

# RNA-Mediated Transfer of the Gene *coxII* from the Mitochondrion to the Nucleus during Flowering Plant Evolution

Jacqueline M. Nugent\*† and Jeffrey D. Palmer†

\*Department of Biology  
University of Michigan  
Ann Arbor, Michigan 48104  
†Department of Biology  
Indiana University  
Bloomington, Indiana 47405

## Summary

The gene *coxII*, normally present in the mitochondrion, was functionally transferred to the nucleus during flowering plant evolution. *coxII* transfer is estimated to have occurred between 60 and 200 million years ago, whereas loss of *coxII* from the mitochondrion occurred much more recently, being restricted to a single genus of legumes. Most legumes have *coxII* in both the nucleus and the mitochondrion; however, no evidence is found for simultaneous *coxII* expression in both compartments. The nuclear *coxII* sequence more closely resembles edited mitochondrial *coxII* transcripts than the genes encoding these RNAs. Hence, gene transfer appears to have involved reverse transcription of an edited RNA intermediate. The nuclear gene contains an intron at the junction of the transit peptide sequence and the mature protein-coding sequence; exon shuffling may have played a role in assembling a functional *coxII* gene in the nucleus.

## Introduction

The idea that some nuclear genes encoding organellar proteins are the result of evolutionary gene transfer from the organelle to the nucleus has gained wide acceptance in recent years (Shih et al., 1986; Froman et al., 1989). Evidence for ongoing gene transfer has been demonstrated for the chloroplast protein genes *tufA* and *rpl22* (Baldauf and Palmer, 1990; J. S. Gantt, S. L. Baldauf, P. Calie, N. Weeden, and J. D. Palmer, unpublished data). Molecular and phylogenetic studies indicate that *tufA* was transferred to the nucleus in a green algal ancestor of land plants (at least 500 million years ago), while *rpl22* transfer occurred later in a common ancestor of all flowering plants (at least 200 million years ago). In addition to these well-characterized gene transfers, a number of gene content differences are known among the three completely sequenced chloroplast genomes (Ohyama et al., 1986; Shinzaki et al., 1986; Hiratsuka et al., 1989); some of these differences may also reflect gene transfer.

Much less is known about the conservation of plant mitochondrial gene content, mainly due to the larger size (200–2400 kb; Newton, 1988) and more complex organization of these organellar genomes. Variability in plant mitochondrial gene content has been demonstrated for three ribosomal protein genes (Bland et al., 1986; Bonen, 1987; Wahleithner and Wolstenholme, 1988; Schuster et al.,

1990a) and for several tRNA genes (Binder et al., 1990; Bartnik and Borsuk, 1986; Izuchi et al., 1990; Maréchal-Drouard et al., 1990). Members of each of these two classes of genes are found in some plant species but not in others. However, whether this variability reflects functional gene transfer to the nucleus is currently unknown.

A central question pertaining to intracellular gene transfer concerns the chemical nature of the nucleic acid molecule involved in the transfer process; are these events mediated solely by DNA or are both DNA and RNA involved? It has been suggested that some interorganellar DNA transfers may occur via an RNA intermediate, followed by local reverse transcription and genome integration (Schuster and Brennicke, 1987). So far, however, there has been little evidence to support this hypothesis; indeed, the first discovered and most dramatic case of gene transfer (albeit one without any apparent transfer of function; Stern and Lonsdale, 1982) almost certainly occurred via a strictly DNA intermediate (Nugent and Palmer, 1988). The recent discovery of plant mitochondrial RNA editing (Covello and Gray, 1989; Gualberto et al., 1989; Hiesel et al., 1989) allows for a more definitive answer to this question than was possible previously. Plant mitochondrial RNA editing involves the selective conversion of C to U residues (and occasionally U to C residues; Gualberto et al., 1990; Schuster et al., 1990b) at various positions along most if not all mRNAs and results in the production of a protein whose sequence is significantly different from that predicted from the DNA sequence. Thus, finding that the sequence of a nuclear gene, recently transferred from the mitochondrion, looks more like an edited transcript of the progenitor mitochondrial gene than the gene itself would provide strong evidence that transfer involved an RNA intermediate. Conversely, finding the transferred sequence more like the mitochondrial DNA sequence would suggest that DNA alone mediated the gene transfer event and would raise the possibility of cytoplasmic RNA editing.

We report loss of the gene *coxII* from the mitochondrion within a single genus of legumes. *coxII* encodes subunit II of the cytochrome c oxidase complex, an essential enzyme involved in the terminal oxidative step in energy metabolism (Capaldi et al., 1983). Evidence is presented for RNA-mediated transfer of *coxII* from the mitochondrion to the nucleus during flowering plant evolution.

## Results

### Loss of *coxII* from the Mitochondrion

Southern blot hybridization was used to examine mitochondrial gene content in 20 flowering plant species using probes for most known mitochondrial protein genes. Based on this survey it appears that the majority of these genes are retained in the mitochondrion over a broad taxonomic range. A striking exception to this, however, is the gene *coxII*. A probe internal to the soybean (*Glycine max*) mitochondrial *coxII* gene failed to hybridize to mung bean

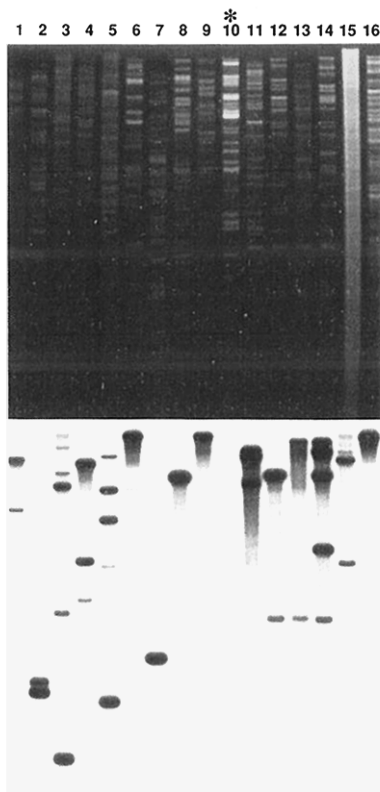


Figure 1. *coxII* Is Present in the mtDNA of All Species Examined except Mung Bean

