrast, fungal mitochondrial genomes are somewhat larger and range from 18 kb in *Schizosaccharomyces pombe* to 176 kb in *Agaricus bitorquis* (reviewed in Clark-Walker 1992). In general, they have a slightly larger set of protein-coding genes in addition to two rRNA genes and 28 tRNA genes (Table 1.1). Many fungal mitochondrial genes contain optional introns and the spacer DNA between genes is larger, both of which contribute to inter- and intraspecific size variations in fungal mitochondrial genomes (reviewed in Clark-Walker 1992). Recombination also results in diversity of genome organization and complexity

Table 1.1 Mitochondrial genome sizes and gene content from animals, fungi, protists and plants.

The symbols + and - indicate the genes are present or absent in the organelle. +* indicates the gene is present in some species, but absent in others and +o indicates part of the gene is present. Data are taken from a. Clayton 1991; b. Clark-Walker 1992; c. Gray 1992; d: Pritchard et al. 1990; e. Wolff et al. 1994; f. Oda et al. 1992; g. Bonen and Brown 1993; h. Grohmann et al. 1993.

	Animals[a]	Fungi[b,c]	Protists[c,d,e]	Plants	
				Liverwort[f]	Flowering plants[g,h]
Genome size	14 - 20 kb	18 - 176 kb	15 - 50 kb	187 kb	200 -2500 kb
LSU rRNA	+	+	+	+	+
SSU rRNA	+	+	+	+	+
5S rRNA	-	-	+*	+	+
tRNA	22	24 - 28	0 - 26	27	≤ 20
coxI	+	+	+	+	+
coxII	+	+	+*	+	+*
coxIII	+	+	+*	+	+
cob	+	+	+	+	+
atp6	+	+	+*	+	+
atp8	-	+	+*	+	+*
atp9	-	+*	+*	+	+
atpa	-	-	+*	+	+
nad1	+	+*	+	+	+
nad2	+	+*	+	+	+
nad3	+	+*	+*	+	+
nad4	+	+*	+	+	+
nad4L	+	+*	+*	+	+
nad5	+	+*	+	+	+
nad6	+	+*	+*	+	+
nad7	-	-	+*	+	+
nad8	-	-	+*	+	+
nad9	-	-	+*	+	?
rps1	-	-	+*	+	+o
rps2	-	-	+*	+	?
rps3	-	-	+*	+	?
rps4	-	-	+*	+	?
rps7	-	-	+*	+	?
rps8	-	-	+*	+	?
rps10	-	-	+*	+	?
rps11	-	-	+*	+	?
rps12	-	-	+*	+	+*
rps13	-	-	+*	+	+*
rps14	-	-	+*	+	+*
rps19	-	-	+*	+	+
rpl2	-	-	+*	+	+o
rpl5	-	-	+*	+	+
rpl6	-	-	+*	+	?
rpl12	-	-	+*	+	?
rpl14	-	-	+*	+	?
rpl6	-	-	+*	+	+
var-1	-	+*	-	-	-
S5	-	+*	-	-	-
RNA maturase	-	+	+*	+	+
ORF	-	2 - 3	0 - 18	27	≥11

(Clark-Walker 1992). In addition, some fungal mitochondria have DNA or RNA plasmids, which apparently play no part in the biogenesis of mitochondria (Meinhardt et al. 1990)

The protist mitochondrial genomes are very divergent in genome organization and gene content. Those examined range in size from 15 to 50 kb (reviewed in Cummings 1992; Stuart and Feagin 1992; Gray 1992). For example, the *Chlamydomonas reinhardtii* mitochondrial DNA, which is very small, codes for two rRNAs, seven protein-coding genes, two ORFs and only three tRNA genes. The complete sequence of the 43 kb mtDNA of *Paramecium* reveals that it encodes four ribosomal protein genes and 16 open reading frames (ORFs) in addition to two rRNA genes and components of the respiratory chain. However it has only three tRNA genes (Pritchard et al. 1990).

The mitochondrial DNA from the order Kinetoplastidae which includes trypanosomes is unique in organization and has two forms of mitochondrial molecules: maxicircles (from 20 to 38 kb) and minicircles which are organized as a network of about 10,000 interlocking circular molecules (reviewed in Stuart and Feagin 1992). Maxicircles encode two rRNA genes, nine protein-coding genes and several open reading frames (ORFs) including some cryptogenes, which have many deletions and insertions in the coding region and show no homology with mitochondrial genes from other organisms. However, their transcripts are corrected by RNA editing (see Chapter 1.3.2.2) to give functional mRNAs (Simpson and Shaw 1989). Interestingly, the trypanosomal mitochondrial DNA has not been found to encode any tRNA genes (Stuart and Feagin 1992).

1.2.2 Plant mitochondria.

The most striking characteristic of flowering plant mitochondrial genomes is that they are very large and complex in structure. The plant mitochondrial genomes are generally believed to be circular and range from 208 kb in *Brassica hirta* (Palmer and Herbon 1987) to 2400 kb in some species of *Cucurbitaceae* (Ward et al. 1980). However, full-size circular molecules have not been observed in flowering plants (Bendich 1993). The wheat mitochondrial genome is estimated to be about 430 kb (Lejeune et al. 1985). Most mitochondrial genomes from flowering plants contain a number of recombinationally-active repeat elements, which result in mitochondrial DNA rearrangements and generate subclasses of DNA molecules through homologous recombination (Hanson and Folkerts 1992). For example, the wheat mitochondrial DNA has 10 recombinationally-active elements. On the other hand, the mitochondrial genome of the bryophyte liverwort (*Marchantia polymorpha*) is about 187 kb and has been completely sequenced (Oda et al. 1992). No recombinationally-active repeats have been found and this DNA has been found to be a circular molecule under the electron microscope (Oda et al. 1992).

Plant mitochondrial DNAs encode two large rRNA genes, but unlike almost all other eukaryotes, they also encode a 5S rRNA gene (Gray 1992). It has been shown through systematic studies that 16 different tRNA genes are present in wheat mitochondria (Joyce and Gray 1989a), whereas the potato mtDNA appears to encode 20 tRNA genes (Dietrich et al. 1992). The set differs slightly among different flowering plants, but in all cases is incomplete (Joyce and Gray 1989b; Dietrich et al. 1992). This suggests that the rest of the tRNA genes have been transferred into the nucleus and there is evidence that certain tRNAs are imported from the cytoplasm (reviewed in Dietrich et al. 1992). For example, in transgenic potato plants carrying a wild type or mutated bean nuclear leucine tRNA gene, such tRNAs were detected in the cytosol and inside mitochondria, providing evidence that the mitochondrial leucine tRNA is derived from a nuclear gene and imported into the mitochondria (Small et al. 1992). In the liverwort mitochondria, however, 29 genes for 27 different species of tRNAs have been identified. Two additional species of tRNA ($tRNA^{Thr}$ and $tRNA^{Ile}$) are required for decoding all the codons in the liverwort mitochondrial genes assuming there are conventional wobble rules and tRNA modifications (Oda et al. 1992). Sixteen protein-coding genes involved in the respiratory chain have been found to be encoded by the mitochondrial genomes from flowering plants, some of which have not been identified in the mitochondria of animals and fungi (Table 1.1). In addition, at least nine ribosomal protein genes, which are not present in either animal or fungal mitochondrial genomes, have been identified (reviewed in Bonen and Brown 1993; Grohmann et al. 1993; Yei et al. 1993). Furthermore, some open reading frames have been found (Table 1.1). Since these genomes are much larger than the others and none have yet been completely sequenced, it is possible that they contain more protein-coding genes. This idea has further been supported by the complete sequence of the liverwort mitochondrial genome (Oda et al. 1992), which encodes 94 possible genes including 30 known protein-coding genes and 29 ORFs, as well as genes given above.

In vitro protein synthesis in isolated plant mitochondria has provided information about gene content. When mitochondria were incubated with α-[^{35}S]-methionine, approximately 15-30 and 30-50 polypeptides were visualized on one-dimensional and two dimensional SDS-polyacrylamide gels, respectively (Lind et al. 1991; Rios et al. 1991). In addition, in transcriptional mapping by Northern blot analysis, Makaroff and Palmer (1987) used clones covering the entire *Brassica campestris* mitochondrial genome (218 kb) and identified 24 abundant non-overlapping transcripts, which cover approximately 30% of the genome (Makaroff and Palmer 1987). Also, many non-abundant transcripts were found (Makaroff and Palmer 1987). If the wheat mitochondrial genome encodes a similar number of protein-coding genes, then approximately 15% of the genome would be actively transcribed. All this information suggests that mitochondrial genomes of plants encode more protein-coding genes than the non-plant ones.

The flowering plant mitochondrial genes that have been identified so far include those coding for subunits of cytochrome c oxidase (*coxI*, *coxII* and *coxIII*), the Fo portion of ATP synthase (*atp6* and *atp9*), F1 portion of the ATP synthase (*atpa*), the subunits of NADH dehydrogenase (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7* and *nad9*), apocytochrome b (*cob*) and ribosomal proteins (*rps1A*, *rps3*, *rps12*, *rps13*, *rps14*, *rps19*, *rpl2*, *rpl5* and *rpl16*). Only a part of the *rps1A* and *rpl2* genes has been found (Gonzalez et al. 1993; Vitart et al. 1992; Schuster 1993), however, the protein encoded by the *rps1A* gene has been detected using antibodies in wheat mitochondria (Gonzalez et al. 1993). Several open reading frames (ORFs), which are common among different plants and transcribed, have also been identified (Bonen and Brown 1993; Sünkel et al 1994; Chanut et al. 1993). In several cases, the proteins encoded by ORFs have been confirmed, for example, maize *orf25* (Prioli et al. 1993) and wheat *orf589* (Gonzalez et al. 1993). However, their functions remain to be determined. It has been suggested that these ORFs may be involved in plant mitochondrial transcriptional and translational mechanisms, in respiratory function or plant specific roles (Bonen 1991).

This idea has been supported by recent findings of some mitochondrial ORFs which share similarity with the bacterial proteins involved in c-type cytochrome biogenesis (Gonzalez et al. 1993; Schuster et al. 1993). In the symbiotic bacterium *Bradyrhizobium japonicum* and in the photosynthetic bacterium *Rhodobacter capsulatus*, at least four genes have been found to be involved in cytochrome c biogenesis by genetic analyses (Ramseier et al. 1991; Beckman et al. 1992). These genes have been designated as *helA*, *helB*, *helC* and *orf25* in *Rhodobacter capsulatus* (Beckman et al. 1992) and *cycV*, *cycW*, *orf263*, and *cycX* in *Bradyrhizobium japonicum* (Ramseier et al. 1991). The second cluster involved in cytochrome c biogenesis has so far been observed only in *Rhodobacter capsulatus* and contains at least two genes, *ccl1* and *ccl2* (Beckman et al. 1992). Derived protein sequence comparisons have shown that the protein sequence of the *orf509* gene in liverwort mitochondria (Oda et al. 1992) shares similarity with that encoded by the *ccl1* gene in *R. capsulatus* (Schuster et al. 1993; Gonzalez et al. 1993). In addition, the liverwort mitochondrial *orf277* and *orf228* genes (Oda et al. 1992) are homologous to *helB/cycW* and *helC/orf263* genes in *R. capsulatus* and *B.japonicum*, respectively (Schuster et al. 1993; Gonzalez et al. 1993).

1.2.2.1 Mitochondrial ribosomal protein-coding genes.

Because a main part of my thesis is involved with ribosomal protein genes, I will here highlight the ribosomal protein genes encoded by mitochondrial genomes. Since mitochondrial mRNAs are translated into proteins by the mitochondrial translational apparatus, understanding the nature of this machinery should help us know more about translational regulation of mitochondrial genes. According to the endosymbiotic theory, chloroplast and mitochondria originated from eubacteria. Therefore, their translational apparatus is expected to be more similar to the bacterial ones than the cytoplasmic ones. In *E. coli*, the ribosomal protein genes are clustered together to form six operons: α (S13-S11-S4-rpoA-L17), *str* (S12-S7-EF-G-EF-Tu), ß (L10-L7/L12-ß-ß'), L11 (L1-L11), *spc* (S14-L14-L5-S14-S8-L6-L18-S5-L30-L15-secY-X) and S10 (S10-L3-L4-L23-L2-S19-L23-S3-L16-L27-L29-S17) (Nomura et al. 1980).

The completely sequenced liverwort, tobacco, and rice chloroplast genomes (Ohyama et al. 1986; Shinozaki et al. 1986; Hiratsuka et al. 1989) show that each of them encodes about 20 ribosomal proteins (Table 1.2). Comparisons show that the *rps16* gene is present in rice and tobacco, but not liverwort and, conversely that the *rpl21* gene is present in liverwort, but not in tobacco or rice chloroplast DNA (reviewed in Palmer 1991).

The chloroplast ribosomal protein genes have organizations similar to the bacterial ones. The rest of the chloroplast ribosomal protein genes are generally believed to have been transferred into the nucleus (Subramanian 1993). Although it is not clear if some ORFs in chloroplast genomes might encode ribosomal proteins, the chloroplast ribosome also contains at least five ribosomal proteins that do not share similarity to any bacterial ones. These novel chloroplast ribosomal proteins are nuclear-encoded (Subramanian 1993).

For yeast mitochondria, only one ribosomal protein gene (*var1*) has been found (Butow et al. 1985). In *Neurospora crassa* mitochondria, the large ribosomal RNA gene has an intron that contains an ORF encoding the ribosomal protein "S5" (Burke and RajBhandary 1982). The completely sequenced mitochondrial genome of *Paramecium aurelia* (Pritchard et al. 1990) reveals that it encodes four ribosomal protein genes (*rps12*, *rps14*, *rpl2* and

Table 1.2 Ribosomal proteins encoded by mitochondria and chloroplasts.

(A) small subunit ribosomal proteins; (B) large subunit ribosomal proteins. The symbols + and - indicate the genes present or absent in the organelle, * represents a pseudogene and +o indicates part of the gene identified. Data are taken from a. Shinozaki et al. 1986; b. Hiratsuka et al. 1989; c. Ohyama et al. 1986; d. Reith and Munholland 1993; e. Michalowski et al. 1990; Löffelhardt et al. 1991; f: Wolf and Del Giudice 1988; g. Clark-Walker 1992; h. Yei et al. 1993; i. Grohmann et al. 1993. Schuster 1993; j. Lang and Forget 1993; k. Wolff et al. 1994; l. Gray 1992; m: Pritchard et al. 1990; n. Stuart and Feagin 1992; o. Oda et al. 1992; p. Bland et al. 1986; Bonen 1987; Gonzalez et al. 1993; q. Hunt and Newton 1991; Gualberto et al. 1991; Bland et al. 1986; r. Suzuki et al. 1991; s. Grohmann et al. 1993; t. Vitart et al. 1992; Bland et al. 1986; u. Schuster 1993; v. Rasmussen and Hanson 1989; Conklin and Hanson 1991; x. Wahleithner and Wolstenholme 1988. y. Bonen 1987. The organisms in this table are *Porphura purpurea*; *Cyanophora paradoxa*; *Aspergillus nidulans*; *Neurospora crassa*; *Podospora anserina*; *Saccharomyces cerevisiae*; *Allomyces macrogynus*; *Phytophthora infestans*; *Prototheca wicherhamii*; *Chlamydomonas reinhardtii*; *Paramecium aurelia*; *Trypanosoma brucei*; tobacco (*Nicotiana tabacum*); liverwort (*Marchantia polymorpha*); wheat (*Triticum aestivum*); rice (*Oryza sativa*), pea (*Pisum sativum*); maize (*Zea mays*); broad bean (*Phaseolus vulgaris*) and soybean (*Glycine max*).

A

	rps 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Chloroplast Genomes																			
Land Plants																			
Angiosperms[a,b]	−	+	+	+	−	+	+	−	+	+	−	+	+	+	−	+	+	+	−
Liverwort[c]	−	+	+	+	−	+	+	−	+	+	−	+	+	−	−	+	+	+	−
Red Alga																			
Porphura purpurea[d]	+	+	+		+		+					+			+	+	+	+	+
Glaucophyta																			
C. paradoxae[e]					+			+											
Mitochondrial Genomes																			
Fungi																			
A. nidulans[f]	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
N. crassa[g]																			
S. cerevisiae[g]																			
A. macrogynus[g]	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
Protists																			
P. infestans[j]	−	+	+	+	−	−	+	+	−	+	+	+	+	+	−	−	−	−	+
P. wickerhamii[k]	−	+	+	+	−	−	+	−	−	+	+	+	+	+	−	−	−	−	+
C. reinhardtii[l]	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
P. aurelia[m]	−	−	−	−	−	−	−	−	−	−	+	−	−	+	−	−	−	+	−
T. brucei[n]	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−
Land Plants																			
Liverwort[o]	+	+	+	+	−	−	+	+	−	+	+	+	+	+	−	−	−	−	+
Angiosperms																			
Wheat[p]	+o											+	+	−					
Maize[q]			+									+	+	−					
Rice[r]												+							
Oenothera[s]			+									*	+	+					+
N. tabacum[t]													+						
Daucus carota[u]													+						+
Petunia hybrida[v]			+									+	+						
Broad bean[x]													−	+					
Pea[y]													−	+					
Soybean[y]													−						
Brassica[h]			+											+					+
Arabidopsis[i]			+										*						

B

	rpl																		Other
	1	2	3	5	6	13	14	16	18	19	20	21	22	23	24	29	33	35	
Chloroplast Genomes																			
Land Plants																			
Angiosperms[a,b]	-	+	-	-	-	-	+	+	-	-	+	-	+	+	-	-	+	-	
Liverwort[c]	-	+	-	-	-	-	+	+	-	-	+	+	+	+	-	-	+	-	
Red Alga																			
Porphura purpurea[d]	-	+	-	+	+	-	+	+	-	-	+		+		+	+	+	+	
Glaucophyta																			
C. paradoxae[e]	+	+	+	+	+	-	+	+	+	+	-	-	-	+	-	-	-	+	
Mitochondrial Genomes																			
Fungi																			
A. nidulans[f]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
N. crassa[g]																			"85"
S. cerevisiae[g]																			var-1
A. macrogynus[g]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Protists																			
P. infestans[j]	-	+	-	+	+	-	-	-	+	-	-	-	-	-	-	-	-	-	
P. wickerhamii[k]	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	
C. reinhardtii[l]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
P. aurelia[m]	-	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	
T. brucei[n]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Land Plants																			
Liverwort[o]	-	+	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	
Angiosperms																			
Wheat[p]																			
Maize[q]								+											
Rice[r]																			
Oenothera[s]				+				-											
N. tabacum[t]		+o																	
Daucus carota[u]				+															
Petunia hybrida[v]				+				•											
Broad bean[x]																			
Pea[y]																			
Soybean[y]																			
Brassica[h]				+				+											
Arabidopsis[i]		+o																	

rpl14). None of these ribosomal protein genes form clusters similar to operons in bacteria and are dispersed throughout the genome. In *Phytophthora infestans*, the mitochondrial genome encodes at least nine small ribosomal protein genes and four large ribosomal protein genes (Table 1.2) (reviewed in Gray 1992; 1993). Only one (*rps12*) has been detected in the maxicircle DNA of the protist kinetoplast mtDNA (Maslove et al. 1992). In contrast, the completely sequenced mouse and human mitochondrial genomes reveal no genes for ribosomal proteins (Attardi and Schartz 1988). In the mitochondrial genome of liverwort, 16 genes for ribosomal proteins have been found (Oda et al. 1992). They form two clusters: *rps12-rps7*, which is very similar to the *E. coli str* operon and *rps10-rpl12-rps19-rps3-rpl16-rpl5-rps14-rps8-rpl6-rps13-rps11-rps1*, which is similar to the fused *E. coli* S10 and *spc* operons. The *rps2* and *rps4* genes are located separately in the liverwort mitochondrial genome (Oda et al. 1992; Takemura et al. 1992).

Nine genes for ribosomal proteins have been detected to date in the mitochondrial genomes of various flowering plants (Table 1.2) (Bonen and Brown 1993; Grohmann et al. 1993; Yei et al. 1993). These genes are usually clustered together with other ribosomal or non-ribosomal protein genes to form co-transcribed RNAs, which in turn are processed to mono-cistronic mRNAs (Schuster 1993). It has been found that some ribosomal protein genes are present in the mitochondrial genomes of some higher plants, but absent in other ones. This topic is discussed further in Chapter 1.5.4.

Mitochondria from different lineages of eukaryotes not only have variations in mitochondrial genome organization and gene content, but also show differences in the genetic code. In animals, yeasts and protists, UGA codes for tryptophan (Wolstenholme 1992), whereas in chloroplasts and plant mitochondria it specifies a stop codon. In mammalian mitochondrial DNAs, AGA and AGG act as termination codons instead of arginine, but specify serine in some invertebrates. In animal and fungal mitochondria, AUA codes for methionine instead of isoleucine and AAA for asparagine rather than lysine. Also in yeast, CUA encodes threonine, rather than leucine. In addition, rates of synonymous substitutions in protein genes are different among different organellar and nuclear genomes (reviewed in Palmer 1991). The rate for angiosperm mitochondrial genes is the most slow and about 0.5 X 10^{-9} substitutions per site, which is 40-100 times lower than that of mammalian mitochondrial genes, 12 and 4-16 times lower than those of angiosperm and mammalian nuclear genes, respectively (Wolfe et al. 1987). This rate is approximately 3 times slower than that of angiosperm chloroplast genes (Wolfe et al. 1987).

From the above discussion, it can be seen that the mitochondria from animals, fungi, protists and plants encode the large and small rRNA genes, a variable number of tRNA genes and a number of proteins associated with the respiratory chain. Some homologous genes such as ATPase and NADH dehydrogenase components have also been found in the chloroplast genomes. However, these genes are very divergent in sequence from the mitochondrial ones. This suggests that mitochondria share a common ancestor which was different from that of chloroplasts. This also supports the idea that mitochondria and chloroplasts originated from different endosymbiotic bacteria. The different mitochondrial gene contents in the different eukaryotic kingdoms may reflect different gene transfer events during the early or more recent stages of the endosymbiosis.

1.3. Regulation of mitochondrial gene expression.

1.3.1. At the transcriptional level.

Since chloroplasts and mitochondria have a bacterial origin as mentioned earlier, their transcriptional machinery should be similar to the bacterial one. It has been shown that many chloroplast genes form similar operons to the *E. coli* ones (Sugiura 1992). The chloroplast genes have -10 and -35 promoters similar to bacterial promoters and can even be transcribed by the bacterial RNA polymerase *in vitro* (Gruissem 1989).

In animal mitochondria, two major promoters have been located within the displacement loop (D-loop) region and each strand is completely transcribed, resulting in polycistronic RNAs from both strands, which are quickly processed to produce discrete mature tRNA, rRNAs and mRNAs (Clayton 1991). In yeast mitochondria, at least 19 transcription initiation sites are dispersed through the yeast mitochondrial genome (Costanzo and Fox 1990). They have a highly conserved promoter, ATATAAGTA, with the last A being the site of transcription initiation (Costanzo and Fox 1990). However, the RNA polymerases in yeast mitochondria have been found to be similar to those from T_3 and T_7 phage (reviewed in Gray 1992).

It is interesting to know whether the transcriptional units of the flowering plant mitochondrial genomes are similar to those found in chloroplasts and bacteria or those in non-plant mitochondria. Some mitochondrial genes from flowering plants are transcribed as monocistronic mRNAs whereas others are co-transcribed, for example, *nad3* and *rps12* in wheat and maize (Gualberto et al. 1988), *cob* and *rps14* in broad bean (Wahleithner and

Wolstenholme 1988), 5S rRNA and *nad5* gene in *Oenothera* (Wissinger et al. 1988). However, in some cases unlike bacterial ones, they are then processed to mono-cistronic mRNAs. It is interesting to note that some ribosomal protein genes like *rps3/rpl16* in maize (Hunt and Newton 1991) and *rps3/rpl16/rpl5/rps14* in *Brassica* (Yei et al. 1993) have organizations similar to the bacterial and/or chloroplast ones. However, no promoters similar to the bacterial and chloroplast ones have been identified (Hunt and Newton 1991; Yei et al. 1993). Furthermore, the mitochondrial mRNAs are much more stable than the bacterial ones. There are believed to be RNA sequences which bind proteins to prolong the half-life of mRNAs (Nagley 1991).

In plant mitochondrial mRNAs, there often are relatively long leader sequences which separate the transcriptional and translational initiation sites (Lonsdale 1989). Guanylyltransferase capping experiments have been used to identify primary transcripts (reviewed in Gray 1992) and in vitro transcriptional systems have been used to identify conserved nucleotides in certain promoter sequences (Hanic-Joyce and Gray 1991; Rapp and Stern 1992; Rapp et al. 1993). Consensus consequences in maize (a/t**CRTA**g/tAa/tAAa) (Mulligan et al. 1991; Rapp and Stern 1992; Rapp et al. 1993), wheat (rAaannG**CRTA**tArtragt) (Covello and Gray 1991) and soybean (YrAAATnn**CRTA**AGAGAAGAAAG) (Brown et al. 1991) have been identified, but they only share a short **CRTA** motif.

The transcriptional profiles of mitochondrial genes differ from plant to plant. For example, those of the mitochondrial *coxII* gene are simple in wheat (Bonen et al. 1984), but complex in maize (Fox and Leaver 1981). In maize, multiple transcription initiation sites are observed for *atp9* and *coxIII* by capping experiments (Mulligan et al. 1988). There are three transcription initiation sites for *coxIII* and at least six transcription initiation sites for *atp9*, resulting in RNA heterogeneity (Mulligan et al. 1991).

1.3.2. At the post-transcriptional level.

1.3.2.1. RNA splicing in fungal and plant mitochondria.

Some chloroplast and mitochondrial genes contain introns which based on their structural characteristics, are classified as group-I or group-II. Both of them can be folded into distinct secondary structures (reviewed by Burke 1988; Michel et al. 1989) and are different from those in nuclear pre-mRNAs or tRNAs (Green 1991). However, it has been suggested that group-II introns are evolutionarily related to nuclear pre-mRNA spliceosomal introns (Cech 1991; Jacquier 1990). Although some group I and group II introns have been shown to undergo self-splicing (Cech 1986) and reverse-splicing *in vitro* (Augustin et al. 1990), splicing *in vivo* relies on specific proteins encoded by intronic ORFs (maturases) and nuclear gene products (Lambowitz 1990).

The mitochondrial introns found so far in flowering plants are classified as group-II (reviewed in Bonen 1993). In contrast, both group-I and group-II introns were found in the liverwort mitochondrial genes. Furthermore, DNA rearrangements in flowering plants have also resulted in segments of genes being dispersed in the genomes. For example, the wheat mitochondrial *nad1* gene contains five exons, which undergo one *cis-* and three *trans-*splicing events to generate a mature mRNA (Chapdelaine and Bonen 1991). To date, a total of three genes (*nad1*, *nad2* and *nad5*) have been found to be *trans*-spliced in the mitochondria of flowering plants (reviewed in Bonen 1993). It has been predicted that *trans*-splicing introns form secondary structures similar to that of *cis*-splicing introns (Chapdelaine and Bonen 1991; Wissinger et al. 1991). Until now no plant mitochondrial introns have been shown to be self-spliced *in vitro* (Bonen 1993).

1.3.2.2. RNA editing in protist mitochondria.

RNA editing is defined in the broad sense as any process which results in the production of an RNA molecule that differs in nucleotide sequence from the DNA template (Simpson and Shaw 1989). RNA editing was first found in the mitochondrial transcripts of trypanosomes, where it involves the addition and deletion of uridine (U) residues at defined sites mostly within coding regions of mRNAs (Benne 1993). A different type of RNA editing has been also observed in RNAs synthesized in mitochondria of the acellular slime mold *Physarum polycephalum*, where insertional editing is seen. It is generally the addition of a single cytidine, but the other three nucleotides are also occasionally inserted in the mitochondrial transcripts to generate mature rRNAs, tRNAs and mRNAs (reviewed in Miller et al. 1993). In addition, RNA editing has resulted in single-nucleotide conversions in mitochondrial tRNAs of *Acanthamoeba* (U to A, U to G, and A to G) (Lonergan and Gray 1993), potato (C-to-U and occasionally U to C) (Marèchal-Drouard et al. 1993) and marsupials (C to U) (Janke and Pääbo 1993). Moreover, RNA editing has been also observed in chloroplast mRNAs (reviewed in Covello and Gray 1993), mammalian nuclear mRNAs and tRNAs, and virus mRNAs with variable modifications including insertions of G or C (reviewed in Hodges and Scott 1993; Kolakofsky et al. 1993). RNA editing in the chloroplast is very similar to that in the mitochondrion of flowering plants (reviewed in Covello and Gray 1993).

RNA editing in trypanosomes is necessary for the generation of functional mitochondrial mRNAs. The *coxIII*, *nad7* and *atp6* genes of *T. brucei* could not be identified based on DNA sequence homology (Feagin and Stuart 1992). Their RNA transcripts have been extensively edited by the addition and deletion of uridines (Feagin and Stuart 1992). It is interesting to note that RNA editing of transcripts for the same gene is more extensive in some species such as *T. brucei* than others such as *L. tarentolae* and *C. fasciulata* (reviewed in Stuart and Feagin 1992). It has been found that small guide RNAs (gRNAs) (Blum et al. 1990) provide information for the addition and deletion of uridines in trypanosomal mitochondrial RNA precursors. These guide RNAs are encoded by either maxicircle or minicircle mitochondrial DNAs. Based on the sequences of gRNAs, the presence of mitochondrial terminal uridylyl transferase (TUTase) and RNA ligase activities, Blum et al. (1990) proposed a mechanism for RNA editing. First, a gRNA forms a partial hybrid with the pre-edited mRNA sequence 3' to an RNA editing site. An endoribonuclease cleaves the pre-edited mRNA at the point where the duplex RNA between gRNA and the pre-edited mRNA ends. The TUTase then inserts U residues which are guided by the gRNA. The mRNA chain is then ligated by RNA ligase to generate a mature mRNA (Sloof and Benne 1993).

1.3.2.3. RNA editing in plant mitochondria.

RNA editing in plant mitochondria was first observed by RNA and cDNA sequence analysis of the *coxII* and *coxIII* transcripts of several flowering plants (Gualberto et al. 1989; Covello and Gray 1989; Hiesel et al. 1989). Some Cs in the DNA sequences have been changed to Us in the RNA sequences. To date, RNA editing, almost exclusively C-to-U changes, has been observed in all flowering plant and gymnosperm mitochondrial transcripts investigated, except for the cytoplasmic-male-sterile specific chimeric gene *T-urf13* (reviewed in Covello and Gray 1993). A few U-to-C changes have also been observed (Lamattina et al. 1989). On the other hand, the mitochondrial genomic sequence data of liverwort suggest that RNA editing is not required in this moss (Oda et al. 1992).

The changes made by RNA editing usually increase the similarity of the mitochondrial protein sequences among different organisms because they mostly occur in the first and second positions of codons. C-to-U changes can result in the generation of translational initiation codons, as for the *nad1* mRNA where ACG is changed into AUG (Chapdelaine and Bonen 1991). However, in a few cases, the C-to-U changes also result in a decrease of the similarity by altering conserved amino acids into non-conserved ones (Schuster and Brennicke 1991). RNA editing is also observed in the non-coding regions of mitochondrial mRNAs and within introns, but at much lower frequency (Wissinger et al. 1992). Mechanisms are not clear for C-to-U type of RNA editing in plant mitochondria. One proposal is the deamination of a specific cytidine in the RNA chain by a cytidine deaminase activity. Such an activity has been observed for the apolipoprotein B mRNA in the cytoplasm of mammalian tissues and is mediated by a single polypeptide, which recognizes the sequence context (Hodges and

Scott 1993). Cytidine mononucleotide deaminase activities have also been observed in flowering plants, but have not been localized to any compartment yet or tested on polynucleotides (Bonnard et al. 1992). So, their functions in RNA editing are speculative. Furthermore, no consensus sequences have been observed in the vicinity of RNA editing sites of plant mitochondrial transcripts to confer specificity (Covello and Gray 1993; Bonnard et al. 1992). The second possibility is that RNA editing could proceed via excision and replacement of the base. Such reactions have been seen in tRNA modification (Wissinger et al. 1992). The third possibility is that there is a mechanism of nucleotide replacement, which is similar to that found in the trypanosome mitochondria. However, no small guide RNAs have yet been identified in the mitochondrial genomes of flowering plants (Gray and Covello 1993). Furthermore, it was shown recently that RNA editing in plant mitochondria did not result in the breakdown of phosphate bonds, suggesting RNA editing in flowering plants does not delete or add nucleotides (Rajasekhar and Mulligan 1993). This experiment does not support the third possibility, but could be consistent with the deamination of a specific cytidine in the RNA chain or excision and replacement of the base.

1.3.3 Differential expression of mitochondrial genes.

We have relatively little understanding about the differential expression of mitochondrial genes. In yeast mitochondria, differences are seen in the relative strengths of promoters when yeast is grown under conditions of respiratory repression or derepression (Costanzo and Fox 1990). Furthermore, polycistronic transcriptional units seem to be subject to transcriptional attenuation (Costanzo and Fox 1990).

It has been shown in *T. brucei* mitochondria that differential RNA editing in the mitochondrial transcripts is associated with developmental regulation (Stuart and Feagin 1992). Similarly, RNA editing of chloroplast transcripts for *psbF* and *psbL* is differentially regulated in spinach seeds and roots (Bock et al. 1993). It is known that alternative respiratory pathways can be dramatically increased when plants are grown under anaerobic environments (Douce and Neuburger 1989). Also, the *atp6* gene and *nad1*b exon show differences in transcript levels relative to rRNAs in 24-hour germinating embryos and 6-day seedlings (Bonen 1987), and the *rps12* gene shows different transcriptional activities in different maize mitochondrial genotypes (Muise and Hauswirth 1992). Differences in mitochondrial polypeptides synthesized in the organelle have been observed among shoots, cobs, tassels and scutella of maize (Newton and Walbot 1985), among the leaves, roots, and flowers of sugar beet (Lind et al. 1991) and between leaves and developing pollen of *Nicotiana sylvestris* (De Paepe et al. 1993). Variations in the *in vitro* synthesis of mitochondrial proteins have been detected in plants with differing nuclear backgrounds (Spangenberg et al. 1992). Furthermore, when *Pharbitis nil*, a photoperiodic flower, was induced by light, three additional mitochondrially-synthesized polypeptides compared to non-induced tissues were observed (Li and Tan 1990). All these data suggest that genes can be differentially expressed in plant mitochondria.

Cytoplasmic male-sterility provides a well documented example of mitochondrial gene differential expression in flowering plants. The phenotype only appears in developmental stages leading to anther or pollen formation, but not in other plant tissues. In sunflower mitochondria, the *orf522* gene associated with cytoplasmic male-sterility encodes a 15 kDa protein, whose levels can be reduced upon restoration of fertility. Interestingly, this reduction only occurs in the male florets from restored hybrid plants (Monéger et al. 1994). It has been suggested that the product of the nuclear restorer gene acts at the post-transcriptional level to destabilize the *orf522* transcript in a tissue-specific manner (Monéger et al. 1994). In cytoplasmic male-sterile *Petunia*, a chimeric gene, *pcf*, is found to be developmentally regulated (Young and Hanson 1987). More *pcf* RNA transcripts were found in anther than other tissues and its protein was co-localized with ATPA and the nuclear-encoded mitochondrial alternative oxidase in sporogenous tissue, where little COXII protein was found (Conley and Hanson 1994).

Generally, there are few studies on post-transcriptional regulation and RNA processing events of plant mitochondrial genes except for RNA editing and splicing events. Transcripts of all flowering plant mitochondrial genes examined except the cytoplasmic-male-sterile specific chimeric gene *T-urf13* have been shown to be edited (reviewed in Covello and Gray 1993). RNA editing in plant mitochondria is very important in generating translational initiation sites and restoring conserved amino acids of mitochondrial proteins. Nuclear genes seem to destabilize some transcripts of plant mitochondrial genes involved in cytoplasmic male-sterility in the sunflower male florets (Monéger et al. 1994)

1.3.4 At the translational level.

The organellar mRNAs are translated by their own translational machinery, but there is important control by nuclear-encoded proteins. It has been shown that translation of specific mRNAs in yeast mitochondria is controlled by a number of specific nuclear-encoded proteins (reviewed in Costanzo and Fox 1990). Different environments will affect the expression of the nuclear genes and, in turn, influence mitochondrial mRNA translation (Costanzo and Fox 1990). Similarly, chloroplast translational control mechanisms are broadly classified into three groups: gene specific controls; class specific controls and general controls (Hauser et al. 1993). Gene specific controls are related to the specific interaction of non-translated leader sequences of transcripts with specific regulatory proteins (Rochaix et al. 1991). Class specific controls refer to that group of chloroplast mRNAs for functionally related proteins which can be preferentially translated. For example, chloroplast ribosomal protein mRNAs are translated preferentially compared to mRNAs coding for photosynthetic proteins in *Chlamydomonas reinhardtii* under conditions of reduced chloroplast protein synthesis (Liu et al. 1989). General controls are factors applying to translation of many or all chloroplast mRNAs irrespective of class (Hauser et al. 1993).

1.4 Recent transfer of genetic information among intracellular compartments in plants.

One of the most important implications of the endosymbiont theory is the massive gene transfer from the endosymbiont genome into the nucleus in early stages of endosymbiosis (Gray 1989). Although mitochondria and plastids have their own DNAs and code for some polypeptides, the majority of enzymes and structural polypeptides are encoded by nuclear genes and imported from the cytoplasm (Glover and Lindsay 1992; Subramanian 1993). It has been found recently that gene transfer from the endosymbiont genomes into the nucleus seems to be an on-going evolutionary process (reviewed in Gray 1992; Brennicke et al. 1993). Gene transfer from the organelle (mitochondrion or plastid) to the nucleus theoretically involves several steps. These include 1) duplication of a region of organellar genetic material; 2) its export from the organelle and import into the nucleus; 3) integration into the nuclear genome; 4) activation of this nuclear-located copy of the organellar gene under appropriate controls; and acquisition of signal peptide sequences to target the protein back into the mitochondrion and plastid; and 6) inactivation of the organellar gene copy and subsequent deletion (Obar and Green 1985; Gray 1992; Brennicke et al. 1993).

1.4.1. DNA transfer between mitochondria and chloroplasts.

DNA transfer in plant cells can occur from one compartment to another one. One type of DNA transfer is chloroplast DNA movement into mitochondrial genomes. DNA sequences with homology to chloroplast DNAs have been found in the mitochondrial DNAs of various flowering plants (Lonsdale 1989). The first characterized example was from the maize mitochondrial genome, where DNA sequences with homology to a 12 kb chloroplast DNA region containing the 16S rRNA gene and its flanking regions were found (Stern and Lonsdale 1982). If this DNA fragment were intact, it should include coding sequences of the maize chloroplast *orf115*, *rps7*, 3' exon of *rps12* and *ndhB* and $tRNA^{Leu}$ gene (Lonsdale 1989). A chloroplast 758 bp DNA fragment containing the *rps4* gene, transcribed spacer region and the $tRNA^{ser}$ gene, has been found in the *Oenothera* mitochondria (Schuster and Brennicke 1987). The reading frame of the *rps4* gene has been disrupted by several short insertion and deletions (Schuster and Brennicke 1987). The most detailed study is for the rice mitochondrial genome, in which sixteen chloroplast DNA sequences, ranging from 32 bp to 6.8 kb in length, were shown to be dispersed throughout the genome (Nakazono and Hirai

1993). The chloroplast DNA sequences total about 22 kb, which is equal to approximately 6% of the rice mitochondrial genome and to 19% of the chloroplast genome (Nakazono and Hirai 1993). Five chloroplast ribosomal protein genes (*rps4, rps19, rpl14, rpl23* and *rpl2*) have been found in the rice mitochondrial genome, but not the *rps7* and 3'-*rps12* genes as in maize (Nakazono and Hirai 1993). No common sequences or structures, which might enhance the transfer of DNA, were found around the termini of the integrated chloroplast DNA sequences. It has also been shown that the transfer of fragments of chloroplast DNA has occurred at different times, both before and after the divergence of rice and maize (Nakazono and Hirai 1993).

From observations of chloroplast DNA sequences found in the mitochondria, it seems that transfer is random and unlikely to play a major biological role (Newton 1988). However, there is an exception in the case of certain chloroplast tRNA genes, which are present and expressed in flowering plant mitochondria (Joyce and Gray 1989b). The tRNA genes of the flowering plant mitochondria can be classified into two groups: chloroplast-like genes which exhibit 90-100% identity to their chloroplast equivalents, and native mitochondrial genes which show approximately 70-80% identity to their equivalent chloroplast genes (Joyce and Gray 1989b). For example, in rice mitochondria, there are nine functional tRNA genes originating from the chloroplast genome (Nakazono and Hirai 1993). Chloroplast-like tRNA genes are believed to have been transferred from the chloroplast to the mitochondrion through direct DNA transfer, or in some cases through an RNA intermediate (Marèchal-Drouard et al. 1993). On the other hand, no chloroplast-like tRNA genes have been found in the liverwort mitochondrial genome (Oda et al. 1992).

There is no direct evidence that nuclear or mitochondrial DNAs have been transferred into plastid genomes. However, there is some indirect evidence that certain DNAs have been transferred into the chloroplast genome. A possible example is the *tufA* gene encoding a translational elongation factor, EF-Tu in chloroplast of *Coleochaete orbicularis*, a charophycear alga (Baldauf et al. 1990). In most lower plants, chloroplast DNAs have a very highly conserved *tufA* gene whereas the *tufA* gene in *Coleochaete orbicularis* chloroplast is much more divergent, suggesting that it was acquired from a foreign source via lateral gene transfer.

1.4.2. DNA transfer from the chloroplast to the nucleus.

The transfer of plastid DNAs to the nucleus has occurred frequently during plastid evolution. Evidence for recent transfer was first observed by direct hybridization studies with chloroplast DNA probes and later by isolation and sequencing of the chloroplast DNA sequences in the nucleus (Ayliffe and Timmis 1992). To date, systematic studies demonstrating recent transfer of plastid DNA to the nucleus have been carried out in flowering plants such as spinach (Pichersky et al 1991), tomato (Pichersky et al 1991) and tobacco (Ayliffe and Timmis 1992). Five clones from the tobacco nuclear genome were found to contain plastid DNA sequences equal to one-third of the tobacco chloroplast genome and to have high sequence similarity (Ayliffe and Timmis 1992). It has been suggested that some of this tobacco plastid DNA was very recently transferred to the nucleus. Apparently, none of these can be expressed in the nuclear genome because they lack appropriate expression signals (Ayliffe and Timmis 1992).

The transfer of plastid DNA to the nucleus has also occurred at different times during evolution. One example is the *tufA* gene. In green algae, the *tufA* gene is a chloroplast gene. However, no chloroplast *tufA* gene has been found in any land plants, but a single copy or multi-gene family of *tufA* genes is present in the nuclear genomes of land plants (Baldauf and Palmer 1990). It has been proposed that the *tufA* gene transfer occurred during evolution prior to the split leading to land plants (Baldauf et al. 1990).

The presence of the *rpl22* gene in the chloroplast was first identified by DNA sequencing of the tobacco chloroplast genome (Shinozaki et al. 1986) and later found in the chloroplast genomes of all plants examined so far except legumes. Gantt et al. (1991) showed that the functional copy of the *rpl22* gene is located in the nucleus of pea. DNA sequencing has shown that the nuclear *rpl22* gene has acquired two additional domains: an exon encoding a putative N-terminal transit peptide and an intron which separates this first exon from the evolutionarily conserved, chloroplast-derived portion of the gene. This gene structure suggests that the transferred gene may have acquired its DNA sequence for a transit peptide by exon shuffling. Furthermore, phylogenetic studies suggest that the *rpl22* gene was transferred into the nucleus from the chloroplast about 100 million years, preceding its loss from the legume chloroplast lineage (Gantt et al 1991).

