J Mol Evol (1994) 39:144–150

Journal of Molecular Evolution

Promiscuous Mitochondrial Group II Intron Sequences in Plant Nuclear Genomes

V. Knoop, A. Brennicke

Institut für Genbiologische Forschung GmbH, Ihnestr. 63, 14195 Berlin, Germany

Received: 7 June 1993 / Revised 2 August 1993

Abstract. Gene translocations from the organelles to the nucleus are postulated by the endosymbiont hypothesis. We here report evidence for sequence insertions in the nuclear genomes of plants that are derived from noncoding regions of the mitochondrial genome. Fragments of mitochondrial group II introns are identified in the nuclear genomes of tobacco and a bean species. The duplicated intron sequences of 75–140 bp are derived from *cis*- and *trans*-splicing introns of genes encoding subunits 1 and 5 of the NADH dehydrogenase. The mitochondrial sequences are inserted in the vicinities of a lectin gene, different glucanase genes and a gene encoding a subunit of photosystem II. Sequence similarities between the nuclear and mitochondrial copies are in the range of 80 to 97%, suggesting recent transfer events that occurred in the basic glucanase genes before and in the lectin gene after the gene duplications in the evolution of the nuclear gene families. Overlapping regions of the same introns are in two instances also involved in intramitochondrial sequence duplications.

Key words: Tobacco — Glucanase genes — Oxygen-evolving enhancer protein gene — Sequence transfer — Endosymbiont theory

Introduction

According to the endosymbiont hypothesis, plastids and mitochondria are derived from predecessors of eubacteria, which have been taken up by the "urkaryote" during endosymbiosis (Gray 1992). The genome of the bacterial endosymbionts has been continually reduced, hand in hand with a net transfer of genetic material into the nucleus of the host cell. This view is generally accepted, and it has lately been shown that functional transfers of genes from the mitochondrion to the nucleus have occurred recently in evolution in the plant kingdom (Nugent and Palmer 1991; Covello and Gray 1992; Grohmann et al. 1992). RNA molecules have been suggested for these and other sequence transfers as physical carriers of the transferred genetic information (Schuster and Brennicke 1987; Nugent and Palmer 1991; Covello and Gray 1992). The presence of multiple copies of the complete plastid genome in the nucleus (Ayliffe and Timmis 1992), however, suggests that other mechanisms involving DNA as carrier-molecule during interorganellar sequence transfers must operate as well.

We have recently reported the serendipitous finding of a mitochondrial intron fragment inserted upstream of the coding region in a lectin gene of a bean species, *Dolichos biflorus* (Knoop and Brennicke 1991). In its nuclear location, the mitochondrial intron sequence may have gained regulatory function, being the only major difference in comparison to a second lectin gene in this species which is differently regulated (Harada et al. 1990).

To identify additional examples of mitochondrial sequence transfers, especially of group II intron fragments, we have carried out a systematic search for traces of such events. Three new examples of interorganellar sequence transfers in tobacco and two intraorganellar sequence duplications are identified. The genuine mito-

Correspondence to: V. Knoop

chondrial intron sequences in *Dolichos* and tobacco have been cloned and sequenced and are compared with the nuclear insertions to estimate the dates of the transfer events.

Materials and Methods

Computer Analysis. Computer analysis was performed using version 7.1 of the UWGCG program package (Devereux et al. 1984). Initial database searching was carried out with the FASTA algorithm (Pearson and Lipmann 1988).

PCR Amplification and Cloning. Dolichos biflorus and tobacco DNA preparations were kindly provided by Drs. M. Etzler, Davis, and C. Gatz, Berlin, respectively. Oligonucleotide primers for PCR amplification were: 5A+ 5′-TGCTCCATGGATCTCATCGG-3′, 5B− 5′− GGGCTATGCGGATCCTCAG-3′, 1A+ 5′-GGGCTACTTTATTTGTTTGC-3′, 1a− 5′− GATGGGGGGTACCTTTCTTTCCC-3′, 1c+ 5′-CAACCAAGCTTAACCGGTCACG-3′, 1c− 5′-GAAGGAGGGAAGGCGCGCTAC-3′, 1de+ 5′−CAGGAGATAGGGAGGATCCG-3′, and 1de− 5′-CGCAAAGGAGGCTCGAGGG-3′. Oligonucleotide nomenclature indicates number of *nad*-subunit, location in (uppercase letters) or proximal to (lowercase letters) the respective exons, and orientation in (+) or against (−) the direction of transcription in the mitochondrion. Oligonucleotides were obtained commercially (TIB Molbiol, Berlin). PCR amplification was performed with a kit (Boehringer Mannheim, Germany) as recommended by the manufacturer. PCR products were cloned in Bluescript SKII+ vectors (Stratagene, USA) and sequenced by the dideoxy method (Sanger et al. 1977) using the T7 polymerase kit (Pharmacia, Sweden). Multiple PCR clones were sequenced to minimize the risk of PCR sequence artifacts. Sequence data were submitted to the databases and accession numbers are given in the respective figure legends.