(Top) Electrophoresis in a 1% agarose gel of mtDNA from *Triticum aestivum* (wheat, lane 1, PstI fragments), *Zea mays* (maize, lane 2, BamHI), *Lactuca sativa* (lettuce, lane 3, XhoI), *Helianthus annuus* (sunflower, lane 4, PstI), *Lycopersicon esculentum* (tomato, lane 5, BamHI), *Nicotiana tabacum* (tobacco, lane 6, PstI), *Oenothera johansen* (evening primrose, lane 7, HindIII), *Pisum sativum* (pea, lane 8, PstI), *Phaseolus vulgaris* (common bean, lane 9, PstI), *Vigna radiata* (mung bean, lane 10, PstI; marked by asterisk), *Capsella bursa-pastoris* (shepherd's purse, lane 11, PstI), *Crambe abyssinica* (lane 12, PstI), *Raphanus sativus* (radish, lane 13, PstI), *Brassica campestris* (cauliflower, lane 14, PstI), *Arabidopsis thaliana* (lane 15, PstI), and *Cucurbita maxima* (squash, lane 16, PvuII). Other mtDNAs included in the survey but not shown here are cucumber, watermelon, spinach, and petunia.

(Bottom) Hybridization to a filter blot of the agarose gel with a 704 bp NcoI-HindIII fragment internal to the soybean mitochondrial *coxII* gene.

(*Vigna radiata*) mitochondrial DNA (mtDNA) (Figure 1, lane 10) but hybridized strongly to all other mtDNAs examined, including those of two other legumes, pea (*Pisum sativum*) and common bean (*Phaseolus vulgaris*) (Figure 1, lanes 8 and 9). The presence of *coxII* in common bean mtDNA and its apparent absence from mung bean mtDNA are particularly interesting since the two species are considered closely related (indeed, were once treated as congeneric). The *coxII* gene loss within *Vigna* therefore appears to be a recent evolutionary event, having occurred since the divergence of *Vigna* and *Phaseolus* from a common ancestor.

#### *coxII* in the Nucleus

Considering the essential nature of the protein encoded

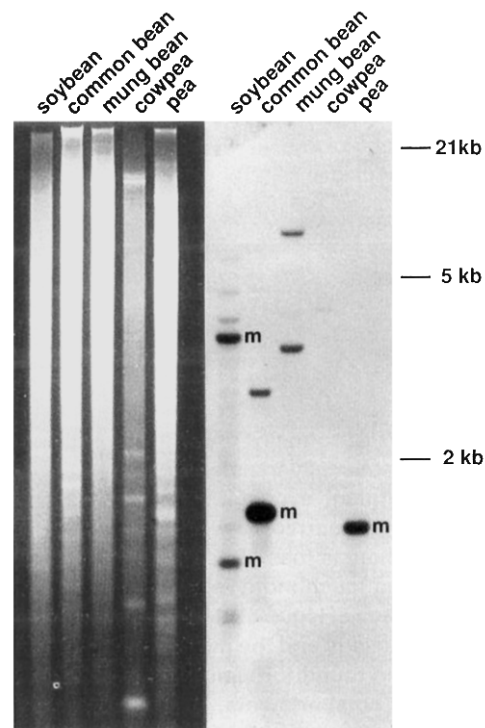


Figure 2. *coxII* Hybridizes to nuDNA from Five Species of Legumes  
Nuclear DNAs were cut with HindIII, electrophoresed in a 1% agarose gel, transferred to a nylon filter, and hybridized with the probe described in Figure 1. The signals labeled "m" denote hybridization to contaminating mtDNA.

by the *coxII* gene, it seemed unlikely that the gene had been lost entirely in mung bean and more likely that the functional gene resided in the nucleus of this species. To test this possibility, nuclear DNA (nuDNA) from mung bean and four other legumes (soybean, common bean, pea, and cowpea [*Vigna unguiculata*]) was probed for the presence of a *coxII* gene (Figure 2). Both mung bean and cowpea lack a mitochondrial *coxII* gene, and consequently their nuDNA hybridization results are complicated by signals due to mtDNA contamination of the nuclear DNA preparations. These signals (marked "m," Figure 2) were diagnosed as such by parallel hybridizations to purified mtDNA (data not shown). We attribute all other signals to hybridization of *coxII* to nuclear DNA. The number of nuclear signals observed suggests one to two nuclear *coxII* sequences in cowpea, mung bean, common bean, and pea (all diploids), and additional nuclear copies in soybean (a tetraploid; Figure 2). These results are reproducible and are supported by Southern blots using other restriction enzymes (data not shown). The weakness of the cowpea signal is attributed to underloading of this lane relative to the others (Figure 2, left). The weak nuclear signals in pea may be explained by genome size differences, the pea genome being 3–10 times larger than that of the other legumes examined (Bennett and Smith, 1976); alternatively, only part of *coxII* may be present in this genome.

### ***coxII* cDNA**

To examine the putative nuclear *coxII* gene, clones were isolated from a cowpea cDNA library using the soybean mitochondrial *coxII* gene as a probe. A polyadenylated cDNA clone was found to contain an open reading frame with 85% nucleotide sequence identity to the soybean mitochondrial *coxII* gene. This reading frame is approximately 318 bp longer at its 5' end than mitochondrial *coxII* genes (Figure 3). If translation begins at the first of the three ATG codons within this extension, the gene would encode an N-terminal transit peptide sequence of approximately 60 amino acids, which could function in directing the protein back to the mitochondrion. This putative transit peptide resembles other known mitochondrial transit peptides in having few negatively charged amino acids, a net positive charge, and an abundance of hydroxylated residues (von Heijne et al., 1989). The 5' sequence of mitochondrial *coxII* genes differs from that of the nuclear gene, and thus the exact processing site for the mature nuclear-encoded protein cannot be determined. The translation stop signal is followed by 295 bp of sequence before the site of poly(A) addition. Possible polyadenylation signals (AATAAA) are located at 196 and 74 bp 5' to the polyadenylation site.