Similarly, the gene for ribosomal protein L21 (*rpl21*) is also present in the chloroplasts of lower plants but absent in flowering plants where corresponding genes have been found in the nucleus (Lonsdale 1989; Smooker et al. 1990; Martin et al. 1990). The gene encoding the small subunit of ribulose biphosphate carboxylase (RuBPCase) is present in the chloroplast genomes of non-green algae, however, in all chlorophyll a/b eukaryotes, it is a nuclear gene (Palmer 1991). The chloroplast ribosomal protein genes are present in some algae, but absent in other algae and land plants (Table 1.2), suggesting these genes have been transferred from the chloroplast into the nucleus.

1.4.3. DNA transfer from the nucleus to the mitochondrion.

DNA transfer from the nucleus to the mitochondrion has also been found in plants. A DNA fragment containing part of the nuclear 18S ribosomal RNA gene was found in *Oenothera* mitochondria (Schuster and Brennicke 1987). Upstream of it, there is an ORF homologous to reverse transcriptase (Schuster and Brennicke 1987). Alternatively the transfer could have occurred the DNA level. The nuclear 18S ribosomal RNA sequence is derived from the transcribed region in the nucleus, suggesting that the 18S cytoplasmic rRNA could have been transferred into the mitochondrion where they were reverse-transcribed into DNAs which were subsequently integrated into the mitochondrial genome (Schuster and Brennicke 1987). Alternatively, the transfer could have occurred at the DNA level. The mitochondria of some flowering plants contain small circular or linear plasmid-like molecules (reviewed in Lonsdale 1989). Many of them are homologous to nuclear DNAs, but not mitochondrial DNAs ([illegible]uchi et al. 1991; Goraczniak and Augustyniak 1991). For example, in lupin mitochondria, there are 1.2 kb (K1) and 1.3 kb (K2) minicircular DNAs (Goraczniak and Augustyniak 1991). It was shown by Southern blot analysis that the K2 plasmid is not homologous to the mitochondrial genome, but to the nuclear DNA. Both plasmid-like DNAs are transcribed in the mitochondria. This suggests that the nuclear DNA might be transferred into mitochondria through a plasmid vector.

1.4.4. DNA transfer from the mitochondrion to the nucleus.

An early example of such gene transfer comes from studying the *atp9* gene in fungi. This gene is not present in animal mitochondrial genomes. In the yeast, *S. cerevisiae*, it is a mitochondrial gene (Weslowski et al. 1981). In *N. crassa*, also an ascomycete, however, the *atp9* gene is found in both the mitochondrial and nuclear genomes. It has been shown that the mitochondrial gene is usually silent and that the ATP9 protein synthesized on cytoplasmic ribosomes is transferred into the mitochondria (van den Boogaart et al. 1982). However, a recent experiment has shown that the mitochondrial *atp9* gene in *N. crassa* is expressed in germinating spores (Bittner-Eddy et al. 1994).

Until recently the cytochrome oxidase subunit II (*coxII*) gene was thought to always be a mitochondrial gene. This gene is highly conserved over a broad taxonomic range and is present in the mitochondrial DNA of all 20 plant species (from monocots to dicots) examined with the exception of two species within the genus *Vigna*, *V. unguiculata* (cowpea) and *V. radiata* (mungbean). There the functional *coxII* gene was surprisingly found to be located in the nucleus (Nugent and Palmer 1990). Because the *coxII* gene is present in the mitochondrial DNA of the closely-related common bean, *Phaseolus vulgaris*, it suggests that gene transfer and gene loss can be very recent events (Nugent and Palmer 1991). It has been shown that the nuclear *coxII* gene is derived from an RNA-edited copy, instead of from the mitochondrial genomic copy. This suggests that the *coxII* mRNA is reverse-transcribed and then *coxII* cDNA is integrated into the nuclear genome. In soybean, the *coxII* gene is present in both mitochondrial and nuclear compartments. However, the mitochondrial copy of the *coxII* gene is not transcribed (Covello and Gray 1991). On the other hand, in pea the *coxII* gene is present in both the mitochondrion and nucleus and the nuclear copy is silent (Nugent and Palmer 1991). The cowpea and soybean nuclear *coxII* genes are both linked to DNA sequences for a transit peptide by an intron which separates the first exon and mitochondrial-derived *coxII* DNA sequences (Nugent and Palmer 1990; Covello and Gray 1991).

Similarly, some mitochondrial ribosomal protein genes have recently been transferred into the nucleus. It has been shown that the *rps13* gene is present in the maize, wheat and tobacco mitochondrial genomes, but not in those of pea, bean, soybean (Bland et al. 1986; Bonen 1987). Furthermore, even though the wheat *rps13* gene is very conserved in sequence, it is silent at least from 24 hour germinating embryos to 12-day seedlings (Bonen 1987). In contrast, the *rps14* gene is present in broad bean mitochondria, but is absent from maize mitochondria (Walheithner and Wolstenholme 1988). The gene for the S12 ribosomal protein has been found in the mitochondrial genomes of a number of flowering plants (reviewed in Grohmann et al. 1993). However, in *Oenothera* mitochondria, the *rps12* gene is a pseudogene (Schuster et al. 1990) and the functional *rps12* gene is in the nuclear genome (Grohmann et al. 1992). Unlike the nuclear *coxII* gene, the nuclear *rps12* gene contains no introns separating the mitochondrial *rps12* domain and sequences for a signal peptide (Grohmann et al. 1992). These results indicate that some plant mitochondrial genes have been transferred recently from the mitochondrion to the nucleus.

Mitochondrial DNA transfer into the nucleus can occur very recently within a species. The DNA sequence analysis of a nuclear polyubiquitin (UBQ13) gene from the Columbia ecotype of *Arabidopsis thaliana* showed that the gene contains a mitochondrial DNA insertion in the coding region (Sun and Callis 1993). However, this insert is only present in the

Columbia ecotype of *Arabidopsis thaliana* but not other ecotypes. It suggests that the integration of the insert occurred after speciation of *Arabidopsis* rather than a deletion event having occurred in all ecotypes except Columbia (Sun and Callis 1993) .

From the above discussion, it is clear that DNA transfer can occur among different genetic compartments in eukaryotic cells during evolution. Such events can be very recent, suggesting that the process is on-going during evolution. However, no experiments have been done to estimate DNA transfer or "successful gene transfer" rates from the mitochondrion or chloroplast to the nucleus during plant evolution.

1.5 Objectives.

In this overview, it has been illustrated that plant mitochondrial genome organization is quite different from other eukaryotic organisms in particular with regard to size and gene content. The latter in part reflects that different genes from these organisms have been transferred from the mitochondrion to the nucleus during early and more recent stages of endosymbiosis. It also shows that plant mitochondrial genomes are very large and complex with recombinationally-active repeat elements that can alter gene structure and expression. RNA processing events are also complex. For example, RNA editing has been observed in virtually all mitochondrial protein-coding transcripts. There is also some evidence showing that some mitochondrial gene expression in flowering plants is differentially regulated.

This work initially aimed to study plant mitochondrial differential gene expression during the early stages of plant development, namely in 24 hour germinating embryos and six-day etiolated seedlings. Mitochondrial biogenesis occurs rapidly during this period and differences were seen in the steady state levels of transcripts for certain genes (Bonen 1987; Bonen, unpublished results). This work involved the construction of wheat mitoplast cDNA banks using RNAs from different stages of plant development.

During this process, the methods used in these experiments led to the identification of new protein-coding genes. Some of these were characterized at the DNA sequence and transcriptional level. In addition, cDNAs of these genes were analyzed to determine whether the mitochondrial protein-coding transcripts are edited. This was carried out for the different developmental stages.

Since mitochondrial gene content can be slightly different from plant to plant, these new mitochondrial genes were checked for their subcellular location in different plants. The location of one of these genes, encoding the S7 ribosomal protein, was found to differ among some plant species and this raised questions about the recent transfer of this gene to the nucleus. Southern hybridization experiments indicated the presence of *rps7* homologous sequences in the nucleus of pea and soybean. The legume nuclear genomic DNA, and cDNA libraries were screened in an attempt to characterize the nuclear copies. This work also provides strong background information for future studies on ribosomal protein gene transfer from the mitochondrion to the nucleus.

CHAPTER 2 MATERIALS AND METHODS

2.1 Isolation of plant mitochondria.

Mitochondria were isolated from 24 hour germinating wheat embryos and six day etiolated plant seedlings as described by Pring and Levings (1978) except for the buffers used (Zhuo et al. 1988). The wheat embryos were dissected from dry wheat grains and incubated for 24 hours at room temperature in the dark on wet filter papers containing 1% glucose. For six-day seedlings, plant seeds were planted in wet vermiculite for six days at room temperature in the dark after they had been treated with diluted Javex (1:6) for 10 min and 10 mM HCl for 10 min, each with repeated rinsing with sterile water. The seeds used were from wheat (*Triticum aestivum* var Frederick), rice (*Oryza sativa*, var V20B), pea (*Pisum sativum* var Thomas Laxton) and soybean (*Glycine max* var. Maple Arrow). After harvesting, diced plant tissue was homogenized in cold Buffer I (0.5 M sucrose, 50mM Tris-HCl (pH 7.5), 5mM EDTA, 0.1% BSA, 0.1% mercapto-2-ethanol) for 5 minutes using ice-cold mortar and pestle. The homogenate was filtered through a 20 μm nylon membrane (Institute of Genetics, Fudan University) and then two layers of Miracloth (CalBiochem). The filtrate was centrifuged for 10 minutes at 4°C at 1500 x g to pellet cell debris and repeated once. The supernatant was centrifuged for 25 minutes at 12,000 x g. The pellet was resuspended in Buffer II (0.3M sucrose, 50mM Tris-HCl (pH 7.5), 1mM EDTA, 0.1% BSA) and centrifuged for 10 minutes at 4°C at 1500 x g. The supernatant was centrifuged for 25 minutes at 12,000 x g to pellet the crude mitochondria.

2.2 Isolation of plant mitochondrial DNA.

Mitochondrial DNAs were isolated as previously described by Wilson and Chourey (1984) with some modifications (Zhuo et al. 1988). The crude mitochondria were suspended in Buffer III (0.3M sucrose, 15mM $MgCl_2$, 50mM Tris-HCl (pH7.5). After DNaseI was added to a final concentration of 100 μg/ml, the solution was incubated for 1 hour at 4°C to digest contaminating nuclear DNA. The reaction was terminated by adding 0.5M EDTA (pH 7.5) to a final concentration of 50mM. The mitochondria were then pelleted for 25 minutes at 12,000 x g, resuspended in Buffer IV (50mM Tris-HCl (pH 8.0), 20mM EDTA) and lysed by adding 10% sarkosyl (International Biotechnologies Inc.) and proteinase K (Boehringer Mannheim) to 0.1% and 100 μg/ml, respectively. The lysate was incubated at 65°C for 1 hour, after which 5M KOAc was added to a final concentration of 1.25M. The mixture was incubated on ice for 30 minutes and centrifuged for 10 minutes at 12,000 x g. The mitochondrial DNA was subsequently extracted with phenol equilibrated with TE Buffer (10mM Tris-HCl (pH7.5), 1mM EDTA), then phenol-chloroform-isoamyl alcohol (24:24:1) and finally chloroform-isoamyl alcohol (24:1). After 0.1 volume of 3M NaOAc (pH 5.6) was added to the aqueous phase, the DNAs were precipitated by 2 volumes of ice-cold ethanol at -20°C. The plant mitochondrial DNAs were centrifuged at 4°C for 10 min and washed sequentially with 70% ethanol and 95% ethanol. After drying under vacuum, the DNA pellets were resuspended in a final concentration of 0.5 to 1.0 μg/μl in TE Buffer. Fifty grams of six-day etiolated seedlings usually yielded about 100 μg mitochondrial DNA and ten grams of 24-hour germinating embryos generated about 50 μg mitochondrial DNA.

2.3 Isolation of plant mitochondrial RNA.

Plant mitochondrial RNAs were isolated from wheat, rice, pea and soybean as described by Bonen (1987). The crude mitochondria were resuspended in Buffer V (10mM Tris-HCl (pH 7.5), 50mM KCl, 10mM $MgCl_2$) and lysed by adding an equal volume of Buffer V containing 8% Triton-X 100 (BDH) and three volumes of 2X detergent mix (2% tri-isopropylnaphthalene sulfanate (Sigma), 12% sodium p-aminosalicylate (Sigma), 20mM Tris-HCl (pH 7.4) and 0.1M NaCl). The lysate was mixed well and placed on ice for 10 min with occasional mixing by vortex. After being centrifuged for 5 minutes at 12000 x g, the upper aqueous phase was extracted by phenol and re-extracted twice with phenol. The mitochondrial RNA was precipitated by ethanol.

In several experiments, the crude mitochondria were resuspended in 0.25M sucrose, 5mM NaCl, 10mM $MgCl_2$, 10mM $CaCl_2$, 10mM Tris-HCl (pH8.0) and micrococcal nuclease was added to final concentrations of 250 U/ml or 500 U/ml. The mitochondria were incubated at 20°C for 20 min, 40 min or 60 min and then lyzed and RNAs were isolated as described above.

2.4 Isolation of wheat mitoplast RNA.

Wheat mitoplasts, that is, mitochondria from which the outer membranes have been removed, were isolated as described by Sparace and Moore (1979) with minor modifications. Crude mitochondria from 24 hour wheat germinating embryos or six-day seedlings were resuspended in hypotonic buffer (10 mM K_2PO_4 (pH 7.4), 1 mM EGTA) and placed on ice for 30-40 min. The mitoplast fraction was recovered by centrifugation at 4°C at 12,000 x g for 25 min and resuspended in 8.6% sucrose in 50 mM Tris-HCl (pH 7.4), 1 mM EGTA, 1% BSA prior to being subjected to centrifugation through discontinuous gradients of 22%, 33%, and 47% sucrose. The mitoplast fraction was recovered from the 33%:47% sucrose interface and diluted slowly with 0.2 M sucrose, 50 mM Tris-HCl (pH 7.4), 1 mM EGTA. This fraction was recovered by centrifugation at 12,000 x g for 25 min and resuspended in 0.6 ml 10 mM Tris-HCl (pH 8.0), 50 mM KCl, 10 mM $MgCl_2$ and lysed with 60 µl of 10% sarkosyl. The RNAs were purified by phenol and phenol-chloroform extraction prior to precipitation overnight at 4° by the addition of LiCl to a final concentration of 2M. After being recovered by centrifugation at 12,000 x g for 10 min, RNA was resuspended in TE buffer or stored in ethanol.

Mitoplast and mitochondrial RNAs were analyzed as described by Locker (1979) on 6M urea-1.5% agarose in TBE buffer. Two volumes of Loading buffer (90% deionized formamide, 1mM EGTA, 0.75mg/ml bromophenol blue) were added to 5-10 µg RNA dissolved in TE buffer, heated to 65°C for 10 min and then put on ice immediately. After electrophoresis, the gels were stained in 10mg/ml ethidium bromide and destained in dH_2O, prior to photography.

2.5 Isolation of plant nuclear DNAs.

Total genomic DNAs were isolated from wheat, pea and soybean as previously described by Ausubel et al. (1990) with some modifications. The etiolated plant tissue (usually about 100 g) was harvested, washed and diced into small pieces. The tissue was frozen with liquid nitrogen and ground to a fine powder. The frozen powder was transferred to a centrifuge bottle and 10ml DNA Extraction Buffer (0.1M Tris-HCl (pH8.0), 0.1M EDTA, 0.2M NaCl, 100μg/ml proteinase K) per g fresh plant tissue were immediately added. After 10% sarkosyl was added to a final concentration of 1%, the solution was incubated at 55°C for 2 hours to lyse the plant cells. The lysate was centrifuged for 10 minutes at 5500 x g at 4°C to pellet cell debris. 0.6 volumes of isopropanol were added to the supernatant and incubated at -20°C for 30 min to precipitate nucleic acids. The nucleic acids were pelleted for 15 minutes at 7500 x g. After the pellet was resuspended in TE buffer, then 5M NaCl and CTAB/NaCl solution (10% CTAB (cetyltrimethylammonium bromide in 0.7M NaCl) were added to final concentrations of 0.6M and 1%, respectively. The solution was incubated for 20 minutes at 65°C and extracted by an equal volume of chloroform/isoamyl alcohol (24:1) to remove polysaccharides. The nucleic acids were re-precipitated by ethanol. The pellet was resuspended in TE buffer and incubated at 55°C to aid re-suspension. After 9.7 g of solid CsCl was added and mixed gently until dissolved, 0.5ml of 10 mg/ml ethidium bromide was added to the CsCl solution, which was transferred to a 30 ml quick-seal ultracentrifuge tube and centrifuged for 24 hour in a Beckman VTi50 rotor at 50,000 rpm (300,000 x g) at 20°C. After the DNA was collected, the ethidium bromide was removed by repeatedly extracting the collected DNA with isopropanol which had been equilibrated with CsCl-saturated aqueous phase. The DNA was precipitated by adding two volumes of water and six volumes of ethanol. The DNA pellet was resuspended in low TE buffer and re-precipitated by adding one-tenth volume of 3M sodium acetate (pH5.6) and two volumes of ethanol. The DNA was collected by centrifugation and resuspended in TE buffer.

In some cases, the nuclear DNAs were isolated from the crude nuclei lacking the mitochondrial fraction. After harvesting, diced plant tissue was homogenized in cold Buffer I (0.5 M sucrose, 50mM Tris-HCl (pH 7.5), 5mM EDTA, 0.1% BSA, 0.1% mercapto-2-ethanol) for 5 minutes using ice-cold mortar and pestle. The homogenate was filtered through a 20 μm nylon membrane (Institute of Genetics, Fudan University) and then two layers of Miracloth (CalBiochem). The filtrate was centrifuged for 10 minutes at 4°C at 1000 x g to pellet nuclei. After the nuclei were resuspended in DNA Extraction Buffer, they were lyzed by adding 10% sarkosyl to a final concentration of 1% and incubated at 55°C for 1 hour. The nuclear DNAs were purified by phenol, phenol-chloroform and chloroform extractions. The DNAs were precipitated by ethanol and resuspended in TE buffer.

2.6 Construction of wheat mitochondrial cDNA banks

Poly(A) tails were added to 5 to 10 μg mitoplast RNA by RNA poly(A) polymerase (Pharmacia) as described by Sippel (1973). The reaction was carried out in 25 μl of 50mM Tris-HCl (pH 7.9), 0.25M NaCl, 10mM $MgCl_2$, 2.5mM $MnCl_2$ at 37°C for 2 hours. Poly(A) tailed mtRNAs were extracted with phenol and precipitated by ethanol. cDNA synthesis was carried out according to Gubler and Hoffman (1983) with some modifications (Rutledge et al 1988). A solution containing 50mM Tris-HCl (pH8.3), 10mM $MgCl_2$, 10mM DTT, 1.25mM each of four dNTPs, 100μg/ml oligo$(dT)_{15}$ (Pharmacia), 400μg/ml random hexadeoxynucleotides (Pharmacia) and 400μg/ml poly$(A)^+$ mRNA, was heated to 65°C for 2 min and then left for 30 min at room temperature. First-strand cDNA synthesis was carried out by adding AMV reverse transcriptase (Life Sciences) at a concentration of 800 units/ml in the same buffer for 1 hour at 43°C. After the volume of the reaction mixture was adjusted with 50mM Tris-HCl (pH7.8), 10mM $MgCl_2$, 200mM DTT, 1mM ATP and 50µg/ml BSA, the cDNA/RNA hybrids were incubated at 15°C for 1 hour with 4 units of T4-DNA ligase. DNA polymerase I (20 units) (GIBCO-BRL) and RNase H (2 units) (Pharmacia) were added to carry out second strand cDNA synthesis for 3 hours at 15°C. After incubation, the samples were extracted with phenol and passed through a 1 ml Sephadex G-50 (Pharmacia) column. 10µl of the above mixture in which dATP was replaced by 1 µl of α- ^{32}P-dATP was used to trace the cDNA size. The eluate was fractionated on a Sepharose-4B (Bio-Rad) column equilibrated with 10mM Tris-HCl (pH 7.5), 1mM EDTA and 10mM NaCl, cDNAs larger than 0.5 kb were ligated into blunt pUC13 vectors for transformation of *E. coli* strain DH5α.

2.7 Methodology to separate mitochondrial mRNAs from mitochondrial rRNAs.

It has been proposed (Sippel 1973) that poly(A) polymerase adds poly(A) tails only to single-stranded RNAs, instead of RNA-DNA duplexes. Therefore, if mitochondrial rRNAs were to form RNA-DNA duplexes with cloned rDNAs, after the polyadenylation step mitochondrial mRNAs should be able to be separated from unpolyadenylated rRNAs by using mAP paper (Amersham), which contains oligo$(dT)_n$. Ten µg mitochondrial RNA were added to 10 - 30 µg M13 single-stranded DNAs containing the 3'-region sequences complementary to wheat mitochondrial rRNAs. The wheat mitochondrial RNAs and M13 DNAs were annealed at 5 X SSC, 50% formamide at 42°C overnight (about 12 to 16 hrs.). The annealed RNA:DNA mixes were precipitated by three volumes of 95% ethanol and pelleted after 8 hours at -20°C. After being resuspended in dH_20, the samples were polyadenylated by poly(A) polymerase (Pharmacia) in 50mM Tris-HCl (pH 7.9), 0.25M NaCl, 10mM $MgCl_2$, 2.5mM $MnCl_2$ and α-^{32}P-ATP at 37°C for 2 hours and then they were transferred onto mAP paper (Amersham) that had been soaked in TE buffer. Then the mAP paper was washed in 0.5 M NaCl for 5 min and blotted on a filter paper. The above wash cycle was repeated twice. The mAP paper was rinsed in 70% ethanol for 2 min and blotted on a filter paper to remove extra ethanol. The poly(A)-containing RNAs were released in dH_2O by heating to 70°C for 5 min and precipitated in three volumes of 95% ethanol.

2.8 Cloning and isolation of recombinant plasmid or phage DNA.

DNAs to be cloned or subcloned were digested and electrophoresed on 0.8 - 1.2% agarose gels in TBE or TEA for 1 to 10 hours (Sambrook et al. 1989). DNA fragments were isolated from the agarose gels by GeneClean II (Bio/Can) as the manufacturer suggested. DNA fragments (0.05 to 1.0 μg) were ligated into M13 (4.0 ng) or pUC vectors (5 ng) by 1.0 unit of T4 DNA ligase in 20mM Tris-HCl (pH 7.6), 10mM $MgCl_2$, 10mM dithiothreitol (DTT), 1mM ATP at 14°C overnight.

Competent cells of *E. coli* strain JM101 for M13 cloning or TB1 for pUC plasmid cloning were prepared as described by Sambrook et al (1989) and plasmid DNA was isolated by the boiling method as described by Sambrook et al (1989) with an additional PEG purification step. After the plasmid DNA was pelleted and resuspended in 50μl TE, 50μl of 20% PEG6000 (Sigma) in 0.5M NaCl were added and left at 4°C overnight. The plasmid DNA was pelleted and resuspended in 50μl TE buffer.

Two legume nuclear DNA libraries and one cDNA library were obtained from Clontech. A pea nuclear library (*Pisum sativum* var. Alaska), had been made by cloning MboI partially digested pea genomic DNA into lambda EMBL3 vectors. The inserts ranged from 8 to 21 kb with an average size of 16 kb. A pea cDNA library (*Pisum sativum* var. Alaska) was constructed by synthesizing cDNA using oligo(dT) and random primers and cloning into lambda gt11. The soybean genomic library (*Glycine max* var Resnik) had been made by cloning soybean genomic DNA, which was partially Sau3A-digested, into lambda EMBL3 SP6/T7 vectors. The insert sizes were about 15 kb. The libraries had been amplified at least once by Clontech.

E. coli host strains (NM538, Y1090r$^-$,or K802) used with these libraries were maintained on magnesium-free LB agar plates containing the appropriate media additives (see Appendix 1) as described by Sambrook et al (1989).

They were diluted to $2.5X10^{-5}$ in 1X Lambda Dilution Buffer (35mM Tris-HCl (pH 7.5), 10mM $MgSO_4$, 0.1M NaCl, 0.2% gelatin). Two µl, 5 µl, 10 µl and 0 µl of the above dilution were added to 100 µl of 1X Lambda Dilution Buffer to infect 200 µl of the overnight bacterial culture obtained as above, incubated in a 37°C water bath for 15 minutes and then plated in 0.75% top agarose. The resulting plaques were counted to determine the titer (pfu/ml) based on the following equation:

$$\text{pfu/ml} = \text{\# of phage/}\mu\text{l used X dilution factor X } 10^3\ \mu\text{l/ml}$$

The titers (pfu/ml) of the Clontech libraries were determined to be between 1×10^9 and 1×10^{10}. For example the pea nuclear DNA library was found to be approximately 8×10^9 pfu/ml using *E. coli* strain NM538 (Appendix A). When the host *E. coli* strain NM539

(Appendix A) was used, smaller plaques were seen with a slightly lower titer. The titer of the pea cDNA library was determined to be 1 X 10^{10} pfu/ml when *E. coli* host strain Y1090r$^-$ (Appendix A) was used. The soybean DNA library was titered using suppressing *E. coli* strain K802 (Appendix A) (Murray et al. 1977) and was determined to be 1 X 10^{10} pfu/ml.

Recombinant lambda DNA was isolated as described by Sambrook et al (1989). A single plaque was picked and re-suspended in 5 ml of LB Broth. Then, 50 µl of an overnight culture of host cells were added to this phage suspension and shaken at 37°C in a water bath for 4-6 hours to allow host cell lysis, as seen when the culture became clear with some solid precipitation. If the host cells were not lysed, fresh LB broth was added in a ratio of 1:3, and the culture incubated for another hour. A few drops of chloroform were added to the lysis solution and then centrifuged at 8,000 rpm for 20 minutes at 4°C to pellet cell debris. 2 ml of supernatant were transferred to a new tube and incubated with 2.5 mg/ml DNaseI at 37 °C for 60 minutes. 300 µl of 5M NaCl and 0.6 ml of 40% of PEG were added, mixed and incubated on ice for one hour. After the phage were pelleted at 5,000 rpm for 10 minutes at 4 °C, they were resuspended in Lambda Dilution Buffer. The phage solutions were chloroform-extracted twice. After 3 µl of 0.5M EDTA, 6 µl of 10% SDS, and proteinase K to 100 µg/ml were added, the solution was incubated at 65°C for 20 min and then extracted sequentially with phenol, phenol/chloroform and chloroform. The lambda DNA was precipitated by ethanol. The pellet was resuspended in 100 µl of low TE (10mM Tris-HCl (pH7.5); 0.1mM EDTA) containing 5 µg/ml RNase A for RNA digestion.

2.9 Polymerase chain reaction (PCR) and reverse-transcriptase PCR amplification.

Polymerase chain reaction (PCR) was used as described by Saiki (1990) and Kawasaki (1990) to amplify different DNAs. PCR and RT-PCR amplifications were carried out in 50mM KCl, 10mM Tris-HCl (pH9.0), 1.5mM $MgCl_2$, 0.1% Triton-100, 0.2 mM for each of the dNTPs, 2 μg/ml each of both oligomers, and 1.25 Units of Taq DNA polymerase (Bio/Can) as well as DNA template. The program for the DNA thermal cycler (Perkin-Elmer Cetus) used was (Segment 1: 94°C, 1 min; Segment 2: 53-58°C, depending on the G-C content of the oligomers used, 1.5 min; Segment 3: 72°C, 2.5 min; cycles: 25; incubation 72°C, 7.0 min). The oligomers shown below are numbered according to the wheat *rps7* gene in Fig 3.7:

LB27: ATCAATTTATCGGCCTCGTC (-41 to -61);

LB28: GCAGGCCTCTTGTGGATTCC (586 to 606)

LB29: ACTGAATGAGGAAGAGCTCC (461 to 480);

LB35: GCATGATCAATGGTAAAAGAAC (55 to 75);

LB36: GCGATAGGAGCCTATGTG (-387 to -369);

LB37: GTTCAGTTCGAGCTAGGCGGTG (107 to 128);

LB38: GCGCGAACTCCGATTGGCG (409 to 428).

The synthetic oligomers were obtained through either the University of Ottawa or Queen's University.

For RT-PCR amplification, the single-stranded cDNAs from mitochondrial RNA templates were made as described in Chapter 2.6 using appropriate synthetic oligomers and the products were passed through a Sephadex G-50 column prior to PCR amplification. For

PCR-amplification of cDNA inserts from lambda phage, the desired number of plaques was picked from plates with micropipette tips. Each plaque was placed into 250 ul of distilled water and briefly vortexed. Such plaques were either frozen-thawed twice on dry ice for 5 min to release DNA, or frozen at -70°C for later analysis. Lambda gt11 5' primer (5'GAC TCC TGG AGC CCG 3') and lambda gt11 3' primer (5'GGT AGC GAC CGG CGC 3'), from Clontech, were used to amplify inserts under conditions as described above. The oligomers LB8: GTTTTCCCAGTCACGAC (universal sequence primer) and LB9: AACAGCTATGACCATG (universal reverse sequence primer) (New England Biolabs) were used to amplify M13 and pUC inserts.

2.10 Generation of ^{32}P-radiolabelled probes.

2.10.1 By PCR amplification.

The method was similar to the PCR amplification described above except for the concentration of the dNTP mix. 1 X Taq polymerase Buffer, 2000ng/ml each of both oligomers, 0.01 mM dATP, 0.11 mM for each of dCTP, dGTP and dTTP, 1.25 Units of Taq DNA polymerase and 50 ng/ml DNA template, were pooled, and then divided into two tubes. 1.5 µl of α-^{32}PdATP (Amersham) (3000Ci/ml) were added into one tube and the second tube was used as a non-radioactive control to check PCR products on agarose gels.

2.10.2. By random oligomers.

About 0.1 - 0.25 µg plasmid DNA was restricted, boiled for 3 min and quick-chilled.

The random hexamer oligomers (Pharmacia) and Klenow DNA polymerase (Pharmacia) were added to final concentrations of 7.5 μg/ml and 60-80 U/ml. Then 2.0 μl of α-^{32}PdATP (3000Ci/ml) was added to the reaction mix. The labelling reaction was carried out in 10 mM Tris-HCl (pH 7.0), 10 mM $MgCl_2$, 0.4 mg/ml BSA, 24 μM each of dGTP, dCTP, and dTTP at room temperature for 3-4 hours. The sample was passed through a Sephadex G-50 column to remove unincorporated radiolabel.

2.10.3 By M13 second strand synthesis.

Approximately 0.3 μg M13 DNA, 5 ng/ml universal sequencing primer and 1.5 μl of α-^{32}PdATP (3000Ci/ml) were pooled in 10 mM Tris-HCl (pH7.0), 10 mM Mg_2Cl and 10 mM NaCl. The mix was incubated at 65°C for 3 min and cooled to room temperature over 30 min. Klenow polymerase and each of dGTP, dCTP, and dTTP were added to final concentrations of 120 U/ml and 17 μM, respectively. The polymerization reaction was carried out in 10 mM Tris-HCl (pH 7.0), 10 mM $MgCl_2$ and 10 mM NaCl at room temperature for 3-4 hours. The sample was passed through a Sephadex G-50 column.

2.11. Southern and Northern hybridization analyses.

Southern and Northern blot analyses were performed using standard procedures (Sambrook et al. 1989). To prepare Southern blots after restricted DNAs were run on agarose gels with HindIII-digested Lambda DNA, or HaeIII-digested ϕX174 DNA, or in some cases both as size markers, the DNAs were partially hydrolyzed in 10 mM HCl for 30 min when high-molecular weight plant DNAs were used prior to denaturation and transfer by capillary action onto nylon membranes (BIOTRANS, ICN) overnight. After being UV cross-linked at room temperature for six minutes, the filters were incubated in Pre-hybridization Buffer (50% formamide, 5X SSC, 5X Denhardt's solution, 0.1% SDS, 100 μg/ml denatured calf thymus DNA) overnight and the filters were hybridized with probes overnight. When heterologous probes were used, hybridizations were conducted in 5 x SSC, 50% formamide at 38°C rather than 42°C for 24 to 48 hours. After hybridization, the filters were washed twice with 2X SSC, 0.1% SDS at room temperature and at 38°C or 42°C. These relatively low-stringent wash conditions were used because the backgound levels were adequately low.

RNAs were electrophoresed through 1.25% agarose, 7M formaldehyde as described by Sambrook et al (1989). 20 μg of RNAs were dissolved in 0.01M $NaPO_4$ (pH 7.5), 6.5% formaldehyde, 50% formamide and RNA Loading Buffer (50% glycerol, 1mM EDTA and 1mM bromophenol blue) was added to a final concentration of 10%. The RNAs were denatured at 60°C for 5 min and then run on agarose gels in 0.01M $NaPO_4$ (pH7.0) solution with *E. coli* rRNAs as size markers. In some cases, the wheat mitochondrial *orf25* hybridizing signal (0.8 kb) (Bonen et al. unpublished data) was also used as a size marker. The RNAs were transferred to nylon membranes overnight at room temperature and UV-cross-linked for six-minutes. Conditions for Northern blot hybridization were similar to those of Southern blot analysis.

2.12 Screening mitochondrial and nuclear genomic or cDNA libraries

Mitochondrial genomic and cDNA libraries were screened by colony hybridization as described by Sambrook et al (1989) with modifications. To prepare the colony hybridization filters, gridded nylon membranes were placed on YT plates containing 50 μg/ml of ampicillin (Sigma), 50 μg/ml of X-gal (GIBCO BRL) and 20 mg/ml of IPTG (GIBCO). White pUC recombinant clones were first patched onto a master plate, then onto filters, using sterile toothpicks. The plates and filters were incubated at 37°C overnight. Filters were floated upright sequentially on 0.5M NaOH, 1.5M NaCl for 7 min and on 1.5M Tris (pH7.5) for 3 min two times and on 1.5M Tris-HCl (pH7.5), 1.5M NaCl for 4 min with blotting between each step, and then air-dried. After being washed with 95% ethanol, the filters were air-dried and then UV cross-linked. The hybridization conditions were as described in Chapter 2.10.

The plant nuclear genomic and cDNA libraries were screened by plaque hybridization as described by Sambrook et al. (1989) with some modifications. Based on the titer obtained previously, dilution was made in 0.1 ml sterile 1X Lambda Dilution Buffer (35 mM Tris-HCl (pH 7.5), 10 mM $MgSO_4$, 0.1 M NaCl, 0.2% gelatin) to yield approximately 50,000 to 60,000 pfu per 150 mm plate. The diluted phage were combined with 600 μl of an overnight bacterial culture, and incubated at 37 °C for 15 to 20 minutes. Seven ml of melted 0.75% LB top agarose were immediately added to the cell suspension. Then it was mixed and poured onto an LB agar plate pre-warmed to 37 °C. The plates were left at room temperature for at least 15 minutes, inverted and incubated at 37 °C for four to twelve hours. The plates were chilled at 4 °C for at least 1 hour to allow the top agarose to harden before preparing plaque lifts. The Biotrans (ICN Biochemical) nylon membranes were placed on the plate for 2 minutes and then floated plaque side up for 5 minutes on 10 ml of the DNA Denaturing Solution (0.5M NaOH, 1.5M NaCl). The filters were briefly dried on paper towels and floated on 10 ml of the Neutralizing Solution (1.5M NaCl, 0.5M Tris-HCl (pH 8.0). The neutralization was repeated once. The filters were briefly rinsed in 3 X SSC and placed on Whatman paper to air dry. The DNA was fixed to the filters by a UV cross-linking at room temperature for 6 minutes.

Plaque hybridization screening of lambda phage libraries was carried out as described by Sambrook et al. (1989) with some modifications. Hybridization probes were amplified by polymerase chain reaction (PCR) in the presence of α-^{32}PdATP and α-^{32}PdCTP (Dupont-NEN). The filters were incubated in Pre-hybridization Solution with slow agitation at 39°C overnight. The Pre-hybridization Buffer was replaced with Hybridization Solution. The DNA probe was denatured at 100 °C for 3 minutes and quickly chilled on ice. The DNA probes were added to the Hybridization Solution and incubated with slow agitation at 39°C overnight. After hybridization, the filters were washed in 2 X SSC, 0.1% SDS at room temperature for one hour, and then at 65 °C for another hour.

Alternatively, PCR products were gel-purified by GeneClean II(Bio/Can) prior to radiolabelling. The DNA probes were radiolabelled using the random oligomer method in the presence of α-^{32}PdATP and α-^{32}PdCTP (Dupont-NEN). After pre-hybridization, the probe was added and hybridized at 37 °C overnight. After hybridization, the filters were washed twice at room temperature for 15 minutes and twice at 37 °C for 15 min.

2.13 S1 nuclease mapping analysis.

DNA regions to be used in S1 nuclease protection assays were generated by PCR amplification of M13 inserts, using forward and reverse sequencing primers with Taq DNA polymerase (Bio/Can) in the presence of α-^{32}P-dATP as described in Chapter 2.9. About 25 μg of wheat mtRNA was allowed to anneal with 0.15 to 0.3 μg of denatured, uniformly radioactively-labelled DNA in the solution containing 0.04M PIPES (pH7.5), 1mM EDTA, 0.4M NaCl and 80% deionized formamide at 48-50°C overnight after incubated at 85°C for 15 min (Sambrook et al 1989). The solution was cooled to 45°C over a period of 30 to 45 min. S1 nuclease digestion was carried out in 1x S1 Buffer (0.25M NaCl, 60mM sodium acetate (pH 4.5), 1mM $ZnSO_4$) at final concentrations of 10-500 units S1 nuclease (Boehringer or GIBCO-BRL) for 30 min at 30°C. The S1 nuclease digestion was stopped by adding 25% Stop Solution (4M ammonium acetate, 50mM EDTA, 50μg/ml yeast tRNA) and 2.5 volumes of 95% ethanol. The protected fragments were resolved on 7% polyacrylamide sequencing gels.

2.14 DNA sequence analysis.

DNAs of interest were subcloned into appropriate M13 vectors and sequenced using the dideoxynucleotide chain termination method (Sanger et al. 1977) with α-^{35}S-dATP (Dupont-NEN) and Sequenase (United States Biochemical) as described by Tabor and Richardson (1987). Sequencing products were separated on 6-7% polyacrylamide sequencing gels. Both strands were completely determined for all sequences presented. Data were analyzed using Microgenie software (Beckman) in conjunction with GeneBank and NBRF Protein Data banks and corrected by visual inspections. The wheat mitochondrial *rps7* gene sequence has been entered in the EMBL, GenBank and DDBJ Nucleotide Sequence Database under the accession number X672042.

CHAPTER 3 CONSTRUCTION OF WHEAT MITOPLAST cDNA CLONE BANKS AND IDENTIFICATION OF THE S7 RIBOSOMAL PROTEIN (*rps7*) GENE

3.1. Isolation of wheat mitoplast RNAs.

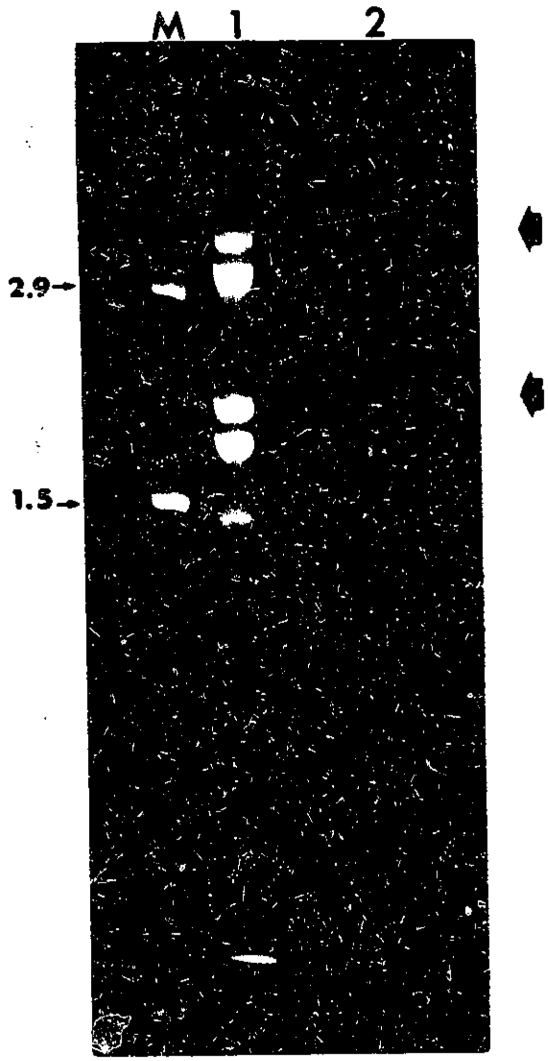
When constructing plant mitochondrial cDNA libraries, high-quality plant mitochondrial mRNAs are needed. It has been shown that animal mitochondrial mRNAs have poly(A) tails of approximately 50 nucleotides (Ojala et al. 1981), so can be separated from the very abundant mitochondrial rRNAs by using oligo$(dT)_{12\text{-}15}$ methods. However, plant mitochondrial mRNAs lack poly(A) tails. Moreover, when plant mitochondria are prepared, cytosolic ribosomes bound on their outer membranes are co-purified (Thompson et al. 1983). Because of this, extra steps of sucrose gradient purification have been used to purify plant mitochondria before lysis (Stern and Newton 1986). Even with this method, the amount of cytoplasmic rRNAs generally varies from 20% to 50% of the total RNA (Stern and Newton 1986). Together the mitochondrial and cytoplasmic rRNAs constitute more than 90% of the total mitochondrial RNA preparations (Stern and Newton 1986). This makes it very difficult to construct plant mitochondrial cDNA libraries enriched for protein-coding transcripts. The only mitochondrial cDNA bank reported in the literature is one for *Oenothera* (Wissinger et ai. 1991). In these experiments, *Oenothera* mitochondrial RNA isolated from cell culture rather than plant tissue, was used because the cytoplasmic and mitochondrial rRNAs are not so abundant based on the fluorescent profile of the total mitochondrial RNAs (Wissinger et al. 1991).

In order to increase the proportion of mRNAs in wheat mt-RNA preparations, several methods were tested to remove both cytoplasmic and mitochondrial rRNAs. To remove cytoplasmic rRNAs: i) micrococcal nuclease treatment of mitochondria was used prior to lysis as has been used for animal mitochondria (Crews and Attardi 1980) and ii) mitoplasts were isolated by disrupting the outer membranes of the mitochondria using digitonin (Ragan et al, 1987) or hypotonic buffer (Sparace and Moore 1979; Bowles et al 1976). In the first method, mitochondria were incubated with 10-500 U/ml of micrococcal nuclease, but no evident differences were found in levels of cytoplasmic rRNAs as monitored by urea-agarose gels (data not shown). The lack of success might reflect differences between the plant and animal mitochondrial systems.

The removal of mitochondrial outer membranes to give mitoplasts was also tested as a way to reduce the amount of the cytosolic rRNAs. Mitoplast preparations have been widely used in studies of mitochondrial metabolism and membrane proteins as well as other mitochondrial membrane components (Sparace and Moore 1979; Bowles et al. 1976; Ragan et al. 1987). Two methods have commonly been used to isolate mitoplasts. One was to use digitonin in amounts equal to that of mitochondrial proteins (Ragan et al. 1987). This method was tried unsuccessfully, perhaps because of an imbalance in the ratio of amounts of digitonin and mitochondrial proteins in my experiments (data not shown). Secondly, the outer membranes of the mitochondria were broken by using hypotonic buffer (Sparace and Moore 1979; Bowles et al. 1976) and the resulting mitoplasts were purified on discontinuous sucrose gradients. Fig 3.1 shows the comparison between a wheat mitochondrial RNA preparation

Figure 3.1 Analysis of RNAs from wheat mitochondria and wheat mitoplasts.

The wheat mitochondrial RNA isolated from sucrose-gradient purification (lane 1) and mitoplast RNA (lane 2) are separated on 1.5% agarose-6M urea gel. (M) shows E. coli rRNAs as size markers. The mitochondrial rRNAs are indicated by arrows.