Results

The Mitochondrial Intron Fragment Insertion in the 5′-Region of a Nuclear Lectin Gene in Dolichos *is of Recent Origin*

To gain insight into the timing and mode of the 116-bp sequence translocation (Knoop and Brennicke 1991) from the mitochondrion into one of two nuclear lectin genes in *Dolichos biflorus* (Harada et al. 1990) the genuine mitochondrial *nad*5 intron sequence between exons a and b in this species was amplified by PCR, cloned, and sequenced. Previously only heterologous comparisons with mitochondrial sequences from other species were possible.

The *Dolichos* mitochondrial intron is 859 nucleotides long and highly conserved in comparison to those of other species with about 90% identical nucleotides and no major deletions or insertions. Alignment of the mitochondrial intron sequence with the nuclear seed lectin gene in *Dolichos* (Fig. 1A) shows only nine point mutations and a single nucleotide deletion in the nuclear copy, resulting in a sequence similarity of 91%. The mitochondrial intron sequences in *Oenothera* (Wissinger et al. 1988), *Arabidopsis* (Knoop et al. 1991), sugar beet (Ecke et al. 1990), and wheat (Pereira de Souza et al. 1991) display additional sequence divergencies (not shown) suggesting a recent sequence transfer after evolutionary branching of these species. The assumption of an evolutionary young transfer, however, still receives its strongest support from the observation that the sequence transfer has necessarily occurred after the duplication of the ancestral lectin gene in the legume.

Analysis of the mitochondrial background of the transferred sequence shows that this intron is also involved in an intramitochondrial sequence duplication event. A sequence of 62 nucleotides overlapping the nuclear-mitochondrial homology (Fig. 1) is also present in the group II intron of the mitochondrial *rps*3 gene of maize (Hunt and Newton 1991). Intron modelling (Michel et al. 1989) suggests that in the two group II introns this sequence encodes in both cases a highly conserved domain II (Fig. 1B). In the *nad*5 intron this sequence block is identical in all five species for which data are available (with the exception of a single G/T exchange in *Oenothera*), indicating a high selective pressure for this secondary structure. Since the duplicated sequence element is also present in the *Oenothera rps*3 intron (H. Bock and W. Schuster, personal communication) the intramitochondrial duplication event has apparently occurred before divergence of mono- and dicotyledonous species and is thus much older than the sequence transfer into the nucleus. The direction of the intramitochondrial sequence transfer cannot be determined while the direction of the interorganellar transposition is clearly from mitochondrion into the nucleus. No additional extensive sequence similarities between the first *nad*5 intron and the *rps*3 intron can be identified and therefore make an evolutionary derivation from a common ancestor intron unlikely. This "shared motif" is thus distinct from the duplicated sequence element observed in two introns of the *nad*1 and *nad*2 genes in *Oenothera* (Wissinger et al. 1991a), which share additional overall similarity besides the 48-bp identical element. Although plant mitochondrial DNA rapidly evolves through recombination via repeated sequences no such recombinational activity is seen for these two common intron motifs. An unrelated recombination event at a different position within the *rps*3 intron gives rise to an NCS phenotype in maize (Hunt and Newton 1991).

Identification of Additional Mitochondrial Intron Fragments in the Nucleus

To test whether other mitochondrial (intron) sequences can be identified in the nuclear genomes of plant species, the available plant mitochondrial group II intron sequences of different species were used to screen the databases with the FASTA algorithm. The high conservation of mitochondrial sequences in higher plants

A

```
  1          mitochondrial nad5 a/b intron > tggcgaccgtgaagtcaccggatgaattgccgagtagatagatcagatccggacgcggc  59
 60 tgttgctccgcggcgatacggactcgacccgctcctacccaccccggggcaccatagcatgtcgggaataagggggacatactggacgtaaccactccc 159
160 ttgggggcggctgtgccgccctgcctttcgatcgatacacagttgagtaggccgatcacgaacgctacaggtgtgggagcgatcctggtcagggaaggct 259
260 aagacggcgcctcgcatatgggtagcaagagggcgcttatgccccgacggtggggccttatggggaagggcccagcccaatagggacagcacaccccca 359
                                               nuclear seed lectin gene > AGCCCAATAGGGACGGCACACCCCCCA 787
360 cttcaagcgcacctctgtatcgactgaatcactctaagagtctagtcggtggaaccggtgaaccacgcgagctggttagatgcgtggggcagagggctcg 459
788 CTTCAAGCGCACCTCTGTGTCAACAGAATCACTCTAAGAGTATAGTC.GTGGAATCGGTGAACCACGCCAGCTGCTTAGATGCGTGGCGCAGAGG      881
460 tagtaccccctttcttgatccagccttttcttcgcttcggtagtgaattacctactaaaaaagaggcaggcctgcacgccctatttgagactactcag 559
560 gcagcttgagtttactaaggggactctttcattaggggaagggaagaaggggcctaagcacggcagatgccgtacacttgagtggcaaaggaaagggaga 659
660 tcgtacctgtttttccaggcctgttcggacatatggttaccgcggaagatcaagttggtgagccgtgtgatgggaaaccttcccgcacggttcggagagc 759
760 actgaattagaatgagaggttcaccaccacatcattgcatgcaaggggagctcgctcgattcgcaaattggtccgactcgtaattcacttctgactccgt 859
```