The *coxII*-encoding region of the cDNA clone was compared with the sequences of mitochondrial *coxII* genes and transcripts from four flowering plants (two monocots and two dicots). The cowpea *coxII* cDNA sequence more closely resembles edited mitochondrial *coxII* transcripts than the genes encoding these mRNAs (Figure 3). The mitochondrial *coxII* RNA sequences are edited via C to U conversion at a combined total of 30 sites. The majority of these are fully edited sites (i.e., within a species all *coxII* transcripts are edited at these positions), while the remainder are partially edited sites (i.e., within a species these positions are edited in some *coxII* transcripts but not others). Twenty-three of these 30 editing sites contain a T in the cowpea cDNA sequence. This result suggests two alternative scenarios, both of which have intriguing implications: the cowpea cDNA clone was derived from an edited nuclear *coxII* transcript; or an edited version of the *coxII* gene is present in the cowpea nuclear genome.

### **Nuclear *coxII* Gene**

Oligonucleotide primers complementary to 5' and 3' regions of the sequenced cDNA clone were used to amplify with the polymerase chain reaction (PCR) the nuclear *coxII* gene from cowpea total DNA. The genomic and cDNA sequences are in complete agreement (data not shown), implying that an edited version of the *coxII* gene is present in the cowpea nucleus.

The cowpea *coxII* gene has a typical nuclear intron (Breathnach and Chambon, 1981), 740 bp in length, located at or near the junction of the putative transit peptide sequence and the mature protein-coding sequence. The alignment presented in Figure 3 places the intron within what is perhaps the first codon of the mature *coxII* protein. However, this placement is speculative since the precise junction of the transit and mature polypeptides is not known.

### **Timing of *coxII* Transfer to the Nucleus**

To address the timing of the *coxII* gene transfer event, we constructed a phylogeny of eight sequenced *coxII* genes (including the cowpea nuclear *coxII* gene) and compared this with an independent "organismal" phylogeny (based on morphological data and chloroplast *rbcl* sequence data) for the same eight species (Figure 4). The organismal phylogeny places cowpea closer to soybean than to any of the other species included in this analysis (Figure 4A). If the gene transfer were a recent evolutionary event that had taken place not long before the loss of *coxII* from the mitochondrion in *Vigna*, we would expect the gene phylogeny to mirror the organismal phylogeny, i.e., the cowpea nuclear sequence should be placed closer to the soybean mitochondrial sequence than to any of the other mitochondrial sequences in the analysis. In fact, however, a phylogenetic tree derived from a distance analysis of nucleotide sequences (Figure 4B) places the cowpea nuclear gene on a very long branch outside the legume mitochondrial lineage as defined by pea and soybean, members of two distantly related tribes in the legume subfamily Papilionoideae. The bootstrap value supporting the pea/soybean node on this tree is 100%.

This result suggests that transfer of *coxII* to the nucleus predated the divergence of the subfamily Papilionoideae. This would place the timing of the transfer event between 60 and 200 million years ago, before the emergence of the Papilionoideae but since the divergence of the monocots and dicots (Crepet and Taylor, 1985; Wolfe et al., 1989b). The nine shortest trees yielded by a parsimony analysis (data not shown) using the same nucleotide sequences all placed cowpea outside soybean and pea, as in the distance analysis. However, three of these nine trees also placed *Oenothera* as the closest relative to pea, a clearly anomalous result that may be related to an accelerated mutation rate in *Oenothera* relative to the other mitochondrial *coxII* sequences (note the long *Oenothera* branch length in Figure 4B). Excluding *Oenothera* from the parsimony analysis gave a single shortest tree (data not shown) that was in agreement with the distance tree presented here, with the pea/soybean node supported by a bootstrap value of 94%. Parsimony analysis using deduced amino acid sequences gave the same result.

### ***coxII* Expression in Legumes**

Southern blot and phylogenetic data suggest that *coxII* might be present in both the mitochondrial and nuclear genomes of most legumes. Northern blot hybridizations were therefore carried out to determine which genes are expressed—mitochondrial, nuclear, or both. A soybean mitochondrial *coxII* probe reveals a 1.3 kb transcript in cowpea that is substantially enriched in poly(A)<sup>+</sup> RNA and absent from mtRNA (Figure 5B). This result indicates that *coxII* mRNA is solely the product of a nuclear gene in cowpea and is consistent with the gene distribution data (Figure 6). The published sequence of the soybean mitochondrial *coxII* gene predicts a functional gene with 97.7% sequence identity to the pea mitochondrial *coxII* gene (Grabau, 1987). Northern blot results, however, suggest that only the nuclear *coxII* gene(s) in soybean is ex-



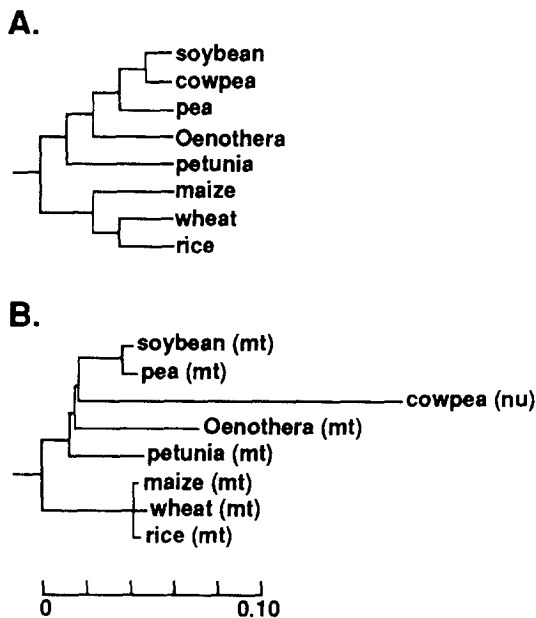


Figure 4. *coxII* Was Transferred to the Nucleus in a Common Ancestor of the Legume Subfamily Papilionoideae