M
1
2
2.9
1.5

(lane 1) and mitoplast RNA preparation (lane 2). Even though in this figure the amounts of rRNAs loaded differ greatly, the differences in relative amounts of mitochondrial and cytoplasmic rRNAs are obvious. The arrows indicate the mitochondrial rRNAs, which because of secondary structural features run slower than their cytosolic counterparts on these gels (Locker 1979). It can be seen that this method efficiently removes at least 90% of the contaminating cytoplasmic rRNAs based on visual inspection of the fluorescence profiles of several experiments.

3.2. Attempt to separate mitochondrial mRNAs and rRNAs.

In an effort to obtain pure mitochondrial mRNAs, mitochondrial rRNAs in total mitochondrial RNA preparations were allowed to anneal with M13 cloned DNAs complementary to the 3'- region of mitochondrial 18S or 26S rRNA sequences prior to the RNAs being polyadenylated by poly(A) polymerase. This was done under the assumption that rRNA-DNA hybrids might not be available for polyadenylation (Sippel 1973). This was monitored by the incorporation of radiolabel in the presence or absence of cloned M13 single-stranded rDNAs. Polyadenylated RNAs were purified by mAP paper (Lizardi 1983; Werner et al. 1984) and used as probes to hybridize with dot blots of M13 clones complementary to the 3' rRNA regions. Dot hybridization showed that the addition of the cloned single-stranded rDNAs did not reduce the amount of mitochondrial rRNAs (data not shown). In contrast, the latter seemed to show slightly increased hybridization signals in the case of the 18S rRNA (data not shown). The experiments were repeated three times. When the amount of cloned single-stranded rDNAs was doubled, the results were still negative. One possibility may be that the rRNAs were more degraded with more 3' termini available in the presence of the cloned single-stranded rDNAs. It is also possible that RNA:DNA hybrids increased poly(A) tail addition by RNA polymerase because of easier accessibility of the 3'-hydroxyl group of the mitochondrial rRNAs.

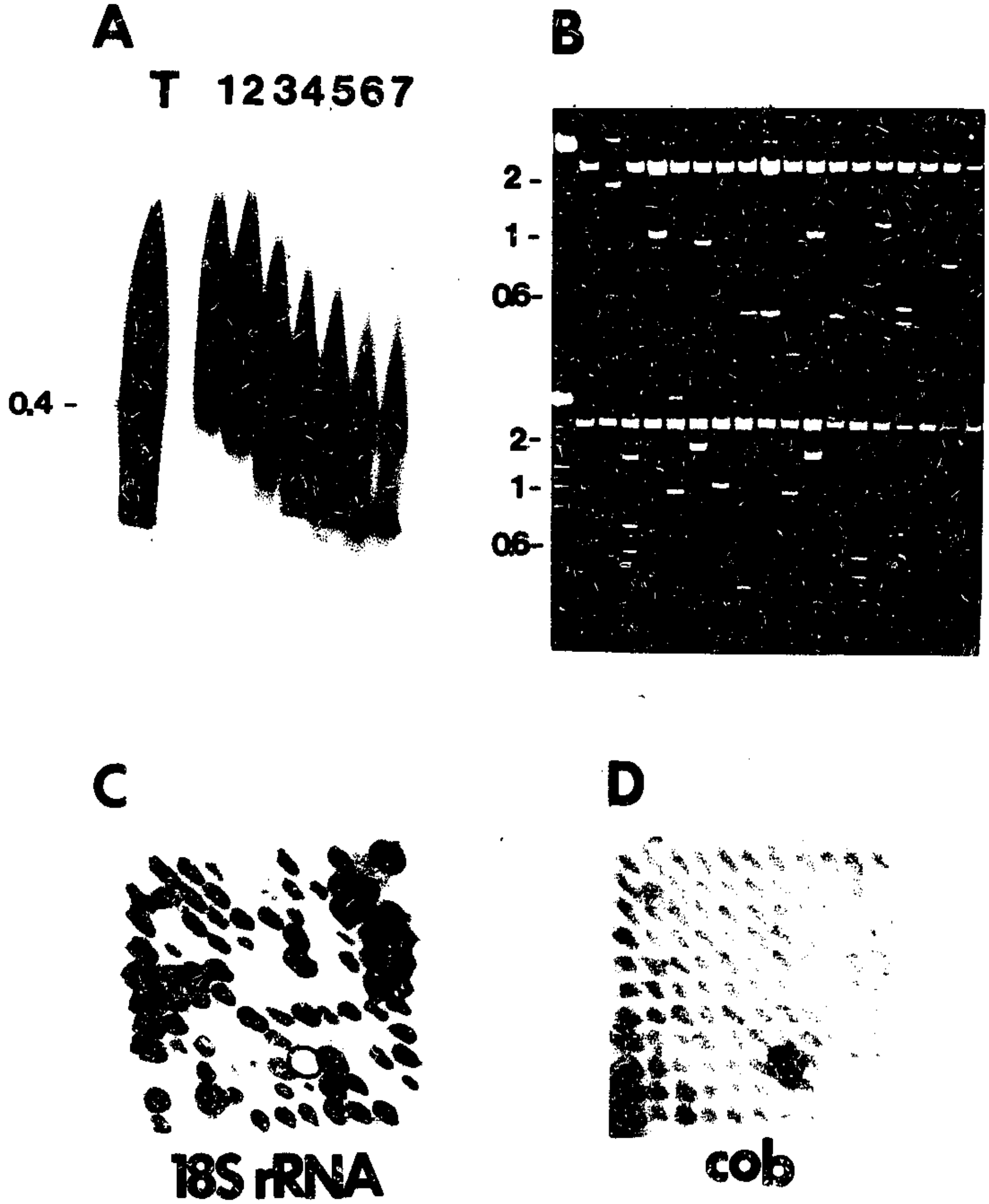
Another method that might be effective could be to generate biotinylated antisense rRNAs in the presence of 5'-(3-amino)allyl UTP(U*) as has been used to study pre-mRNA splicing (Ruby and Abelson 1988). The biotinylated rRNAs would be added to the mitoplast RNA to anneal with the mitochondrial rRNAs. Once biotinylated agarose and avidin are added to the solution, the biotinylated RNA:RNA would be subsequentially removed with the mitoplast mRNAs being left in the solution.

3.3 Construction of wheat mitoplast cDNA clone banks.

cDNA clone banks were constructed from mitoplast RNAs using pUC plasmids as vectors. In order to clone only larger cDNAs, the cDNAs were separated by Sepharose-4B columns (Fig 3.2A). In a pilot experiment, about 0.3 µg cDNA which was made from 2.5 µg mpRNA yielded about 360 cDNA clones. In this experiment, the transformation efficiency was very low. The reasons may be that the ratio of blunt vectors to cDNAs was inappropriate or that a high proportion of the cDNAs were not blunt and could not be ligated into the vector. The cDNA clone bank was first screened by colony hybridization using wheat mitochondrial rRNA gene clones and cytoplasmic rRNA probes and then using 12 known wheat mitochondrial protein-coding gene probes: *rps12*, *rps13*, *nad1*b/c, *nad3*, *nad5*a/b, *coxI*, *coxII*, *cob*, *orf25*, *atpa*, *atp6* and *atp9* inserts isolated from pUC plasmids available in our

Figure 3.2 Construction and screening of the wheat mitoplast cDNA clone banks.

(A) Electrophoresis of cDNA fractions from Sepharose-4B column. T: total unfractionated cDNA; 1. eluate 12-13; 2 eluate 14-15; 3. eluate 16-17; 4.eluate 18-19; 5. eluate 19-20; 6. eluate 21-22. (B). Agarose gel electrophoresis of cDNA clones after digestion with EcoRI and HindIII to liberate inserts. (C) and (D). Colony hybridization with the wheat mitochondrial 18S rRNA (C) and cytochrome b (*cob*) gene probes (D). The positive *cob* clone is circled in (C).



A
T 1234567
0.4 -
B
2 -
1 -
0.6-
2-
1-
0.6-
C
18S rRNA
D
cob

laboratory. Fig 3.2C and D show colony hybridization results using the wheat mitochondrial 18S rRNA and *cob* genes, respectively. These data suggested that the majority of the mitoplast cDNA clone library contains cDNA copies of the mitochondrial rRNAs while a small proportion is derived from the mitochondrial mRNAs. Fig 3.2B shows the insert sizes of a number of cDNA clones, which vary from 150 bp to 1.2 kb; the average size was about 450 bp. The rRNA containing clones were larger (the average size of fourteen cDNA clones being 850 bp). About 91% of cDNAs were from the mitochondrial rRNA genes whereas only 4 clones were derived from cytoplasmic rRNAs. About 31 cDNA clones did not hybridize with probes containing mitochondrial rRNA genes. Of these cDNA clones, 20 hybridized with the 12 known mitochondrial gene probes. The remaining 11 cDNA clone were of potential interest. Six of these clones (TE39, TE46, TE135, TES46, TES73 and TES96) were partially or completely sequenced (Table 3.1 and Appendix C). The longest clone TE46 (1.2 kb) showed homology to the mitochondrial *nad4* gene (Lamattina and Grienenberger 1991). TE135 showed homology to the 3' terminus of the wheat mitochondrial *nad7* gene, which was being characterized in the laboratory (Bonen et al. 1994). It showed three C-to-U changes, indicative of RNA editing, but contained part of the terminal intron, suggesting this clone comes from a precursor RNA molecule. TES73 also showed homology to the *nad7* gene, but most of its sequence was from 3' flanking region beyond its major 3' terminus. TE39 showed homology at the protein level to the carboxyl-terminal 39 residues of bacterial and chloroplast S7 ribosomal proteins (Reinbolt et al. 1979; Shinozaki et al. 1986). Because this gene had not been identified in the mitochondrion of any flowering plant and the S7 protein is known to be one of two major ribosomal assembly initiation protein and play a role

Table 3.1 Analysis of selected wheat mitochondrial cDNA clones

cDNA clones	insert size (kb)	Homology with
TE39	0.5	E. coli and chloroplast S7 ribosomal protein
TE46	1.2	Mitochondrial NAD4 protein
TE135	0.4	Wheat mitochondrial nad7 gene
TES46	0.6	None
TES73	0.7	Wheat mitochondrial nad7 gene
TES96	0.6	None

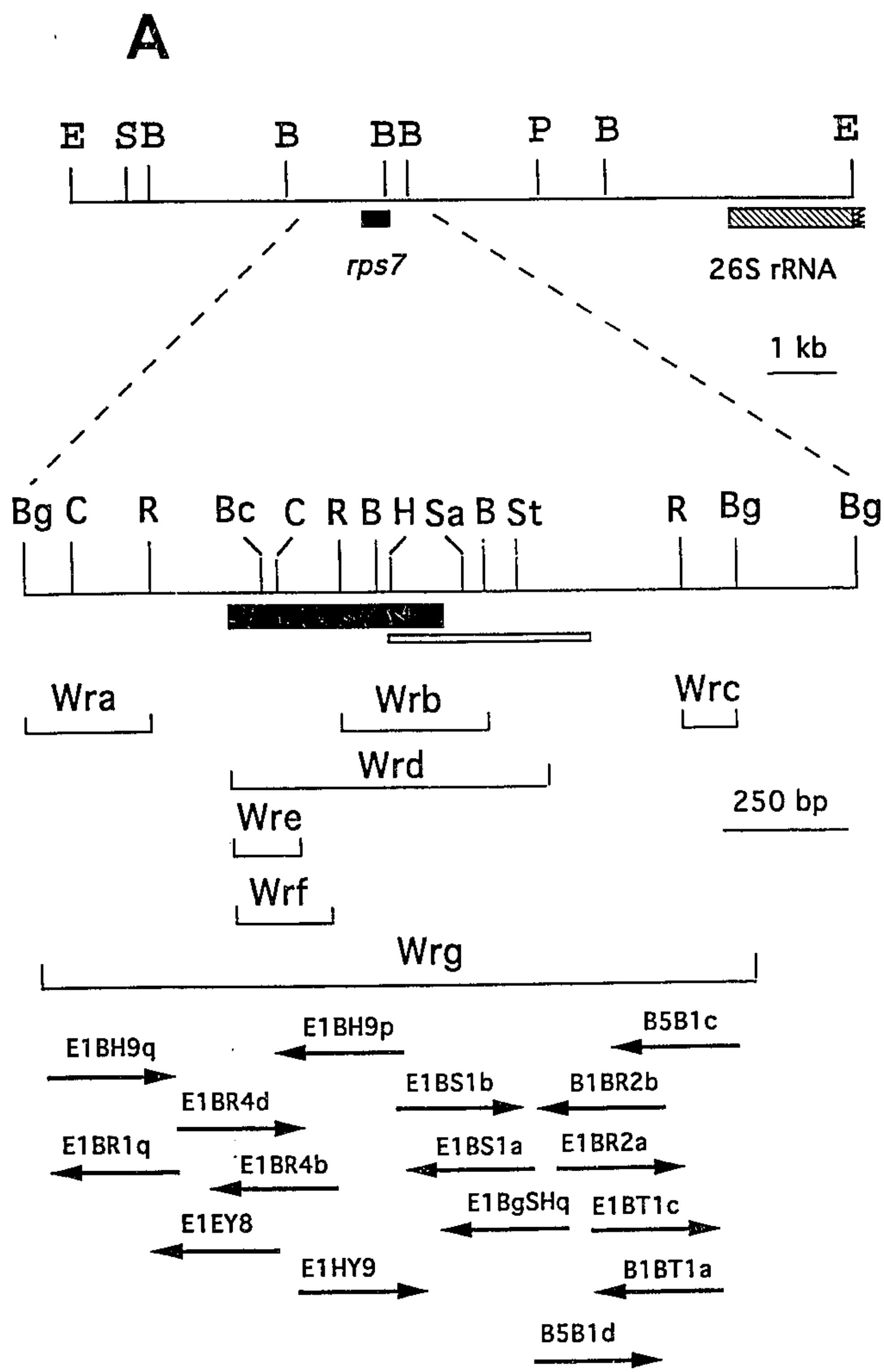
of translation initiation in bacteria (Abdurashidova et al. 1991; Dontsova et al. 1991), it was characterized further. Both TES46 and TES96 (data not shown) share no homology with any plant mitochondrial DNA sequences or have no long open reading frames. These clones may be derived from transcribed non-coding regions. Plant mitochondrial gene transcripts often have long leader and tail sequences, which can account for about 50 to 85% of nucleotides in the some plant mitochondrial RNAs, such as maize *cob* and *atp9* transcripts (Mulligan et al. 1988).

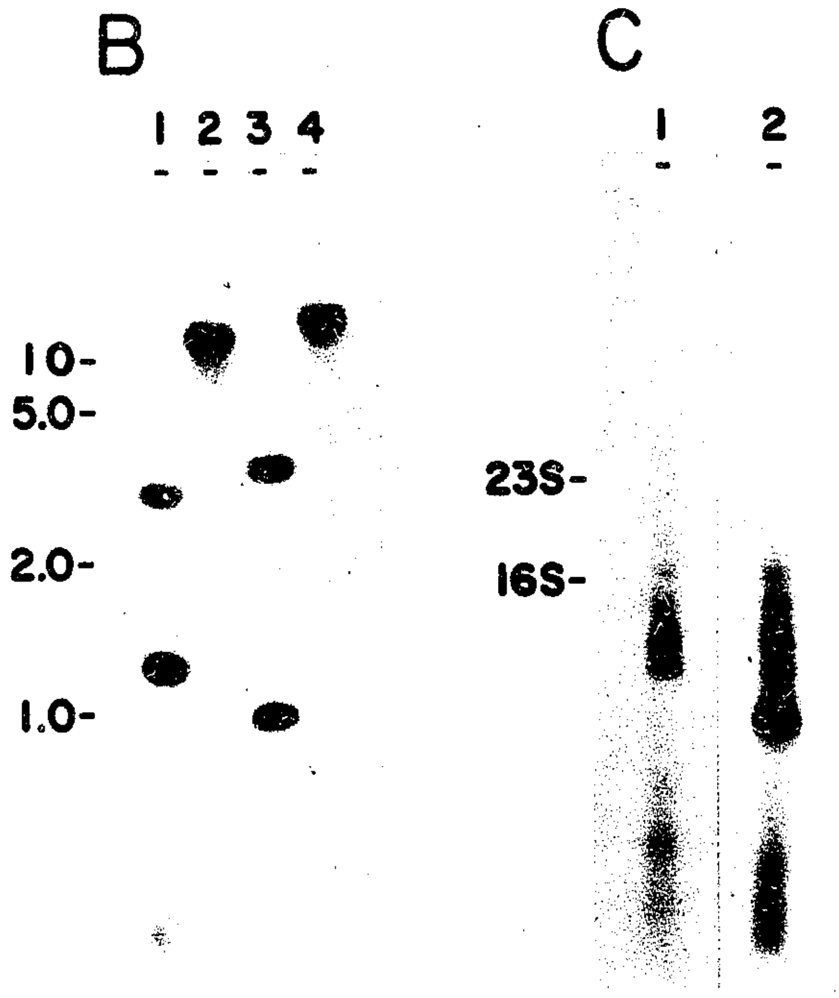
3.4 Isolation of the wheat mitochondrial *rps7* gene.

To obtain the complete wheat mitochondrial *rps7* gene, this partial cDNA clone (TE39) was used to probe wheat mitochondrial DNA digested by various restriction enzymes. Southern blot analysis (Fig 3.3B) showed that the cDNA hybridized with a 12.1 kb EcoRI fragment and this fragment was cloned in a pUC plasmid vector. Its restriction map is shown in Figure 3.3A. This region overlaps a Sal I fragment of 11.7 kb which contains one of the two copies of the 26S rRNA gene (designated K1 by Falconet et al. 1985). The *rps7* gene is therefore located 4.9 kb upstream of one copy of the 26S rRNA gene and about 8 kb downstream of one of three copies of 18S-5S rRNA genes (named G2) (Lejeune et al. 1987). This was further verified by restriction analysis and partial DNA sequence analysis of flanking sequences (data not shown). The wheat mitochondrial *rps7* gene is about 1.5 kb upstream of the end of recombination repeat #5 (Falconet et al. 1985) and 7 kb downstream of the breakpoint of the repeat #3 (Lejeune et al. 1987). Based on Southern hybridization analysis (Fig 3.3B) with the probe, Wrg, there is only a single copy of the *rps7* gene in the

Figure 3.3. Restriction map of the wheat mitochondrial *rps7* gene region and Southern blot and Northern blot analysis.

(A) Restriction map of the wheat mitochondrial *rps7* gene region. The positions of the *rps7* gene (black bar), 26S rRNA gene (hatched bar) and *rps7* partial cDNA clone (open bar) are shown. Restriction sites: B: BamHI; Bc: BclI; Bg: BglII; C: ClaI; E: EcoRI; H: HindIII; P: PstI; R: RsaI; Sa: SacI; St: StuI; S: SalI. Wra - Wrg show the positions of probes used in hybridization experiments. (B) Southern blot of wheat mtDNA restricted with BamHI (lane 1); EcoRI (lane 2); HindIII (lane 3); XhoI (lane 4) and hybridized with probe "Wrg". Size markers are in kb. (C) Northern blot of wheat mtRNA hybridized with probes "Wra" (lane 1) and "Wrb" (lane 2).


A
E
SB
B
BB
P
B
E
rps7
26S rRNA
1 kb
Bg
C
R
Bc
C
R
B
H
Sa
B
St
R
Bg
Bg
Wra
Wrb
Wrc
Wrd
250 bp
Wre
Wrf
Wrg
E1BH9q
E1BH9p
B5B1c
E1BS1b
B1BR2b
E1BR4d
E1BR1q
E1BR4b
E1BS1a
E1BR2a
E1BgSHq
E1BT1c
E1EY8
E1HY9
B1BT1a
B5B1d



B
1 2 3 4
10-
5.0-
2.0-
1.0-
C
1 2
23S-
16S-

Figure 3.4. Nucleotide sequence of the wheat mitochondrial *rps7* gene and flanking regions.

RNA editing sites are shown by asterisks. Restriction sites shown in Fig 3.3 are underlined. The 80 bp upstream repeat is blocked and short arrows indicate an inverted repeat. Stretches homologous to 26S rRNA gene segments are shown by broken underline and the chloroplast DNA-related segment is overlined. Solid arrowheads indicate positions of the 5' and 3' termini of *rps7* transcripts and the open arrowhead shows the location of the 3' end of the cDNA clone. The EMBL Data Library accession number for this sequence is X67242.

-501 AGATCTTGGCTCGCCCTACAAAAAAACAACTATATGCCCTATAAACTTGCTTCTTCGAATCTCGAAATAACATATAGAAAGTGTTTCTGTGATGAGACC

-402 ATTCTCGATCGATAAATGCGATAGGAGCCTATGTGTTTCACGGGCAGCCGGCCAATAGGGGAAGGTTGTGAATCCGGCGAACCGACCGCTTTAAGAACG

-303 TGATTGTCTTCTCTCACTCAGTTATCAATTGACAAAGATGACCCATTCTTTTTCGCCCTAAAAACGACAATGGTATGGTACTTTTTCTTCAAATCGAGA

-204 TTGTGTGGGTGTCCGCTCATGTTCACGTTACATGCTAAATCAGGCTTTCCTTGGAAAAACCAAGGACAACCCCTATCTCAGTCTCCTTTCCTTTCTCTT

-105 TTCGGGAGCAGAGCTGAAAAAGATGGACAGTAACGATCGCGTAATATCAATTTATCGGCCTCGTCATCGAAAGCGGCTTCCAATTGCTCGGAAATTCTC

-6 AGCTAT ATG GGG GAC TTT GAT GGT GAG CAA AAA GAA TTG ATC AAG AAA TTG GTA AAC TTT CGC ATG ATC GAT GGT
M G D F D G E Q K E L I K K L V N F R M I D G

70 AAA AGA ACG AGA GTT CGT GCT ATT GTT TAT AAA ACT TTT CAC CGC CTA GCT CGA ACT GAA CGC GAT GTA ATA AAA
K R T R V R A I V Y K T F H R L A R T E R D V I K

145 CTT ATG GTT GAC GCC GTA GAT AAT ATA AAG CCA ATA TGC GAA GTG GTC AAA GTA GGA GTC GCA GGT ACT ATT TAT
L M V D A V D N I K P I C E V V K V G V* A G T I Y

220 GAT GTT CCT GGG ATT GTA GCC AGG GAT CGT CAA CAA ACC TTA GCT ATT CGT TGG ATC CTT GGA GCA GCT TTC AAA
D V P G I* V A R D R Q Q T* L A I R W I L→F G A A F K

295 CGA CGT ATA AGC TAC AGG ATA AGC TTA GAG AAA TGT TCA TTT GCT GAG ATA CTG GAT GCT TAC CGA AAG AGG GGA
R R I S Y R I S L E K C S→L F A E I L D A Y R K R G

370 ATT TCA CGT AAG AGA AGG GAG AAT CTT CAT GGA CTG GCT TCC ACC AAT CGG AGT TTC GCG CAT TTC AGA TGG TGG
I S R K R R E N L H G L A S T N R S* F A* H F R W W

445 TAA AGTGATACCACATAAGGAGCTCTTCCTCATTCAGTCATACTCAACAAAAGTAAGAAATGTTTGACCCGGATCCTTTTTTCATCTTCATATAGAAA

543 GAAAATCGGCCTTCCTCATACTCCCCCCTTCATTCATAGAGTTGGAGGAATCCACAAGAGGCCTGCCCGTTCATAATTGCATAAAAGAACCATTCTTTT

642 TATGAAAACTCTTGTTCCAACCTCACCTCAGGTCGAATGAATACGAAAGGGGGATCAATCAAATCAATAAGCCATGAATGAAGAAGTAGTGGGCCTTTC

741 GCCTTATTTCGTTTGACTCGTGGAGCTGAGAAATCTACTTCAAAGAGAGGGAAAGAGTGCTAGTCCGAAAGAACAAGCAAGGGATGAGCGCACGGGAGC

840 GAAATCCGTTGCGCCCACGCGCATACGTTTTCTTGCTGGGTAATTTCTCAGAAAGTTTTTGTGTTGTTGATTCCTATTCCTAGGTGTTGTGCTTTTTCC

939 CCTATGCCGCCTATTGGTACTAGTGGAGTAGGATTGGCCTGTAATACAGAACCTATAGGTGGTGTAACCTTTCGCTCAATACTCAAATCTACAATTGAA

1038 GCATCTGAGGCTGCATCAATCGAGATACACGACATAAGGAATTGTTAGATCTCCAAACTGAACTTCGCCTTCACCAAGCGTGGGTTTTCACTGGTCCTA

1137 ACTAACGGTTGGCAAGCGATCCCCTATTCTGCATCCCGAAGATCAGACGGCCCAGAGCCGGAGCGATCATTCAAGCAAGTTACTCAGCTTGAGTTTTTC

1236 AGTCAATCAATAGATTGTGCAGTTTCTTAGTCTTTTGGTCCACGTCTTCATGACTTATAGAAAGGGTTTCTTTCTTTTACAAGATCATGGAAGCTT

Figure 3.5. Alignment of derived S7 ribosomal protein sequences.

Data are given from wheat mitochondria (this work), liverwort mitochondria (Oda et al. 1992), tobacco chloroplast (Shinozaki et al. 1986), *H. halobium* (Leffers et al. 1989) and *E. coli* S7 ribosomal protein (Reinbolt et al. 1979). Positions shared between wheat and liverwort are blocked and ones conserved in at least four organisms including wheat are indicated by asterisks. The amino acids altered by RNA editing in wheat mitochondria are shown by the open arrowheads. Numbers indicate additional amino acids at the amino or carboxyl terminus in liverwort, *E. coli* and *H. halobium.*

```
WHEAT MITOCHONDRIA                 MGDFDGEQKELIKKLVNFRMIDGKRTRV
LIVERWORT MITOCHONDRIA         M-81-IRGLNGKQKQLIKKLVHICMIDGKKTRS
E.COLI                  MPRRRVIGQRKILPDPKFGSELLAKFVNILMVDGKKSTA
TOBACCO CHLOROPLAST     MSRRGTAEKKTAKSDPIYRNRLVNMLVNRILKHGKKSLA
H.HALOBIUM           M-57-MGRHAQKQFKKSEISIVERLANRLMKTGANAGKKQQA
                                             *    **     **

RAIVYKTFHRLA-RTERDVIKLMVDAVDNIKPICE-VVKVGVAGTIYDVPGIVARDRQQTLAI
RAIVYKTFHRLA--PHGDVIKLLVNAIENVKPICE-VKKVRISGTTRLVPSIIATNRQETLAI
ESIVYSALETLAQRSGKSELEAFEVALENVRPTVE-VKSRRVGGSTYQVPVEVPVRRN-ALAM
YQIIYRAVKKIQQKTETNPLSVLRQAIRGVTPDIT-VKARRVGGSTHQVPIEIGSTQGKALAI
LKIVRDAFDIVHERTDENPIQVLVSAVENAAPREETVRLKYGGISVPQAVDTAPQRRVDQALK
  ***                    *  *  *  * *      *    **      *   **
   ▽                  ▽
RWIEG-AAFKRRISYRISLEKCIFAEIILDAYRKR--GISRKRRENLHGLASTNRSFAHFRWW
RWMLESAAKRRMGKKSISLDQCLYAEILEASQKM--GIARKKRDDLHKLAEANRSFSHYRWW
RWIVE-AARKRGD---KSMALRLANELSDAAENK--GTAVKKREDVHRMAEANKAFAHYRWL22
RWLLA-ASRKRPG---RNMAFKLSSELVDAAKGS--GDAIRKKEETHRMAEANRAFAHFR
FLADG-AHSASFK-TPTDAAEALANQLAGAADYNVQTYAIGQKKEKERVAAAAR
**    *   *           *  *   *      *         *  ** * * *
```

wheat mitochondrial genome. The presence of restriction sites for BamHI and HindIII within the *rps7* structural gene results in multiple hybridizing fragments; however, the 250 bp BamHI fragment shows only a weak hybridization signal (Fig 3.3B, lane 1), probably because small DNA fragments were diffuse.

3.5 DNA sequence analysis of the wheat mitochondrial *rps7* gene region.

The DNA sequence of the wheat mitochondrial *rps7* gene, which is 444 bp in length, potentially encodes a protein of 148 amino acids as shown in Figure 3.4 and that sequence has been aligned with its counterparts from liverwort mitochondria (Oda et al. 1992), *E. coli* (Reinbolt et al. 1979), tobacco chloroplast (Shinozaki et al.1986) and the archaebacterium, *Halobacterium halobium* (Leffers et al. 1989) in Figure 3.5. The wheat and liverwort mitochondrial sequences are the most similar with 61% amino acid identity, while the *E. coli*, chloroplast and archaebacterial sequences show 36%, 28% and 16% identities, respectively, with the derived wheat mitochondrial S7 ribosomal protein. There are 34 positions shared among four out of the five sequences including that of wheat mitochondria. At the nucleotide level, the wheat and liverwort sequences share 78% identity and they show high conservation over the full length plus 43 bp upstream of the putative initiation codon in wheat mitochondria. However, the liverwort mitochondrial *rps7* gene is 246 bp longer than its wheat counterpart at the 5' terminus (Oda et al. 1992). In addition, the S7 ribosomal protein from *Halobacterium halobium* is 52 amino acids longer at the amino terminus than the wheat counterpart (Leffers et al. 1989). The S7 ribosomal proteins appear to be variable in length at both amino and carboxyl termini. For example, the S7 ribosomal proteins from different strains of *E. coli* vary by 22 amino acids at the carboxyl terminus (Reinbolt et al. 1979). Even though they are different in length, the core sequences of the S7 ribosomal proteins from different organisms are very conserved. The similarity between the wheat mitochondrial and *E. coli* deduced S7 ribosomal proteins is about 36%, which is lower than that between the *E. coli* and chloroplast S7 ribosomal proteins (44%). The low similarities between the wheat mitochondrial and tobacco chloroplast ribosomal protein S7 sequences suggest that wheat mitochondrial *rps7* gene is not a chloroplast copy which was recently transferred from the chloroplast into the mitochondrion as found in the maize mitochondrial genome (Stern and Lonsdale 1982; Lonsdale 1989).

Relative to the *E.coli* S7 ribosomal protein, the deduced S7 ribosomal protein from wheat mitochondria has two insertions and one deletion. Compared to the liverwort mitochondrial sequence it has one insertion and one deletion at different positions. One of the two insertions is a duplication of three amino acids, Arg-Ile-Ser between the 112 and 113 amino acids (numbered on the *E. coli* S7 ribosomal protein), which is just in the front of the amino acid sequence Ser-Met-Ala-Leu-Arg that is known to interact with 16S rRNA in E.coli and can be cross-linked with 16S rRNA (Zwieb and Brimacombe 1979; Möller et al 1978). However, the comparable region is not highly conserved among the S7 ribosomal proteins from organelles.

The flanking regions of the wheat mitochondrial *rps7* gene have blocks that are very similar to sequences found elsewhere in the genome. Upstream there is a stretch of 80 bp (blocked, Fig. 3.4) homologous to ones preceding the genes for *coxII* (Bonen et al. 1984), *orf25* (Bonen et al. 1990), *atp6-1* and *atp6-2* (Bonen and Bird 1988). Interestingly, the *rps7* and *atp6-2* copies show a breakpoint in homology relative to the other three at precisely the same position. Fig 3.6 shows that the *coxII*, *orf25*, and *atp6-1* upstream repeats extend an additional 110 bp and the 5' termini of their mRNAs map within that region (Covello and Gray 1991; Bonen, unpublished data) to a motif similar to the consensus promoter motif for mitochondria (Covello and Gray 1991; Brown et al. 1991; Rapp and Stern 1992). These repeats immediately precede the translational initiation codon of the genes except for the wheat mitochondrial *rps7* gene, where these sequences and the putative translational initiation codon are separated by 77 nucleotides. The possible significance of this extra sequence is discussed further in Chapter 6.

In addition to this upstream repeat element, several short stretches of homology with other mitochondrial genes are also found in the flanking regions. Immediately upstream of the 80 bp repeat, there is a stretch (broken underline, Fig 3.4) related to the sequence preceding the 26S rRNA gene in maize mitochondria (Dale et al. 1984), but not in wheat mitochondria (Falconet et al. 1988). It is interesting to note that this sequence widely exists in the upstream and/or downstream regions of a number of mitochondrial genes of flowering plants (data not shown). Since this sequence is part of the maize mitochondrial 26S rRNA precursor transcript and immediately precedes the position of the mature 26S rRNA 5' terminus (Mulligan et al. 1988), it may be involved with RNA processing.

Downstream of the wheat *rps7* gene, there are two stretches (broken underline, Fig. 3.4) homologous to internal pieces of the 26S rRNA gene (Falconet et al. 1988). One is a stretch of 25 bp (from 735 to 760 bp), which is homologous to wheat 26S rRNA sequences. Another one, which is homologous to the maize mitochondrial 26S rRNA coding region (Dale

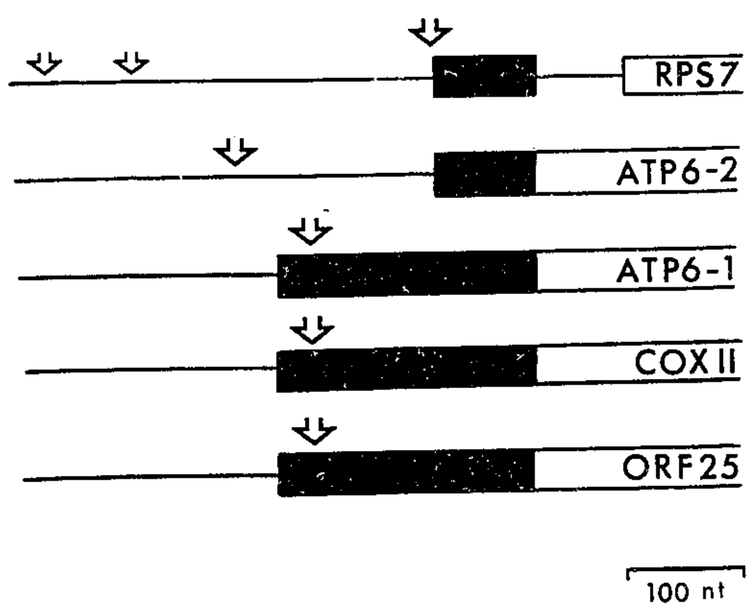
Figure 3.6 Comparison of the repeat elements preceding the wheat mitochondrial *coxII*, *orf25*, *atp6-1*, *atp6-2* and *rps7* genes.

(A). Sequence comparison of the repeat element. Capital and small letters indicate homologous and non-homologous regions respectively. Asterisks represent identical nucleotides in repeat elements within the repeated region. The translation initiation codons are underlined. (B). Schematic comparison of the repeat elements. The black and open bars represents the repeat elements and protein-coding regions respectively. The larger and small arrows indicates major 5' termini and minor 5' termini respectively.

A

```
orf25   CACGAAATGCTACTC-GAAAACGCGTATAGTAAGTAGTCTTCGTCGATGGGAC-----CTCCAGTGTATG
coxII   CCGTTGTTGCTCTTCAGAAAACGCGTATAGTAAGTAGTCTTCGTCGATGGGAC-----CTCCAGTGTATG
atp6-1  GCAATGTTGCTCTTCAGAAAACGCGTATAGTGG----CCTTCGTCGATGGGACAAACGCTCCAGTGTATG
atp6-2  aacacaaaggaatcgtcgaatttCGTATAGAAAGaaaaggaaccacttctattctcttttcttaagtata
rps7    tataaacttgcttcttcgaatctcgaaataacatatagaaagtgtttctgtgatgagtccattctcgatc


        CGTTACGAGGCAACTAGCATTTTGTCATTAAAGTTCGTGAAAGAATGTTTCGTTGGAAAAACCAACGCCG
        CGTTACGAGGCAACTAGCATTTTGTCATTAAAGTTCGTGAAAGAATGTTTCGTTGGAAAAACCAACGCCG
        CGTTACAAGGCAACTAGCATTTTGTCATGGAAGTTCGTGAAAGAATGTTTCGTTGGAAAAACCAACGCCG
        aga-61-gtgcaatgcttatgtgaaatctcaatccaaaacttattcGTTTCGTTGGAAAAACCAACGCCG
        gata-194-ggtgtccgctcatgttcacgttacatgctaaatcaggCTTTCCTTGGAAAAACCAAGGACA
                                                      **** ************** * *


        ATCTAAAACTAAGTCTTCCTTT---CAAAAGTGAGCGAGCAGAGCTGAAAAAGATGGAGTTACCTGGAGATG
        ACGTCAAGATCAGTCTCCTTTCCTTTCTCTTTTCGGGAGCAGAGCTGAAAAAGATGGGAAATTCCAATG
        ACGTCAAGATCAGTCTCCTTT----CAAAAGTGAGCGAGCAGAGCTGAAAAAGATGGAGTTACCTGGAGATG
        ATCTAAAACTAAGTCTTCCTTT---CAAAAGTGAGCGAGCAGAGCTGAAAAAGATGGAGTTACCTGGAGATG
        ACCCCTATCTCAGTCTCCTTTCCTTTCTCTTTTCGGGAGCAGAGCTGAAAAAGATGGAcagtaac-70-ATG
        *    *  * ***** * **         *  *  *********************
```

B



et al. 1984), is a longer stretch from +781 to +880. This element is also present in the 3'-flanking region of the maize *rps13* and wheat *rps12* genes. However, in the latter two cases, these sequences are closer to or even within the 3'-protein-coding region. These upstream and downstream repeats may be generated through inter- or intra- molecular recombination or by insertion/deletion events .

Further downstream there is a 240 bp region (solid overline, Figure 3.4) which is related to sequences from the chloroplast genome located between the *psaA* and $tRNA^{Ser}$ genes of the rice and tobacco chloroplast DNA with 85% and 82% similarity (Hiratsuka et al. 1989; Shinozaki et al. 1986), suggesting a relatively recent transfer from the chloroplast into the mitochondrion.

3.6. Northern blot and S1 nuclease mapping analysis

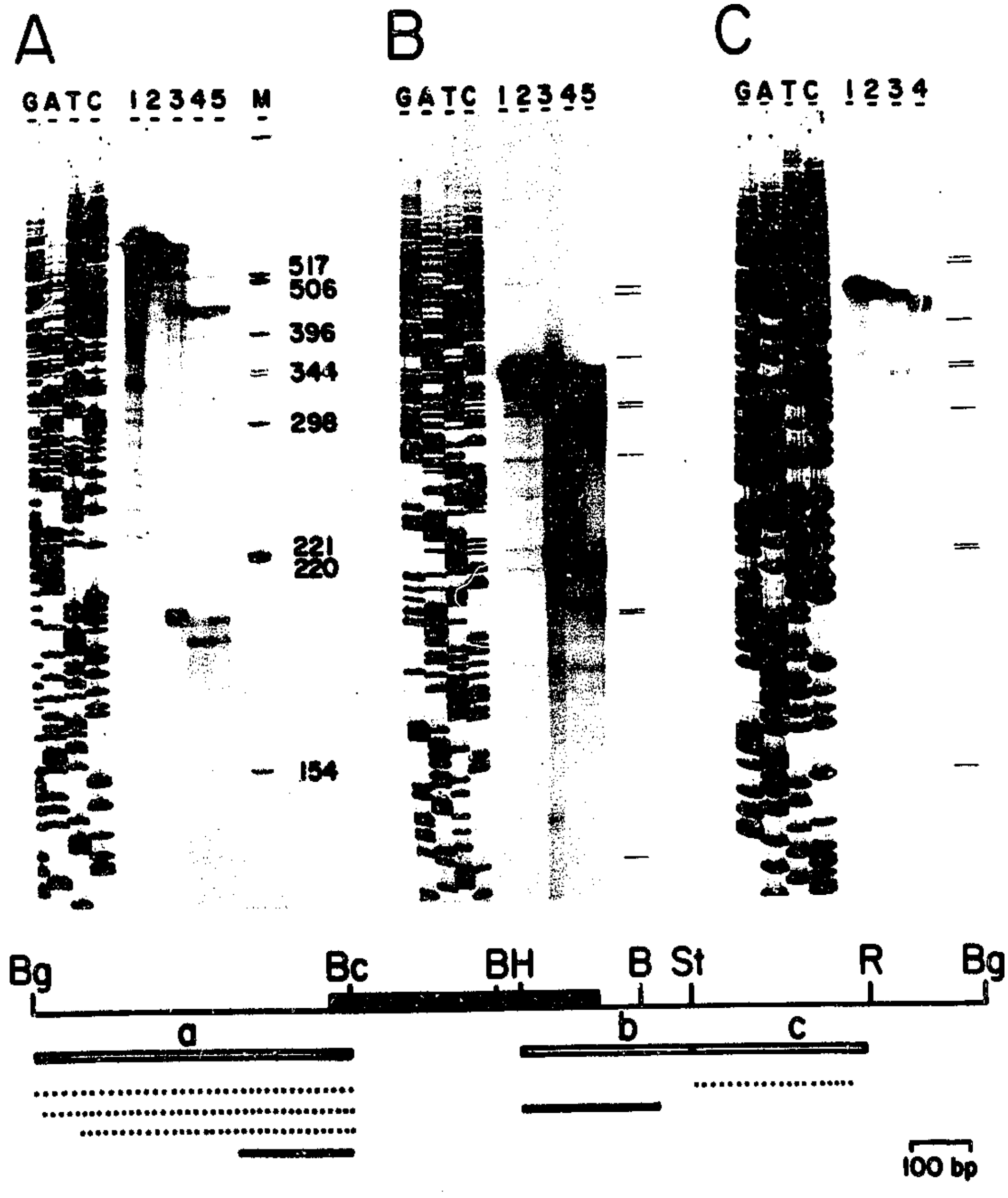
In order to establish whether the wheat mitochondrial *rps7* gene is actively transcribed, RNA blot analysis was carried out using hybridization probes Wra, Wrb and Wrc as shown in Figure 3.3. Coding sequences showed hybridization to an abundant transcript of approximately 0.7 kb and several minor species of 1.0 kb and larger (Fig.3.3C, lane 2). When probes located upstream of the 80 bp repeat element were used, only the larger transcripts were seen (Fig.3.3C, lane 1) and no signals were seen when a 0.3 kb probe located 640 bp downstream of the *rps7* gene was used (data not shown).

To determine the precise positions of the 5' and 3' termini of wheat *rps7* transcripts, S1 nuclease protection assays were carried out. Selected mtDNA regions were amplified by PCR in the presence of α-^{32}P-dATP. When the 5' terminal BglII-BclI segment (Figure 3.7A) was used, a major protected region of 195-205 nt was observed (Fig.3.7A, lane 3). This is consistent with 5' termini being located 160-170 nt upstream of the *rps7* initiation codon (bracketed arrowheads, Fig.3.4). This corresponds to the position at which the *atp6-2* and *rps7* upstream repeat elements diverge from those preceding the *coxII*, *orf25* and *atp6-1* copies. Three larger protected fragments are also seen, with 5' termini predicted to occur at approximately -425 and -485 nt (small arrowheads, Fig. 3.4), as well as ones extending upstream of the BglII site, which are fully protected and consistent with longer transcripts (Fig. 3.4). In addition, in lane 3 of Fig 3.7A, the largest fragments is a undigested DNA fragment. This heterogeneity in *rps7* transcript length may reflect RNA processing and/or multiple transcription initiation sites. As noted above, DNA sequences preceding position -170 resemble those upstream of the RNA processing site in the maize 26S rRNA precursor (Mulligan et al. 1988). Therefore, there is a strong possibility that the transcripts undergo processing at their 5' termini. This also means that the stem-loop structure may determine the processing site by the endonuclease and/or stop further degradation. However, direct confirmation of whether this is a processed or primary transcript will require guanylyltransferase capping experiments.

To determine the position of the wheat *rps7* mRNA 3' terminus, a HindIII-StuI segment (Fig.3.7B) was used in S1 nuclease protection assays. A protected fragment of approximately 245 nt was seen, consistent with the major 3' terminus being located approximately 110 nt downstream of the *rps7* termination codon. Taken together, the S1 mapping data predict *rps7* mRNAs of 715-725 nucleotides in length, and minor transcripts of about 980 and 1040 nt, in agreement with the Northern analysis. Transcripts extending

Figure 3.7. S1 nuclease mapping analysis of the wheat *rps7* gene region.