(A) Organismal phylogeny based on morphological data (Cronquist, 1988; Donoghue and Doyle, 1989) and chloroplast *rbcL* sequence data (R. Olmstead, K. Sytsma, and J. Palmer, unpublished data). (B) Phylogenetic tree derived from distance analysis of *coxII* DNA sequence data. Mitochondrial sequences were taken from Fox and Leaver (1981) (maize); Hiesel and Brennicke (1983) (*Oenothera*); Bonen et al. (1984) (wheat); Kao et al. (1984) (rice); Moon et al. (1985) (pea); Grabau (1987) (soybean); and Pruitt and Hanson (1989) (petunia). The tree was constructed by the method of Olson (1988a, 1988b) and is rooted using the monocots as the outgroup. Horizontal branches are drawn proportional to length; the scale bar below the tree indicates extent of sequence divergence. Positions known to be edited in any species were excluded from the analysis.

pressed. A transcript of approximately 1.2 kb is observed in the soybean poly(A)<sup>+</sup> RNA lane, no transcript is detected in the poly(A)<sup>-</sup> RNA lane, and only a faint broad smear (with "ghost" areas in regions of rRNA bands) is seen in the mtRNA lane. Part of the *coxII* gene is known to be repeated within the soybean mitochondrial genome, and it is possible that this hybridization smear is the result of transcription through this region (Grabau, 1987). In common bean the mitochondrial *coxII* gene is also silent and only the nuclear gene(s) is expressed (data not shown; Figure 6). Conversely, in pea the mitochondrial gene is expressed and the nuclear *coxII* sequences are silent (Figure 5B). Rehybridization of the same filter with a pea probe specific for the nuclear H-protein gene (encoding the smallest subunit of the mitochondrial glycine decarboxylase complex; Kim and Oliver, 1990) demonstrates that a nuclear transcript can be readily detected in the pea poly(A)<sup>+</sup> RNA lane (Figure 5C).

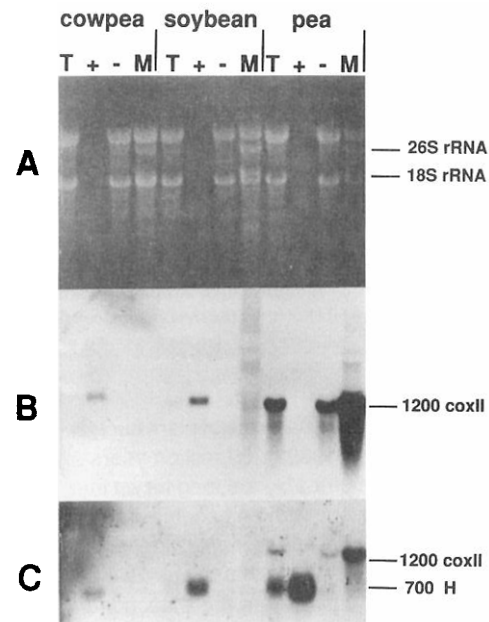


Figure 5. *coxII* Is Expressed in the Nucleus of Cowpea and Soybean but in the Mitochondrion of Pea

(A) Electrophoresis in a 1.2% agarose gel of total RNA (T), poly(A)<sup>+</sup> RNA (+), poly(A)<sup>-</sup> RNA (-), and mitochondrial RNA (M) from cowpea, soybean, and pea. (B) Hybridization to a filter blot of the RNA gel with the soybean mitochondrial *coxII* probe described in Figure 1. (C) The Northern blot was stripped and rehybridized with a probe specific for the nuclear H-protein gene (encoding the smallest subunit of the mitochondrial glycine decarboxylase complex). *coxII* signals in this panel represent carryover from the previous *coxII* hybridization.

## Discussion

### Differential Timing of Events in Gene Transfer

Transfer of a functional *coxII* gene to the nucleus can be resolved into at least three temporally distinct phases (Figure 6): physical transfer of the gene to the nucleus; activation of the nuclear gene and concomitant inactivation of the mitochondrial gene; and loss of the mitochondrial gene. Phylogenetic analyses indicate that physical transfer of the *coxII* gene to the nucleus occurred 60–200 million years ago, after the emergence of the monocots and dicots, but before the origin of the legume subfamily Papilionoideae. This finding is supported by Southern blot data showing nuclear *coxII* sequences in a broad range of species within the Papilionoideae. The nuclear *coxII* gene represents the most recent example of organellar gene transfer known. The two other well-characterized cases of "modern" gene transfer both involve chloroplast genes that were transferred to the nucleus before the emergence of flowering plants (Baldauf and Palmer, 1990; Gantt et al. unpublished data).

*coxII*-coding region are underlined. Nucleotide 1375 is followed by a poly(A) tract in the cowpea cDNA. The amino acid sequence inferred from the cDNA is denoted by the single-letter code above the cDNA sequence. Underlined 18-mer oligonucleotides in the 5' and 3' untranslated regions of the cDNA sequence indicate sequences of primers used to PCR amplify the cowpea nuclear *coxII* gene. The arrow between positions 320 and 321 in the cowpea sequence indicates the location of a 740 bp intron present in the PCR-amplified nuclear *coxII* gene. The intron sequence is presented below the sequence alignments.

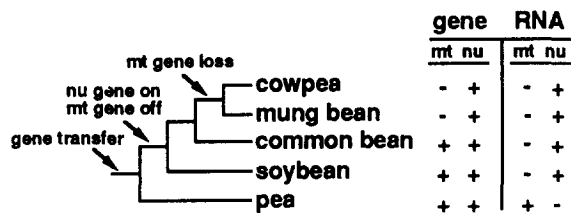


Figure 6. Schematic Representation of Timing of Events in Functional *coxII* Gene Transfer during Angiosperm Evolution

Organismal phylogeny is derived from morphological data (reviewed in Polhill, 1981). The table on the right denotes the cellular distribution of *coxII* genes and indicates which genes are being expressed.

Given the relatively fast rate of plant nuclear gene evolution (Wolfe et al., 1989a), 60 million years (minimum for gene transfer) is probably too long for an inactive gene to reside in the nucleus without becoming so highly mutated as to lose potential functionality. Indeed, Northern blot results (Figure 5) indicate that the nuclear *coxII* gene was active for some time before the mitochondrial gene was lost in the lineage leading to *Vigna*. As summarized in Figure 6, activation of the nuclear gene and inactivation of the mitochondrial gene occurred after the divergence of the lineage leading to pea but prior to the divergence of the lineages leading to soybean and common bean.

Presumably, before mitochondrial gene loss can occur, both the mitochondrial and nuclear genes must be expressed within the same lineage for at least a short period of time. However, despite apparent coexistence of nuclear and mitochondrial *coxII* genes in several lineages (Figures 2 and 6), Northern blot experiments show no evidence of dual mitochondrial and nuclear *coxII* expression in legumes (Figure 5). *coxII* expression needs to be examined in additional legumes, of phylogenetic relationship corresponding to the apparent switchpoint in mitochondrial to nuclear expression, to determine more precisely when this switch may have occurred. Dual expression cannot, however, be totally discounted from our current results. For example, posttranscriptional regulation may cause either a mitochondrial or nuclear *coxII* transcript to be degraded rapidly. Alternatively, *coxII* expression may be developmentally regulated in either the mitochondrion or the nucleus; investigation of different developmental stages may produce evidence of simultaneous nuclear and mitochondrial *coxII* transcription within these species.

The nuclear *coxII* gene is evolving considerably faster, at both nucleotide (Figure 4B) and amino acid sequence levels (data not shown), than mitochondrial *coxII* genes. This acceleration at both levels may be largely, if not entirely, due to mutation pressures in the nucleus that are 12-fold greater than in the mitochondrion (Wolfe et al., 1989a). In addition, it is possible that some amino acid changes, e.g., those aiding transport of the protein through the cytoplasm and across the mitochondrial membrane, have been selected for since the nuclear gene became functional.

#### Gene Activation Mechanism

For transferred genes to become functional, several se-

quences must be acquired, including transcription and translation signals as well as an N-terminal transit peptide sequence to direct the protein back to the organelle. The nuclear *coxII* gene has acquired two sequences since its transfer from the mitochondrion: a putative transit peptide-encoding domain and an intron that appears to divide the gene into one exon encoding this transit peptide and a second exon encoding the mature *coxII* protein. It is tempting to speculate that the intron was instrumental in establishing a functional *coxII* gene in the nucleus by promoting recombination events in the region upstream of the transferred gene, resulting in recruitment of expression sequences and the transit peptide sequence—in essence, that the nuclear gene was assembled by a form of exon shuffling (Gilbert, 1978). Several other nuclear genes, encoding either mitochondrial or chloroplast proteins, also have introns located at transit peptide/mature polypeptide junctions (Tsuzuki et al., 1987; Takeshima et al., 1988; Gantt et al., unpublished data).

#### Gene Transfer Mechanism

The existence of extensive RNA editing in plant mitochondria has allowed us to make a compelling argument for the involvement of an RNA intermediate in intracellular gene transfer. This is based on the striking sequence similarity of the cowpea nuclear *coxII* gene to edited mitochondrial *coxII* transcripts rather than to mitochondrial *coxII* genes. Twenty-three of a combined total of 30 positions known to be edited via C to U conversion in several mitochondrial *coxII* transcripts have the "edited" nucleotide equivalent in cowpea. A total of 15 edited positions are common to both monocot and dicot mitochondrial *coxII* transcripts, indicating that plant mitochondrial RNA editing predates the emergence of monocots and dicots. Since phylogenetic studies indicate that *coxII* gene transfer occurred after the divergence of monocots and dicots (Figure 4B) we can conclude that gene transfer occurred after the advent of mitochondrial RNA editing in flowering plants.

Gene transfer involving an edited RNA intermediate is therefore the likely scenario, since transfer involving DNA alone would necessitate many changes in the nuclear sequence before the establishment of a functional gene. At least eight fully edited positions in mitochondrial *coxII* transcripts may represent essential alterations of the RNA molecule. These editing events change the amino acid specificity of a codon, are observed in all mitochondrial *coxII* transcripts examined to date (Figure 3), and the edited nucleotide equivalent is found at all eight positions in the cowpea nuclear gene. At the very least, therefore, a DNA-mediated transfer event would likely require mutations at these eight positions in the nuclear *coxII* sequence before a potentially functional gene could be achieved. It is difficult to envision how such selective changes could occur in the absence of any obvious selection pressure, since until such time as a functional gene is established in the nucleus, a functional gene must also be maintained in the mitochondrion. One possibility is that these changes have occurred as a result of general enzymatic deamination of 5-methylcytosine to thymine in the nuDNA sequence, similar to the RIP process that occurs in some filamentous

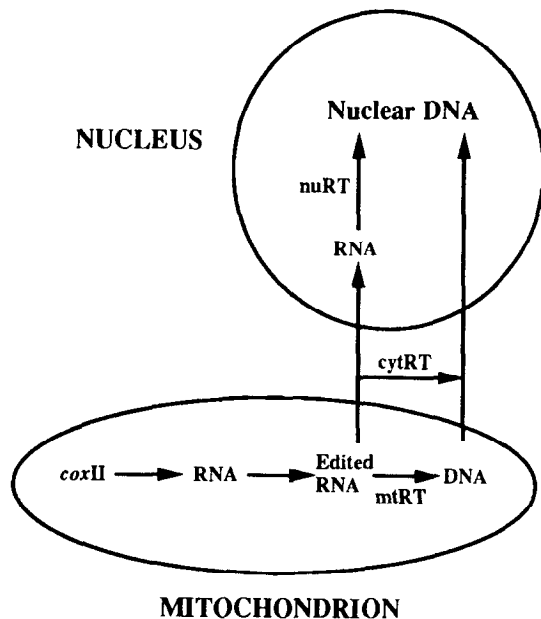


Figure 7. Several Avenues of Gene Transfer Involving an Edited RNA Intermediate May Be Possible within Plant Cells

mtRT, mitochondrial reverse transcriptase; cytRT, cytoplasmic reverse transcriptase; nuRT, nuclear reverse transcriptase.

ascumycetes (Selker, 1990). In this case one would expect potentially methylated cytosines (CpG and CpXpG) to be much rarer in the nuclear *coxII* sequence than in a corresponding mitochondrial sequence; in fact, however, mitochondrial *coxII* sequences and the cowpea nuclear *coxII* sequence contain similar numbers of potentially methylated cytosines (e.g, 35 and 37 for *Oenothera* and cowpea, respectively).