The black bar represents the *rps7* coding region and the open bars are DNA regions amplified/radiolabelled by PCR, corresponding to panel letters. The positions of major (dark lines) or minor (broken lines) protected fragments are indicated on the schematic. Sequencing ladders and HinfI-restricted pBR322 DNA are used as size markers. (A) Lanes 1-5 correspond, respectively, to DNA "a" control, 0, 50, 200 and 500 units S1 nuclease (Boehringer). (B) Lanes 1-5 correspond to DNA "b" control, 0, 10, 50 and 200 units S1 (GIBCO-BRL). (C) Lanes 1-4 correspond to DNA "c", 0, 50 and 200 units S1 (Boehringer). The autoradiogram in (C) is of a longer exposure time than in (A) or (B).


A
GATC 12345 M
517
506
396
344
298
221
220
154
B
GATC 12345
C
GATC 1234
Bg
Bc
BH
B
St
R
Bg
a
b
c
100 bp

farther downstream were also anticipated because the initial cDNA clone was derived from an RNA with a 3' non-translated tail of 347 nucleotides (open arrowhead, Fig.3.4). When the adjacent StuI-RsaI region was used for S1 mapping, minor protected fragments were detected (Fig 3.7C), consistent with heterogeneous transcripts terminating about 430 nt downstream of the *rps7* structural gene. These minor signals are located at the 3' end of a stretch (from +781 to +880), which is homologous to the maize mitochondrial 26S rRNA coding region (Dale et al. 1984). The sequences around this block can be folded into double stem-loop secondary structures (data not shown) preceded by GAGA, similar to those found in the 3'-flanking regions of the broad bean *rps14* gene and maize and tobacco *atp9* and *cob* genes (Wahleithner and Wolstenholme 1988). This stem-loop secondary structure may be involved with either transcript 3' processing or transcription termination (Wahleithner and Wolstenholme 1988). The original cDNA clone TE39, which ends at the position 794 (Fig 3.4), may represent a partially processed *rps7* mRNA.

3.7 RNA editing in wheat rps7 transcripts

To study RNA editing of wheat *rps7* transcripts, cDNAs of the coding region were generated with reverse transcriptase and oligomer LB29, located downstream of the *rps7* coding sequence and then amplified by PCR using the oligomers LB27 and LB29 which flank the *rps7* coding region. The products were cloned and sequenced. Three C-to-U changes were found in the *rps7* coding region in seedling RNA (asterisks in Fig 3.4; black dots in Fig 3.8), two of which alter amino acids. One of the changes (position 321 in Fig 3.4), from serine to leucine, increases the similarity to other S7 ribosomal proteins. However, the other one (position 276 in Fig 3.4), from leucine to phenylalanine, decreases the similarity between the wheat and liverwort mitochondrial S7 ribosomal proteins (Fig 3.5). In addition, when the wheat *rps7* gene was compared to the segment covered by the previous partial cDNA clone (open bar, Fig 3.3A), two additional C-to-U changes in the 3' non-translated region were observed (asterisks, Fig 3.4). One editing site occurs within the BamHI site and restriction of the RT-PCR products from both germinating embryos and seedlings showed that at least 50% of the RNA population is edited at this site (data not shown).

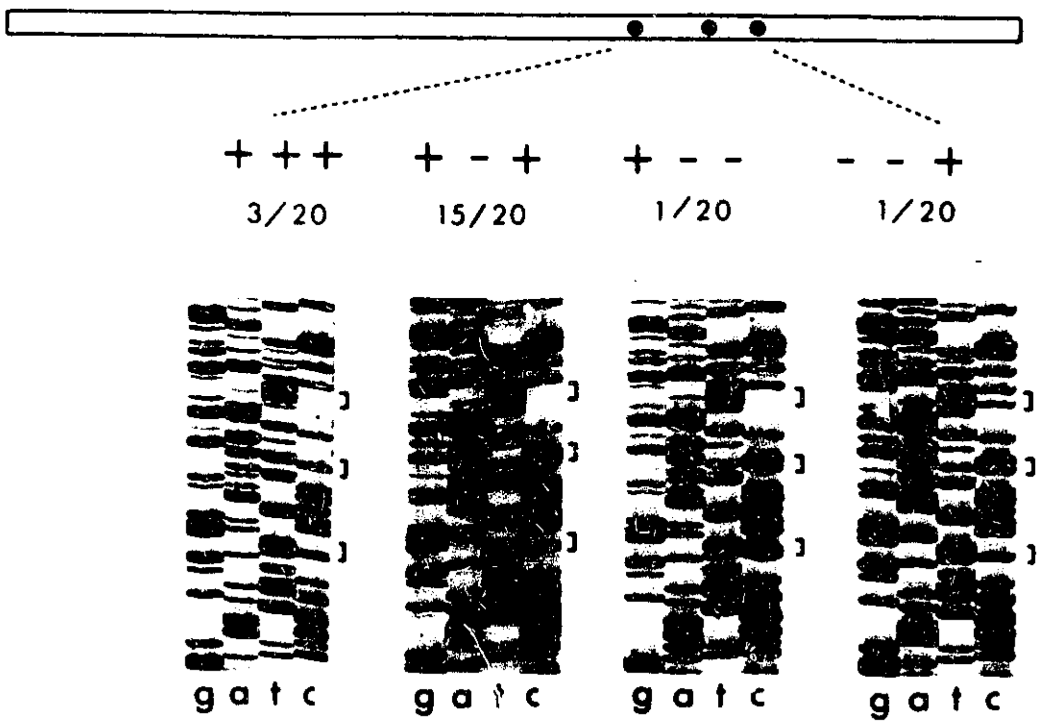
To know whether there might be differences in RNA editing between the wheat developmental stages, *rps7* cDNAs were made from mRNAs of the 24 hour wheat germinating embryos and six-day etiolated seedlings, amplified by PCR, cloned and sequenced. Eighteen and 20 cDNA clones from the six-day etiolated seedlings and 24 hour germinating embryos respectively were examined. In 24 hour germinating embryos, all cDNA clones were edited at the non-silent sites, but not at the silent site. In six-day etiolated seedlings, 75% of the cDNA clones were edited at the non-silent sites but none at the silent site. 10% of the cDNA clones were found to be edited at only one of the two non-silent sites and a silent C-to-U change was observed in 3 out of these 20 cDNA clones (Fig 3.8). Such cDNA clones (1) lack the middle site, (2) lack the 5' site and the middle site, or (3) lack the middle and 3' sites. These data support the idea that RNA editing in plant mitochondria lacks polarity (Gray and Covello 1993). This is in contrast to trypanosomal mitochondrial RNA editing, which has a 3'-to-5' directional polarity (Gray and Covello 1993).

In addition, 12 non C-to-U differences between genomic and cDNA *rps7* sequences were seen in the RT-PCR sequence experiments but each site was observed in only one

Figure 3.8. RNA editing analysis of wheat *rps7* transcripts.

The open bar represents the *rps7* protein-coding region. The black dots indicate C-to-U editing sites in the cDNA clones. The four blocks are sequencing gels of cDNA clones with brackets showing positions of potential C-to-U editing sites.

RPS7



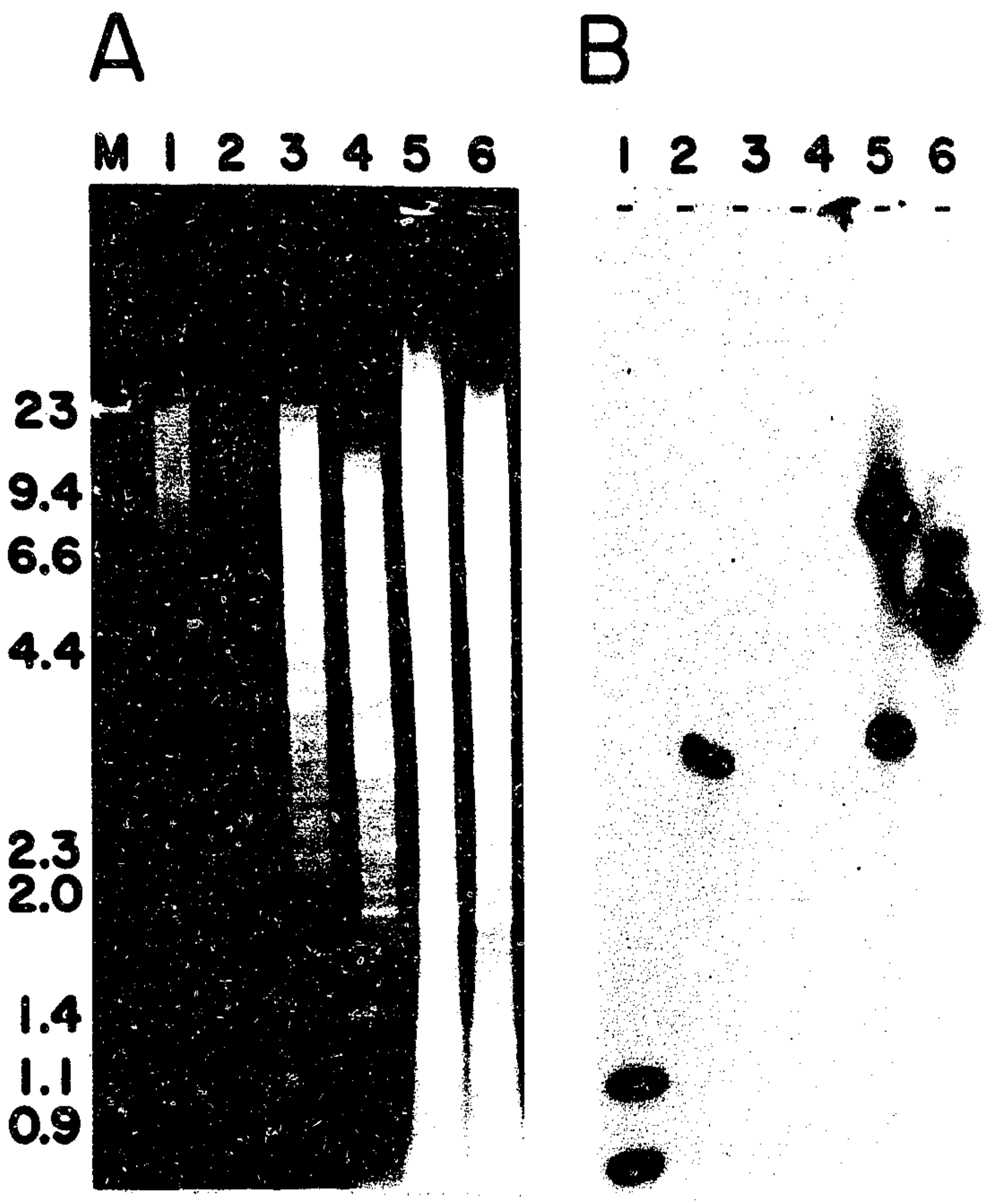
cDNA clone. There were three G-to-A, four A-to-G, three T-to-C changes and two extra T insertions observed out of 14,000 bp from 38 cDNA clones. It has been estimated that reverse transcriptase and Taq DNA polymerase have combined error frequencies of about 4 X 10^{-4} (Kunkel and Eckert 1989), which is somewhat lower than the rate of changes observed here (approximately 8.6 X 10^{-4}). This may reflect differences in conditions of reverse transcriptase reactions and Taq DNA polymerase or it may be related to the compositions and secondary structures of RNAs used. However, we can not rule out the possibility that these differences may reflect in vivo differences of plant mitochondrial gene transcripts. Such low-frequency changes have also been reported in analysis of other plant mitochondrial cDNAs (reviewed by Schuster et al. 1991; Bonnard et al. 1992).

3.8 Search for mitochondrial *rps7* gene homologues in other plants.

Since some ribosomal protein genes, such as *rps12*, *rps13* and *rps14*, are present in the mitochondria of some plants, but absent in those of others (Bland et al. 1986; Bonen 1987; Walheithner and Wolstenholme 1988; Schuster and Brennicke 1991), an obvious question is whether or not the *rps7* gene is present in the mitochondria of other flowering plants. To address this issue, the probe Wrd, a 542 bp PCR amplification product covering the wheat mitochondrial *rps7* gene generated by primers LB27 and LB29, was used to hybridize with BamHI- and EcoRI-restricted mitochondrial DNA from rice, pea and soybean. Fig 3.9 shows that there are strong hybridization signals in the mitochondrial DNA from rice and pea, consistent with a single copy of this gene being present in those genomes. In contrast, virtually no hybridization signals were seen with soybean mtDNA under low

Figure 3.9 Southern blot analysis of mitochondrial DNA from rice, soybean and pea hybridized with a *rps7*-specific probe.

A. UV fluorescence profile of BamHI- and EcoRI-restricted mtDNA from rice (lanes 1, 2), soybean (lanes 3, 4) and pea (lanes 5, 6). B. Southern blot after hybridization with the wheat mitochondrial rps7 probe, Wrd (as shown in Fig 3.3). The minor 7 kb hybridization signal is believed to be due to partially digested mtDNA.


A
B
M 1 2 3 4 5 6
1 2 3 4 5 6
23
9.4
6.6
4.4
2.3
2.0
1.4
1.1
0.9

stringency conditions. Therefore, the soybean mitochondrial DNA either does not have mitochondrial *rps7* gene, or has a very divergent *rps7* gene, for example, a copy of *rps7* gene transferred from the chloroplast. Since pea and soybean are closely related legumes, it seems unlikely that the soybean mitochondrial genome has a functional *rps7* gene that is very divergent from the pea counterpart. Another possibility is that the soybean mitochondrial S7 ribosomal protein may be not needed for plant mitochondrial translation. As mentioned in Chapter 3, in *E. coli*, the S7 ribosomal protein is believed to be very important for ribosomal assembly and translation initiation. Therefore, the most reasonable explanation is that the *rps7* gene is not present in the soybean mitochondrial genome and has been transferred from the mitochondrion to the nucleus during plant evolution. If so, it is very interesting to characterize this gene and also to know whether the pea mitochondrial *rps7*-homologous DNA sequence encodes a functional gene.

CHAPTER 4 CHARACTERIZATION OF THE PEA MITOCHONDRIAL GENOMIC REGION CONTAINING *rps7*-HOMOLOGOUS SEQUENCES AND THE *orf228-r* GENE

4.1. Isolation of *rps7*-homologous DNA sequences from pea mitochondria.

As shown in Fig 3.9B, a 5 kb EcoRI-restricted fragment of pea mitochondrial DNA hybridized with the wheat *rps7* gene probe (Wrd). A partial mitochondrial DNA library of EcoRI fragments of this size was constructed and screened by colony hybridization with Wrd. Positive clones were characterized further and the cloned DNA was used to construct a restriction map as shown in Fig 4.1.

After subcloning fragments into M13 vectors, sequencing showed that the *rps7*-homologous region is an open reading frame of 117 codon (Fig 4.2). It has 92% identity with the wheat mitochondrial *rps7* gene at the nucleotide level, but it is 119 bp shorter at the 5' terminus (Fig 3.4 and Fig 4.3). At the 3' end, their homology ends 80 bp downstream of the termination codon (Fig 4.3).

An important question is whether the 5' *rps7* region is present anywhere in the pea mitochondrial genome. A functional gene might have a *cis*-spliced or *trans*-spliced intron. If there is a *cis*-spliced intron, it would be expected to be less than 4 kb in length based on other organellar group-II introns (Michel et al. 1989). To address this question, the BamHI- and EcoRI-restricted mitochondrial DNAs from pea, rice and soybean were hybridized under conditions of reduced stringency with a probe (Wre), specific to the 5' region of the wheat

Figure 4.1 Physical map of the pea mitochondrial *rps7* pseudogene and *orf228-r* gene and flanking regions.

The black and open bars represent the *rps7* pseudogene and *orf228-r* gene, respectively. The Po1 and Po2 are two probes used in the Southern and Northern blot analysis. B: BamHI, Bg: BglII; E: EcoRI; F: SfuI; H: HindIII; Sa:SacI; X: XhoI. Within the *rps7* pseudogene, the BamHI site is 37 bp upstream of the HindIII site. The orientation and names of M13 clones are shown below.

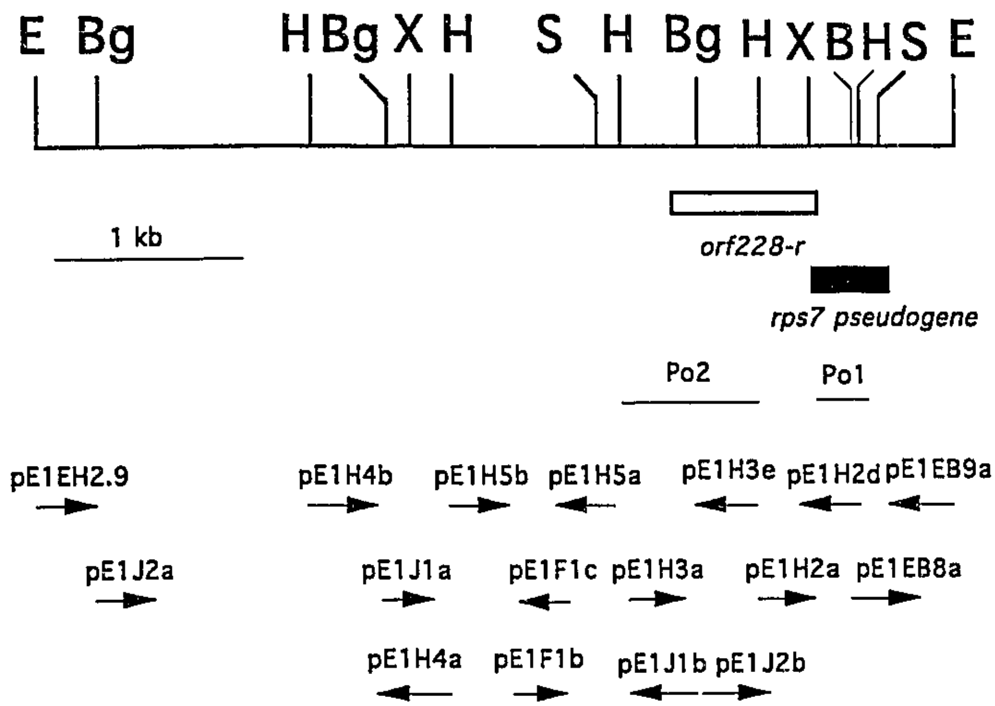

E Bg H Bg X H S H Bg H X B H S E
1 kb
orf228-r
rps7 pseudogene
Po2
Po1
pE1EH2.9
pE1H4b
pE1H5b
pE1H5a
pE1H3e
pE1H2d
pE1EB9a
pE1J2a
pE1J1a
pE1F1c
pE1H3a
pE1H2a
pE1EB8a
pE1H4a
pE1F1b
pE1J1b
pE1J2b

Figure 4.2. Nucleotide sequence of the pea mitochondrial *rps7* pseudogene and *orf228-r* gene and flanking regions.

The restriction sites Fig 4.1 are underlined. A stretch homologous to the downstream DNA sequence of the wheat *atp6* gene is shown by broken underline and sequences homologous to those upstream of the *Petunia atp9*-2 gene are overlined. The overlapping region between the *rps7* pseudogene and *orf228-r* gene is blocked. Arrowheads indicate the breakpoints in homology between the pea and wheat *rps7* regions.

-769 TTCGAATAACATACATAATCTCGTTCATTTTTCCTCCAGCAAGGAGCAGAGCCAGTTCAACTCATGAAATGGATGCTTTCGTAGGGGCTGTATGCCTTT

-670 TTATTCCTAAAGCTAGAGTTGGCCTGACCGGCTTCGCAGGATCATTTCCACTCTGTAGGCCGCTCATTCCTAAATAGGAAATAAAGTGCCTTTTCGATT

-571 CTTCCTTCCTGCTTGTGTGCTTTCCCTAGGAAAATTGACTCATTCTAGCTCTCCTCTACTTCAACCACTAGGAGTTCGCTCCTTTAATTAAGAATAGAA

-472 AATATCGCCTTTTTGGCATGAACATATAGCATGTGTGCCGTGAGGTCAATCACTGTCAGTTCTTCTTTCGAAATTGTCGAATATGATGAAGAGCTCTTA

-373 TCGGCTTGAGGCTTCTTTTCAAGCTGAGTATCAATTTCATAAGAGAAGAAAGTCGCGTAGAAAGGATTTGGATTCGAGGCTTGCTAAGTGAAGTCAACA

-274 AAGAAAAAAAGCTTTCTCCTGAGAAAGCGTCTGTTCACGTCAGGGTCCGCAGAGATGTAGTCACTTTGACTTTTATCGAGATTGGGTTGGTATTAAGTA

-175 TACTCATGATCTTTCTTTTTTTTTCCGGTATGCCGCTCCGCCAGCAAGGAGTGCCACGCCGAAGCGGACAGANANACAAAGCAGGGGGAATTATTCGTT

-76 GGGAGAAATCCTTTGATTGCGTATTGAATATAGATCCATGTCTTTCTTGTTCCACTAGCTAGGAGAAAATTATCAG ATG TCC TTA CAA CCT
M S L S L

16 TCT CTT TCG TTA TCG TTA TTT TTG ATG TCA AAG ACC AGA AGC TGT GCG CTA ATT CTC ATT GGA TCT CGG TTG TTC
S L L Q ·P S F L M S K T R S C A L I L I G S R L F

91 TTA ACA GCG ATG GCT ATT CAT TTA AGT CTT CGG GTA GCA CCA TTA GAT CTT CAA CAA GGT GGA AAT TCT CGT ATT
L T A M A I H L S L R V A P L D L Q Q G G N S R I

156 CTT TAT GTA CAT GTT CCT GCG GCT CGG ATG AGT ATT CTT GTT TAT ATC GCT ACG GCT ATA AAC ACT TTC TTA TTC
L Y V H V P A A R M S I L V Y I A T A I N T F L F

241 CTA TTA ACA AAA CAT CCC CTT TTT CTT CGC TCT TCC GGA ACC GGT ACA GAA ATG GGT GCT TTT TTT ACG TTG TTT
L L T K H P L F L R S S G T G T E M G A F F T L F

316 ACC TTA GTT ACT GGG GGG TTT CGG GGA AGA CCT ATG TGG GGC ACC TTT TGG GTG TGG GAT GCT CGT TTA ACC TCT orf228-r
T L V T G G F R G R P M W G T F W V W D A R L T S

391 GTA TTC ATC TCG TTT CTG ATT TAC CTG GGT GCA CTG CGT TTT CAA AAG CTT CCT GTC GAA CCG GCT CCT ATT TCA
V F I S F L I Y L G A L R F Q K L P V E P A P I S

466 ATC CGT GCT GGA CCG ATC GAT ATA CCA ATA ATA AAG TCT TCA GTC AAC TGG TGG AAT ACT TTG CAT CAA CCT GGG
I R A G P I D I P I I K S S V N W W N T L H Q P G

541 AGC ATT AGC CGA TCT GGT ACA TCA ATA CAT GTT CCT ATG CCC ATT CCA ATC TTG TCT AAC TTT GCT AAC TTC CCC
S I S R S G T S I H V P M P I P I L S N F A N F P

616 CTC TCA ACC CGT ATC TTG TTT GTT CTG GAA ACA CGT CTT CCT ATT CTA TCT TTT CCC GAA TCT CCT TTA AGG TAT
L S T R I L F V L E T R L P I L S F P E S P L R Y

U Q N I V Y F P A Q T E R D V I K
691 GAA ATA GAA GCT CGA GAA GGA ATA GCA AAA CCT AGT TTA CTT CCC AGC TCA AAC TGA A CGC GAT GTA ATC AAA
E I E A R E G I A K P S L L P S S N

L M V D A V E N I K P I C E V E K V G V A G T I Y
764 CTT ATG GTT GAC GCC GTA GAG AAT ATA AAG CCC ATA TGC GAA GTG GAA AAA GTA GGA GTA GCA GGT ACT ATT TAT

D A P G I V A R D R Q R T L A I R W I L E A A F K
839 GAT GCC CCT GGG ATT GTA GCC AGG GAT CGT CAA CGA ACC TTA GCT ATT CGT TGG ATC CTT GAA GCA GCT TTC AAA ψrps7

R R I S Y R I S L E K C S F A E I L D A Y R K R G
914 CGA CGT ATA AGC TAC AGG ATA AGC TTA GAG AAA TGT TCA TTT GCT GAG ATA CTG GAT GCT TAC CGA AAG AGG GGA

I A R K K R D N L H G L A S T N R S F A H F I W W
989 ATT GCA CGT AAG AAA AGG GAT AAT CTT CAT GGA CTG GCT TCC ACC AAT CGA AGT TTC GCG CAT TTC ATA TGG TGG

1064 TAAAGTGAGACCACATAAAGAGCTCTCTGATCCTCATTCAGTCAGATTCTTCAGTAAGATATGGTTTTACCCTTTTCCTTTTTGTTTGAATCTTCATT

1162 CATTGAGTATGAGGAATCCAAAAGCAGTAAGACCCCAATACAACAATAAGACAATAAAATTTTGTTACCTACACCTAATTCAAGACCGGTTAATGCAAG

1261 AACTATAAATAATAATATATACAGAAAAAGGCCATGACCGAAGGAGACTACAACCTAGTTATGATTTTAAGCCCGGCTTTATTTTTTACCTTCTTACCT

1360 ATTATGGTAAGGGCAACCCCTGTTGTCGACGGAATTC

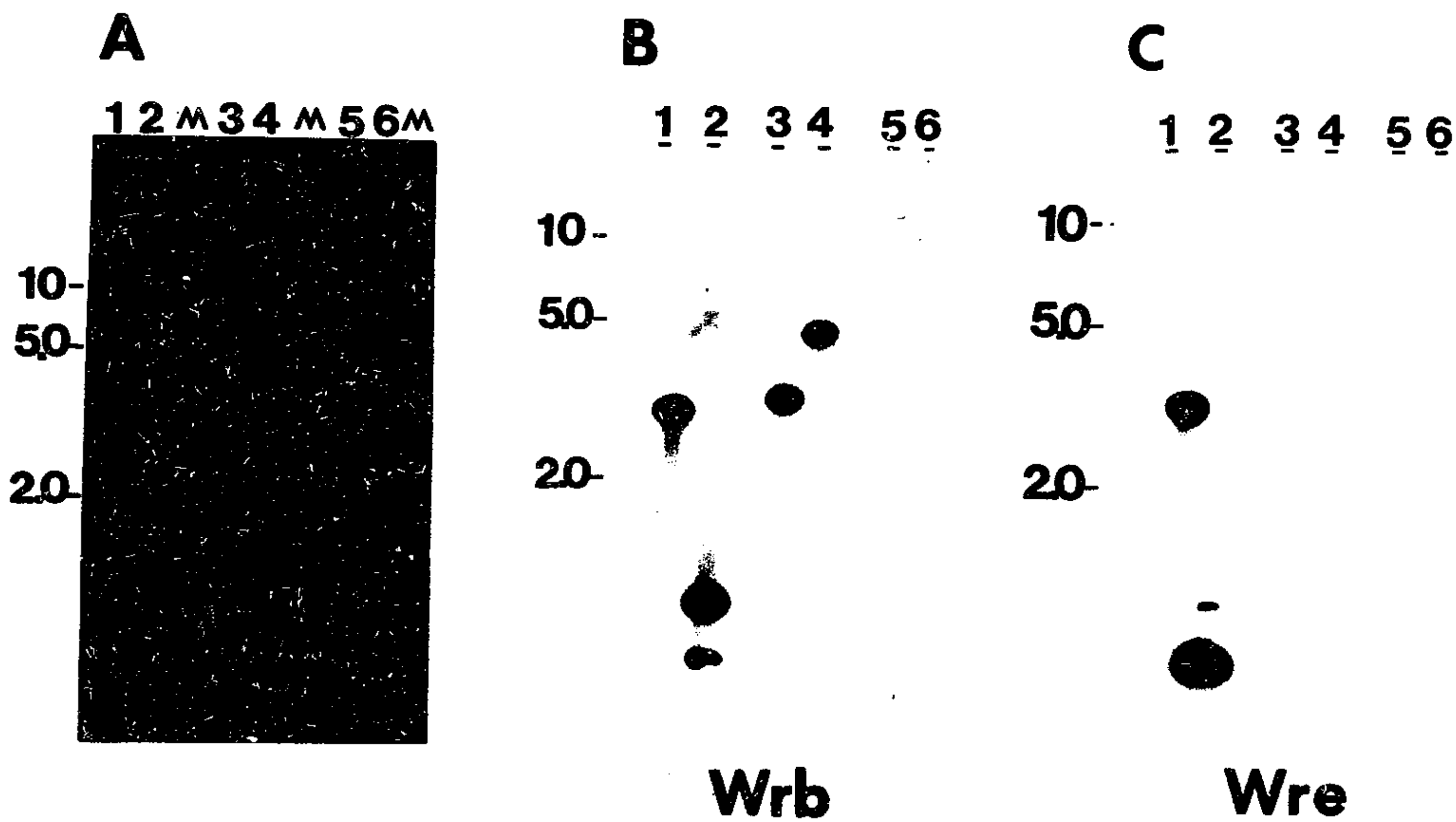
Figure 4.3. Alignment of the wheat *rps7* gene and pea mitochondrial *rps7* pseudogene.

1: the wheat *rps7* gene; 2: pea mitochondrial *rps7* pseudogene. Black arrowheads indicate the non-conservative amino acid changes. Asterisks indicate the identical positions.

```
               ▼   ▼   ▼   ▼
   K   T       F   H   R   L   A   R   T   E   R   D   V   I   K   L   M   V   D   A   V   D   N   I
1 AAA ACT -   TTT CAC CGC CTA GCT CGA ACT GAA CGC GAT GTA ATA AAA CTT ATG GTT GAC GCC GTA GAT AAT ATA
  ***  *       **  **    * * * *** * * *** *** *** *** *** **  *** *** *** *** *** *** *** **  *** ***
2 AAA CCT A   GTT TAC TTC CCA GCT CAA ACT GAA CGC GAT GTA ATC AAA CTT ATG GTT GAC GCC GTA GAG AAT ATA
               V   Y   F   P   A   Q   T   E   R   D   V   I   K   L   M   V   D   A   V   E   N   I

                           ▼
   K   P   I   C   E   V   V   K   V   G   V   A   G   T   I   Y   D   V   P   G   I   V   A   R   D
  AAG CCA ATA TGC GAA GTG GTC AAA GTA GGA GTC GCA GGT ACT ATT TAT GAT GTT CCT GGG ATT GTA GCC AGG GAT
  *** **  *** *** *** *** *   *** *** *** **  *** *** *** *** *** *** *   *** *** *** *** *** *** ***
  AAG CCC ATA TGC GAA GTG GAA AAA GTA GGA GTA GCA GGT ACT ATT TAT GAT GCC CCT GGG ATT GTA GCC AGG GAT
   K   P   I   C   E   V   E   K   V   G   V   A   G   T   I   Y   D   A   P   G   I   V   A   R   D

                                               ▼
   R   Q   Q   T   L   A   I   R   W   I   L   G   A   A   F   K   R   R   I   S   Y   R   I   S   L
  CGT CAA CAA ACC TTA GCT ATT CGT TGG ATC CTT GGA GCA GCT TTC AAA CGA CGT ATA AGC TAC AGG ATA AGC TTA
  *** *** * * *** *** *** *** *** *** *** *** * * *** *** *** *** *** *** *** *** *** *** *** *** ***
  CGT CAA CGA ACC TTA GCT ATT CGT TGG ATC CTT GAA GCA GCT TTC AAA CGA CGT ATA AGC TAC AGG ATA AGC TTA
   R   Q   R   T   L   A   I   R   W   I   L   E   A   A   F   K   R   R   I   S   Y   R   I   S   L

                                                                       ▼
   E   K   C   S   F   A   E   I   L   D   A   Y   R   K   R   G   I   S   R   K   R   R   E   N   L
  GAG AAA TGT TCA TTT GCT GAG ATA CTG GAT GCT TAC CGA AAG AGG GGA ATT TCA CGT AAG AGA AGG GAG AAT CTT
  *** *** *** *** *** *** *** *** *** *** *** *** *** *** *** *** ***  ** *** *** * * *** **  *** ***
  GAG AAA TGT TCA TTT GCT GAG ATA CTG GAT GCT TAC CGA AAG AGG GGA ATT CCA CGT AAG AAA AGG GAT AAT CTT
   E   K   C   S   F   A   E   I   L   D   A   Y   R   K   R   G   I   A   R   K   K   R   D   N   L

                                                       ▼
   H   G   L   A   S   T   N   R   S   F   A   H   F   R   W   W
  CAT GGA CTG GCT TCC ACC AAT CGG AGT TTC GCG CAT TTC AGA TGG TGG TAA AGTGATACCACATAAGGAGCTCTT----CCTC
  *** *** *** *** *** *** *** *** *** *** *** *** *** * * *** *** *** ***** ********* *******     ****
  CAT GGA CTG GCT TCC ACC AAT CGA AGT TTC GCG CAT TTC ATA TGG TGG TAA AGTGAGACCACATAAAGAGCTCTCTGATCCTC
   H   G   L   A   S   T   N   R   S   F   A   H   F   I   W   W

  ATTCAGTCATACTCAACAAAAGTAAGAAATG-TTTGACCCGGATCCTTTTT
  ********* * **  **    ********** *** ****    ********
  ATTCAGTCAGATTCTTCA---GTAAGATATGGTTTTACCCTTTTCCTTTTT
```

Figure 4.4. Southern blot analysis of mitochondrial DNA from rice, pea and soybean hybridized with the 5' and 3' specific *rps7* probes.

A. UV fluorescence profile of BamHI and EcoRI-restricted mtDNA from rice (lanes 2, 1), pea (lanes 3, 4) and soybean (lanes 5, 6). B. Southern blot after hybridizing with the probe, Wrb, specific to the 3' region of the wheat *rps7* gene. C. Southern blot after hybridization with the Wre probe, specific to the 5' *rps7* gene.


A
1 2 M 3 4 M 5 6 M
10-
5.0-
2.0-
B
1 2 3 4 5 6
10-
5.0-
2.0-
Wrb
C
1 2 3 4 5 6
10-
5.0-
2.0-
Wre

mitochondrial *rps7* gene (Fig 3.3). Fig 4.4C shows that there is virtually no hybridization signal for pea and soybean but strong signals for rice. In contrast when the same Southern blot was hybridized with a probe, Wrb, specific to the 3' region of the wheat mitochondrial *rps7* gene, as expected, strong hybridization signals were detected in the mitochondrial DNA from pea and rice, but not in that from soybean (Fig 4.4B). These results support the idea that the pea mitochondrial *rps7* homologue is shorter than the wheat counterpart. However, there is the possibility that pea mitochondrial *rps7*-homologous DNA sequence may contain a trans-spliced exon with a very divergent 5' terminus. If this is the case, such a functional gene should be transcribed. Northern blot analysis shows that the pea *rps7*-homologous DNA sequence is not transcribed (Fig 4.8C lane 1). Furthermore, DNA sequence analysis shows that the sequences upstream of the *rps7*-homologous sequence can not form the domain V and VI structures of a group-II intron (Michel et al. 1989). In addition, an open reading frame of 744 bp is present immediately upstream of the pea *rps7*-homologous DNA sequence.

The third possibility is that the pea mitochondrial *rps7*-homologous DNA sequence encodes a shorter protein and that translation involves an alternative initiation codon. As shown in Fig 4.2, no AUG or ACG codons were observed in the expected region. In contrast, there are at least two stop codons immediately upstream of the breakpoint in homology with wheat *rps7* gene. There is an AUG codon 51 bp downstream of the breakpoint between the wheat and pea *rps7*-homologous sequences, but the pea S7 ribosomal protein would then be one-third shorter than the wheat counterpart and lack some conserved amino acid regions.

Compared to the wheat *rps7* gene, there are 20 differences within the potential pea *rps7* coding region. Five differences occur in the first nucleotide of the codon; nine in the second; six in the third. The pea mitochondrial *rps7*-homologous DNA sequence also has a one nucleotide insertion which may result in potential frameshift. Eight non-conservative amino acid changes were seen and seven (Phe→Val; His→Phe; Val→Gln; Gly→Gln; Arg→Phe; Ser→Ala and Arg→Ile) can not be corrected by RNA editing in the pea *rps7*-homologous DNA sequence. Such differences indicate that the *rps7*-homologous DNA has drifted because of lack of constraint. Furthermore, the *rps7*-homologous DNA sequence is not transcribed. Taken together, these data support the idea that the pea mitochondrial *rps7*-homologous DNA sequences represents a truncated pseudogene.

4.2. Characterization of the pea mitochondrial *orf228-r* gene.

DNA sequence analysis showed that an open reading frame was located upstream of the pea mitochondrial *rps7* pseudogene. When these DNA sequences were searched against data banks, homology was found with the liverwort mitochondrial *orf228* gene (Oda et al. 1992). To avoid confusion, this open reading frame of 248 amino acids has been designated as the pea mitochondrial *orf228-r* gene because it is related to the liverwort *orf228* gene. Figure 4.2 shows the pea mitochondrial *orf228-r* gene sequence and flanking regions. The gene is 744 bp in length, which is 60 bp longer at its 3' terminus than its liverwort counterpart because it overlaps by 50 bp with the 5' region of the pea mitochondrial *rps7* pseudogene (blocked in Fig. 4.2). They share 80% similarity at the nucleotide level and 63% at the protein level (Fig 4.5 and Table 4.1). The relatively low protein sequence similarity suggests that the pea mitochondrial *orf228-r* transcripts might be heavily edited. No C-to-U RNA editing is expected in liverwort mitochondria (Oda et al. 1992). In the pea *orf228-r*

Figure 4.5 Alignment of the nucleotide and protein sequences of the *orf228-r* gene from pea mitochondria and *orf228* liverwort mitochondria.

1: the *orf228-r* gene from pea mitochondria; 2: the *orf228* gene from liverwort mitochondria (Oda et al. 1992). The potential amino acid-altering editing sites in the pea *orf228-r* sequence are double-underlined and the altered amino acids are indicated by horizontal arrows. Asterisks represent identical nucleotides between the wheat and liverwort mitochondrial sequences.

```
   M   S   L   S   L   L   Q   P   S→F F   L   M   S   K   T   R   S   C   A   L   I   L   I   G   S→F
1  ATG TCC CTT TCG TTA TTA CAA CCT TCT TTT TTG ATG TCA AAG ACC AGA AGC TGT GCG CTA ATT CTC ATT GGA TCT
   *** *    *  **  *** *** *   *** * * *** **  *** *       *       **  * * *** * * *** *** *** *** * *
2  ATG TAT GTC TCC TTA TTA CGT CCT TTT TTT TTT ATG TGC TGC TCT TTT CGC TAT GCG CAA ATT CTC ATT GGA TTT
   M   Y   V   P   L   L   R   P   F   F   F   M   C   C   S   F   R   Y   A   Q   I   L   I   G   F

   R   L   F   L   T   A   M   A   I   H→Y L   S   L   R→W V   A   P   L   D   L→F Q   Q   G   G   N
   CGG TTG TTC TTA ACA GCG ATG GCT ATT CAT TTA AGT CTT CGG GTA GCA CCA TTA GAT CTT CAA CAA GGT GGA AAT
    *  * * **  *** *** *** *** *** ***  ** *** ***  **  ** *** *** **  *   ***  ** *** *** *** * * ***
   TGT TGG TTT TTA ACA GCG ATG GCT ATT TAT TTA AGT ATT TGG GTA GCA CCC TCC GAT TTT CAA CAA GGT GAA AAT
   C   W   F   L   T   A   M   A   I   Y   L   S   I   W   V   A   P   S   D   F   Q   Q   G   E   N

   S   R   I   L   Y   V   H   V   P   A   A   R→W M   S   I   L   V   Y   I   A   T→M A   I   N   T
   TCT CGT ATT CTT TAT GTA CAT GTT CCT GCG GCT CGG ATG AGT ATT CTT GTT TAT ATC GCT ACG GCT ATA AAC ACT
   * * **  ***  ** *** **  *** *** *** **  ***  ** *** ***  *  ***  ** *** **  **  * * *** **  * * * *
   TAT CGC ATT ATT TAT GTG CAT GTT CCT GCA GCT TGG ATG AGT TTA CTT ATT TAT ATT GCA ATG GCT ATT AGC AGT
   Y   R   I   I   Y   V   H   V   P   A   A   W   M   S   L   L   I   Y   I   A   M   A   I   S   S

   F   L   F   L   L   T   K   H   P   L   F   L   R   S→F S   G   T   G   T   E   M   G   A   F   F
   TTC TTA TTC CTA TTA ACA AAA CAT CCC CTT TTT CTT CGC TCT TCC GGA ACC GGT ACA GAA ATG GGT GCT TTT TTT
    *  *** ***  ** *** *** *** *** *** *** *** *       * * ***   * *** * *  *   ** **  *** *** **  ***
   GTG TTA TTC TTA TTA ACA AAA CAT CCC CTT TTT CAG TTA TTT TCC AAA ACC GCT GCC AAA ATA GGT GCT TTG TTT
   V   L   F   L   L   T   K   H   P   L   F   Q   L   F   S   K   T   A   A   K   I   G   A   L   F

   T   L   F   T   L   V   T   G   G   F   R→W G   R   P   M   W   G   T   F   W   V   W   D   A   R
   ACG TTG TTT ACC TTA GTT ACT GGG GGG TTT CGG GGA AGA CCT ATG TGG GGC ACC TTT TGG GTG TGG GAT GCT CGT
   *** *** *** ***  ** **  **  *** **  ***  ** *** * * *** *** *** **  *** *** *** *** *** *** *** ***
   ACG TTG TTT ACC CTA GTA ACC GGG GGT TTT TGG GGA AAA CCT ATG TGG GGT ACC TTT TGG GTG TGG GAT GCT CGT
   T   L   F   T   L   V   T   G   G   F   W   G   K   P   M   W   G   T   F   W   V   W   D   A   R

   L   T   S   V   F   I   S→L F   L   I   Y   L   G   A   L   R   F   Q   K   L→F P→S V   E   P   A
   TTA ACC TCT GTA TTC ATC TCG TTT CTG ATT TAC CTG GGT GCA CTG CGT TTT CAA AAG CTT CCT GTC GAA CCG GCT
   *** *** *** *** **  *** * * ***  *  *** *** *** *** **  *** *** *** ***  **  **  * * **      ***
   TTA ACC TCT GTA TTA ATC TTG TTT TTT ATT TAC CTG GGT GCA CTG CGT TTT CAA GAG TTT TCT GCG GAT GTT GCT
   L   T   S   V   L   I   L   F   L   I   Y   L   G   A   L   R   F   Q   E   F   S   A   D   V   A

   P→S I   S   I   R→C A   G   P→L I   D   I   P   I   I   K   S→F S   V   N   W   W   N   T   L   H
   CCT ATT TCA ATC CGT GCT GGA CCG ATC GAT ATA CCA ATA ATA AAG TCT TCA GTC AAC TGG TGG AAT ACT TTG CAT
    ** *** *   **   **     **  *   ***  ** *** *** *** **  *** * * **  *** *** *** *** *** **  *** ***
   TCT ATT TTT ATA TGT ATA GGG CTA ATC AAT ATA CCA ATA ATT AAG TTT TCT GTC AAC TGG TGG AAT ACA TTG CAT
   S   I   F   I   C   I   G   L   I   N   I   P   I   I   K   F   S   V   N   W   W   N   T   L   H

   Q   P   G   S   I   S   R   S→F G   T   S   I   H   V   P→S M   P→L I   P   I   L   S   N   F   A
   CAA CCT GGG AGC ATT AGC CGA TCT GGT ACA TCA ATA CAT GTT CCT ATG CCC ATT CCA ATC TTG TCT AAC TTT GCT
   *** ***     *** *** *** * * * * *** **  *** *** ***  **  ** *** * * *** *** *** **  *   * * *** ***
   CAA CCT TCA AGC ATT AGC CAA TTT GGT ACT TCA ATA CAT ATT TCT ATG CTC ATT CCA ATC TTT TTA ATC TTT GCT
   Q   P   S   S   I   S   Q   F   G   T   S   I   H   I   S   M   L   I   P   I   F   L   I   F   A

   N   F   P→F L→F S→L T   R   I   L   F   V   L   E   T   R   L   P   I   L   S   F   P   E   S   P
   AAC TTC CCC CTC TCA ACC CGT ATC TTG TTT GTT CTG GAA ACA CGT CTT CCT ATT CTA TCT TTT CCC GAA TCT CCT
   * * ***   *  *  * * **   *  **  **  ***  **  ** *** *** *** *       *** *** *** ***             *
   AGC TTC TTC TTT TTA ACT GGG ATT TTT TTT ATT TTG GAA ACA CGT CAA ATA ATT CTA TCT TTT TAT TTT CAG CGA
   S   F   F   P   L   T   G   I   F   F   I   L   E   T   R   Q   I   I   L   S   F   Y   F   Q   R

   L   R   Y   E   I   E   A   R   E   G   I   A   K   P   S   L   L   P   S   S   N
   TTA AGG TAT GAA ATA GAA GCT CGA GAA GGA ATA GCA AAA CCT AGT TTA CTT CCC AGC TCA AAC TGA
    *  **   *
   AAA AGC CAA TGA
   K   S   Q
```

sequence, at least 22 C-to-U changes are expected to be edited (double underlined in Fig 4.5) because they would lead to identical amino acids between pea and liverwort at those positions (Fig 4.5). Sequence analysis of *orf228-r* cDNA is needed to establish the actual RNA editing sites. Sequences downstream of the pea *rps7* pseudogene show homology to the 3' protein-coding region and flanking sequences of the *atp6* gene (broken underline in Fig 4.2), with 85% identity over 66 nucleotides to the *Oenothera* one (Schuster and Brennicke 1987). Shorter versions of this sequence have been also found downstream of the wheat and liverwort mitochondrial *rps7* genes. Upstream of the *orf228-r* gene, there is a stretch of 192 bp DNA sequence (overlined in Fig 4.2) showing 67% identity to the sequence upstream of the *Petunia* mitochondrial *atp9*-2 gene (Rothenberg and Hanson 1987). The *Petunia* mitochondrial *atp9*-1 and *apt9*-2 genes are divergent around this point and the 5' RNA termini of the *Petunia atp9*-2 transcripts are located in this region (Rothenberg and Hanson 1987).