Figure 7 shows several possible avenues of gene transfer mediated by an RNA molecule. All of these involve a reverse transcription step but differ by where in the cell this step takes place and by the type of molecule that is transferred across membranes. In one scheme reverse transcription occurs in the mitochondrion and a DNA copy of the gene is transferred to the nucleus. In alternative schemes, an edited RNA molecule leaves the mitochondrion, is reverse transcribed in either the cytoplasm or the nucleus, and a DNA copy is subsequently incorporated by the nuclear genome. Reverse transcriptase-like sequences have been found in several plant mitochondrial genomes (Schuster and Brennicke, 1987; Wahleithner et al., 1990), and the existence of processed pseudogenes in plant nuclear genomes (Drouin and Dover, 1987) suggests that reverse transcriptase activity is also present in the cytoplasm and/or nucleus. Thus, plant cells could contain the reverse transcriptases necessary to facilitate any of the pathways outlined in Figure 7. Other examples of expressed genes derived by retrotransposition (but not involving intracellular gene transfer) have been found in mammals and include the preproinsulin I gene (Soares et al., 1985), the *Pgk-2* gene (McCarrey and Thomas, 1987; Boer et al., 1987), and the *Pdha-1* gene (Dahl et al., 1990).

It is possible that all successful transfers of plant mito-

chondrial genes to the nucleus have involved an edited RNA intermediate. This transfer mechanism precludes the necessity for specific DNA mutations to restore the universal genetic code and may allow for the more rapid establishment of a functional transferred gene in the nucleus. One prediction from this is that successful gene transfers may occur more readily for plant mitochondria than for either animal or fungal mitochondria, which use nonuniversal genetic codes and appear not to employ RNA editing.

#### Experimental Procedures

##### Nucleic Acid Preparations

mtDNAs were isolated from either green leaves or etiolated seedlings according to the DNAase I procedure of Kolodner and Tewari (1972). Total DNA was isolated from green leaves by the CTAB isolation procedure of Doyle and Doyle (1987).

Nuclear DNA was prepared from etiolated seedlings by the procedure of Watson and Thompson (1986). Total RNA was prepared from etiolated shoot tissue according to a scaled-up version of the guanidinium isothiocyanate extraction procedure described by Hamby et al. (1988). Poly(A)<sup>+</sup> RNA was fractionated by batch procedure using oligo(dT)-cellulose (type 3, Collaborative Research Inc.) as described by Jacobson (1987). mtRNA was prepared according to the method of Stern and Newton (1988); after lysis of the mitochondria the RNA was pelleted through a cushion of 5.7 M CsCl solution, and the remaining treatments were as described by Hamby et al. (1988) for total RNA preparation.

##### Southern/Northern Blots

DNAs were electrophoresed through 1% agarose gels before transfer to Zetabind membranes by standard blotting procedures (Maniatis et al., 1982). The membranes were prehybridized and hybridized at 60°C in Blotto solution (0.5% nonfat dry milk, 1% SDS, 4 × SSC, preheated to 60°C). Following hybridization, the filters were washed in 2 × SSC, 0.5% SDS (three 5 min washes at room temperature followed by three 30 min washes at 60°C). RNA was electrophoresed through a 1.2% agarose gel after denaturation with glyoxal and dimethyl sulfoxide (Sambrook et al., 1989) and transferred to Zetabind membranes. Filters were prehybridized and hybridized at 65°C in 5% SDS, 10% dextran sulfate, and 1 M NaCl. Following hybridization, filters were washed in 2 × SSC, 0.5% SDS as described above (high temperature wash 65°C). Hybridization probes were labeled using random oligonucleotide primers (Sambrook et al., 1989); for Northern blots unincorporated nucleotides were removed by passing the probe over a Sephadex G-50 column.

##### cDNA and Genomic *coxII* Gene

A cowpea cDNA library, constructed in the  $\lambda$  Zap cloning vector (Stratagene), was provided by Dr. Narendra Singh and Dr. Ray Bressan. The library was screened by plaque hybridization on nitrocellulose filters with a 704 bp NcoI-HindIII fragment internal to the soybean mitochondrial *coxII* gene (Grabau, 1987). This fragment was subcloned from a 3.5 kb HindIII clone provided by Dr. Elizabeth Grabau. Hybridization conditions were as described for Southern blots. The genomic *coxII* gene was amplified using Taq polymerase-mediated PCR. Two synthetic 18-mer oligonucleotides were designed for use as amplification primers based on the 5' and 3' sequence of the cDNA clone (positions 100-117 and 1214-1231; Figure 3). Within the 3' primer the sequence GAATTG was changed to GAATTC, creating an EcoRI restriction site. All sequencing was by the dideoxynucleotide chain termination method (Sanger et al., 1977) using a combination of synthetic oligonucleotide primers and a series of nested deletions created by exonuclease III digestion (Erase-A-Base, Promega). All DNAs were sequenced completely on both strands.

##### Computer Analyses

Nucleotide sequences were analyzed using the distance matrix program developed by Olsen (1988a, 1988b) and by maximum parsimony, using the exhaustive search option (branch and bound) of PAUP 3.0 (David Swofford, Illinois Natural History Survey, Champaign, IL). Maxi-

mum parsimony analysis was also used to analyze deduced amino acid sequences. Edited positions were excluded from all analyses. Bootstrap analysis (Felsenstein, 1985) was performed to assess the relative support of clades identified by both parsimony and distance; 100 replications were used for distance analysis and 1000 replications for maximum parsimony analysis.