The derived protein sequence of the pea mitochondrial *orf228-r* gene was found to share 38% and 34% similarity with the deduced protein of the *orf263* gene from *Bradyrhizobium japonicum* (Ramseier et al. 1991), a nitrogen fixing root nodule endosymbiont of soybean, and the *helC* gene from *Rhodobacter capsulatus* (Beckman et al. 1992), a gram-negative photosynthetic bacterium (Fig 4.6). In addition, the pea mitochondrial *orf228-r* protein also shares 17-25% similarity with the derived proteins from Paramecium mitochondrial *orf238* (Pritchard et al. 1990), liverwort chloroplast *orf320* (Ohyama et al. 1988), tobacco chloroplast *orf318* (Shinozaki et al. 1986) and rice chloroplast *orf321* (Hiratsuka et al. 1989) (Table 4.1). If the pea *orf228-r* transcript is edited, it should

Figure 4.6 Alignment of the deduced *orf228-r* protein homologues from pea mitochondria, liverwort mitochondria, *Bradyrhizobium japonicum* and *Rhodobacter capsulatus*.

Data are taken from the pea mitochondrial *orf228-r* gene (this work), the liverwort mitochondrial *orf228* gene (Oda et al. 1992), the *orf263* gene from *Bradyrhizobium japonicum* (Ramseier et al. 1991) and the *helC* gene from *Rhodobacter capsulatus* (Beckman et al. 1992). The asterisks indicate identical amino acid positions in at least three sequences. Hyphens indicate deletions of amino acid sequences.

```
Pea mt orf228         M S L S L L Q - P S F L M S K T R S C A L I L I G S R L F L T A M A
Lverwort mt orf228    M Y V P L L R - P F F F M C C S F R Y A Q I L I G F C W F L T A M A
helo                  M S I W E Y A N P V K F M Q T S G R L L P W V V A A T V L T L L P G
orf263-bja            M T L I D L A N P T R F L A L T A R V L P W L A A A T V I L L A I G
                      *         *     *     * *         *         *             *

I H L S L R V A P L D L Q Q G G N S R I L Y V H V P A A R M S I L V Y I A T A I N T F L F L L T K H
I Y L S I W V A P S D F Q Q G E N Y R I I Y V H V P A A W M S L L I Y I A M A I S S V L F L L T K H
L V W G F F F T P V A A E F G A T V K V I Y V H V P A A T L A I N I W V M M L V A S L I W L I R R H
L Y Q S A L A   P D D Y Q Q G A T V K I M F I H V P N A W L S M P V W G V M S I A S L G T L V M R H
      *         *       * * *         * * * * * * * * *     *             *   *   *       *       *

P L F L R S S G T G T E M G A F F T L F T L V T G G F R G R P M W G T F W V W D A R L T S V F I S F
P L F Q L F S X T A A K I G A L F T L F T L V T G G F W G K P M W G T F W V W D A R V L I L F L T S
H V S A L A A K A A A P I G M V M T L I A L I T G A F W G Q P H W G T W W E W D P R L T S F L I L F
P L A D V A A K A A A P I G A A F T F L A L L T G S L W G R P M M G T Y W E W D A R L T S V L I L F
* *           *   * *   * * *   * * *     *   * *   * * *   * * * * *   *   * * * * * * *     *   *

L I Y L G - - A L R F Q K L P V E P A P I S I R A - - - - G P I D I P I I K S S V N W W N T - L H Q
F I Y L G - - A L R F Q E F S A D V A S I F I C I - - - - G L I N I P I I K F S V N W W N T - L H Q
L-F Y L G Y M A L W E A I E N P D T A A D L T G V L C L V G S V F A V L S R Y A A I F W N Q G L H Q
L M Y L G L M A L W R A V D D P S R A A R A A A V L T L V G A I N L P I I K F S V D W W N T - L H Q
*   * * *     * *                   *                     *   *     * * * *   * *   * * * *   * * *

G T S I H V P M P I P G S I S R S - P I L S N F A N F P L S T R I L F V L E T R L P I L S F P E S P
P S S I S Q F G T S I H I S M L I - P I F L I F A S F F F L T G I F F I L T T R Q I I L S F Y F Q R
H I A D V Y W Q P L V L S I A G F G K L F V A L L L L R T R T E I R A R R L K A L E - - - - - - - -
P A S V M R M G G S S L D K S F L I P L L V M A I A F T L L F V T L H L A A M R N E I L R R R V R S
*   *                 *             *   *           *       *   *             *     * *

L R Y E I E A R E G I A K P S L L P S S
K G S T L S L D K E E S Q
- - - - Q R E R M A - - -
L Q M N Q A S R M A F S S E M G A G S R Q N H A S N E V G A A
```

Table 4.1 Comparison of the deduced proteins from the *orf228-r* gene homologues in mitochondria, chloroplast and bacteria.

M. p.m: *Marchantia polymorpha* mitochondrial *orf228* (Oda et al. 1992); B. j: *Bradyrhizobium japonicum orf263* (Ramseier et al. 1991); R. c: *Rhodobacter capsulatus helC* (Beckman et al. 1992); P. a: *Paramecium aurelia* mitochondrial *orf238* (Pritchard et al. 1990); O. s: *Oryza sativa* chloroplast *orf321* (Hiratsuka et al. 1989); N. t: *Nicotiana tabacum* chloroplast *orf318* (Shinozaki et al. 1986) and liverwort (*M. polymorpha*) chloroplast *orf320* (Ohyama et al. 1988). Values were obtained by using Microgenie comparison programs.

	M. p.m	B.j	R.c	P. a	O. s	N. t	*M. polymorpha* ct ORF320
P. sativum mt ORF228-*r*	63	38	34	21	20	18	22
M. polymorpha mt ORF228	-	43	34	25	19	22	21
B. japonicum ORF263		-	45	18	20	19	24
R. capsulatus helC protein			-	19	18	19	21
P. aurelia mt ORF238				-	20	19	21
O. sativa ct ORF321					-	72	56
N. tabacum ct ORF313						-	57

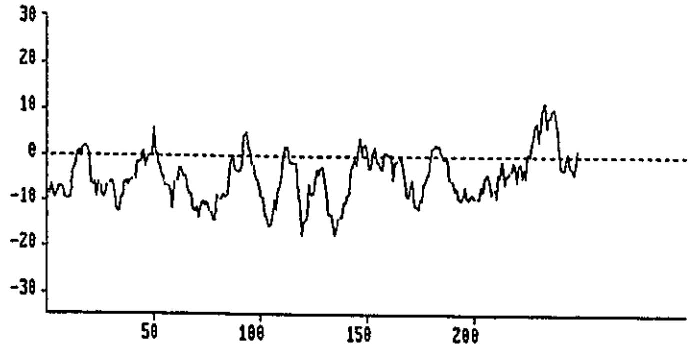
significantly increase the similarity of the pea deduced *orf228-r* protein and its homologues in other organisms. The *orf228* gene homologues encode highly hydrophobic proteins, typical of integral membrane proteins with trans-membrane helices, even without considering changes due to RNA editing (Fig 4.7).

The pea mitochondrial *orf228-r* gene is present as a single genomic copy based on Southern blot analysis (Fig 4.8B) as are the rice, soybean and wheat (data for wheat not shown) homologues. When the same probe was used to hybridize with the pea mitochondrial RNA, there was a strong hybridization signal of about 1.2 kb (Fig 4.8C, lane 2). Absence of transcripts from the pea *rps7* pseudogene region (Fig 4.8C, lane 1) places the 5' terminus of the *orf228-r* mRNA approximately at the position of sequences related to those at the 5' termini of the *Petunia* mitochondrial *atp9*-2 mRNA (Rothenberg and Hanson 1987). Since the pea mitochondrial *orf228-r* gene is actively transcribed, highly conserved and present in the mitochondrial genomes of many flowering plants, it is likely to be a functional mitochondrial gene.

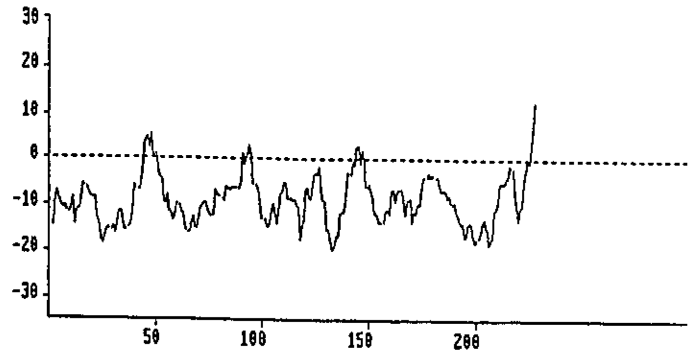
Figure 4.7 Hydropathy profiles of the deduced *orf228-r* proteins from pea mitochondria, the liverwort mitochondria, *Rhodobacter capsulatus* and *Bradyrhizobium japonicum*.

(A). The pea *orf228-r* protein; (B) liverwort mitochondrial *orf228* protein; (C) *helC* protein from *Rhodobacter capsulatus* and (D). *orf263* protein from *Bradyrhizobium japonicum*. The algorithm of Hopp and Woods (1981) was used to plot the hydropathy profiles of the deduced *orf228-r* proteins.

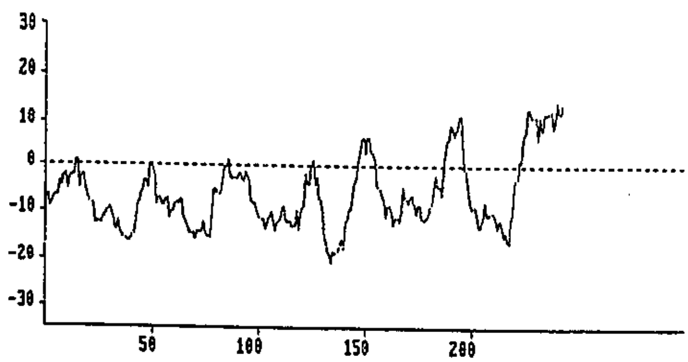
A


30
20
10
0
-10
-20
-30
50
100
150
200

B


30
20
10
0
-10
-20
-30
50
100
150
200

C


30
20
10
0
-10
-20
-30
50
100
150
200

D

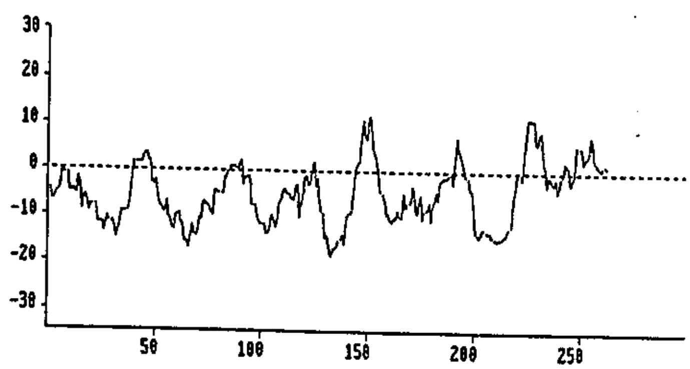
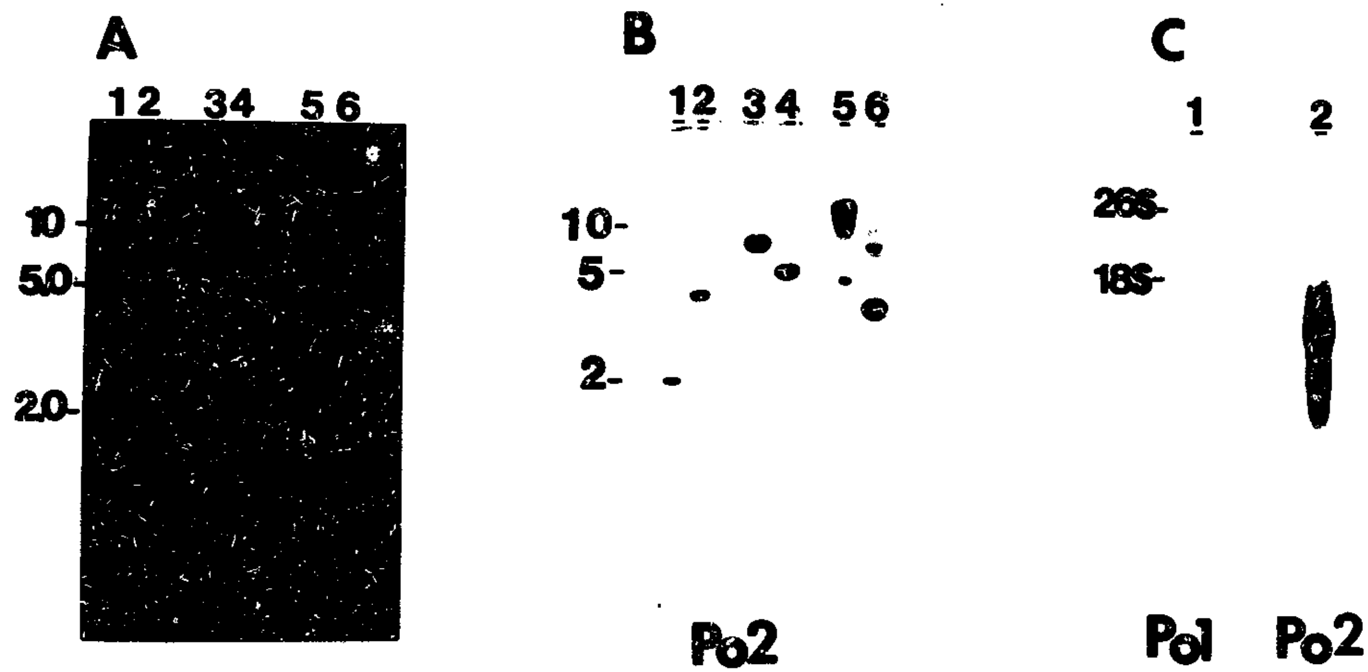

30
20
10
0
-10
-20
-30
50
100
150
200
250

Figure 4.8 Southern and Northern blot analysis of the pea mitochondrial *orf228-r* gene region.

(A) UV fluorescence profile of BamHI and EcoRI-restricted mt-DNA from rice (lanes 2,1) pea (lanes 3, 4) and soybean (lanes 5, 6). (B) Southern blot after hybridization with the Po2, the pea mitochondrial *orf228-r* gene. This blot was also used in the experiment shown in Fig 4.4. (C) Northern blot analysis of pea mitochondrial RNA with the probes, Po2 (lane 2) and Po1 (lane 1) in Fig 4.1.


A
1 2 3 4 5 6
10
5.0
2.0
B
1 2 3 4 5 6
10-
5-
2-
Po2
C
1
2
26S-
18S-
Po1
Po2

CHAPTER 5 SEARCH FOR PEA AND SOYBEAN NUCLEAR GENES ENCODING THE MITOCHONDRIAL S7 RIBOSOMAL PROTEIN

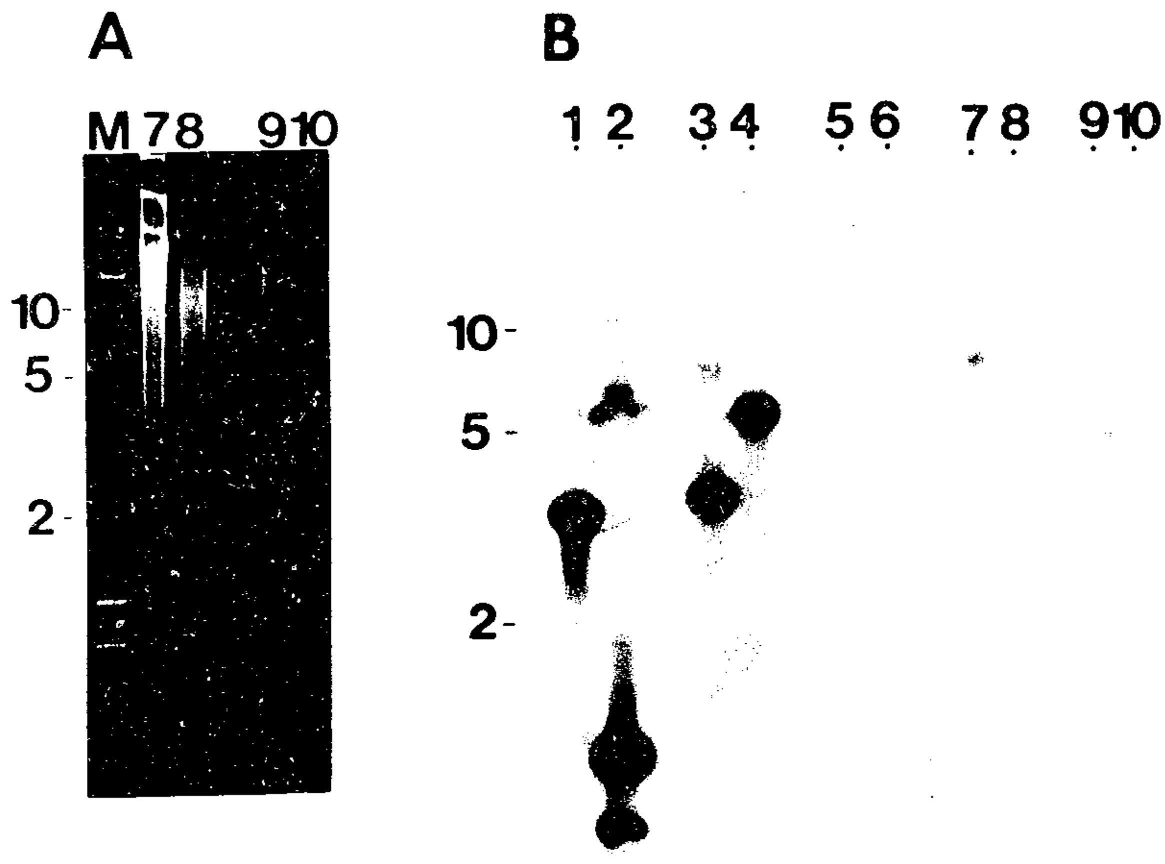
5.1 Southern blot analysis of legume nuclear DNA with the wheat mitochondrial *rps7* probe.

The S7 ribosomal protein is expected to be present in plant mitochondrial ribosomes because *rps7* gene homologues are found in wheat and liverwort mitochondria and the S7 ribosomal protein is essential for 30S subunit assembly and plays a very important role in translation initiation in bacteria (Moore and Capel 1988; Abdurashidova et al. 1991; Dontsova et al. 1991). Although Southern blot analysis showed that an *rps7*-homologous sequence is present in the pea mitochondrial genome, it was shown by sequence analysis to be a truncated pseudogene. No hybridization signals were detected in the mitochondrial genomes of soybean (Fig 5.1) and bean (*Phaseolus vulgaris*) in experiments repeated twice (data not shown). One reasonable explanation is that the *rps7* gene is present in the nuclear genome in these legumes.

When a wheat mitochondrial *rps7* probe (Wrb) was used in heterologous Southern hybridization experiments with soybean and pea nuclear DNAs, hybridization signals (to 8.5 kb BamHI and 5 kb EcoRI fragments for pea; 4.8 kb BamHI and 5.2 kb EcoRI fragments for soybean) which differed from the mitochondrial ones were seen (Fig 5.1), suggesting that *rps7*-homologous DNA sequences are present as a single copy in the nuclear genomes of pea and soybean. However, it is uncertain whether these nuclear sequences represent functional

Figure 5.1 Southern blot analysis of nuclear genomic DNA from pea and soybean hybridized with the 3' specific *rps7* probe.

(A) UV fluorescence profile of BamHI and EcoRI-restricted nuclear DNA from pea (lanes 7, 8) and soybean (lanes 9, 10). (B) Southern blot after hybridization with the Wrb probe (See schematic, cf. Fig 3.3A for its position). The UV fluorescence profiles and Southern blot for lanes 1-6 are also shown in Fig 4.4.


A
M 7 8 9 10
10
5
2
B
1 2 3 4 5 6 7 8 9 10
10
5
2

genes or pseudogenes. To verify whether the pea and soybean nuclear genomes contain the 5' *rps7* region, a PCR probe (Wre), specific to the wheat 5' region, was allowed to hybridize under the same conditions with pea and soybean nuclear DNAs. Again, hybridization signals showed the same patterns. No hybridization signals were detected in the wheat nuclear DNA (data not shown), but because of its large genome size it remains an open question whether *rps7* homologous sequences are present in the nucleus in wheat.

5.2 Screening a pea genomic library for *rps7*-homologous sequences.

In an attempt to isolate the nuclear-encoded *rps7* gene in pea, a pea nuclear DNA library (Clontech) was screened. Approximately 1.65 X 10^6 plaques were screened with complete *rps7* and 5' specific probes Wrd and Wre, respectively (Table 5.1). The initial 6 X 10^5 plaques were screened with the Wrd probe generated using PCR amplification. Since the pea genome is approximately 4.5 X 10^9 bp (Murray et al. 1977), it was expected that there would be one positive clone per 8.4 X 10^5 pfu or 1.29 X 10^6 pfu at 95% and 99% confidence levels, respectively. However, more than 0.7% plaques were found to be positive. This number appears too high to represent nuclear *rps7*-homologous DNA sequences, suggesting that the majority might be derived from contaminating pea mitochondrial DNA in the clone bank.

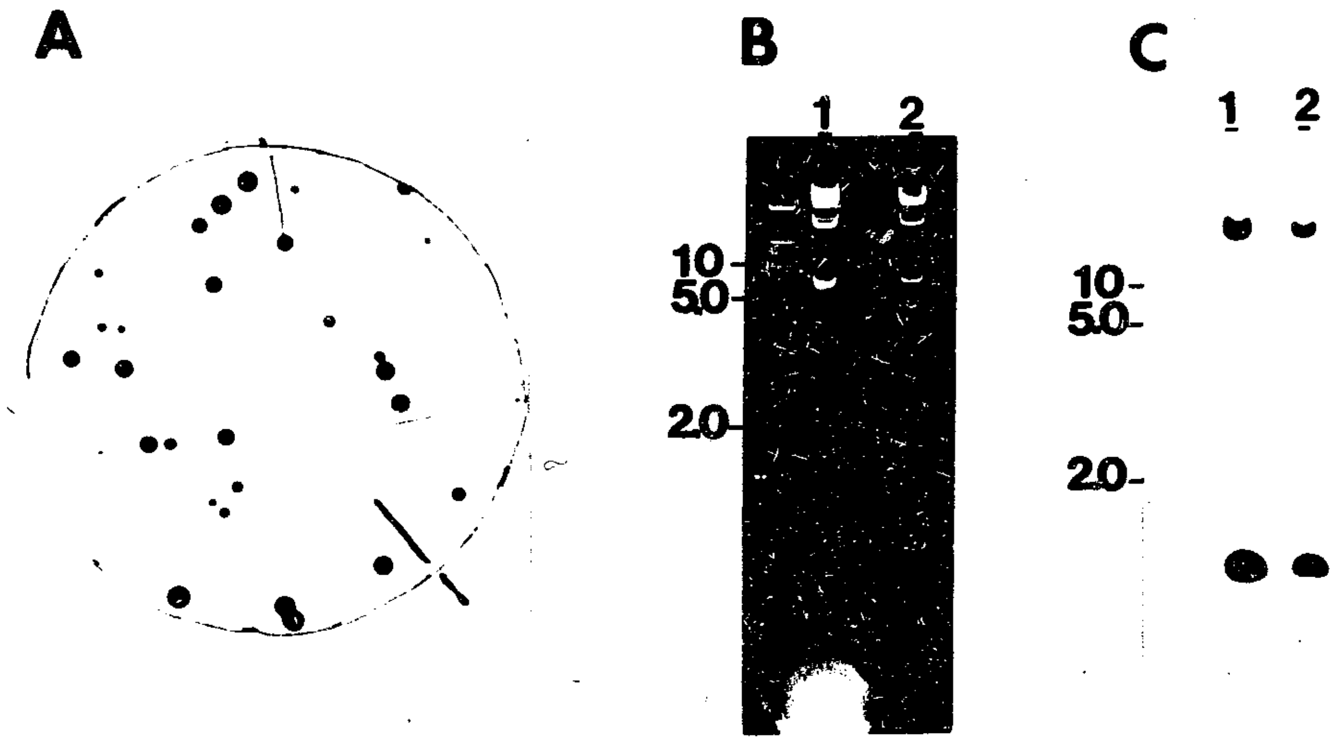
In an attempt to distinguish between pea mitochondrial *rps7* pseudogene and nuclear *rps7*-homologous clones, a PCR-generated probe (Wre) corresponding to the missing 5' region of the pea mitochondrial *rps7* pseudogene, was also used. Three clones giving positive hybridization signals to both Wre and Wrd were identified and two were taken

Table 5.1 Summary of results of screening pea nuclear, pea cDNA and soybean nuclear DNA libraries.

Libraries	Probes	No of Plaques	Results	Conclusion	Comments
Pea nuclear genomic library	Wrd	$6X10^5$	0.7% clones are positive.	Positive clones are from mitochondrial DNA.	One positive clone out of $8.4 X 10^5$ at 95% possibility
	Wrd	$6X10^5$	Three potentially positive clones. Two clones analyzed. They have same restriction pattern. 2.5 kb fragment partially sequenced.	Positive clones are from yeast-bacterial shuttle vector DNA.	
	Wre Wrd	$1.65X10^6$	Sixteen clones positive with Wre, none with both Wrd and Wre.	As above.	The PCR probe templates were gel-purified.
Pea nuclear cDNA library	Wrd	$4X10^5$	Fifteen potentially positive clones. Inserts: 1.35 -1.85 kb. Four clones analyzed and one clone sequenced.	Positive clones are from mitochondrial DNA: *orf228* plus *rps7* pseudogene (1.3 kb EcoRI fragment) and homology to the sunflower mitochondrial *orfH522* (0.5 kb EcoRI fragment).	
	Wrd Wrf	$5X10^5$	Sixteen potentially positive clones. Six clone inserts amplified by PCR. None hybridized with Wrf.	Positive clones are from mitochondrial DNA.	
Soybean nuclear genomic library	Wrd	$1.2X10^6$	Three potentially positive clones. Two analyzed and one sequenced.	Positive clones are from mitochondrial DNAs: *orf228* gene plus 100 bp *rps7*-homologous sequences. Downstream of them, there is a $tRNA^{phe}$ gene.	One positive clone out of $7X10^5$ at 99% possibility. *E. coli* K802 used.
	Wrd	$3.6X10^5$	0.8% clones are positive with the Wrd probe. Three clones analyzed.	As above.	Screened at low stringency conditions. *E. coli* MN538 used.

Figure 5.2 Analysis of the pea nuclear DNA library to identify the *rps7* gene.

(A) Plaque hybridization with the Wrd probe showing the result of the second round screening. (B) UV fluorescence profile of SalI-restricted pLR7-6 (Lane 1) and pLR7-9 DNA (Lane 2). (C) Southern blot analysis of SalI-restricted pLR7-6 and pLR7-9 DNA with the PCR-amplified Wrd probe. (D) Homology between preliminary DNA sequence of positive clone and (i) the 5' *GAL1-GAL10* promoter region, (ii) 3' *GAL1-GAL10* promoter region and (iii) the *URA3* gene of *Saccharomyces cerevisiae* as determined by Microgenie data searches. The underlined sequence indicates the multiple cloning site. Asterisks represent sequence differences.


A
B
1
2
10
5.0
2.0
C
1
2
10
5.0
2.0

D

(i)

```
1     TTATCAGCAACAACACAGTCATATCCATTCTCAATTAGCTCTACCACAGTGTGTGAACCA
13    TTATCAGCAACAACACAGTCATATCCATTCTCAATTAGCTCTACCACAGTGTGTGAACCA

61    ATGTATCCAGCACCACCTGTAACCAAAACAATTTTAGAAGTACTTTCACTTTGTAACTGA
73    ATGTATCCAGCACCACCTGTAACCAAAACAATTTTAGAAGTACTTTCACTTTGTAACTGA

121   GCTGTCATTTATATTGAATTTTCAAAAATTCTTACTTTTTTTT GGATGGACGCAAAGAA
133   GCTGTCATTTATATTGAATTTTCAAAATTTCTTACTTTTTTTTTGGATGGACGCAAAGAA
                                *                *
180   GTTTAATAATCATATTACATGGCATTACCACCATATACATATCCATAT         227
193   GTTTAATAATCATATTACATGGCATTACCACCATATACATATCCATAT         240
```

(ii)

```
1     GCTCCGAACAATAAAGATTCTACAATACTGACTTTTATGGTTATGAAGAGGAAAAATTGG
461   GCTCCGAACAATAAAGATTCTACAATACTAGCTTTTATGGTTATGAAGAGGAAAAATTGG
                                   **
61    CAGTAACTTGGCCCCACAAACCTTCAAATTAACGAATCAAATTAACAACCATAGGACGAT
521   CAGTAACTTGGCCCCACAAACCTTCAAATTAACGAATCAAATTAACAACCATAGGACGAT

121   AATGCGATTAGTTT TTAGCCTTATTTCTGGGGTAATTAATCAGCGAAGCGATGATTTTT
581   AATGCGATTAGTTTTTTAGCCTTATTTCTGGGGTAATTAATCAGCGAAGCGATGATTTTT
                    *
180   GATCTATTAACAGATATATAAATGCAAAAGCTGCATAACCACTTTAACTAATACTTTCAA
641   GATCTATTAACAGATATATAAATGCAAAAGCTGCATAACCACTTTAACTAATACTTTCAA

240   CATTTTCGGTTTGTGTTACTTCTTATTCAAATGTCATAAAAGTATCAACAAAAAATTGTT
721   CATTTTCGGTTTGTGTTACTTCTTATTCAAATGTCATAAAAGTATCAACAAAAAATTGTT

300   AATATACCTCTATACTTTAACGTCAAGGAGAAAAATTCGAGCTCGGTACCCGGGGATCC
781   AATATACCTCTATACTTTAACGTCAAGGAGAAAAAACTATA
                                         ******
```

(iii)

```
1     CTTTTCAATTCATCATTTTTTTTTTATTCTTTTTTTTTGATTTCGGTTTCCTTGAAATTTT
3     CTTTGCAATTCATCATTGGTTTTTTATTCTTTTTTTTTGATTTCGGTTTCTTTGAAATTTT
          *            **                              *
61    TTTGATTCGGTAATCTCCGAACAGAAGGAAGAACGAAGGAAGGAGCACAGACTTAGATTG
63    TTTGATTCGGTAATCTCCGAACAGAAGGAAGAACGAAGGAAGGAGCACAGACTTAGATTG

121   GTATATATACGCATATGTAGTGTTGAAGAAACATGAAATTGCCCAGTATTCTTAACCCAA
123   GTATATATACGCATATGTAGTGTTGAAGAAACATGAAATTGCCCAGTATTCTTAACCCAA

181   CTGCACAGAACAAAAACCTGCAGGAAACGAAGATAAATCATGTGCAAAGCTACATATAAG
183   CTGCACAGAACAAAAACCTGCAGGAAACGAAGATAAATCATGTCGAAAGCTACATATAAG

241   GAAGCGTGCTGCTACTCATCCTAGTCCTGTTGCTGCCAAGCTATTTAATATCATGCACGA
243   GAAGCGTGCTGCTACTCATCCTAGTCCTGTTGCTGCCAAGCTATTTAATATCATGCACGA

301   AAAGCAAACAAACGTTGTGT CTTCATTGGATGTTGCTACCACCAAGGAATTACTGGAGT
303   AAAGCAAACAAAC TTGTGTGCTTCATTGGATGTTCGTACCACCAAGGAATTACTGGAGT
                   *      *              **
350    AGTTGAAGCATTAGGTCC AAA TTTGTTTCCATAAAA        398
352   TAGTTGAAGCATTAGGTCCCAAAATTTGTTTACTAAAAA        404
      *                  *   *       * **
```

through further plaque purification screenings (Fig. 5.2). Recombinant lambda DNAs were isolated and the two clones were seen to have the same restriction pattern based on two enzymes, but are quite different from digested EMBL-3 lambda vector DNA patterns. The clones were also shown to hybridize with the wheat *rps7* probe Wrd (Fig 5.2).

A 2.5 kb HindIII hybridizing fragment was used for subcloning and partial sequencing. The sequences showed no homology to the wheat mitochondrial *rps7* DNA sequences (Fig 5.2). One of these regions was identical to part of the *URA3* gene from yeast (*Saccharomyces cerevisiae*) (Fig 5.2D iii) (Yocum et al. 1984); two clones are from the *GAL1-GAL10* promoter region of *Saccharomyces cerevisiae* (Fig 5.2D i, ii) and a 26 bp stretch had 100% identity to the multiple cloning site of pUC19 vectors (Fig 5.2D ii). The *GAL1-GAL10* promoter and the *URA3* gene of *Saccharomyces cerevisiae* have been widely used to construct yeast-bacterial shuttle vectors (West et al. 1984). The simplest explanation is that the pea genomic library is contaminated with yeast-bacterial shuttle vector DNA and such clones were detected because the PCR-generated probes contain small amounts of polylinker sequence. This could occur because Taq polymerase can synthesize DNA sequences in the M13 vector beyond both oligomers during early stages of amplification. It appears that the vector sequences containing the multiple cloning site hybridized with the polylinker sequence of the cloned yeast DNA. It was later confirmed by Clontech that the pea genomic library had been contaminated with yeast tRNA, yeast genomic DNA and some other plant genomic DNAs. However, this contamination is different from those mentioned above.

Successive hybridizations with two different probes were carried out to try to overcome the contamination problem. The plaque filters were first screened with a PCR probe

(Wre) from the 5' wheat *rps7* region which because it was PCR-generated by LB27 and reverse vector flanking primers, will hybridize with both the nuclear *rps7* gene and cloned yeast DNA. The filters were then restripped and screened with the *rps7* probe (Wrd), which can hybridize with the pea mitochondrial and nuclear-encoded *rps7* genes, but not with the yeast DNA because it was made using mtDNA-specific PCR primers. Before PCR probes were prepared, template DNAs were gel-purified. If a clone was positive with both screenings, this clone should contain the nuclear-encoded mitochondrial *rps7*-homologous sequences. Approximately 1.65 X 10^6 plaques were screened as described above. However, no plaques hybridized with both probes.

There are three possibilities to explain the lack of clones hybridizing with both probes. The first is that the screening methods used in this experiment were not sensitive enough because of inappropriate hybridization conditions or low sequence similarity between the wheat mitochondrial and pea nuclear-encoded *rps7* genes. However, clear signals were seen in Southern hybridization experiments (Fig 5.1). Also, since the short stretches of contaminating bacteria-yeast vector DNA did hybridize with the wheat mitochondrial *rps7* PCR-generated probe, it supports the belief that the methods used to screen the pea genomic library were sensitive enough to isolate the desired clones.

The second is that 1.65 X 10^6 plaques which were screened were not enough to contain the whole pea nuclear genome. Given that the total pea genome including organellar DNAs is approximately 4.5 X 10^9 bp (Murray et al. 1977), when only 1.29 X 10^6 plaques were screened, we should have had a 99% chance of finding a clone containing the nuclear-encoded *rps7* gene, but we can not rule out that there may be bias in this library with some sequences being over-represented.

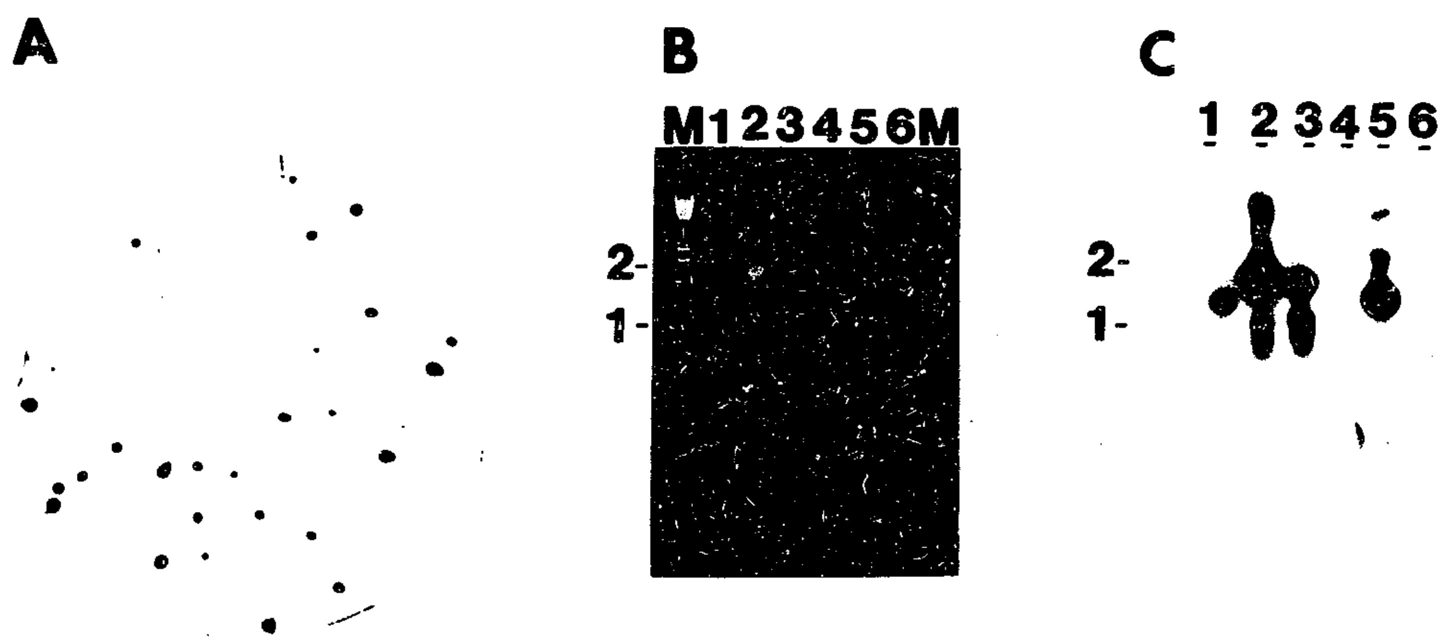
Lastly, the pea genomic library might not be large enough to contain all regions of the genome. Minor errors during titering and dilution of the library might have affected the titer values (pfu/ml) of genomic libraries because of small volumes used.

5.3 Searching for *rps7*-homologous sequences by screening the pea cDNA library.

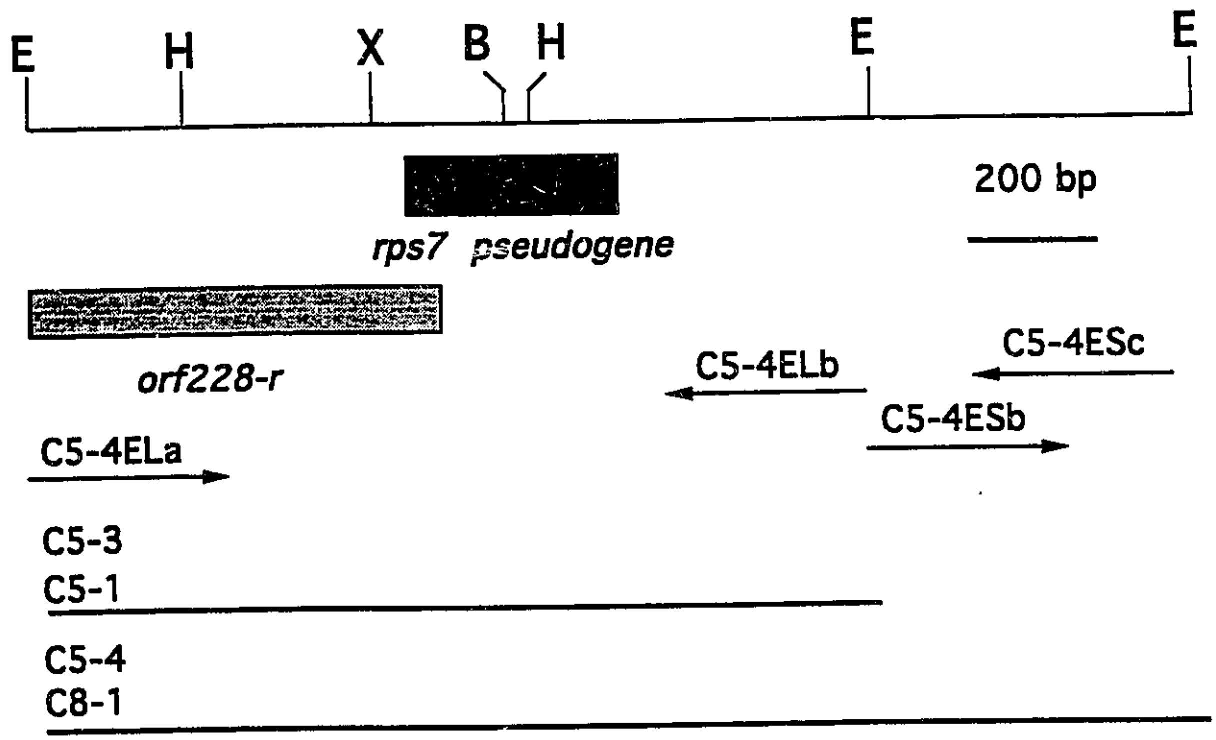
Because of the problems with the pea nuclear DNA bank, Clontech provided a pea cDNA library. If the *rps7* cDNA clone could be identified, its sequence could be used to design primers for PCR experiments to amplify the corresponding genomic sequences. A total of approximately 4 X 10^5 plaques from the pea cDNA library were screened with the probe Wrd (Fig 3.3 and Table 5.1) and fifteen potentially-positive clones were subjected to second and third screenings (Fig 5.3). Six of these cDNA clones were PCR-amplified using the lambda gt11 reverse and forward flanking primers (see Chapter 2.9). Fig 5.3B showed that four clones (C5-1, C5-2, C5-3 and C8-1) have inserts of either 1.35 kb or 1.85 kb (Fig 5.3B). Two (C5-4 and C8-4) did not have inserts in this particular experiment due to the method used. In some of the PCR amplification experiments, some minor products (Fig 5.3B lane 2 and 4) were also observed, which may be due to the buffers used and high GC content of the primers. The Southern blot in Fig 5.3C shows that the Wrd probe hybridizes with the major product but not the minor one. Two clones of each size were further analyzed. When the 1.85 kb clones were restricted with EcoRI, they yielded 1.35 kb and 0.5 kb fragments. These were subcloned into M13 vector and partially sequenced. The large EcoRI fragment contains the pea mitochondrial *orf228-r* and *rps7* pseudogene region (Fig 5.3D and data not shown).

Figure 5.3. Analysis of the pea cDNA library to identify the *rps7* gene.

(A) Plaque hybridization with the Wrd probe. (B) UV fluorescence profile of C5-1 (1), C5-2 (2), C5-3 (3), C5-4 (4), C8-1 (5) and C8-4 (6) PCR amplification products. (C) Southern blot analysis with the Wrd probe. (D) Physical map of the pea cDNA clones. The shaded and solid bars indicate the pea mitochondrial *orf228-r* gene and the pea mitochondrial *rps7* pseudogene, respectively. The dark lines represent cDNA clones analyzed by restriction enzymes. E: EcoRI; B: BamHI; H: HindIII; X: XhoI. (E) DNA sequence of the 450 bp EcoRI fragment from the cDNA insert. The asterisks indicate the nucleotide differences. (1). The pea mitochondrial DNA sequence. (2). The sunflower mitochondrial *orfH522* gene.


A
B
M123456M
2-
1-
C
1 2 3 4 5 6
2-
1-

D


E
H
X
B
H
E
E
200 bp
rps7 pseudogene
orf228-r
C5-4ESc
C5-4ELb
C5-4ESb
C5-4ELa
C5-3
C5-1
C5-4
C8-1

E

```
1 AATTCACGATAGATTTCTGAATTGCGGCCGCTTAGGTTTACTTGTTATATAAAGAACGTC
2 AATTCTT ATAGATTTCGGAATTTCTCCCATTTTGGGTTAATTCCTTGATAAAGAAATTC
       ***         *     * **  **  *  *   *  ** **         *

  AGACTTCCATCCATGAGATGGCGCGGATAGGCATGTTCAAGGGAGCTAAGCGGGAA ACC
  AAAGTACCTTCCATCAGATCCTTCTGAAAAGGATGCTCTAAAGCGCTCACCGGCTCCACA
   * * *  *     *    **** *  * *   *  * ** *   * *   ****  *

  TTCTCCTGAATGAAGTCGTAGGCTTCC TACCAACTATTTGAATAGGGAGATAGCTCCAT
  GAAGCCTCAAGGAAATTATAGGCCTCCATACGGA TATTTTGATATGGAGAAAGTTCCAT
  ****   *  *   * **     *   *   ** *     **   *     *  *

  AATTCGAGCAAGATTCTGCTCATCAAGCATCAGATTGAAGAGTTTATCCTGCAG ACCGG
  TATACGAGAAATCTTTTGCTCACCAAGCATAAGATGGAAGAGCCTATCTTGCAATAGACC
  *  *    *  ** *      *       *    *      **    *    ** ****

  CCTTTTTATCTAATACGCGTAGTTCGAGATACTCTCGATCAAAGATTTCCCAGTAGTGAT
  CCTTTCTTTCTAATACTGTGAGTTCAAAATATTCTCTATCCAGGACCGTTCTATAATGGG
       * *        ****     * *   *    *   * *  ***** **  *  **

  CTACGTTTAGGGGCTTGATCGAAGTGCTCTCGAACCAGCCTCGCGTAATCTCCCGGGTTA
  GAACATCGAGAG CTTGCTCGAATTCTTCCCTAACAAGAACCTCATACTC CCCAGG  A
  **  * **  * *    *     * **  * *   *  *** * *  *  *   *  **

  GGTTATTATGAGGGGGGATATTGTAGTAATGCTGGTTCTCGAGAACCCA ATTCGCTCAA
    TTATTCTGGGGAGGAAGGTTGTAGTAATCCTGGTCTCAAGATTGCGTATGCGAGTAA
  **     *  *  *  * **          *    *    *   *** **  *  *** *

  ATATTTCCGCTTCATTTTCTGCTGCGATAACATTTCGAAAAGTAGCCAGAGACTC
  AAATTTCAGCCTCGTTCTCGGCTGCTATTTCGGCTCTAAACGTAGCCAATGATTC
   *     *  *  *  *  *     *  ** ***  *   *       **  *
```

Therefore, this cDNA clone comes from contaminating mitochondrial DNA instead of cytoplasmic mRNA. The small insert was seen to share 60% identity with part of the mitochondrial *orfH522* gene in sunflower (Fig 5.3E) (Köhler et al. 1991), where it is co-transcribed with the *atpa* gene. The *orfH522* gene encodes a 15 kDa protein which is believed to be associated with cytoplasmic male-sterility (CMS) (Monéger et al. 1994). It is not clear whether this small EcoRI fragment is directly linked to the larger one in pea mitochondria or they have been ligated together in the process of constructing the cDNA library.

To determine whether any of the potentially positive clones were not mitochondrial *rps7* pseudogene ones, the filters containing positive clones were stripped and rehybridized at 36 °C, a condition of very low stringency with the probe Wrf, which was generated from a purified PCR product by random oligomer labelling (Table 5.1). No hybridization signals were detected. Therefore these positive clones were concluded to be due to contaminating mitochondrial DNA.

Because the pea nuclear DNA and cDNA libraries were seen to be contaminated, reports from other laboratories were checked. A soybean cDNA library made by Clontech had also been reported to be contaminated with soybean mitochondrial DNA (Covello and Gray 1992). However, in that case nuclear-encoded *coxII* cDNA clones had been successfully isolated.

When an additional 5 X 10^5 plaques were screened with the probe Wrd, six clones were found to be positive after the second screening. The PCR-amplified inserts from these six clones were from 1 kb to 1.8 kb, but no hybridization signals with the 5' *rps7* Wrf probe were detected (data not shown). Furthermore, these clones failed to generate the PCR amplification products expected when primers specific to the *rps7* protein-coding region were used. These data suggest that the pea cDNA library was heavily contaminated with pea mitochondrial DNA.

In the pea cDNA library, all 21 potentially-positive clones analyzed turned out to be clones containing the mitochondrial *rps7* pseudogene and a region of the mitochondrial *orf228-r* gene. Northern blot analysis of poly(A) mRNAs with the Wrd probe showed strong non-discrete hybridization signals (data not shown), which may be due to partially degraded *rps7* mRNA. Once again, this supports the Southern blot data that a functional *rps7* gene is present in the nucleus.

The cDNA clone library prepared from cytoplasmic poly(A) mRNAs was contaminated with mitochondrial DNAs. One reason may be the method used in isolation of the cytoplasmic poly(A) mRNAs. If only the nuclei were pelleted, the mitochondria would remain in the supernatant with the cytosol RNAs, and be co-purified. Any sheared, linearized mitochondrial DNAs could then be ligated into vectors. The mitochondrial DNAs might be difficult to separate from the poly(A) mRNAs because of their size or (dA)-rich regions. In order to remove the contaminating mitochondrial DNAs, the poly(A) mRNAs could be treated with RNase-free DNase.

A total of 9 X 10^5 plaques screened should be large enough to represent the population of cytosolic mRNAs. Another explanation is that reverse transcriptase did not generate full-length cDNAs from mRNAs. Since only cDNAs larger than 0.5 kb were ligated into the lambda gt11 vector during the library construction, this results in the under-representation of some cDNAs.

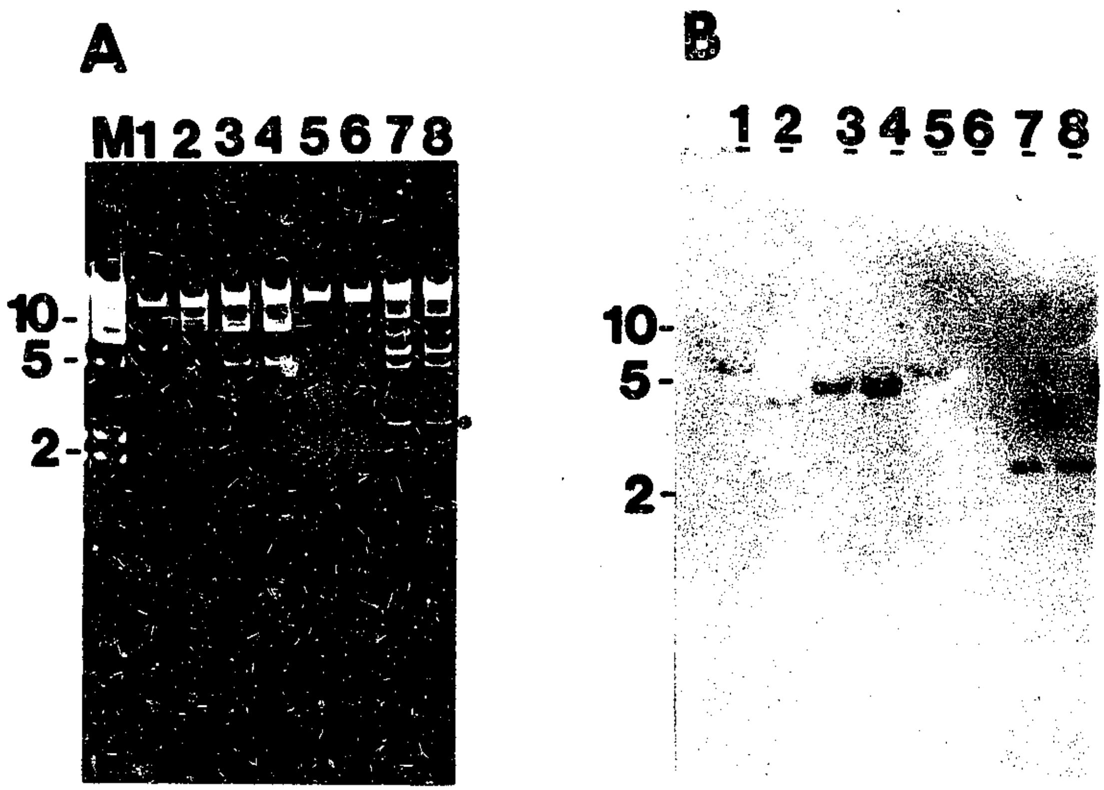
5.4 Searching for the soybean *rps7* gene by screening the soybean nuclear DNA library.

Since both pea nuclear DNA and cDNA libraries were contaminated, an attempt was made to identify the *rps7* gene by screening a soybean nuclear DNA library, again provided by Clontech. Since no hybridization signals had been detected in the soybean mitochondrial DNA with the *rps7*-specific Wrd probe (Fig 3.3), it was used as a probe to screen this bank. More than 1.2 X 10^6 plaques were screened as described above (Table 5.1). Since the soybean nuclear DNA is about 1 - 2 X 10^9 bp (Goldberg 1979; Curley et al. 1979), one positive clone out of 0.7 x 10^6 is expected with 99% confidence. Six weakly-hybridizing clones were observed. Two of these clones underwent second and third screenings with the Wrd probe (data not shown). DNA from one of these potentially-positive clones was isolated and a 4.5 kb SacI fragment giving a weak hybridization signal using the Wrd probe, was subcloned into pUC13 vector (Fig 5.4A). Fig 5.4B shows the Southern blot analysis of these lambda DNA clones. Subsequent sequence analysis (Fig 5.4C) showed that the clone contained a soybean mitochondrial DNA fragment containing a 92 bp DNA region which is homologous to the wheat mitochondrial *rps7* gene. These clones also contained the soybean mitochondrial *orf228-r* gene which overlaps with the *rps7* short segment, but immediately after the stop codon there is a deletion of nine nucleotides compared to the pea one. Further downstream of the soybean *rps7* pseudogene, there is a stretch which is identical to the *Oenothera* mitochondrial $tRNA^{Phe}$ gene (Binder et al. 1990).

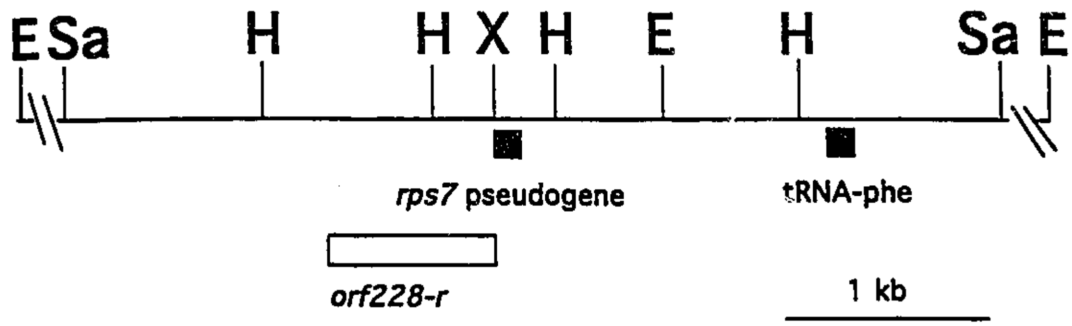
Since the hybridization signals in the experiments were very weak and in fact due to the short stretch of mitochondrial *rps7* homologous DNA sequence, there may have been a problem with the host strain and/or the techniques used to screen the soybean nuclear DNA

Fig 5.4. Analysis of the soybean nuclear DNA library to identify the *rps7* gene.

(A) UV fluorescence profile of SacI- and XhoI-restricted S2a (lanes 1 and 5), S2b (lanes 2 and 6), S8 (lanes 3 and 7) and S10 (lanes 4 and 8) clones. The black dot indicates the hybridizing DNA fragment. (B) Southern blot analysis of SacI- and XhoI-restricted S2a, S2b, S8 and S10 clones with the Wrd probe. (C). Physical map of the soybean *rps7*-homologous sequence. The solid and open bars represents a stretch of 92 bp DNA sequence homologous to the wheat mitochondrial *rps7* gene and a region homologous to the pea mitochondrial *orf228-r* gene, respectively. The shadowed bar indicates DNA sequences identical to *Oenothera* mitochondrial *tRNA^{phe}* gene (Binder et al. 1990). E: EcoRI; H: HindIII; Sa: SacI; X: XhoI. (D). Comparison of the mitochondrial *rps7*-homologous DNA sequences between pea and soybean. The asterisks indicate identical nucleotides. (E) Preliminary DNA sequence containing the *tRNA-$_{Phe}$* gene which is underlined. Preliminary DNA sequence of the soybean mitochondrial orf228-r and rps7 pseudogene region is seen in Appendix D.


A
M 1 2 3 4 5 6 7 8
10-
5-
2-
B
1 2 3 4 5 6 7 8
10-
5-
2-

C



D

```
Wheat     AAAAC-TT-TTCACCGCCTAGCTCGAACTGAACGCGATGTAATAAAAC
Pea       AAAACCTAGTTTACTTCCCAGCTCAAACTGAACGCGATGTAATCAAAC
Soybean   AAAACCTAGTTTACTTCCCAGCTCAAACTGAACGC----------AAAC
          ***** *  ** **  ** ***** ***********          ****

          TTATGGTTGACGCCGTAGATAATATAAAGCCAATATGCGAAGTGGTCAAAGTA
          TTATGGTTGACGCCGTAGAGAATATAAAGCCCATATGCGAAGTGGAAAAAGTA
          TTATGGTTGACGCCGTAGAGAATATAAAGCCCATATGCGAAGTGGAAAAAGTA
          ******************* ************ **************  ******
```

E

CCAAGCTTGACGAACGAAAAAAGCCCTTTCTTTTAAGTTATTAGTAAAGGGCTTTTCCCT

TGGTAAGTAGGGCGCTATTCGATAAGAAAACATACTTTAGGAAAGTGGTTCAGGTAGCTC

AGCTGGTTAGAGCAAAGGACTGAAAATCCTTGTGTCAGTGGTTCGAATCCACTTCTAAGC

GGGCTTTGGGTCCAAAGGCAAGGTAGCCCGAGCCGTATTTTTTTAATTCAAGTGAAAGA

GGAAGCAAAAAAATGTGAGCGCTCCTGCACTATACGGCTTTTTTGGCGTCGCCTATCTGG

AGGCAGATTTTTCTAAAAAAAGCGGTTGATTACACGGATTGG library because of modifications by Clontech. Approximately 0.36 X 10^6 plaques from the soybean nuclear genomic library were used to infect the *E. coli* host strain NM538, a permissive strain (Kuziel and Trucker 1987; Ovchinnikov et al. 1979) and screenings with the wheat mitochondrial *rps7* Wrd probe at low stringency conditions of 36 °C were carried out. One of the previous twenty plaque lifts screened above was used as a control. About 0.8% plaques were found to show weak or strong hybridization signals, but no hybridization signals were found for the control (data not shown). The difference in hybridization signals between the two host stains might reflect different levels of lambda DNA generated. This high number of positive plaques further supported the idea that they were the soybean mitochondrial *orf228-r* and *rps7* pseudogene clones instead of ones containing the soybean nuclear DNA. DNAs were isolated from three positive clones. One of them has the same restriction patterns as the one analyzed above. Fig 5.4 shows that two clones (S8 and S10) have the same restriction patterns, but the other two (S2a and S2b) have different restriction patterns. The Southern blot analysis in Fig 5.4B shows that they all have very weak hybridization signals. It is believed that the differences in restriction patterns may reflect cloning sites which were generated by partial restriction during construction of the genomic library.

The earlier Southern blot analysis had indicated that the mitochondrial genome of soybean (var. Maple Arrow) has no *rps7*-homologous DNA sequences (Fig 3.8). However, in this experiment, the mitochondrial genome from the soybean (var. Resnick) has a 92 bp DNA sequence homologous to the wheat mitochondrial *rps7* gene. One possibility is that the differences reflect different varieties of soybean. The second less likely possibility is that because the conditions for Southern blot analysis of the soybean mitochondrial DNA were more stringent than those of the plaque hybridizations, the short mitochondrial *rps7*-homologous region was not detected.

Problems with these clone banks arose because (1) nuclear DNAs were contaminated with mitochondrial DNA and other sources of DNAs; (2) the cytosolic poly(A) mRNAs were heavily contaminated with mitochondrial DNA and (3) perhaps the clone banks were too small to contain the entire population of DNA fragments or mRNAs. Another approach to find the nuclear-encoded *rps7* homologous DNA sequence would be to use PCR amplification. A pair of oligomers could be designed based on the conserved coding region of the mitochondrial *rps7* gene and used to amplify legume nuclear template DNA. Preliminary experiments were conducted and an expected PCR product of 0.3 kb was found. Another set of reverse oligomers could be designed for inverse-PCR to obtain flanking sequences. However, there was not enough time to pursue this approach further.

CHAPTER 6 DISCUSSION

6.1 Bacterial nature of the plant mitochondrial translation machinery.

Because mitochondria and chloroplasts have originated from bacteria, their gene expression machinery is expected to more closely resemble bacterial ones than eukaryotic nuclear ones (Gray 1983, Gray 1992). Indeed the comparison of genes encoding components involved in transcription and translation show clear similarities between organellar and bacterial machinery.

Plant organellar ribosomal RNAs, tRNAs and ribosomal proteins share strong sequence similarity with eubacterial ones (Gray 1992). Chloroplast ribosomes have a very similar size to the bacterial ones, being 70S with 50S and 30S subunits. The 23S and 16S ribosomal RNAs not only have a conserved secondary structure, but also conserved nucleotide sequences (reviewed in Gray 1990). On the other hand, mitochondrial ribosomes are variable among different eukaryotes. Plant mitochondrial rRNAs are more similar to eubacterial ones in sequence and secondary structure, but much larger in size (Gray 1990). The wheat mitochondrial 18S and 26S rRNAs share about 84% similarity with the *E. coli* 16S rRNA secondary structure and 78% similarity with the 23S secondary structure. Within these conserved regions, they show 77% and 73% sequence identity with their *E. coli* homologues, respectively (Gray 1990).

In *E. coli*, genes for 21 small ribosomal proteins and 31 large ribosomal proteins are clustered together to form six operons. In the chloroplast genomes of land plants, a total of

21 ribosomal protein genes have been identified and they are clustered together to form operons similar to bacterial ones. In addition, their transcriptional signals can be recognized by *E. coli* RNA polymerase (Subramanian 1993; Sugiura 1993).

An interesting question is whether the ribosomal protein genes are organized in a bacterial fashion in the mitochondria of land plants as they are in the chloroplast. Another question is whether all bacterial counterparts of the ribosomal proteins are present in plant mitochondrial ribosomes. In the liverwort mitochondrial genome, genes for 12 small subunit ribosomal proteins and 4 large subunit ribosomal proteins have been found. Except for the *rps2* and *rps4* genes, they are clustered together to form two operons similar to the *E. coli* ones (Oda et al. 1992; Takemura et al. 1992). This further supports the idea of a mitochondrial endosymbiotic origin and suggests that these ribosomal protein genes have maintained an ancestral organization even though some ribosomal protein genes are no longer present because of transfer to the nucleus (Oda et al. 1992; Takemura et al. 1992).

The mitochondrial genomes of flowering plants differ from those of liverwort in that they have recombinationally-active repeat elements (Hanson and Folkerts 1992), which make them more flexible and may disrupt gene organizations that were present in their ancestor. Ribosomal protein genes are found in some cases to be clustered with other ribosomal protein genes or with respiratory chain genes and the organization can differ among plants (Table 6.1). The example of the largest operon identified so far is the *Brassica napus rps3, rpl16, rpl5* and *rps14* genes clustered with the *trnD, trnK* and *cob* genes (Yei et al. 1993). The four ribosomal protein genes have an organization similar to that which would be found by the fusion of the *E. coli* S10 and *spc* operons (Nomura et al. 1980). The mitochondrial *rps3/rpl16*

Table 6.1 Genomic organization of mitochondrial ribosomal protein genes in flowering plants.

The symbol Φ represents a pseudogene. Data are taken from (a) Hunt and Newton 1991, (b) Gualberto et al. 1991, (c) Bland et al. 1986, (d) Gonzalez et al. 1993, (e) Schuster 1993, (f)Wissinger et al. 1990, (g) Bock et al. 1993, (h) Schuster et al. 1990, (i)Rasmussen and Hanson 1989, (j) Sünkel et al. 1994 (k) Conklin and Hanson 1991, (l) Grohmann et al 1993, (m) Vitart et al. 1992, (n) Wahleithner and Wolstenholme 1988, (o) Yei et al 1993, (p) Suzuki et al. 1991.

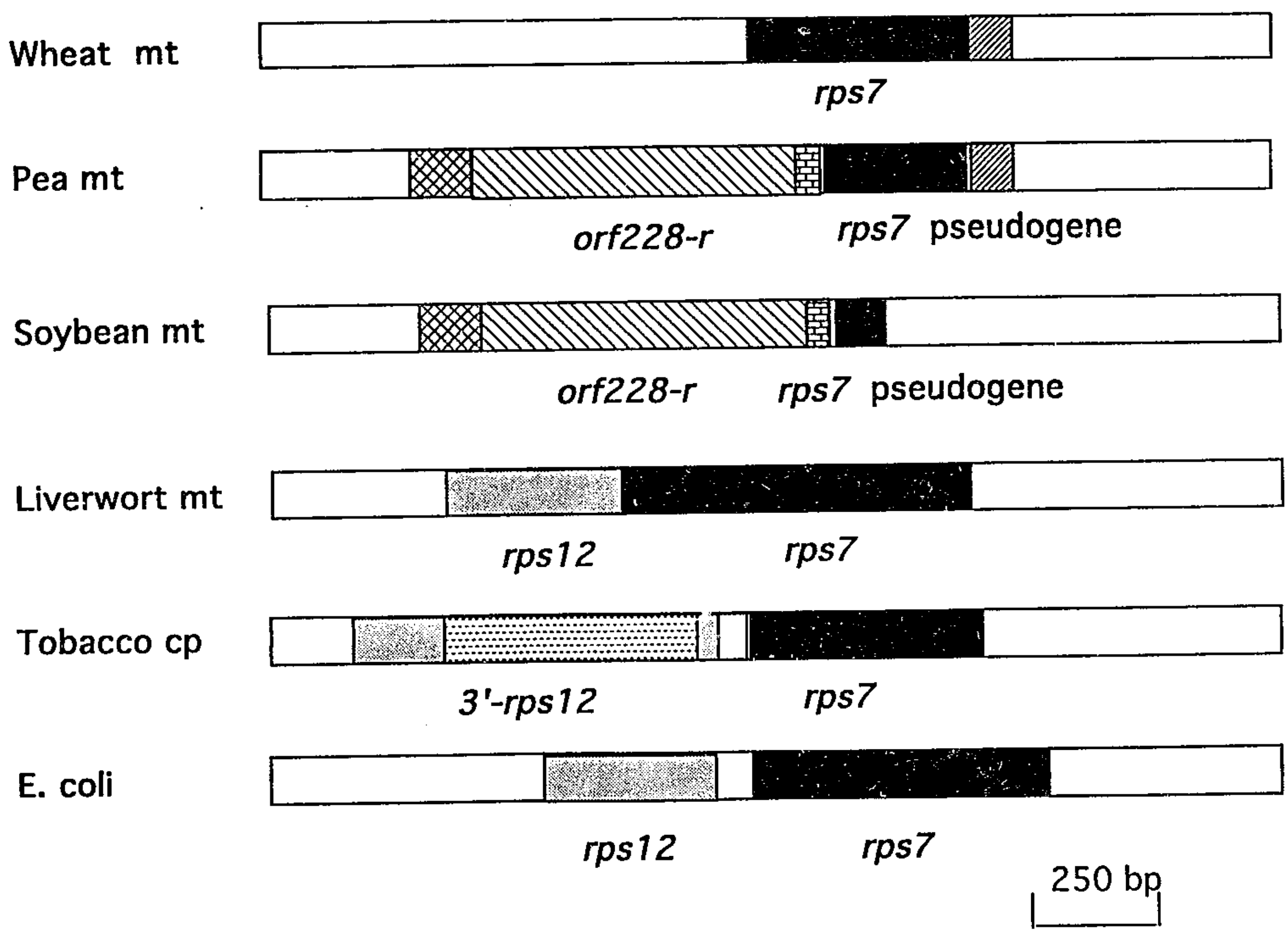
Plants	Co-transcribed mitochondrial ribosomal protein genes	References
Maize	**rps19 - rps3 - rpl16*	(a)
	atp9 - nad3 - rps12	(b)
	rps13 - nad1 b/c	(c)
Wheat	*nad3 - rps12*	(b)
	orf589 - rps1A	(d)
Oenothera	*rps15 - nad3 - *rps12*	(e)
	coxI - rps13 - nad1 b/c	(f)
	**rps19 - rps3 - rpl16*	(g)
	**rps12 - orf206 - rps14 - cob*	(h)
Petunia	*nad3 - rps12*	(i)
	rps19 - rps3 - rpl16 - cob	(k)
Arabidopsis	*nad3 - rps12*	(l)
	rpl2 - orfx	(j)
	*rpl5 - *rps12 - cob*	(l)
tobacco	*orf1 - rpl2 - orf2*	(m)
	atp9 - rps13 - nad1 b/c	(c)
Broad bean	**rpl5 - rps14 - cob*	(n)
Brassica	*tnaD - tnaK - rps3 - rpl16 - rpl5 - rps14 - cob*	(o)
Rice	*nad3 - rps12*	(p)

genes in several flowering plants have also been found to have a similar arrangement to *E. coli* ones (Brennicke et al. 1993). Why are ribosomal protein genes located close to other non-ribosomal protein genes? One possibility is that this arrangement has some advantage to control mitochondrial gene expression at the transcriptional level. Another advantage might be to co-ordinate mitochondrial and nuclear mitochondrial gene expression since components of respiratory chain and transcriptional and translational machinery are encoded by both the mitochondrion and the nucleus.

In this research project, the *rps7* gene has been characterized in wheat mitochondria and its status was examined in the mitochondrion of several legumes. The wheat mitochondrial *rps7* gene is not clustered together with other genes. Instead it is located 4.9 kb upstream of one copy of the 26S rRNA gene and about 8 kb downstream of one of three 18S-5S rRNA gene copies (named G2) (Lejeune et al. 1987). The pea mitochondrial genome has a truncated *rps7* pseudogene of 326 bp and the soybean mitochondrial truncated *rps7* pseudogene is even shorter at 92 bp (Fig 6.1). Apparently, the fusion of the *orf228-r* gene and *rps7* pseudogene in legumes is a very recent event. However, it is not clear whether the *orf228-r* and *rps7* genes were co-transcribed before this deletion event. In *E. coli*, the S7 ribosomal protein gene is separated by 96 nt from the S12 ribosomal protein gene and located in the str (S12-S7-fus-tufA) operon (Post and Nomura 1980). Interestingly in liverwort mitochondria, the *rps7* and *rps12* genes have maintained the ancestral bacterial gene order (Oda et al. 1992). However, the termination codon of the liverwort mitochondrial *rps12* gene overlaps two nucleotides with the initiation codon of the *rps7* gene because the *rps7* gene is unusually long (Fig 6.1). On the other hand, it can not be ruled out that the liverwort

Figure 6.1. Schematic comparison of the mitochondrial, chloroplast and bacterial *rps7*-homologous DNA sequences.

Data include liverwort mitochondrial *rps7* gene (Oda et al. 1992), tobacco chloroplast *rps7* gene (Shinozaki et al. 1986) and *E. coli rps7* gene (Post and Nomura 1980). The black and hatched bars represent the *rps7*-homologous DNA sequence and the mitochondrial *orf228-r* gene, respectively. The cross-hatched and narrow-hatched bars indicate the homologous sequences from the upstream and downstream region, respectively. The grey bars indicate the *rps12* gene; the stippled bar represent an intron in *rps12* gene.



Wheat mt
rps7
Pea mt
orf228-r
rps7 pseudogene
Soybean mt
orf228-r
rps7 pseudogene
Liverwort mt
rps12
rps7
Tobacco cp
3'-rps12
rps7
E. coli
rps12
rps7
250 bp

mitochondrial one uses another codon as a initiation codon as found in certain non-plant mitochondria (Wolstenholme 1992). In the chloroplasts of most land plants, the 3' region of the *rps12* gene, which is a trans-spliced gene, and the *rps7* gene are clustered together to form a similar operon to the *E. coli* one (Fig 6.1). However, in *Chlamydomonas reinhardtii*, the chloroplast *rps7* and *rps12* genes are far apart in different transcriptional units. It should be noted that until recently the chloroplast *rps7* gene in *Chlamydomonas reinhardtii* was thought to be divided into two pieces, which might be independently translated (Gilham et al. 1991). However, it has been shown that several mistakes led to the wrong conclusion (Hauser et al. 1993). In *Paramecium* mitochondria, the *rps12* gene is closely linked to the *nad2* and *nad7* genes but may not be co-transcribed (Pritchard et al. 1990). The presence of recombinationally-active repeat elements downstream and upstream of the wheat mitochondrial *rps7* gene is consistent with recent DNA rearrangements (Fauron et al. 1991; Hanson and Folkerts 1992). It would be very interesting to know the organization of the *rps7* gene in the maize and rice mitochondrial genomes

6.2 Potential regulatory sequences for the wheat mitochondrial *rps7* gene.

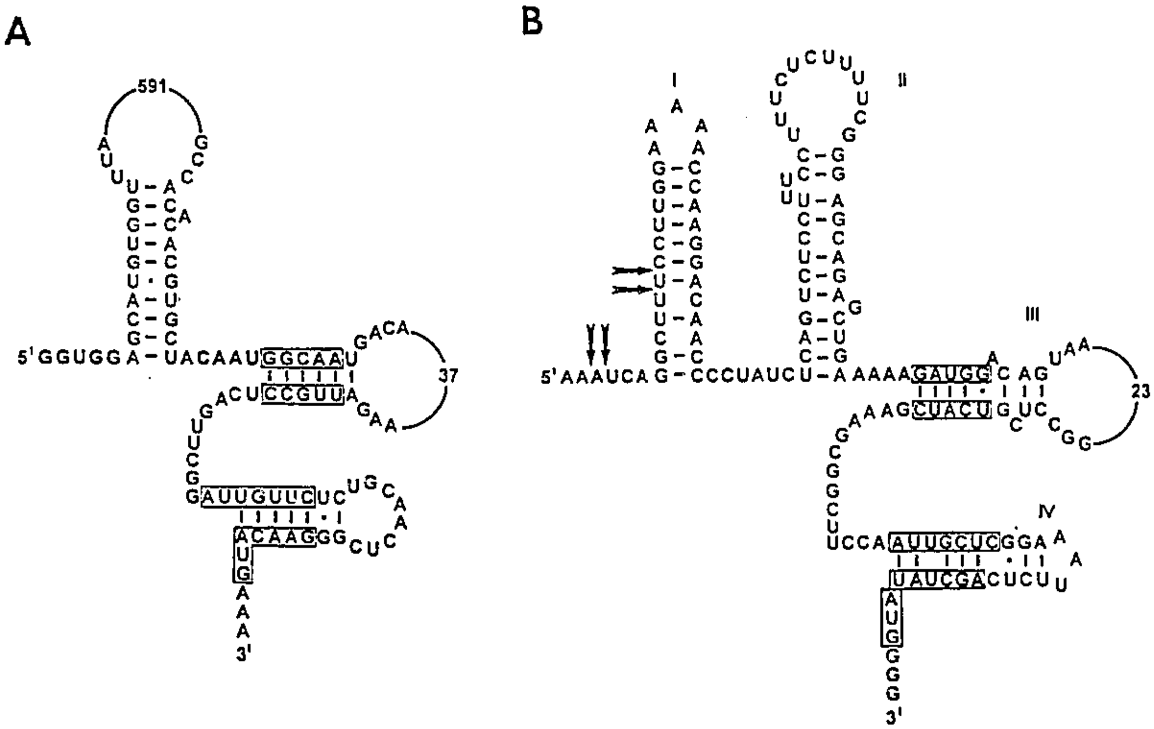
Sequences upstream of the wheat *rps7* gene are closely related to ones upstream of at least four other wheat mitochondrial genes (Bonen et al. 1984; Bonen et al. 1990; Bonen and Bird 1988). These repeat elements contain potential regulatory elements for translational and /or mRNA stability. Different plants have different repeat elements preceding various mitochondrial genes. A 657 bp sequence block is located upstream of each of the *Oenothera coxI* and *coxIII* genes and is believed to regulate gene expression (Hiesel et al. 1987). In soybean mitochondria, a stretch of DNA sequence containing the *tRNA*$^{f\text{-}met}$ gene is highly repeated and located upstream of several mitochondrial genes (Grabau 1987; Grabau et al. 1988; Zhuo D, Bruneau B, Ell M, Leclaire B and Bonen L, unpublished data). On the other hand, different mitochondrial genes may sometimes be regulated by different nuclear genes (Nagley 1991). This might be similar to processes in yeast mitochondria where specific nuclear-encoded proteins are known to mediate interactions between specific mRNA leader sequences and components of the mitochondrial translation system (Costanzo and Fox 1991). In the chloroplast in *Chlamydomonas reinhardtii*, the 5' leaders of both the *rps12* and *atpB* mRNAs are predicted to form stem-loop structures. Under normal growth conditions, only one nuclear-encoded protein binds to the 5' leader of the *rps12* mRNA, while three different proteins interact specifically with the 5' leader of the *atpB* mRNAs (Hauser et al. 1993). Under conditions of reduced protein synthesis, one of these latter proteins no longer binds to the 5' *atpB* leader. In contrast, three other proteins are now found to interact with the 5' leader of the *rps12* mRNA (Hauser et al. 1993).

When the 5' leader of the wheat mitochondrial *rps7* mRNA is examined, 77 nucleotides are found to separate the coding region from the repeat element, rather than it preceding the initiation codon directly as for the wheat *coxII*, *orf25* and *atp6* genes. One possibility is that this sequence may play a role in regulation of the *rps7* mRNA translation. In *E. coli*, the S7 ribosomal protein can function as a repressor and is involved in feedback regulation of ribosomal protein translation (Nomura et al. 1980). This is done through the binding of the S7 protein to the leader region which can form a secondary structure similar to those in 16S rRNA (Nomura et al. 1980). In *E. coli*, one common characteristic of such repressor ribosomal proteins is that they are among the "initial binding proteins" in the in vitro reconstitution of ribosomes and have strong and specific binding to rRNAs (Nomura et al. 1980; Moore and Capel 1988).

To address the question whether these *rps7* stem-loop structures might have functions similar to those in *E. coli*, I compared the potential secondary structures (Fig 6.2). When the RNA sequences upstream of the major *rps7* transcripts are folded into secondary structures using Microgenie programs, they form four hairpin structures (Fig 6.2B). The last three hairpins (II, III, IV) are similar to the wheat mitochondrial 18S rRNA stem-loop structures (Spencer et al. 1984) (Fig 6.2A) corresponding to those which interact with the S7 ribosomal protein in *E. coli* (Nomura et al. 1980). Even though the sequences are not identical to the corresponding regions, they are relatively conserved. The predicted thermodynamic stability of RNA folding is -2.6 units (KCal) for the stem-loop IV structure, compared to -6.0 of the corresponding 18S rRNA stem-loop structure. Differences may reflect that the S7 ribosomal protein binds to the 18S rRNA sequences more strongly than the sequences from the *rps7* transcripts. This suggests that the stem-loop structure in the *rps7* mRNA is less stable than that of the wheat 18S rRNA. A similar secondary structure is also observed in the upstream sequences of the tobacco chloroplast *rps7* mRNA and its functions were to be tested *in vitro* (Bourque et al. 1991). Therefore, the sequences upstream of the wheat *rps7* gene may play a role in regulating *rps7* gene expression. In addition, the hairpin I structure is very similar to those in the 5' regions of *E. coli* transcripts, which stabilize the transcript (Emory et al. 1992). Therefore, the leader sequences in the *rps7* transcripts may play a role in translational regulation and transcript stabilization in wheat mitochondria.

Fig 6.2 Predicted secondary structures of the wheat mitochondrial S7 binding sites on 18S rRNA and sequences upstream of the wheat *rps7* coding region.

(A) The secondary structure of the predicted S7 binding sites of 18S rRNA. (B) The secondary structure of the sequence upstream of the *rps7* coding region. The arrows indicate the 5' termini of the major *rps7* mRNA. The blocked sequences are shown to be conserved.


A
B
591
37
23
I
II
III
IV
5'GGUGGA
5'AAAUCA
3'

6.3 RNA editing in the wheat mitochondrial rps7 transcripts.

RNA editing in plant mitochondria and chloroplasts usually improves protein similarity. It was surprising to find that one of the three C-to-U changes in the *rps7* coding region decreases the similarity between the deduced wheat and liverwort S7 ribosomal proteins. This phenomenon has been called "wrong editing" and has been observed in several other ribosomal protein transcripts such as *rps3* and *rps19* (Schuster et al. 1991; Conklin and Hanson 1991). It is possible that RNA editing of conserved positions does not alter protein function because of compensating changes at other positions. In petunia mitochondria, RNA editing of *rps3/rpl16* transcripts creates a premature stop codon in the *rpl16* reading frame (Sutton et al. 1993). Therefore, in this case the mitochondrial *rpl16* gene becomes a pseudogene. It will be of interest to know whether this "wrong-editing" site is present in the *rps7* transcripts from other plants like rice and maize.

Editing at the silent position of a codon has been called "meaningless editing" (Conklin and Hanson, 1991; Bonnard et al. 1992). RNA editing is sometimes seen in only a fraction of the RNA molecules and is then called "partial editing". The silent *rps7* RNA editing site is edited only in some cDNA clones. Similarly, in wheat mitochondrial *nad7* transcripts, four out of five silent C-to-U changes are edited in only about 50% of the transcripts (Bonen et al. 1994). These partial editing sites may be "over-edited" and this may reflect older transcripts that are accessible to RNA editing machinery (Ishikawa and Kodowaki 1993). This idea has also been indirectly supported by observing that CGA codons have been edited to UGA stop codons in *coxI* and *rps3* coding sequences, although that was seen in only one out of six cDNA clones (Schuster and Brennicke 1991). It seems that the

RNA editing machinery might only recognize certain RNA secondary structure. If older transcripts have more chance to form certain secondary structures, they would be "over-edited".

Partial editing of non-silent sites has also been observed in the cDNA clones of the wheat *rps7* transcripts with different sites being unedited in different clones. Various partially edited transcripts have also been observed in transcripts of a number of flowering plant mitochondrial genes (Covello and Gray 1993; Grienenberger 1993). This observation of different partially edited cDNA clones supports the idea that RNA editing lacks polarity (Bonnard et al. 1992). This is in contrast to RNA editing in trypanosome mitochondria where RNA editing proceeds along the RNA molecule from the 3'-to-5' direction (reviewed in Benne 1993).

Minor differences in RNA editing have been observed between the wheat *rps7* transcripts from 24 hour germinating embryos and six-day etiolated seedlings. Partial editing is observed only in the cDNA clones from seedling RNA and the silent site is not edited in any cDNA clones from the germinating embryo RNA. It is generally believed that RNA editing mechanisms in plant chloroplasts and mitochondria are very similar (Gray and Covello 1993; Covello and Gray 1993). It has been reported that RNA editing of the spinach plastid *psbF* and *psbL* transcripts is associated with various developmental stages (Bock et al. 1993). Completely edited transcripts were observed only in etioplasts, whereas reduced editing was observed in seeds and roots (Bock et al. 1993). However, no experiments have been done to determine whether there is a relationship between RNA editing activity and the metabolic activity of plant tissues. RNA editing is also found to be developmentally regulated in trypanosome mitochondria (Stuart and Feagin 1992).

In order to know more about differences in RNA editing between newly-transcribed and mature RNA transcripts, oligomers specific for the precursor RNA transcripts were used in RT-PCR amplification experiments. However, these experiments were unsuccessful perhaps because of the low level of long unprocessed RNA transcripts or RNA secondary structure, which made it difficult to generate full-length cDNAs.

6.4 The pea mitochondrial *orf228-r* gene potentially has a role in cytochrome *c* biogenesis.

The *rps7* gene studies also led to identification of another plant mitochondrial gene *orf228-r* which is related to a bacterial gene involved in cytochrome c biogenesis. Most studies on cytochrome c/c1 biogenesis in eukaryotes have been carried out in *Saccharomyces cerevisiae* and *Neurospora crassa* where there are two mechanistically distinct routes (Hartl and Neupert 1990). The first pathway is that apocytochrome c is imported from the cytoplasm into the mitochondria via a direct translocation across the mitochondrial membrane to the intermembrane space. In the second pathway, apocytochrome c1 biogenesis includes two consecutive translocation steps. Initially, the apoprotein is directly transported into the mitochondrial matrix and the N-terminal part of the bipartite transmembrane signal peptide sequence is removed. The apoprotein is re-exported from the matrix to the outer surface of the inner mitochondrial membrane. This is followed by the formation of a complex between cytochrome c1 heme lyase and apocytochrome c1, the subsequent covalent linkage of (reduced) heme to the apoprotein and the removal of the second part of transmembrane signal peptide sequence (Nicholson et al. 1989). The first translocation pathway is thought to be simpler and evolutionarily novel whereas the second one, which seems to be gaining support now, has retained some characteristics of its bacterial origin (Nicholson et al. 1989). second model

The pea and soybean mitochondrial *orf228-r* deduced gene products share similarity with the *helC* protein from *Rhodobacter capsulatus*, a gram-negative photosynthetic bacteria (Beckman et al. 1992), and the *orf263* protein from *Bradyrhizobium japonicum* (Ramseier et al. 1991). The *helC/orf263* gene is believed to be involved in cytochrome c assembly along with *helA/cycV, helB/cycW, orf52/x, ccl1* and *ccl2* genes in *Rhodobacter capsulatus* and *Bradyrhizobium japonicum* (Ramseier et al. 1991; Beckman et al. 1992). The *helc/orf263* proteins are very hydrophobic membrane-binding proteins. In *Rhodobacter capsulatus* Hel$^-$ strains, all c-type cytochromes including the cytochrome bc_1 complex are missing. However, the Hel$^-$ strains synthesize normal amounts of b-type cytochromes except for cytochrome b (Beckman et al. 1992). It is thought that Hel gene products are necessary for translocation of heme to the bacterial periplasm (Beckman et al. 1992).