#### Acknowledgments

We gratefully acknowledge R. Bressan and N. Singh for providing us with the cowpea cDNA library, E. Grabau for the soybean mitochondrial *coxII* gene probe, D. Oliver for the H-protein gene probe, and M. Hanson, K. Newton, and B. Sears for providing several of the mtDNAs used in Figure 1. We thank S. Barns and C. Turbeville for computer assistance, S. Baldauf and K. Wolfe for critically reading the manuscript, P. Pukkila for helpful discussion, and two anonymous reviewers for helpful comments. This work was supported by NIH grant GM-35087 to J. D. P. and a Midwest Biotechnology Consortium Fellowship to J. M. N.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC Section 1734 solely to indicate this fact.

Received April 29, 1991; revised June 7, 1991.

#### References

Baldauf, S. L., and Palmer, J. D. (1990). Evolutionary transfer of the chloroplast *tufA* gene to the nucleus. *Nature* 344, 262–265.

Bartnik, E., and Borsuk, P. (1986). A glycine tRNA gene from lupine mitochondria. *Nucl. Acids Res.* 14, 2407.

Bennett, M. D., and Smith, J. B. (1976). Nuclear DNA amounts in angiosperms. *Phil. Trans. Roy. Soc. (Lond.) B* 274, 227–274.

Binder, S., Schuster, W., Grienemberger, J.-M., Weil, J.-H., and Brennicke, A. (1990). Genes for tRNA<sup>Gly</sup>, tRNA<sup>His</sup>, tRNA<sup>Phe</sup>, tRNA<sup>Ser</sup> and tRNA<sup>Tyr</sup> are encoded in *Oenothera* mitochondrial DNA. *Curr. Genet.* 17, 353–358.

Bland, M. M., Levings, C. S., III, and Matzinger, D. F. (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.

Boer, P. H., Adra, C. N., Lau, Y.-F., and McBurney, M. W. (1987). The testis-specific phosphoglycerate kinase gene *pgk-2* is a recruited retroposon. *Mol. Cell. Biol.* 7, 3107–3112.

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

Bonen, L., Boer, P. H., and Gray, M. W. (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.

Breathnach, R., and Chambon, P. (1981). Organization and expression of eukaryotic split genes coding for proteins. *Annu. Rev. Biochem.* 50, 349–383.

Capaldi, R. A., Malatesta, F., and Darley-Usmar, V. M. (1983). Structure of cytochrome c oxidase. *Biochim. Biophys. Acta* 726, 135–148.

Covello, P. S., and Gray, M. W. (1989). RNA editing in plant mitochondria. *Nature* 341, 662–666.

Covello, P. S., and Gray, M. W. (1990). Differences in editing at homologous sites in messenger RNAs from angiosperm mitochondria. *Nucl. Acids Res.* 18, 5189–5196.

Crepet, W. L., and Taylor, D. W. (1985). The diversification of the *Leguminosae*: first fossil evidence of the *Mimosoideae* and *Papilionoideae*. *Science* 228, 1087–1089.

Cronquist, A. (1988). *The Evolution and Classification of Flowering Plants, Second Edition* (New York: The New York Botanical Garden).

Dahl, H. H. M., Brown, R. M., Hutchison, W. M., Maragos, C., and Brown, G. K. (1990). A testis-specific form of the human pyruvate dehydrogenase E1a subunit is coded for by an intronless gene on chromosome 4. *Genomics* 8, 225–232.

Donoghue, M. J., and Doyle, J. A. (1989). Phylogenetic studies of seed plants and angiosperms based on morphological characters. In *The Hierarchy of Life: Molecules and Morphology on Phylogenetic Analysis* (Amsterdam, New York, Oxford: Elsevier Science Publishers), pp. 181–193.

Doyle, J. J., and Doyle, J. L. (1987). A rapid isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19, 11–15.

Drouin, G., and Dover, G. A. (1987). A plant processed pseudogene. *Nature* 328, 557–558.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using bootstrap. *Evolution* 39, 783–791.

Fox, T. D., and Leaver, C. J. (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.

Froman, B. E., Tait, R. C., and Gottlieb, L. D. (1989). Isolation and characterization of the phosphoglucose isomerase gene from *Escherichia coli*. *Mol. Gen. Genet.* 217, 126–131.

Gilbert, W. (1978). Why genes in pieces? *Nature* 271, 501.

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

Gualberto, J., Lamattina, L., Bonnard, G., Weil, J.-H., and Grienemberger, J.-M. (1989). RNA editing in wheat mitochondria results in the conservation of protein sequences. *Nature* 341, 660–662.

Gualberto, J. M., Weil, J.-H., and Grienemberger, J.-M. (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.

Hamby, R. K., Sims, L. E., Issel, L. E., and Zimmer, E. A. (1988). Direct ribosomal RNA sequencing—optimization of extraction and sequencing methods for work with higher plants. *Plant Mol. Biol. Rep.* 6, 175–192.

Hiesel, R., and Brennicke, A. (1983). Cytochrome oxidase subunit II gene in mitochondria of *Oenothera* has no intron. *EMBO J.* 2, 2173–2178.

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

Hiesel, R., Wissinger, B., and Brennicke, A. (1990). Cytochrome oxidase subunit II mRNAs in *Oenothera* mitochondria are edited at 24 sites. *Curr. Genet.* 18, 371–375.

Hiratsuka, J., Shimada, H., Whittier, R. F., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.-R., Meng, B.-Y., Li, Y.-Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., and 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.

Izuchi, S., Teracji, T., Sakamoto, M., Mikami, T., and Sugita, M. (1990). Structure and expression of tomato mitochondrial genes coding for tRNA<sup>Cys</sup> (GCA), tRNA<sup>Asn</sup> (GUU) and tRNA<sup>Tyr</sup> (GUA): a native tRNA<sup>Cys</sup> gene is present in dicot plants but absent in monocot plants. *Curr. Genet.* 18, 239–243.

Jacobson, A. (1987). Purification and fractionation of poly(A)<sup>+</sup> RNA. *Meth. Enzymol.* 152, 254–261.

Kao, T., Moon, E., and Wu, R. (1984). Cytochrome oxidase subunit II gene of rice has an insertion sequence within the intron. *Nucl. Acids Res.* 12, 7305–7315.