In addition, the *ccl1* and *ccl2* genes encode hydrophobic periplasmic proteins involved in c-type cytochrome biogenesis (Beckman et al. 1992), perhaps acting as specific heme chaperons, instead of the lyase that catalyze heme ligation (Beckman et al. 1992). It is interesting to note that a *ccl1* homologous gene has recently been observed in wheat (*orf589*), Oenothera (*orf577*), carrot (*orf579*) and liverwort (*orf509*) mitochondria (Schuster et al. 1993; Gonzalez et al. 1993). In addition, the liverwort mitochondrial *orf277* deduced protein shares similarity with those of the *helB* gene from *Rhodobacter capsulatus* and the *cycW* gene from

Bradyrhizobium japonicum (Schuster et al 1993). The chloroplasts also require c-type cytochrome in photosynthesis. Chloroplast *orf228* homologues have been identified, but they are very divergent (Table 4.1). This may be related to the fact that mitochondria have originated from the proteobacterial lineage including *Rhodobacter capsulatus* and *Bradyrhizobium japonicum.* All these data suggest that these genes which are involved in c-type cytochrome biogenesis are highly conserved in a wide range of species and organelles, including bacteria, chloroplasts and mitochondria from plants and protists. However, apparently *orf228* homologues are not present in the animal and fungal mitochondria and may have been transferred into the nucleus during very early stages of endosymbiosis.

6.5 Gene transfer of the mitochondrial *rps7* gene into the nucleus during evolution.

Southern hybridization data suggest that the mitochondrial genomes of soybean and bean (data not shown) do not have a functional *rps7* gene. The presence of a truncated *rps7* pseudogene (326 bp) in pea mitochondria and a very short (92 bp) *rps7* homologous region in soybean mitochondria was confirmed by sequence analysis. Southern blot analysis suggests that *rps7*-homologous DNAs are present in the nuclear genomes of pea and soybean. It has been mentioned that the S7 ribosomal protein is one of only two proteins (S7 and S4 ribosomal proteins) that are assembly-initiator proteins of the 30S subunit in *E. coli* (Nowotny and Nierhaus 1988; Dragon and Braker-Gingras 1993) and play a role in protein synthesis initiation (Dontsova et al. 1991; Abdurashidova et al. 1991). An apparently functional *rps7* gene has been found in wheat mitochondria, but not in pea or soybean mitochondria. This is similar to observations for the *coxII* (Nugent and Palmer, 1991; Covello and Gray, 1992) and *rps12* genes (Grohmann 1992). The *coxII* gene in cowpea and soybean has been transferred from the mitochondrion into the nucleus (Nugent and Palmer, 1991; Covello and Gray, 1992). It is very interesting to note that the *coxII* gene is present in both the mitochondrion and the nucleus in pea, but the nuclear copy seems to be silent (Nugent and Palmer 1991). The presence of a *coxII* gene in the pea and soybean mitochondrial genomes (Covello and Gray 1992; Nugent and Palmer 1991), but not in the mung bean mitochondria (Nugent and Palmer 1991) suggests that this reflects different stages in an evolutionary process of inter-compartment gene transfer. A mitochondrially-encoded *rps12* gene has been found in a number of plants from liverwort to maize, but this gene in *Oenothera* mitochondria is a pseudogene (Schuster and Brennicke 1991; Grohmann et al. 1992). The possibility that the mitochondria of pea, soybean and bean import the S7 ribosomal protein from the chloroplast is extremely unlikely. There are no reports that any organellar proteins are exported except perhaps for a short portion of a subunit of NADH dehydrogenase in mouse (Loverland et al. 1990). On the other hand, many mitochondrial and chloroplast proteins are imported from the cytoplasm (Glover and Lindsay 1992). All these data strongly suggest that the *rps7* gene has been transferred from the mitochondrion into the nucleus during plant evolution.

Models of organelle-to-nucleus gene transfer invoke gene duplication and functional genetic redundancy as essential steps (Obar and Green 1985). The pea *coxII* case represents this stage. If the soybean *coxII* case is equivalent to an early stage of selection in the model of Obar and Green (1985), the pea *rps7* pseudogene would represent the stage in which the mitochondrial gene has lost its function and has begun to degrade and the shorter *rps7*-

homologous sequence from soybean mitochondria (Resnick) would indicate further genetic drift due to lack of functional constraint. The cowpea *coxII* case (Nugent and Palmer 1991) would represent that in which the mitochondrial sequences have completely been deleted from the mitochondrial genome.

It is not clear why mitochondrial DNAs have been transferred into the nucleus. Hadler et al. (1983) have proposed that it occurs as a result of pressure on mitochondria by carcinogens, which results in more DNA mutations, and their metabolites. The mitochondrial genes may be under pressure to be transferred into the nucleus because of regulation advantages. If the gene transfer from one compartment to another is random, mtDNA is expected to show higher frequency because of occasional breakdown of mitochondrial membranes and the high copy number of the relatively low complexity mitochondrial genome. In fact, only a few cases of nuclear DNAs have been found in the mitochondrion (Schuster et al. 1988). On the other hand, many factors, such as codon usage differences between mitochondrial and nuclear genomes, are expected to limit successful mitochondrial gene transfer into the nucleus (Brennicke et al. 1993). There are two models to explain the mechanism of gene transfer from the mitochondrion to nucleus: direct gene transfer and RNA-intermediate gene transfer (Gellisen and Michaelis 1987). A mechanism similar to that of transposable elements has been proposed for the integration of mitochondrial DNA into nuclear DNA based on the finding of Ty elements present in the vicinity of some transferred sequences (Gellison and Michaelis 1987). However, DNA sequence analysis has revealed no characteristic structures such as direct/inverted repeats or duplication of nuclear sequences (Gellisen and Michaelis 1987). Like the mitochondrial *rps14*, *rps13* and *rps12* genes, the *rps7* gene is present in the mitochondrion in some land plant species but absent in others. These data suggest that gene transfer can be a very recent and on-going evolutionary event. I would like to suggest that direct mitochondrial and chloroplast DNA transfer is more common and wide-spread than RNA intermediate gene transfer during evolution.

Since RNA editing has been found in mitochondria from flowering plants but not lower plants such as liverwort, RNA editing is believed to represent a derived rather than retained primitive trait (Covello and Gray 1993). It has been shown that the nuclear *coxII* and *rps12* gene sequences correspond to the RNA edited sequences (Nugent and Palmer 1991; Covello and Gray 1992; Grohmann et al. 1992). Therefore, gene transfer from the mitochondrion to the nucleus seems to include reverse transcription of an edited RNA intermediate. Any functional mitochondrial gene transfer into the nucleus in flowering plants must involve an RNA intermediate. Therefore, I propose a two-step model (Fig 6.4) of transfer of mitochondrial DNA into the nucleus based in part on models proposed by Gellisen and Michaelis (1987). This model is different from the "Lucky Insertion" Scenario (Fig 6.3). In this model, mitochondrial mRNAs are reverse-transcribed into cDNA which are luckily integrated into the favourable nuclear genome environments (Fig 6.3). The nuclear-encoded mitochondrial genes become functional. I propose that mitochondrial DNAs have first been transferred and integrated into the nuclear genomes during evolution. The transferred mitochondrial DNAs are shuffled and recombined within the nuclear genome. By chance, they may be inserted into pre-existing nuclear-encoded mitochondrial gene sequences. At this stage, mitochondrial genes would function as introns and occasionally they could generate mRNAs due to alternative RNA splicing. However their proteins could not function properly because their mRNAs are not from the edited copies. When mitochondrial mRNAs are reverse-transcribed into cDNAs and occasionally recombined with the homologous DNAs in the nucleus, the functional mitochondrial genes are generated by "copy correction". In some cases such as the mitochondrial *rps7* gene, a few mutations can generate a functional domain. Some of these DNAs have more chance of expression. I call this DNA transfer mechanism a "Wait and See" model (Fig 6.4).

In this model, the first step is mitochondrial DNA transfer to the nucleus. Constant transfer into the nucleus and direct integration into nuclear chromosomes can occur by direct fusion of the mitochondrion and nucleus, or occasional breakdown of mitochondria and release of mtDNA that can enter the nucleus. This is supported by the finding that intact mitochondrial DNA is present in the nuclei of Hela cells (Kristensen and Prydz 1986). It has been shown that transfer of yeast mitochondrial DNAs into the nucleus occurs at a surprisingly high frequency (about 2×10^{-5} per cell per generation) (Thorsness and Fox 1990). It has also been reported that at least one-third of the chloroplast genome is integrated into the nuclear genomes of some plants (Alewife and Timmis 1992; Pichersky et al. 1991). Another way is that mitochondrial DNA can be transferred by plasmid DNAs functioning as shuttle vectors (Fukuchi et al. 1991; Goraczniak and Augustyniak 1991). Plant mitochondrial plasmids have been found to be integrated into the nuclear genomes with fragments of some mitochondrial genes (Fukuchi et al. 1991). Such DNA fragments have been sequenced and they show no specific patterns that would support a homologous recombination mechanism (Fukuchi et al. 1991). Mitochondrial DNAs seem to be integrated into the nuclear chromosomes by random insertions or non-homologous recombination (Sun and Callis 1993).

Figure 6.3 "Lucky Insertion" model of mitochondrial gene transfer into the nucleus.

The twisted circle is the mitochondrial DNA. The dashed and solid lines represent the nuclear and mitochondrial DNA sequences, respectively. This is adapted from that of Gellisen and Michaelis (1987).

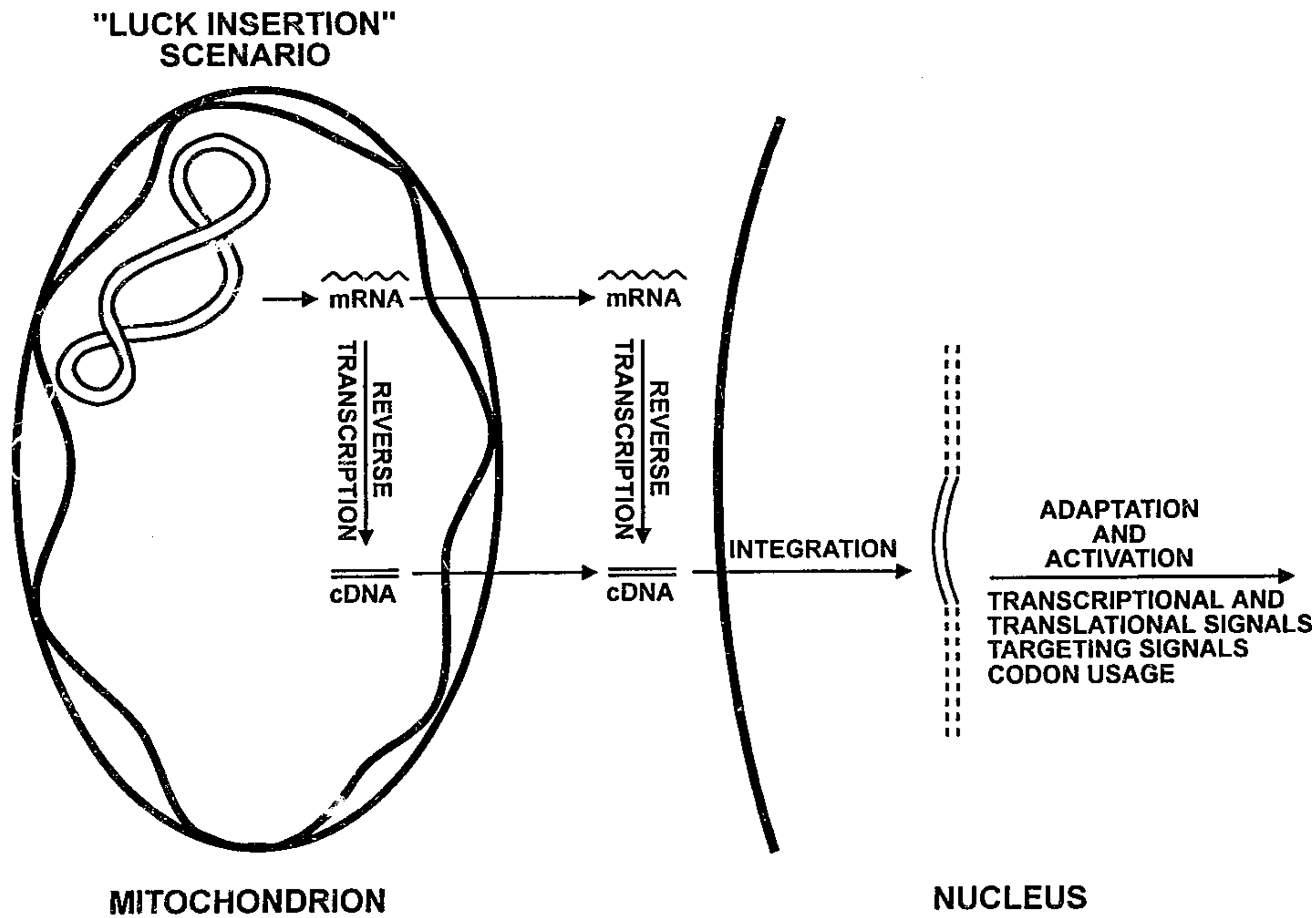
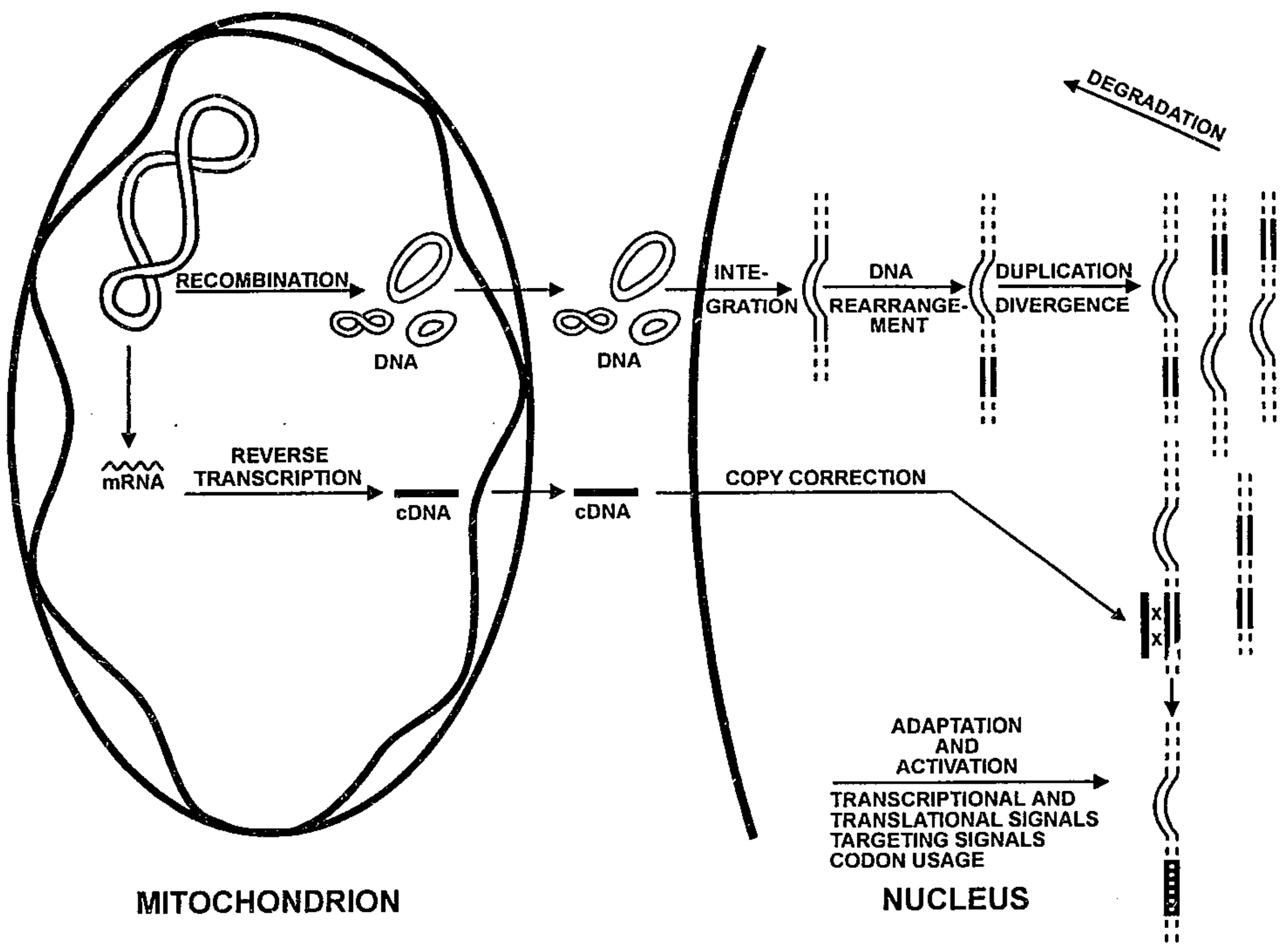

"LUCK INSERTION"
SCENARIO
mRNA
REVERSE
TRANSCRIPTION
cDNA
mRNA
REVERSE
TRANSCRIPTION
cDNA
INTEGRATION
ADAPTATION
AND
ACTIVATION
TRANSCRIPTIONAL AND
TRANSLATIONAL SIGNALS
TARGETING SIGNALS
CODON USAGE
MITOCHONDRION
NUCLEUS

Figure 6.4 "Wait and See" model of mitochondrial gene transfer into the nucleus.

The twisted circle is the mitochondrial DNA. The small circles are mitochondrial plasmids or subclasses of mitochondrial DNA molecules generated by recombination. The dashed and solid lines represent the nuclear and mitochondrial DNA sequences, respectively.



DEGRADATION
RECOMBINATION
DNA
DNA
INTE-
GRATION
DNA
REARRANGE-
MENT
DUPLICATION
DIVERGENCE
mRNA
REVERSE
TRANSCRIPTION
cDNA
cDNA
COPY CORRECTION
ADAPTATION
AND
ACTIVATION
TRANSCRIPTIONAL AND
TRANSLATIONAL SIGNALS
TARGETING SIGNALS
CODON USAGE
MITOCHONDRION
NUCLEUS

Therefore, entry of mitochondrial DNA to nucleus is not a major limiting factor. During evolution, the integrated mtDNAs can be amplified, diversified and also copy-corrected by newly-transferred mitochondrial DNAs via homologous recombination to reduce genetic drift. It has been reported that the rice mitochondrial B4 plasmid DNA has been transferred into the nucleus and four diversified copies are present in the nuclear genome (Fukuchi et al. 1991). The integrated mtDNAs could acquire plant nuclear signals by exon domain shuffling, or occasionally be inserted into introns of pre-existing nuclear genes encoding mitochondrial proteins. Because there can be alternative splicing in these genes, the integrated mitochondrial genes might be expressed.

The next step is that the mitochondrial DNAs adapt to the nuclear environment (Brennicke et al. 1993). In certain environments like dramatic changes in temperature and UV intensity, plants become unstable genetically and reverse transcriptases become very active. It has been shown that retrotransposons exist in a variety of higher plants (reviewed in Grandbastien 1992). The copy number ranges from several copies to ten thousands. Most of these retrotransposons encode a reverse transcriptase gene, which can be actively expressed under certain conditions (Grandbastien 1992). For example, in tobacco, the retrotransposons (To1 - To3) become activated during tissue culture (Hirochika 1993). Copy correction (Fig 6.4) of mitochondrial gene copies in the nucleus could occur at this stage. The copy correction could occur through mutation, or through recombination. It has been demonstrated in *Saccharomyces cerevisiae* that copy correction can occur through recombination between a reverse transcript and the homologous chromosomal allele (Derr and Strathern 1993). This experiment provides a potential mechansim for copy correction of transferred mitochondrial

DNAs. During mitosis, if broken mitochondria release mitochondrial mRNAs and at same time retrotransposons become active, RNA-mediated recombination between mitochondrial mRNAs and nuclear-encoded mitochondrial genes would result in "copy correction". Then both the mitochondrial and nuclear gene copies would be functional, but if the mitochondrial copies are not under functional constraint, they are under pressure to be deleted. However, the mitochondrial gene copies would be important if the nuclear gene copies do not have necessary DNA sequences to be expressed in all different developmental stages or at various tissues. In *Neurospora crassa*, the functional *atp9* gene is located in the nucleus while the mitochondrial copy is silent (van der Boogaart et al. 1982). Recently it has been found that the mitochondrial *atp9* copy is actively transcribed in germinating spores (Bittner-Eddy et al. 1994). This is consistent with the model in which gene transfer from the mitochondrion to the nucleus is not a one-step of lucky insertion (Fig 6.3), but long-term on-going evolution.

One potential problem with this model is that the nuclear *coxII* and *rps7* genes seem to be present in single rather than multiple copies in the nucleus as found in case of the rice mitochondrial B4 plasmid (Fukuchi et al. 1991). As Southern blots have shown, the hybridization signals of these nuclear genes were relatively low (Nugent and Palmer 1991; Covello and Gray 1992) and the nuclear *coxII* gene has about 85% identity with the mitochondrial counterpart suggesting that it was transferred about 200 million years ago (Nugent and Palmer 1991). This nuclear *coxII* gene is under functional constraint during evolution, while other nuclear *coxII*-homologous DNAs are not. In addition, plant nuclear genes have evolved at a rate about 12 times faster than that of plant mitochondrial genes (Wolfe et al. 1987). Therefore, plant mitochondrial genes and their nuclear non-functional copies would diverge rapidly and other nuclear mitochondrial *coxII*-homologous DNAs may be difficult to detect at stringent hybridization conditions when the mitochondrial gene copy is used as a probe. In the chloroplast, multiple copies of the *tufA* gene have been observed in the nucleus (Baldauf and Palmer 1990). This suggests that systematic studies of mitochondrial gene transfer may find mitochondrial DNAs dispersed throughout the nuclear genome with the degree of sequence similarity depending on the time of DNA transfer.

Even though eukaryotic organisms have a tendency to transfer mitochondrial genes into the nucleus, it is not clear why mitochondria keep some mitochondrial genes. Several possible reasons are: 1) proteins encoded by most mitochondrial genes are highly hydrophobic, therefore, import of these proteins from the cytoplasm is relatively difficult and may be energy-consuming; 2) in some organisms the codon usage is different between the mitochondrion and nucleus, for example, in mammalian and protist mitochondria, TGA specifies tryptophan instead of termination codons; 3) RNA editing is necessary to generate functional proteins in plant and some protist mitochondria; 4) some mitochondrial gene expression signals are different from those of nuclear genes; 5) transferred mitochondrial genes are expected to have targeting signals to re-direct proteins to the mitochondrion and 5) keeping genes in the mitochondria can be beneficial to both the mitochondrion and nucleus because it is easier to co-ordinate the mitochondrial biogenesis.

6.6 Future directions

The methods for constructing cDNA libraries from plant mitoplasts will be useful for more detailed studies on the expression of wheat mitochondrial genes. cDNA libraries can be constructed from mitochondrial mRNAs of additional stages of plant development and followed by differential screening to find out which genes are differentially expressed and then how this occurs. This method can also be used to study the effects of environmental factors on plant mitochondrial gene expression, for example under conditions of limited respiration and anaerobic stress.

To understand mechanisms of mitochondrial gene transfer, systematic studies in various plant species are needed. DNA clones covering the entire mitochondrial genome can be used to screen nuclear DNA libraries and to systematically analyze the mitochondrial sequences in the nuclear genomes. If most of the nuclear mitochondrial gene copies are from mitochondrial DNAs, instead of from RNA intermediates, it would be more plausible that the functional nuclear-encoded genes are generated by recombination of mitochondrial DNA with mitochondrial cDNAs. Alternatively, the RNA-intermediate DNA direct integration into the nuclear genome could be an extremely rare but still possible event.

To study how many mitochondrial ribosomal protein genes the nuclear genome encodes in various flowering plants as well as to understand what features the mitochondrial ribosomal proteins have, mitochondrial ribosomes can be purified, the individual ribosomal proteins isolated and antibodies generated as for chloroplast studies (Subramanian 1991; Schmidt et al. 1992). These antibodies could be used to screen a nuclear gene expression cDNA library and ribosomal proteins compared with those from other organisms. As yet about half of the known nuclear-encoded mitochondrial proteins lack any relationship to known proteins of other organisms (Harrer et al. 1993).

6.7 Concluding remarks.

In this research project, it has been shown that the wheat mitochondrial genome encodes a potentially functional S7 ribosomal protein (*rps7*) gene, but those of soybean and pea do not. Unlike certain other plant ribosomal protein genes, which are physically close to other genes, the wheat *rps7* gene is far away from other genes and its mRNA is monocistronic. However, because a 80 nt sequence element upstream of it is also found to precede at least four other wheat mitochondrial genes, these repeats may play a common role in gene regulation.

The most likely possibility is that the *rps7* gene has been transferred from the mitochondrion into the nucleus in pea and soybean. This idea is further supported by Southern blot analysis of the pea and soybean nuclear DNAs, showing positive hybridization signals. During this project, the complete sequence of the liverwort mitochondrial genome became available and it shows that the *rps7* gene is present in that mitochondrial genome (Oda et al 1992). The information suggests that the *rps7* gene transfer from the mitochondrion into the nucleus may have occurred during legume evolution. This idea needs further systematic analysis of whether the *rps7* gene is present in the mitochondrial genomes of other monocots and dicots.

Gene transfer from the mitochondrion to the nucleus may take a long time to generate an optimal nuclear-encoded mitochondrial gene. During this process, the mitochondrial gene copy may go through different stages. The pea mitochondrial *rps7* pseudogene could reflect a late stage where the mitochondrial copy is under genetic drift and the soybean *rps7* pseudogene is even more degenerate. One interesting aspect comes from the studies of in vitro mitochondrial protein synthesis in different tissues (Lind et al. 1991). Because variable proteins have been observed in some plant tissues and different stages of plant development, some of these proteins may actually reflect expression of both nuclear-encoded and mitochondrial genes. Usually the mitochondrial genes would be silent, however, in some stages of plant development or some tissues, the nuclear-encoded mitochondrial gene copies might not be expressed because they may lack some expression signals or because of their positions in the nuclear chromosomes. This would be analogous to *Neurospora crassa atp9* gene situation (Bittner-Eddy et al 1994). This raises the possibility that some ribosomal protein genes which have been recently transferred from the mitochondrion into the nucleus have both mitochondrial and nuclear copies being functional at certain stages of plant development.

REFERENCES

Abdurashidova GG, Tsvetkova EA, Budowsky EI (1991) Direct tRNA-protein interactions in ribosomal complexes. *Nucl Acids Res* 19:1909-1915.

Attardi G, Schatz G (1988) Biogenesis of mitochondria. *Ann Rev Cell Biol* 4:289-333.

Aubert D, Bisanz-Seyer C, Herzog M (1992) Mitochondrial *rps14* is a transcribed and edited pseudogene in *Arabidopsis thaliana*. *Plant Mol Biol* 20:1169-1174.

Auer J, Spicker G, Mayerhofer L, Pühler G, Böck A (1991) Organisation and nucleotide sequence of a gene cluster comprising the translation elongation factor 1α from *Sulfolobus acidocaldarius*. *System Appl Microbiol* 14:14-22.

Augustin S, Muller MW, Schweyen RJ (1990) Reverse self-splicing of group II intron RNA in vitro. *Nature* 343:383-386.

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA and Struhl K (1990) Current protocols in molecular Biology. Current Protocols, Boston.

Ayliffe MA, Timmis JN (1992) Plastid DNA sequence homologies in the tobacco nuclear genome. *Mol Gen Genet* 236:105-112.

Baldauf, SL, Manhart JR, Palmer JD (1990) Different fates of the chloroplast *tufA* gene following its transfer to the nucleus in green algae. *Proc Natl Acad Sci USA* 87:5317-5321.

Beckman DL, Trawick DR, Kranz RG (1992) Bacterial cytochromes c biogenesis. *Genes & Dev* 6:268-286.

Bendich AJ (1993) Reaching for the ring: the study of mitochondrial genome structure. *Curr Genet* 24:279-290.

Benne R (1992) RNA editing, an overview. In: RNA editing: The alteration of protein coding sequences of RNA. by Benne R ed. Springer-Verlag Berlin Heidelberg. p1-10.

Binder S, Schuster W, Grienenberger JM, Weil JH, Brennicke A (1990) Genes for $tRNA^{Gly}$, $tRNA^{His}$, $tRNA^{Lys}$, $tRNA^{Phe}$, $tRNA^{Ser}$ and $tRNA^{Tyr}$ are encoded in *Oenothera* mitochondrial DNA. *Curr Genet* 17:353-358

Bittner-Eddy P, Monroy AF, Brambl R (1994) Expression of mitochondrial genes in the germinating conidia of *Neurospora crassa*. *J Mol Biol* 237:881-897

Bland MM, Levings CS, Matzinger DF (1986) The tobacco mitochondrial ATPase subunit 9 gene is closely linked to an open reading frame for a ribosomal protein. *Mol Gen Genet* 204:8-16

Blum B, Bakalara N, Simpson L (1990) A model for RNA editing in kinetoplastid mitochondria: "Guide" RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* 60:189-198

Bock R, Hagemann R, Kössel H, Kudla J (1993) Tissue- and stage-specific modulation of RNA editing of *psbF* and *psbL* transcript from spinach plastids - a new regulatory mechanism? *Mol Gen Genet* 240:238-244.

Boer PH, McIntosh JE, Gray MW, Bonen L (1985) The wheat mitochondrial gene for apocytochrome b: absence of a prokaryotic ribosome binding site. *Nucl Acids Res* 13:2281-2292.

Bömmer D, Haberhausen G, Zetsche K (1993) A large deletion in the plastid DNA of the holoparasitic flowering plant *Cuscuta reflexa* containing two ribosomal proteins (*rpl1, rpl23*), one transfer RNA (*trnI*) and an ORF 2280 homologue. *Curr Genet* 24:171-176.

Bonen L (1987) The mitochondrial S13 ribosomal protein gene is silent in wheat embryos and seedlings. *Nucl Acids Res* 15:1393-10404.

Bonen L (1991) The mitochondrial genome: so simple yet so complex. *Curr Opinion in Genet Dev* 1:515-522.

Bonen L (1993) Trans-splicing of pre-mRNA in plants animals and protists. *FASEB J* 7:40-46.

Bonen L, Bird S (1988) Sequence analysis of the wheat mitochondrial *atp6* gene reveals a fused upstream reading frame and markedly divergent N termini among plant ATP6 proteins. *Gene* 73:47-56

Bonen L, Brown GG (1993) Genetic plasticity and its consequences: perspectives on gene organization and expression in plant mitochondria. *Can J Bot* 71:645-660.

Bonen L, Boer PH, Gray MW (1984) The wheat cytochrome oxidase subunit II gene has an intron insert and three radical amino acid changes relative to maize. *EMBO J* 3:2531-2536

Bonen L, Boer PH, McIntosh JE, Gray MW (1987) Nucleotide sequence of the wheat mitochondrial gene for subunit I of cytochrome oxidase. *Nucl Acids Res* 16:6734

Bonen L, Bird S, Belanger L (1990) Characterization of the wheat mitochondrial *orf25* gene.

Plant Mol Biol 15:793-795

Bonen L, Chapdelaine Y, Zhuo D (1993) Maturation of messenger RNAs in wheat mitochondria. In: Plant Mitochondria. Brennicke A, Kück U (eds), VCH, Weinheim. p.199-205.

Bonen L, Williams K, Bird S, Wood C (1994) The NADH dehydrogenase subunit 7 gene is interrupted by four group II introns in the wheat mitochondrial genome. *Mol Gen Genet* 243:81-89

Bonhomme S, Bird S, Bonen L (1989) Comparison of the wheat mitochondrial *atp9* gene sequence with mitochondrial and chloroplast homologues from other plants. *Plant Mol Biol* 13:395-397.

Bonnard G, Gualberto JM, Lamattina L, Grienenberger JM (1992) RNA editing in plant mitochondria. *Critical Rev in Plant Sci* 10:503-524.

Bourque DP, Elhag G, Bonham-Smith P, Thomas F, McCreery T, Glinsmann-Gibson (1991) Expression of the nuclear and chloroplast gene for tobacco chloroplast ribosomal proteins. In: The Translational Apparatus of Photosynthetic Organelles. Mache R, Stutz E, Subramanian AR (eds). Springer-Verlag Berlin Heidelberg p.95-105.

Bowles Dj, Schnarrenberger C, Kauss H (1976) Lectins as membrane components of mitochondria from *Ricinus communis*. *Biochem J* 160:375-382

Bowman CM (1991) The evolution of genes and pseudogenes for some chloroplast ribosomal proteins. Transposition and recombination lead to different fates in different genomes. In: The Translational Apparatus of Photosynthetic Organelles. Mache R, Stutz E, Subramanian AR (eds) Springer-Verlag Berlin Heidelberg 55:71-83

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

Brennicke A, Grohmann L, Hiesel R, Knoop V, Schuster W (1993) The mitochondrial genome on its way to the nucleus: different stages of gene transfer in higher plants. *FEBS Lett* 325:140-145.

Burger G, Lang FB, Wolff Kück U (1993) The mitochondrial genome of *Prototheca wickerhamii* (Strain 263-11). In Genetic maps 6th ed Book 3. O'Brien SJ (eds), Cold Spring Harbor Laboratory Press, New York. p.172-173.

Burke JM, Raj Bhandary VL (1982) Intron within the large rRNA gene of *N. crassa* mitochondria: a long open reading frame and a consensus sequence possibly important in splicing. *Cell* 31:509-520.

Burke JM (1988) Molecular genetics of Group I introns: RNA structures and protein factors required for splicings -- a review. *Gene* 73:273-294

Butow TA, Perlman PS, Grossman LI (1985) The unusual *var1* gene of yeast mitochondrial DNA. *Science* 228:1496-1501.

Cantwell R, McEntee CM, Hudson AP (1992) Regulation of mitochondrial transcription during the stringent response in yeast. *Curr Genet* 21:241-247

Carlson JE, Erickson LR, Kemble RJ (1986) Cross hybridization between organelle RNAs and mitochondrial and chloroplast genomes in *Brassica*. *Curr Genet* 11:161-163

Cattaneo R (1991) Different types of messenger RNA editing. *Annu Rev Genet* 25:71-88.

Cavalier-Smith T (1993) Evolution of the eukaryotic genome. In: The eukaryotic genome: organization and regulation. Broda P, Oliver SG, Sims PF (eds). Cambridge University Press. p333-385.

Cech TR (1988) Conserved sequences and structures of group I introns: building an active site for RNA catalysis -- a review. *Gene* 73:259-271.

Cech TR (1991) RNA editing: world's smallest introns? *Cell* 64:667-669.

Chanut FA, Grabau EA, Gesteland RF (1993) Complex organization of the soybean mitochondrial genome: recombination repeats and multiple transcripts at the atpA loci. *Curr Genet* 23:234-247.

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

Clark-Walker GD (1992) Evolution of mitochondrial genomes in Fungi. *Int Rev Cytol* 141:65-88

Clayton DA (1991) Replication and transcription of vertebrate mitochondrial DNA. *Annu Rev Cell Biol* 7:453-478

Clayton DA (1992) Transcription and replication of animal mitochondrial DNA. *Int Rev Cytol* 141:217-232

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

Conley CA, Hanson MR (1994) Tissue-specific protein expression in plant mitochondria. *Plant Cell* 6:85-91

Costanzo MC, Fox TD (1990) Control of mitochondrial gene expression in *Saccharomyces cerevisiae*. *Ann Rev Genet* 24:91-113

Covello PS, Gray MW (1989) RNA editing in plant mitochondria. *Nature* 341:662-666

Covello PS, Gray MW (1992) Silent mitochondrial and active nuclear genes for subunit 2 of cytochrome *c* oxidase (*cox2*) in soybean: evidence for RNA-mediated gene transfer. *EMBO J* 11:3815-3820.

Covello PS, Gray MW (1991) Sequence analysis of wheat mitochondrial transcripts capped in vitro: definitive identification of transcription initiation sites. *Curr Genet* 20:245-251

Covello PS, Gray MW (1993) On the evolution of RNA editing. *TIG* 9:265-268.

Crews S, Attardi G (1980) The sequences of the small ribosomal RNA gene and the phenylalanine tRNA gene are joined end to end in human mitochondrial DNA. *Cell* 19:775-784.

Cummings DJ (1992) Mitochondrial genomes of the ciliates. *Int Rev Cytol* 141:1-64

Dale RMK, Mendu N, Ginsburg H, Kridl JC (1984) Sequence analysis of the maize mitochondrial 26S rRNA gene and flanking regions. *Plasmid* 11:141-150

De Paepe R, Forchioni A, Chetrit P, Vedel F (1993) Specific mitochondrial proteins in pollen: presence of an additional ATP synthesis ß subunit. *Proc Natl Acad Sci USA* 90:5934-5938.

De Paepe R, Chetrit P, Vitart V, Ambard-Bretteville F, Vedel F (1990) Several nuclear genes control both male sterility and mitochondrial protein synthesis in *Nicotiana sylvestris* protoclones. *Mol Gen Genet* 222:206-219.

Derr LK, Strathern JN (1993) A role for reverse transcripts in gene conversion. *Science* 361:170-173.

Dibb NJ (1993) Why do genes have introns? *FEBS Lett* 325:135-139

Dietrich A, Small I, Desprez T, Masson J, Weber F, Ramamonjisoa D, Souciet G, Cosset A, Green G, Guillemaut P, Pelletier G, Weil JH, Marèchal-Drouard L (1993) The plant mitochondrial transfer RNA population: a mosiac of species with different genetic origins. In: The Translational Apparatus. Nierhaus KH, Franceschi F, Subramanian AR, Erdmann VA, Wittmann-Liebold B (eds). p617-627

Dietrich A, Weil JH, Marèchal-Drouard L (1992) Nuclear-encoded transfer RNAs in plant mitochondria. *Annu Rev Cell Biol* 8:115-131

Dohmen G, Tudzynski P (1994) A DNA-polymerase-related reading frame (pol-r) in the mtDNA of *Secale cereale*. *Curr Genet* 25:59-65

Dontsova O, Kopylov A, Brimacombe R (1991) The location of mRNA in the ribosomal 30S initiation complex; site-directed cross-linking of mRNA analogues carrying several photo reactive labels simultaneously on either side of the AUG start codon. *EMBO J* 10:2613-2620

Douce R, Neuburger M (1989) The uniqueness of plant mitochondria. *Ann Rev Plant Physiol Plant Mol Biol* 40:371-414

Dragon F, Brakier-Gingras L (1993) Interaction of Escherichia coli ribosomal protein S7 with 16S rRNA. *Nucleic Acid Res* 21:1199-1203.

Elthon TE, McIntosh L (1987) Identification of the alternative terminal oxidase of higher plant mitochondria. *Proc Natl Acad Sci USA* 84:8399-8403

Emory SA, Bouvet P, Belasco JG (1992) A 5'-terminal stem-loop structure can stabilize mRNA in *Escherichia coli*. *Genes & Dev* 6:135-148

Falconet D, Delorme S, Lejeune B, Sevignac M, Delcher E, Bazetoux S, Quetier F (1985) Wheat mitochondrial 26S ribosomal RNA gene has no intron and is present in multiple copies arising by recombination. *Curr Genet* 10:169-174

Falconet D, Sevignac M, Quetier F (1988) Nucleotide sequence and determination of the extremities of the 26S ribosomal RNA gene in wheat mitochondria: evidence for sequence rearrangements in the ribosomal genes of higher plants. *Curr Genet* 13:75-82

Fauron CMR, Havlik M, Casper M (1991) Organization and evolution of the maize mitochondrial genome. In: Plant Molecular Biology 2. Herrmann RG, Larkins B (eds). p345-363.

Feinberg AP, Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochem* 132:6-13

Fox TD, Leaver CJ (1981) The *Zea mays* mitochondrial gene coding cytochrome oxidase subunit II has an intervening sequence and does not contain TGA codons. *Cell* 26:315-323.

Fukuchi M, Shikanai T, Kossykh VG, Yamada Y (1991) Analysis of nuclear sequences homologous to the B4 plasmid-like DNA of rice mitochondria: evidence for sequence transfer from mitochondria to nuclei. *Curr Genet* 20:487-494

Gantt JS, Baldauf S, Calie PJ, Weeden NF, Palmer JD (1991) Transfer of *rpl22* to the nucleus greatly preceded its loss from the chloroplast and involved the gain of an intron. *EMBO J* 10:3073-3078

Gass DA, Makaroff CA, Palmer JD (1992) Variable intron content of the NADH dehydrogenase subunit 4 gene of plant mitochondria. *Curr Genet* 21:423-430

Gellissen G and Michaelis G (1987) Gene transfer: mitochondria to nucleus. *Ann New York Acad Sci.* 503:392-401

Gillham NW, Harris EH, Randolph-Anderson BL, Boynton JE, McEclwain, Newman SM (1991) Molecular genetics of chloroplast ribosomes in *Chlamydomonas.* In: The translational apparatus of photosynthetic organelles. Mache R, Stutz E, Subramanian AR (eds). Springer-Verlag, Berlin Heidelberg p107-115.

Glinsmann-Gibson B (1991) Expression of nuclear and chloroplast genes coding for tobacco chloroplast ribosomal proteins. 55:95-105.

Glover LA, Lindsay JG (1992) Targeting proteins to mitochondria: a current overview. *Biochem J* 284:609-620

Goldberg RB (1978) DNA sequence organization in the soybean plant. *Biochem Genet* 16:45-68.

Goldberg RB, Barker SJ, Perez-Grau L (1989) Regulation of gene expression during plant embryogenesis. *Cell* 56:149-160

Gonzalez DH, Bonnard G,Grienenberger JM (1993) A gene involved in the biogenesis of cytochromes is co-transcribed with a ribosomal protein gene in wheat mitochondria. *Curr Genet* 24:248-255

Goraczniak RM, Augustyniak H (1991) Properties of lupin mitochondrial plasmid-like DNAs and nucleotide sequence of a new minicircular DNA. *Plant Sci* 79:173-179

Graack HR, Grohmann L, Choli T (1988) Mitochondrial ribosomes of yeast: isolation of individual proteins and N-terminal sequencing. *FEBS Lett* 242:4-8.

Grabau A (1987) Cytochrome oxidase subunit II gene is adjacent to an initiator methionine tRNA in soybean mitochondrial DNA. *Curr Genet* 11:287-293

Grabau E, Havlik M, Gesteland R (1988) Chimeric organization of two genes for the soybean mitochondrial ATPase subunit 6. *Curr Genet* 13:83-89

Grandbastien MA (1992) Retrotransposons in higher plants. *TIG* 8:103-108.

Gray MW (1983) The bacterial ancestry of plastids and mitochondria. *BioScience* 33:693-699

Gray MW (1989a) The evolutionary origins of organelles. *TIG* 5:25-50.

Gray MW (1989b) Origin and evolution of mitochondrial DNA. *Ann Rev Cell Biol* 5:25-50.

Gray MW (1990) The mitochondrial genome and its expression. In: Plant physiology, biochemistry and molecular biology. Dennis DT, Turpin DH (eds). Longman Scientific & Technical, Harlow. p145-159.

Gray MW (1991) Origin and evolution of plastid genome and genes. In: Cell culture and somatic cell genetics of plants 7A:303-331

Gray MW (1992) The endosymbiont hypothesis revisited. *Int Rev Cytol* 141:233-357.

Gray MW (1993) Origin and evolution of organelle genomes. *Curr Opinion in Genet and Dev* 3:884-890

Gray MW, Covello PS (1993) RNA editing in plant mitochondria and chloroplasts. *FASER J* 7:64-71.

Gray MW, Cedergren R, Abel Y, Sankoff D (1989) On the evolutionary origin of the plant mitochondrion and its genome. *Pro Natl Acad Sci USA* 86:2267-2271

Green MR (1991) Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annu Rev Cell Biol* 7:559-99.

Grienenberger JM (1992) RNA editing in plant organelles. In: RNA editing: The alternation coding sequences of RNA. Benne R ed. Springer-Verlag Berlin Heidelberg p153-177.

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

Grohmann L, Brennicke A, Schuster W (1993) Genes for mitochondrial ribosomal proteins in plants. In: The translational apparatus. Nierhaus KH, Franceschi F, Subramanian AR, Erdmann VA, Wittmann-Liebold B (eds). Plenum Press, New York, p599-607

Gruissem W (1989) Chloroplast gene expression: how plants turn their plastids on. *Cell* 56:161-170

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 wheat and maize mitochondrial genomes and are co-transcribed. *Mol Gen Genet* 215:118-127

Gualberto JM, Lamattina L, Bonnard G, Weil JH, Grienenberger JM (1989) RNA editing in wheat mitochondria results in the conservation of protein sequences. *Nature* 341:660-662

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. *Nucl Acids Res* 18:3771-3776

Gualberto JM, Bonnard G, Lamattina L, Grienenberger JM (1991) Expression of the wheat mitochondrial *nad3-rsp12* transcription unit: correlation between editing and mRNA maturation. *Plant Cell* 3:1109-1120

Gubler U, Hoffman BJ (1983) A simple and very efficient method for generating cDNA libraries. *Gene* 25:263-269

Hack E, Hendrick CA, Al-Janabi SM, Crane VC, Girton LE (1994) Translation in a wheat germ cell-free system of RNA from mitochondria of the normal and Texas male-sterile cytoplasms of maize (*Zea mays* L.). *Curr Genet* 25:73-79.

Hadler HI, Dimitrijevic B, Mahaligam R (1983) Mitochondrial DNA and nuclear DNA from normal rat liver have a common sequence. *Proc Natl Acad Sci USA* 80:6495-6499

Hallick RB, Christopher DA, Copertino DW, Dragere RG, Jenkins KP, Stevenson JK (1991) Chloroplast ribosomal protein operons of *Euglena gracilis*. In: The translational apparatus of photosynthetic organelles. Mache R, Stutz E, Subramanian AR (eds). Springer-Verlag, Berlin Heidelberg. p145-153.

Hanic-Joyce PJ, Gray MW (1991) Accurate transcription of a plant mitochondrial gene *in vitro*. *Mol Cell Biol* 11:2035-2039

Hanson MR, Folkerts O (1992) Structure and function of the higher plant mitochondrial genomes. *Int Rev Cytol* 141:129-167.

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.

Harington A, Thornley AL (1982) Biochemical and genetic consequences of gene transfer from endosymbiont to host genome. *J Mol Evol* 18:287-292

Harrer R, Schwank S, Schüller HJ, Schweizer E (1993) Molecular cloning and analysis of the nuclear gene *MRP-L6* coding for a putative mitochondrial ribosomal protein from *Saccharomyces cerevisiae*. *Curr Genet* 24:136-140

Harris M, Decker C, Sollner-Webb B, Hajduk S (1992) Specific cleavage of pre-edited mRNAs in trypanosomal mitochondrial extracts. *Mol Cell Biol* 12:2591-2598.

Hartl FU, Neupert W (1989) Protein sorting to mitochondria: evolutionary conservations of folding and assembly. *Science* 247:930-938.

Hasegawa M, Kishino H, Hayasaka K, Horai S (1990) Mitochondrial DNA evolution in primates: transition rate has been extremely low in the lemur. *J Mol Evol* 31:113-121.

Hauser CR, Randolph-Anderson BL, Hohl TM, Harris EH, Boynton JE, Gillham N (1993) Molecular genetics of chloroplast ribosomes in Chlamydomonas reiinhardtii. In: The translational apparatus. Niehaus KH, Franceschi F, Subramanian AR, Erdmann VA, Wittmann-Liebold B (eds). Plenum Press, New York. p545-554.

Hiesel R, Schobel W, Schuster W, Brennicke A (1987) The cytochrome oxidase subunit I and III genes in *Oenothera* mitochondria are transcribed from identical promoter sequences. *EMBO J* 6:29-34

Hiesel R, Wissinger B, Schuster W, Brennicke A (1989) RNA editing in plant mitochondria. *Science* 246:1632-1634.

Hiratsuka J, Shimada H, Whittier RF, Ishibashi T, Sakamoto M, Mori M, Koto C, Honji Y, Sun CR, Meng BY, Li Y, Kanno A, Nishizawa Y, Hirai A, Shinozaki k, Sugiura M (1989) The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol Gen Genet* 217:185-194.

Hirochika H (1993) Activation of tobacco retrotransposons during tissue culture. *EMBO J* 12:2521-2528

Hodges P, Scott J (1993) Editing of mammalian apolipoprotein B mRNA by site-specific RNA deamination. RNA editing: The alteration of protein coding sequences of RNA. Benne R (ed) Springer-Verlag Berlin Heidelberg. p125-151.

Hopp TP, Woods KR (1981) Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci USA* 78:3824-3828.

Hunt MD, Newton KJ (1991) The NCS3 mutation: genetic evidence for the expression of ribosomal protein genes in *Zea mays* mitochondria. *EMBO J* 10:1045-1052

Ishikawa M, Kadowaki K (1993) Excess RNA editing in rice mitochondrial *atp9* transcripts. *Plant Cell Physiol* 34:959-963

Jacquier A (1990) Self-splicing group II and nuclear pre-mRNA introns: how similar are they? *TIBS* 15:351-354.

Janke A, Pääbo S (1993) Editing of a tRNA anticodon in marsupial mitochondria changes its codon recognition. *Nucleic Acids Res* 21:1523-1525.

Joyce PBM, Gray MW (1989a) Aspartate and asparagine tRNA genes in wheat mitochondrial DNA: a cautionary note on isolation of tRNA genes from plants. *Nucleic Acids Res* 17:7865-7878.

Joyce PBM, Gray MW (1989b) Chloroplast-like transfer RNA genes expressed in wheat mitochondria. *Nucleic Acids Res* 14:5461-5476.

Joyce PBM, Spencer DF, Gray MW (1988) Multiple sequence rearrangements accompanying the duplication of a $tRNA^{Pro}$ gene in wheat mitochondrial DNA. *Plant Mol Biol* 11:833-843.

Kawasaki ES (1990) Amplification of RNA. In: PCR protocols: a guide to methods and amplifications. Innis MA, Gelfand DH, Sninsky, White TJ (eds). Academic Press, San Diego. p21-27.

Klaff P, Gruissem W (1991) Changes in chloroplast mRNA stability during leaf development. *Plant Cell* 3:517-529.

Knoop V, Schuster W. Wissinger B, Brennicke A (1991) *Trans* splicing integrates an exon of 22 nucleotides into the *nad5* mRNA in higher plant mitochondria. *EMBO J* 10:3483-3493.

Köhler RH, Horn R, Lössl A, Zetsche K (1991) Cytoplasmic male sterility in sunflower is correlated with the co-transcription of a new open reading frame with the *atpA* gene. *Mol Gen Genet* 227:369-376

Kolakofsky D, Curran J, Pelet T, Jacques JP (1993) Paramyxovirus P gene mRNA editing. In: RNA editing: The alteration of protein coding sequences of RNA. Benne R ed, Springer-Verlag, Berlin Heidelberg. p105-123.

Kössel H, Hoch B, Igloi GL, Maier RM, Ruf S (1993) Editing creates the initiator codon of the *rpl2* transcript from maize chloroplasts. In: The translational apparatus. Ed Nierhaus KH, Franceschi F, Subramanian AR, Erdmann VA, Wittmann-Liebold B, Plenum Press, New York. p609-616.

Krishnasamy S, Makaroff CA (1993) Characterization of the radish mitochondrial *orfB* locus: possible relationship with male sterility in Ogura radish. *Curr Genet* 24:156-163.

Kristensen T, Pzydz H (1986) The presence of intact mitochondrial DNA in Hela cell nuclei. *Nucleic Acids Res* 14:2597-2609.

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

Kunkel TA, Eckert KA (1989) Fidelity of DNA polymerase used in polymerase chain reactions. Curr Comm Mol Biol: Polymerase chain reaction. Erlich HA, Sibbs R, Kazazian HH eds. Cold Spring Harbor Laboratory Press.

Kuziel W, Trucker PW (1987) Determination of vector:insert junctions in lambda gt10 cDNAs that do not recut with EcoRI. Nucleotide sequence of the lambda *imm*434 HindIII-EcoRI DNA fragment encoding part of *c*I protein. *Nucleic Acids Res* 15:3181

Lagrange T, Carol P, Bisanz-Sayer C, Mache R (1991) Comparative analysis for four different cDNA clones encoding chloroplast ribosomal proteins. In: The translational apparatus of photosynthetic organelles. Mache R, Stutz E, Subramanian AR (eds). Springer-Verlag Berlin Heidelberg p107-115

Lamattina L, Weil JH, Grienenberger JM (1989) RNA editing at a splicing site of NADH dehydrogenase subunit IV gene transcript in wheat mitochondria. *FEBS Lett* 258:79-83

Lamattina L, Gonzalez D, Gualberto J, Grienenberger JM (1993) Higher plant mitochondria encode an homologue of the nuclear-encoded 30-kDa subunit of bovine mitochondrial complex I. *Eur J Biochem* 217:813-838.

Lambowitz AM (1989) Infectious introns. *Cell* 56:323-326

Lang BF, Forget L (1993) The mitochondrial genome of *Phytophthora infestans*. In Genetic maps 6th ed Book 3. O'Brien SJ (eds), Cold Spring Harbor Laboratory Press, New York. p.133-135.

Laurent GS, Yoshie O, Floyd-Smith G, Samanta H, Sehgal PB, Lengyel P (1983) Interferon action: two (2'-5') $(A)_n$ synthetases specified by distinct mRNAs in Ehrlich ascites tumor cells treated with interferon. *Cell* 33:95-102.

Leffers H, Gropp F, Lottspeich F, Zillig W, Garrett RA (1989) Sequence, organization, transcription and evolution of RNA polymerase subunit genes from the archaebacterial extreme halophiles *Halobacterium halobium* and *Halococcus morrhuae*. *J Mol Biol* 206:1-17

Lejeune B, Delorme S, Delcher E, Quetier (1987) Recombination in wheat mitochondrial DNA: occurrence of nine different genomic contexts for the 18S-5 S genes. *Plant Physiol Biochem* 25:227-233.

Lejeune B, Quetier F (1988) Structure de l'ADN mitochondrial du blè. In Variabilitè Gènetique cytoplasmique et Stèrilitè mâle cytoplasmique (Paris: INRA) p229-295.

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

Levings CS, Brown GG (1989) Molecular biology of plant mitochondria. *Cell* 56:171-179

Li LW, Tan KH (1990) Changes in the protein synthesis of isolated mitochondria from cotyledons of *Pharbitis nil* caused by photoperiodic flower induction. *Chin J Bot* 2:130-134.

Liaud MF, Zhang DX, Cerff R (1990) Differential intron loss and endosymbiotic transfer of chloroplast glyceraldehyde-3-phosphate dehydrogenase gene to the nucleus. *Proc Natl Acad Sci USA* 89:8918-8922

Lind C, Hallde C, Moller IM (1991) Protein synthesis in mitochondria purified from roots leaves, and flowers of sugar beet. *Physiol Plant* 83:7-16.

Lizardi PM (1983) Methods for the preparation of messenger-RNA *Methods Enzymol* 96:24-38

Locker J (1979) Analytical and preparative electrophoresis of RNA in agarose-urea. *Analytical Biochem* 98:358-377

Löffelhardt W, Michalowski C, Kraus M, Pfanzagl B, Neumann-Spallart C, Jakowitsch J, Brandtner M, Bohnert HJ (1991) RPS10 and 6 other ribosomal protein genes from the S10/*spc*-operon not encountered on higher plant plastid DNA are located on the cyanelle genome of *Cyanophora Paradoxa.* In: The translational apparatus of photosynthetic organelles, ed. Mache R, Stutz E, Subramanian AR (eds). Springer-Verlag Berlin Heidelberg, p155-165

Lonergan KM, Gray MW (1993) Editing of transfer RNAs in *Acanthamoeba castellanii* mitochondria. *Science* 259:812-816.

Lonsdale DM (1989) The plant mitochondrial genome. In: The biochemistry of plants: a comprehensive treatise. Marcus A (ed). vol 15. Academic Press, New York. p229-295.

Lonsdale DM (1989) The plant mitochondrial genome. In: The biochemistry of plants: a comprehensive treatise. vol 15. Marcus A (ed) Academic Press, New York, pp 229-295

Loveland B, Wang CR, Yonekawa H, Hermel E, Lindahl KF (1990) Maternally transmitted histocompatibility antigen of mice: a hydrophobic peptide of mitochondrially encoded protein. *Cell* 60:971-980.

Makaroff CA, Palmer JD (1987) Extensive mitochondrial specific transcription of the *Brassica campestris* mitochondrial genome. *Nucl Acids Res* 15:5141-5155

Maloney AP, Traynor PL, Levings III CS, Walbot V (1989) Identification in maize mitochondrial 26S rRNA of a short 5'-end sequence possibly involved in transcription initiation and processing. *Curr Genet* 15:187-192

Maniatis T, Goodboun S, Fischer (1987) Regulation of inducible and tissue-specific gene expression. *Science* 236:1237-1245.

Marèchal-Drouard L, Ramamonjisoa D, Cosset A, Weil JH, Dietrich A (1993) Editing corrects mispairing in the acceptor stem of bean and potato mitochondrial phenylalanine transfer RNAs. *Nucleic Acid Res* 21:4909-4914.

Marèchal-Drouard L, Weil JH, Dietrich (1993) Transfer RNAs and transfer RNA genes in plants. *Annu Rev Plant Physiol Plant Mol Biol* 44:13-32.

Margulis L (1981) Symbiosis in cell evolution. Freeman, San Francisco.

Martin W, Lagrange T, Li YF, Bisanz-Seyer C, Mache R (1990) Hypothesis for the evolutionary origin of the chloroplast ribosomal protein L21 of spinach. *Curr Genet* 18:553-556

Maslov DA, Sturm NR, Niner BM, Gruszynski ES, Peris M, Simpson L (1992) An intergenic G-rich region in *Leishmania tarentolae* kinetoplast maxicircle DNA is a pan-edited cryptogene encoding ribosomal protein S12. *Mol Cell Biol* 12:56-67

Meinhardt F, Kempken F, Kamper J, Esser K (1990) Linear plasmids among eukaryotes - Fundamentals and applications. *Curr Genet* 17:89-95.

Michalowski CB,Pfanzagi B, Löffelhardt W, Bohnert HJ (1990) The cyanelle S10 *spc* ribosomal protein gene operon from *Cyanophora paradoxa*. *Mol Gen Genet* 224:222-231

Michel F, Umesono K, Ozeki H (1989) Comparative and functional anatomy of group II catalytic introns - a review. *Gene* 82:5-30.

Miller D, Mahendran R, Spottswood M, Ling M, Wang S, Yang N, Costandy (1993) RNA editing in mitochondria of *Physarum polycephalum*. In: RNA editing: The alteration of protein coding sequences of RNA. Benne eds Springer-Verlag Berlin Heidelberg. p87-103.

Möller K, Zwieb C, Brimacombe R (1978) Identification of the oligonucleotide and oligopeptide involved in an RNA-protein crosslink induced by ultraviolet irradiation of *Escherichia coli* 30S ribosomal subunits. *J Mol Biol* 126:489-506

Monéger F, Smart CJ, Leaver CJ (1994) Nuclear restoration of cytoplasmic male sterility in sunflower is associated with the tissue-specific regulation of a novel mitochondrial gene. *EMBO J* 13:8-17.

Moore PB, Capel MS (1988) Structure-function correlations in the small ribosomal subunit from *Escherichia coli*. *Ann Rev Biophys Chem* 17:349-367

Muise RC, Hauswirth WW (1992) Transcription in maize mitochondria: effects of tissue and mitochondrial genotype. *Curr Genet* 22:235-242.

Mullet JE (1988) Chloroplast development and gene expression. *Ann Rev Plant Physiol Plant Mol Biol* 39:475-502.

Mulligan RM, Leon P, Walbot V (1991) Transcriptional and posttranscriptional regulation of maize mitochondrial gene expression. *Mol Cell Biol* 11:533-543

Mulligan RM, Lau GT, Walbot V (1988b) Numerous transcription initiation sites exist for the maize mitochondrial genes for subunit 9 of the ATP synthase and subunit 3 of cytochrome oxidase. *Proc Natl Acad Sci USA* 85:7998-8002.

Mulligan RM, Maloney AP, Walbot V (1988a) RNA processing and multiple transcription initiation sites result in transcript size heterogeneity in maize mitochondria. *Mol Gen Genet* 211:373-380

Murray NE, Brammar WJ, Murray K (1977) Lambdoid phages that simplify the recovery of in vitro recombinants. *Mol Gen Genet* 150:53-61

Murray MG, Cuellar RE, Thompson WF (1978) DNA sequence organization in the pea genome. *Biochem* 17:5781-5790.

Nagley P (1991) Coordination of gene expression in the formation of mammalian mitochondria. *TIG* 7:1-4

Nakanozo M, Hirai A (1993) Identification of the entire set of transferred chloroplast DNA sequences in the mitochondrial genome of rice. *Mol Gen Genet* 236:341-246.

Neale DB, Marshall KA, Sederoff (1989) Chloroplast and mitochondrial DNA are paternally inherited in Sequoia sempervirens D. Don Endl. *Proc Natl Acad Sci USA* 86:9347-9349.

Newton KJ, Walbot V (1985) Maize mitochondria synthesize organ-specific polypeptides. *Proc Natl Acad Sci USA* 82:6879-6883

Newton KJ (1988) Plant mitochondrial genomes: organization, expression and variation. *Ann Rev Plant Physiol Plant Mol Biol* 39:503-532

Nicholson DW, Stuart RA, Neupert W (1989) Biogenesis of cytochrome c_1. *J Biol Chem* 264:10156-10168.

Nomura M, Yates JL, Dean D, Post LE (1980) Feedback regulation of ribosomal protein gene expression in *Escherichia coli*: Structural homology of ribosomal RNA and ribosomal protein mRNA. *Proc Natl Acad Sci USA* 77:7084-7088

Nowotny V, Nierhaus KH (1988) Assembly of the 30S subunit from *Escherichia coli* ribosomes occurs via two assembly domains which are initiated by S4 and S7. *Biochem* 27:7051-7055

Nugent JM, Palmer JD (1991) RNA-mediated transfer of the gene *coxII* from the mitochondrion to the nucleus during flowering plant evolution. *Cell* 66:473-481

O'Brien TW, Denslow ND, Faunce WH, Anders JC, Liu J, O'Brien BJ (1993) Structure and function of mammalian mitochondrial ribosomes. In: The translational apparatus. Nierhaus KH, Franceschi F, Subramanian AR, Erdmann VA, Wittmann-Liebold B (eds). Plamun Press, New York. p575-586

Obar R, Green J (1985) Molecular archaeology of the mitochondrial genome. *J Mol Evol* 22:243-251

Ochman H, Medhora MM, Garza D, Hartl DL (1990) Amplification of flanking sequences by inverse PCR. In: PCR protocols: A guide to methods and applications. Innis MA, Gelfend DH, Sninsky JJ, White TJ. Academic Press p219-227

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

Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Zhang Z, Aota S, Inokuchi H, Ozeki H (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* 322:572-574.

Ojala D, Merkel C, Gelfand R, Attardi G (1980) The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA. *Cell* 22:393-403.

Ovchinnikov TA, Guryev SO, Krayev AS, Monastyrskaya GS, Skryabin KG, Sverdlov ED, Zakharyev VM, Bayev AA (1979) Primary structure of an EcoRI fragment of lambda *imm*434 DNA containing regions *cI-cro* of phage 434 and *cII-O* of phage lambda. *Gene* 6:235-249

Palmer JD (1990) Contrasting modes and tempos of genome evolution in land plant organelles. *TIG* 6:115-120.

Palmer JD (1991) Plastid chromosomes: structure and evolution. In: Cell culture and somatic cell genetics of plants. 7A:5-53.

Palmer JD, Herbon LA (1986) Tricircular mitochondrial genomes of *Brassica* and *Raphanus*: reversal of repeat configurations by inversion. *Nucleic Acids Res* 14:9755-9765.

Palmer JD, Soltis D, Soltis P (1992) Large size and complex structure of mitochondrial DNA in two nonflowering land plants. *Curr Genet* 21:125-129

Pfanner N, Neupert W (1990) The mitochondrial protein import apparatus. *Ann Rev Biochem* 59:331-353

Pichersky E, Logsdon Jr JM, McGrath M, Stasys RA (1991) Fragments of plastid DNA in the nuclear genome of tomato: prevalence, chromosomal location, and possible mechanism of integration. *Mol Gen Genet* 225:453-458

Post LE, Nomura M (1980) DNA sequences from the *str* operon of *Escherichia coli*. *J Biol Chem* 255:4660-46666.

Pring DR, Levings III CS (1978) Heterogeneity of maize cytoplasmic genomes among male-sterile cytoplasms. *Genetics* 89:121-136

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

Pritchard AE, Seilhamer JJ, Mahalingam R, Sable CL, Venuti SE, Cummings DJ (1990) Nucleotide sequence of the mitochondrial genome of *Paramecium*. *Nucl Acids Res* 18:173-180

Raff RA, Mahler HR (1972) The non symbiotic origin of mitochondria: the question of the origin of the eucaryotic cell and its organelles is reexamined. *Science* 177:575-582.

Ragan CI, Wilson MT, Darley-Usmar VM, Lowe PN (1987) Sub-fractionation of mitochondria and isolation of the proteins of oxidative phosphorylation. Mitochondria: A Practical Approach. Darley-Usmar VM, Rickwood D, Wilson MT (eds). IRL Press. Oxford. p79-112

Rajasekhar VK, Mulligan RM (1993) RNA editing in plant mitochondria: α-phosphate is retained during C-to-U conversion in mRNAs. *Plant Cell* 5:1843-1852

Ramseier TM, Winteler HV, Hennecke H (1991) Discovery and sequence analysis of bacterial genes involved in the biogenesis of c-type cytochromes. *J Biol Chem* 266:7793-7803.

Rapp WD, Stern DB (1992) A conserved 11 nucleotide sequence contains an essential promoter element of the maize mitochondrial *atp1* gene. *EMBO J* 11:1065-1073

Rapp WD, Lupold DS, Mack S, Stern DB (1993) Architecture of the maize mitochondrial *atp1* promoter as determined by linker-scanning and point mutagenesis. *Mol Cell Biol* 13:7232-7238.

Rasmussen J, Hanson MR (1989) A NADH dehydrogenase subunit gene is co-transcribed with the abnormal *Petunia* mitochondrial gene associated with cytoplasmic male sterility. *Mol Gen Genet* 215:332-336

Reinbolt J, Tritsch D, Wittmann-Liebold B (1979) The primary structure of ribosomal protein S7 from *E. coli* strains K and B. *Biochimie* 61:501-522

Reith M, Munholland J (1993) A high-resolution gene map of the chloroplast genome of the red alga *Porphyra purpurea*. *Plant Cell* 5:465-475.

Rios R, de Buyser J, Henry Y, Ambard-Bretteville F, Rémy R (1991) Two-dimensional electrophoretic comparison of mitochondrial polypeptides from different wheat (*Triticum aestivum* L.) tissues. *Plant Sci* 76:159-166.

Rivera MC, Lake JA (1992) Evidence that eukaryotes and eocyte prokaryotes are immediate relatives. *Science* 257:74-76.

Rochaix JD, Goldschmidt-Clermont M, Choquet Y, Kuchka M (1991) Nuclear and chloroplast genes involved in the expression of specific chloroplast genes of Chlamydomonas reiinhardtii. In: Plant molecular biology 2. Herrmann RG, Larkins B (eds). Plenum Press, New York. p401-410.

Rothenberg M, Hanson MR (1987) Different transcript abundance of two divergent ATP synthase subunit 9 genes in the mitochondrial genome of *Petunia hybrida*. *Mol Gen Genet* 209:21-27

Ruby SW, Abelson J (1988) An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. *Science* 242:1028-1035.

Rutledge RG, Seligy VL, Côté MJ, Dimock K, Lewin LL, Tenniswood MP (1988) Rapid synthesis and cloning of complementary DNA from any RNA molecule into plasmid and phage M13 vectors. *Gene* 68:151-158

Saiki RK (1990) Amplification of genomic DNA. In: PCR protocols: a guide to methods and amplifications. Innis MA, Gelfand DH, Sninsky, White TJ (eds). Academic Press, San Diego. p13-20. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press, New York.

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463-5467.

Schmidt J, Herfurth E, Subramanian AR (1992) Purification and characterization of seven chloroplast ribosomal proteins: evidence that organelle ribosomal protein genes are functional and that NH_2-terminal processing occurs via multiple pathways in chloroplasts. *Plant Mol*

Biol 20:459-465

Schmidt J, Weglöhner W, Subramanian AR (1993) The nuclear genes for chloroplast ribosomal protein L11 and L12 in higher plants. In: The translational apparatus. Nierhaus KH, Franceschi F, Subramanian AR, Erdmann VA, Wittmann-Liebold B (eds). Plamun Press, New York. p719-737

Schuster W (1993) Ribosomal protein gene *rpl5* is co-transcribed with the *nad3* gene in Oenothera mitochondria. *Mol Gen Genet* 240:445-449.

Schuster W, Brennicke A (1987) Plastid, nuclear and reverse transcriptase sequences in the mitochondrial genome of *Oenothera*: is genetic information transferred between organelles via RNA? *EMBO J* 6:2857-2863

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 *coxI* and *rps3* mRNAs of *Oenothera. Nucl Acids Res* 19:6923-6928

Schuster W, Hiesel R, Issac PG, Leaver CJ, Brennicke A (1986) Transcript termini of messenger RNAs in higher plant mitochondria. *Nucl Acids Res* 14:5943-5954

Schuster W, Unseld M, Wissinger B, Brennicke A (1990a) Ribosomal protein S14 transcripts are edited in *Oenothera* mitochondria. *Nucl Acids Res* 18:229-233

Schuster W, Wissinger B, Unseld M, Brennicke A (1990b) Transcripts of the NADH-dehydrogenase subunit 3 gene are differentially edited in *Oenothera* mitochondria. *EMBO J* 9:263-269

Schuster W, Combettes B, Flieger K, Brennicke A (1993) A plant mitochondrial gene encodes a protein involved in cytochrome c biogenesis. *Mol Gen Genet* 239:49-57.

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J* 5:2043-2049

Siemeister G, Buchholz C, Hachtel W (1990) Genes for ribosomal proteins are retained on the 73Kb DNA from *Astasia longa* that resembles *Euglena* chloroplast DNA. *Curr Genet* 17:457-464

Simpson L, Shaw J (1989) RNA editing and the mitochondrial cryptogenes of kinetoplastid protozoa. *Cell* 57:355-366

Sippel AE (1973) Purification and characterization of adenosine triphosphate: Ribonucleic acid adenyltransferase from *Escherichia coli*. *Eur J Biochem* 37:31-40

Sloof P, Benne R (1993) RNA editing in trypanosome mitochondria: guidelines for models. *FEBS Lett* 325:146-151

Small I, Maréchal-Drouard L, Masson J, Pelletier G, Cosset A, Weil JH, Dietrich A (1992) *In vitro* import of a normal and mutagenized heterologous transfer RNA into the mitochondria of transgenic plants: towards novel ways of influencing mitochondrial gene expression. *EMBO J* 12:1291-1296.

Smooker PM, Kruft V, Subramanian AR (1990) A ribosomal protein is encoded in the chloroplast DNA in a lower plant but in the nucleus in Angiosperms. *J Biol Chem* 265:16699-16703

Spangenberg G, Pèrez-Vicent R, Oliveira MM, Osusky M, Nagel J, Pais MS, Potrykus I (1992) CMS system in *Nicotiana*: Flower development, patterns of mitochondrial DNA and mitochondrial gene expression. I. Analysis of parental *Nicotiana* species and alloplasmic CMS analogs of tobacco. *Sex Plant Reprod* 5:13-26.

Sparace SA, Moore Jr TS (1979) Phospholipid metabolism in plant mitochondria. *Plant Physiol* 63:963-972

Spencer DF, Schnare MN, Gray MW (1984) Pronounced structural similarities between the small subunit ribosomal RNA genes of wheat mitochondria and *Escherichia coli*. *Proc Natl Acad Sci USA* 82:493-497

Stahl DJ, Rodermel SR, Bogorad L, Subramanian AR (1993) Co-transcription pattern of an introgressed operon in the maize chloroplast genome comprising four ATP synthase subunit genes and the ribosomal *rps2*. *Plant Mol Biol* 21:1069-1076

Stegink SJ, Siedow JN (1986) Binding of butyl gallate to plant mitochondria. *Plant Physiol* 80:196-201

Stern DV, Lonsdale DM (1982) Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. *Nature* 299:698-702.

Stern DV, Newton KJ (1986) Isolation of plant mitochondrial RNA. *Methods in Enzymol.* 118:488-496.

Stuart K, Feagin JE (1992) Mitochondrial DNA of Kinetoplastids. *Int Rev Cytol* 141:65-88

Subramanian AR (1991) Nuclear-coded chloroplast r-proteins, precursor cDNA clones and transit sequences. In: The translational apparatus of photosynthetic organelles. Mache R, Stutz

E, Subramanian AR (eds). Springer-Verlag, Berlin Heidelberg p107-115

Subramanian AR (1993) Molecular genetics of chloroplast ribosomal proteins. *TIBS* 18:177-181.

Sugiura M, Torazawa K, Wakasugi T (1991) Chloroplast genes coding for ribosomal proteins in land plants. In: The translational apparatus of photosynthetic organelles. Mache R, Stutz E, Subramanian AR (eds). Springer-Verlag, Berlin Heidelberg p59-69

Sugiura M (1992) The chloroplast genome. *Plant Mol Biol* 19:149-168.

Sun CW, Callis J (1993) Recent stable insertion of mitochondrial DNA into an *Arabidopsis* polyubiquitin gene by nonhomologous recombination. *Plant Cell* 5:97-107

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

Sutton BCS, Flanagan DJ, Gawley JR, Newton CH, Lester DT, El-kassaby (1991) Inheritance of chloroplast and mitochondrial DNA in *Picea* and composition of hybrids from introgression zones. *Theor Appl Genet* 82:242-248

Sutton CA, Conklin PL, Pruitt KD, Calfee AJ, Cobb AG, Hanson MR (1993) Editing of *rps3/rpl16* transcripts creates a premature truncation of the *rpl16* open reading frame. *Curr Genet* 23:472-476.

Suzuki T, Kazama S, Hirai A, Akihama T, Kadowaki K (1991) The rice mitochondrial *nad3* gene has an extended reading frame at its 5' end: nucleotide sequence analysis of rice *trnS*, *nad3*, and *rps12* genes. *Curr Genet* 20:331-337

Tabor S, Richardson CC (1987) DNA sequence analysis with a modified bacteriaophage T7 DNA polymerase. *Proc Natl Acad Sci USA* 84:4767-4771

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

Thompson WF, Everett M, Polans NO, Jorgensen RA, Palmer JD (1983) Phytochrome control of RNA levels in developing pea and mung-bean leaves. *Planta* 158:487-500.

Thorsness PE, Fox TD (1990) Escape of DNA from mitochondria to the nucleus in *Saccharomyces cerevisiae*. *Nature* 364:376-379 van den Boogaart P, Samallo J, Agsteribbe E (1982) Similar genes for a mitochondrial ATPase subunit in the nuclear and mitochondrial genomes of *Neurospora crassa*. *Nature* 298:187-189.

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

Wahleithner JA, Wolstenholme DR (1988) Ribosomal protein S14 genes in broad bean mitochondrial DNA. *Nucl Acids Res* 16:6897-6913

Walberg MW, Clayton DA (1983) *In vitro* transcription of human mitochondrial DNA. *J Biol Chem* 258:1268-1275

Ward BL, Anderson RS, Bendich AJ (1980) The size of the mitochondrial genome is large and variable in a family of plants (*Cucurbitanceae*). *Cell* 25:793-803.

Weglöhner W, Subramanian AR (1993) Nucleotide sequence of maize chloroplast *rpl32*: completing the apparent set of plastid ribosomal protein genes and their tentative operon organization. *Plant Mol Biol* 21:543-548

Werner D, Chemla Y, Herberg M (1984) Isolation of poly $(A)^+$ RNA by paper affinity-chromatography. *Anal Biochem* 141:329-336

Weslowski M, Algeri A, Fukuhara (1981) Gene organization of the mitochondrial DNA of yeasts: *Kluyveromyces lactis* and *Saccharomycopsis lipolytica*. *Curr Genet* 3:157:162.

West RW, Yocum RR, Ptashne M (1984) *Saccharomyces cerevisiae GAL1-GAL10* divergent promoter region: location and function of the upstream activating sequence UAS_G. *Mol Cell Biol* 4:2467-2478.

Wilson AJ, Chourey PS (1984) A rapid inexpensive method for the isolation of restrictable mitochondrial DNA from various plant sources. *Plant Cell Rep* 3:237-239.

Wintz H, Grienenberger JM, Weil JH, Lonsdale DM (1988) Location and nucleotide sequence of two tRNA genes and a tRNA pseudo-gene in the maize mitochondrial genome: evidence for the transcription of a chloroplast gene in mitochondria. *Curr Genet* 13:247-254

Wissinger B, Schuster W, Brennicke A (1990) Species-specific RNA editing patterns in the mitochondrial *rps13* transcripts of *Oenothera* and *Daucus*. *Mol Gen Genet* 224:389-395

Wissinger B, Schuster W, Brennicke A (1991) *Trans* splicing in Oenothera mitochondria: *nad1* mRNAs are edited in exon and *trans*-splicing group II intron sequences. *Cell* 65:473-482.

Wissinger B, Brennicke A, Schuster W (1992) Regenerating good sense: RNA editing and *trans* splicing in plant mitochondria. *TIG* 8:322-328

Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA* 87:4576-4579.

Wolf K, Del Giudice L (1988) Variable mitochondrial genome of ascomycetes: organization, mutational alternations and expression. *Adv Genet* 25:185-308

Wolfe KH, Gouy M, Yang YW, Sharp PM, Li WH (1989) Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. *Proc Natl Acad Sci USA* 86:6201-6205

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.

Wolff G, Plante I, Lang BF, Kück U, Burger G (1994) Complete sequence of the mitochondrial DNA of the chlorophyte alga *Prototheca wickerhamii*: gene content and genome organization. *J Mol Biol* 237:75-86

Wolstenholme DR (1992) Animal mitochondrial DNA: structure and evolution. *Int Rev Cytol* 141:65-88

Wolter AMA, Koornneef M, Gilissen LW (1993) The chloroplast and mitochondrial DNA types are correlated with the nuclear composition of somatic hybrid calli of *Solanum tuberosum* and *Nicotiana plumbaginifolia*. *Curr Genet* 24:260-267

Yei F, Bernhardt J, Abel WO (1993) Genes for ribosomal proteins S3, L16, L5 and S14 are clustered in the mitochondrial genome of *Brassica napus* L. *Curr Genet* 24:223-329.

Yocum RR, Hanley S, West R, Ptashne M (1984) Use of *lacZ* fusions to delimit regulatory elements of the inducible divergent *GAL1-GAL10* promoter in *Saccharomyces cerevisiae*. *Mol Cell Biol* 4:1985-1998.

Young EG, Hanson MR (1987) A fused mitochondrial gene associated with cytoplasmic male sterility is developmentally regulated. *Cell* 50:41-49

Zhuo DG, Ye ZX and Wang XM (1988) Restriction endonuclease analysis of mitochondrial DNA from normal and cytoplasmic male-sterile rice. *J Fudan University (Natural Science)* 27:367-373.

Zhuo DG, Bonen L (1993) Characterization of the S7 ribosomal protein gene in wheat mitochondria. *Mol Gen Genet* 236:395-401

Zwieb C, Brimacombe R (1979) RNA-protein cross-linking in *Escherichia coli* 30S ribosomal subunits: precise localisation of the nucleotide in 16S RNA which is coupled to protein S7 by ultraviolet irradiation. *Nucl Acids Res* 6:1775-1790

Appendix A: Vectors and host strains/ LB broth media additives.

Vectors	*E. coli* Hosts	Media Additives
EMBL3	NM538*	0.2% maltose
	NM539*	0.2% maltose
Lambda gt11	Y1090r$^-$	50 µg/ml ampicillin
EMBL3 SP6/T7	K802**	0.2% maltose
pUC	TB1	50 µg/ml ampicillin; 50 µg/ml X-gal; 20 mg/ml IPTG
M13	JM101	50 µg/ml ampicillin; 50 µg/ml X-gal; 20 mg/ml IPTG

* NM538: a permissive E.coli strain.

** K802: suppressing E. coli strain

Appendix B: Buffers often used in these experiments.

Buffer I:	0.5M sucrose
	50mM Tris-HCl; pH 7.5
	5mM EDTA
	0.1% BSA
	0.1% mercapto-2-ethanol
Buffer II:	0.3M sucrose
	50mM Tris-HCl; pH 7.5
	1mM EDTA
	0.1% BSA
Buffer III:	0.3M sucrose
	15mM $MgCl_2$
	50mM Tris-HCl; pH7.5
Buffer IV:	50mM Tris-HCl, pH 8.0
	20mM EDTA
Buffer V:	10mM Tris-HCl, pH 7.5
	50mM KCl

10mM $MgCl_2$

DNA Extraction Buffer:	0.1M Tris-HCl, pH8.0
	0.1M EDTA
	0.2M NaCl
	100ug/ml proteinase K
TE Buffer:	10mM Tris-HCl, pH7.5
	1mM EDTA
Low TE Buffer:	10mM Tris-HCl, pH7.5
	0.1mM EDTA
2X detergent mix:	2% tri-isopropylnaphtalene sufanate
	12% sodium p-aminosalicylate
	0.1M NaCl

Appendix C: Preliminary DNA sequences of the cDNA TE46.

te46lqa1

TTTTAAAATTGGGAACCTACCGAGTTTTTAAGATTTTCCATACCCATGTTTCCCGAGG

CGACACTTTGTTTCACTCCTTTGATTTATACTCTAGCCGGCGTTTGGTATAATATATAC

CATCCTTGACCACTTTAAGAGAGGTCGGTCTTAAGAAGATCATTGCCTAGACCTG

GAGTAGGCCCGATATG

te46lqza

AATTTTTTTAAAATTGGGAACCTACGGGTTTTTAAGATTTTCCATACCCATGTTTCCCG

AGGCGAGAGTTTGTGTAACTCCTTTCATTTATACTCTAAGCGCGATTGCTATAATATATA

CTTCCTTGACCACTTTAAGACAGATCGATCTTAAGAAGATCATTGCCTACTCCTCAGTAG

CCCATATGAANTTTGGTGACTATTCGTATGTTAAGTCTGAACATACAAGGGATCGGAAGG

GAGCATNTTACTATTGTTAAGTCATGGACTTGGTTTNTAAGCCCTT

te46sa

AGCACATCAGTATTTCCTATATACTTTACTGGGATCNGTTTTTATGCTATTAGCTATTCT

GTTGATTCTTCTCCAAACAGGAACCACCGATTTACAAATTTTATTAACCATTGAAT

TTAGTGAGCGGCGCCAANTCCTTCTATGGATTTCTTTTTTTGCCTCTTTTGCCGTTAAAA

GCCTATTGGTACCAGTTNATATTTGGTTACCCGAAGCCCATGTAGAGGCAACTACGGAT

GGATCCGTCATCTTGGATGG

Appendix D: DNA sequence comparison between the soybean (upper stand) and pea (lower strand) mitochondrial *orf228-r* gene and *rps7* pseudogene region.

```
ATTAACTTTGTCACATGAATTGACTTTTTTTTTTTCGGTATGCCGCTCCGCCAGCAAGGA
ATTAAGTATA CTCATGATCTTTCTTTTTTTTTC CGGTATGCCGCTCCGCCAGCAAGGA
     * * ** *     ** **          **

GTGCCACGCCGAAGCGGA AGACA ACTAAGCAGGGGGAATTATTCGTTGGGAGAAATCC
GTGCCACGCCGAAGCGGACAGANANACAAAGCAGGGGGAATTATTCGTTGGGAGAAATCC
                  *   * *  *

TTTGATTGCGTATT AATATA ATCCATGTCTTTCTTGTTCCACTAGCTAGGACAAAATT
TTTGATTGCGTATTGAATATAGATCCATGTCTTTCTTGTTCCACTAGCTAGGAGAAAATT
              *      *                               *

CTACTATGTCCCTTT G    C TTATTACAACCTTCTTTTTTGAGTTCAAAGACCAGA
ATCAGATGTCCCTTTCGTTATCGTTATTACAACCTTCTTTTTTGATGTCAAAGACCAGA
* ***          * **** *                      **

AGCTATGGCCTAATTCATCTTGGATCTCGGTTGTTCTTAACAGCGATCGGCTATTCATTT
AGCTGTGCGCTAATTCTCATTGGATCTCGGTTGTTCTTAACAGCGAT GGCTATTCATTT
    *  **       ***                            *

AAGTCTTCGGGTAGCACCATTAGATCTTCAACAAGGTGGAAATTCTCGTATTCTGTATGT
AAGTCTTCGGGTAGCACCATTAGATCTTCAACAAGGTGGAAATTCTCGTATTCTTTATGT
                                                      *

ACATGTTC TGCGGCTCGGATGAGTATTCTTGTTTATAGTCCTACGGCTATAAACACTTT
ACATGTTCCTGCGGCTCGGATGAGTATTCTTGTTTATATCGCTACGGCTATAAACACTTT
        *                             ***

CTTATTCCTATTAACAAAACATCCCCTTTTTCTTCGCTCTTCCGGAACCGGTACAGAAAT
CTTATTCCTATTAACAAAACATCCCCTTTTTCTTCGCTCTTCCGGAACCGGTACAGAAAT

GGGTACTTTTTTTACGTTGTTTACCTTAGTTACTGGGGGGTTTCGGGGAAGACCTATGTG
GGGTGCTTTTTTTACGTTGTTTACCTTAGTTACTGGGGGGTTTCGGGGAAGACCTATGTG
    *

GGCACCTTTTGGGTGTGGGATGCTCGTTTAACCTCTGTATTCATCTCGTTTCTGATTTA
GGCACCTTTTGGGTGTGGGATGCTCGTTTAACCTCTGTATTCATCTCGTTTCTGATTTA

CCTGGGTGCACTGCGTTTTCAAAAGCTTCCTGGTCGAACCGGCTCCTATTTCAATCCGTG
CCTGGGTGCACTGCGTTTTCAAAAGCTTCCTG TCGAACCGGCTCCTATTTCAATCCGTG
                                        *
```

```
CTGGACCGATCGATATACCAATAATAAAGTCTTCAGTCAACTGGTGGAATACATTACATC
CTGGACCGATCGATATACCAATAATAAAGTCTTCAGTCAACTGGTGGAATACTTTGCATC
                                                    *  *

AACCTGGGAGCATTAGCCGATCTGGTACATCAATACATGTTCCTATGCCCATTCCAATCT
AACCTGGGAGCATTAGCCGATCTGGTACATCAATACATGTTCCTATGCCCATTCCAATCT

TGTCTAACTTTGCTAACTTCCCCCTCTCAACCCCTATCTTGTTTGTTCTGGAAACACGTC
TGTCTAACTTTGCTAACTTCCCCCTCTCAACCCGTATCTTGTTTGTTCTGGAAACACGTC

TTCCTATTCTATCTTTTCCCGAATCTCCTTTAAGGGATGAAATCAGAAGCTCGAGAAGGA
TTCCTATTCTATCTTTTCCCGAATCTCCTTTAAGGTATGAAAT AGAAGCTCGAGAAGGA

ATAGCAAAACCTAGTTTACTTCCCAGCTCAAACTGAACGC A   AA C    TTATGGT
ATAGCAAAACCTAGTTTACTTCCCAGCTCAAACTGAACGCGATGTAATCAAACTTATGGT
                                        * ***  * ****

TGACGCCGTAGAGAATATAAAGCCCATATGCGAAGTGGAAAAAGTA
TGACGCCGTAGAGAATATAAAGCCCATATGCGAAGTGGAAAAAGTA
```