Kim, Y., and Oliver, D. J. (1990). Molecular cloning, transcriptional characterization, and sequencing of cDNA encoding the H-protein of the mitochondrial glycine decarboxylase complex in peas. *J. Biol. Chem.* 265, 848–853.

Kolodner, R., and Tewari, K. K. (1972). Physicochemical characterization of mitochondrial DNA from pea leaves. *Proc. Natl. Acad. Sci. USA* 69, 1830–1834.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Maréchal-Drouard, L., Guillemaut, P., Cosset, A., Arbogast, M.,



- Weber, F., Weil, J.-H., and Dietrich, A. (1990). Transfer RNAs of potato (*Solanum tuberosum*) mitochondria have different genetic origins. *Nucl. Acids Res.* 18, 3689–3696.
- McCarrey, J. R., and Thomas, K. (1987). Human testis-specific PGK gene lacks introns and possesses characteristics of a processed gene. *Nature* 326, 501–505.
- Moon, E., Kao, T., and Wu, R. (1985). Pea cytochrome oxidase subunit II gene has no intron and generates two mRNA transcripts with different 5'-termini. *Nucl. Acids Res.* 13, 3195–3212.
- Newton, K. J. (1988). Plant mitochondrial genomes: organization, expression and variation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 39, 503–532.
- Nugent, J. M., and Palmer, J. D. (1988). Location, identity, amount and serial entry of chloroplast DNA sequences in crucifer mitochondrial DNAs. *Curr. Genet.* 14, 501–509.
- Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S.-I., Inokuchi, H., and Ozeki, H. (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* 322, 572–574.
- Olsen, G. J. (1988a). The earliest phylogenetic branchings: comparing rRNA-based evolutionary trees inferred with various techniques. *Cold Spring Harbor Symp. Quant. Biol.* 52, 825–838.
- Olsen, G. J. (1988b). Phylogenetic analysis using ribosomal RNA. *Meth. Enzymol.* 164, 793–812.
- Polhill, R. M. (1981). *Papilionoideae*. In *Advances in Legume Systematics*, R. M. Polhill and P. H. Raven, eds. (Kew, England: Royal Botanic Gardens), pp. 191–208.
- Pruitt, K. D., and Hanson, M. R. (1989). Cytochrome oxidase subunit II sequences in *Petunia* mitochondria: two intron-containing genes and an intron-less pseudogene associated with cytoplasmic male sterility. *Curr. Genet.* 16, 281–291.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Second Edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- Schuster, W., and 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., Wissinger, B., Unseld, M., and Brennicke, A. (1990a). Transcripts of the NADH-dehydrogenase subunit 3 gene are differentially edited in *Oenothera* mitochondria. *EMBO J.* 9, 263–269.
- Schuster, W., Hiesel, R., Wissinger, B., and Brennicke, A. (1990b). RNA editing in the cytochrome b locus of the higher plant *Oenothera lamarckiana* includes a C-to-U transition. *Mol. Cell. Biol.* 10, 2428–2431.
- Selker, E. U. (1990). Premeiotic instability of repeated sequences in *Neurospora crassa*. *Annu. Rev. Genet.* 24, 579–613.
- Shih, M.-C., Lazar, G., and Goodman, H. M. (1986). Evidence in favor of the symbiotic origin of chloroplasts: primary structure and evolution of tobacco glyceraldehyde-3-phosphate dehydrogenases. *Cell* 47, 73–80.
- 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, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Sugiura, M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J.* 5, 2043–2049.
- Soares, M. B., Schon, E., Henderson, A., Karathanasis, S. K., Cate, R., Zeitlin, S., Chirgwin, J., and Efstratiadis, A. (1985). RNA-mediated gene duplication: the rat preproinsulin I gene is a functional retroposon. *Mol. Cell. Biol.* 5, 2090–2103.
- Stern, D. B., and Lonsdale, D. M. (1982). Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. *Nature* 299, 698–702.
- Stern, D. B., and Newton, K. J. (1988). Isolation of plant mitochondrial RNA. In *Methods for Plant Molecular Biology*, A. Weissbach and H. Weissbach, eds. (New York: Academic Press), pp. 309–317.
- Takehima, H., Joh, T., Tsuzuki, T., Shimada, K., and Matsukado, Y. (1988). Structural organization of the mouse mitochondrial malate dehydrogenase gene. *J. Mol. Biol.* 200, 1–11.
- Tsuzuki, T., Obaru, K., Setoyama, C., and Shimada, K. (1987). Structural organization of the mouse mitochondrial aspartate aminotransferase gene. *J. Mol. Biol.* 198, 21–31.
- von Heijne, G., Steppuhn, J., and Herrmann, G. (1989). Domain structure of mitochondrial and chloroplast targeting peptides. *Eur. J. Biochem.* 180, 535–545.
- Wahleithner, J. A., and Wolstenholme, D. R. (1988). Ribosomal protein S14 genes in broad bean mitochondrial DNA. *Nucl. Acids Res.* 16, 6897–6913.
- Wahleithner, J. A., Macfarlane, J. L., and Wolstenholme, D. R. (1990). A sequence encoding a maturase-related protein in a group II intron of a plant mitochondrial *nad7* gene. *Proc. Natl. Acad. Sci. USA* 87, 548–552.
- Watson, J. C., and Thompson, W. F. (1986). Purification and restriction endonuclease analysis of plant nuclear DNA. *Meth. Enzymol.* 118, 57–74.
- Wolfe, K. H., Sharp, P. M., and Li, W.-H. (1989a). Rates of synonymous substitution in plant nuclear genes. *J. Mol. Evol.* 29, 208–211.
- Wolfe, K. H., Gouy, M., Yang, Y.-W., Sharp, P. M., and Li, W.-H. (1989b). Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. *Proc. Natl. Acad. Sci. USA* 86, 6201–6205.

#### Note Added in Proof

The work referred to as J. S. Gantt, S. L. Baldauf, P. Calie, N. Weeden, and J. D. Palmer, unpublished data, can now be updated: Transfer to the nucleus of *rpl22* greatly preceded its loss from the chloroplast and involved gain of an intron. *EMBO J.*, in press.