B

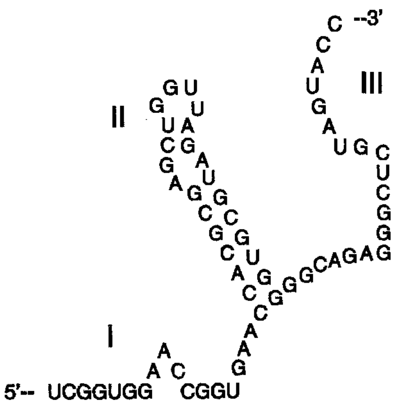

Fig. 1. Inter- and intraorganellar sequence duplications of the mitochondrial *nad*5 a/b intron. **A** The *Dolichos biflorus* intron sequence (*lowercase letters*, accession No. X72287) is aligned with the nuclear seed lectin gene (*uppercase letters*, Harada et al. 1990) in the region of homology (numbering is according to the respective database entries). A sequence of 62 nucleotides (*underlined*) is also present in the *rps*3 intron of maize. **B** The sequence of 62 nucleotides identical between the *nad*5 intron and the *rps*3 intron of maize constitutes domain II and parts of domains I and III in both introns (modeling according to Michel et al. 1989).

with more than 90% sequence similarity even in intron sequences should allow the detection of homologous sequences also in different species.

In this analysis three additional group II intron sequences were identified in the nuclear genome of tobacco flanking known reading frames. These intron fragments are in size and degree of interorganellar sequence conservation similar to the intron sequence in the lectin gene of *Dolichos biflorus* (Fig. 2). The three tobacco transfer events involve intron sequences derived from the mitochondrial *nad*1 gene. The five exons of this gene are scattered over the mitochondrial genome in higher plants, and mature mRNA molecules are obtained by *cis*- and *trans*-splicing (Chapdelaine and Bonen 1991; Wissinger et al. 1991b).

A Mitochondrial Intron Fragment is Inserted in Inverted Orientation Upstream of oee2-A

A sequence derived from the (*trans*-splicing) intron part downstream of exon c of the *nad*1 gene is present upstream of the *oee*2-A gene (Hua et al. 1991), encoding a 23-kDa polypeptide of the oxygen-evolving complex of photosystem II in tobacco. The intron sequence is inserted 772 bp upstream of the *oee*2-A start codon in inverted orientation. The inserted fragment is at least 97 bp long and may extend further into the as-yet-unknown 5′-region of the *oee*2 gene. Only three nucleotide transitions are observed in the transferred sequence, which increase the AT content of the nuclear copy (Fig. 3A). The conservation of this promiscuous sequence thus either suggests an evolutionarily even more recent transfer event than for the *nad*5 intron fragment in *Dolichos* and/or a selective pressure on the transferred sequence element in the nucleus. The sequence transfer has clearly occurred after the speciation of tobacco, *Oenothera* (Wissinger et al. 1991b), and watermelon (Stern et al. 1986), since sequence divergencies are observed in the respective intron sequences of the two latter species (not shown).

The *oee*2-A gene belongs to a family of four *oee*2

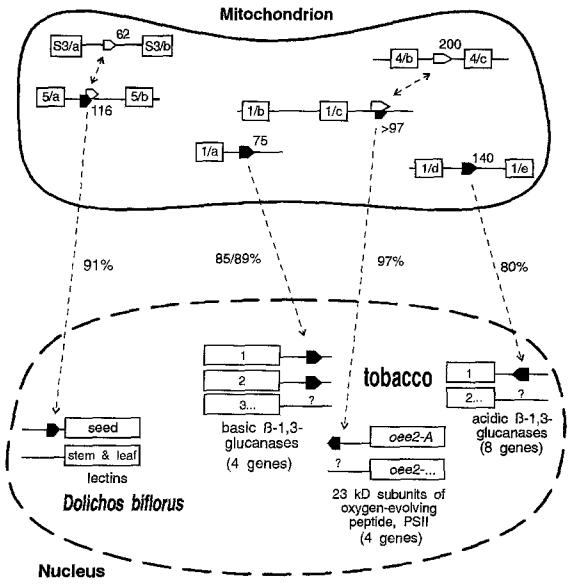

Fig. 2. Group II intron sequences involved in intra- and intermitochondrial sequence duplications. Exons 1/a through 1/e of the *nad*1 gene in higher plant mitochondria are separated by group II introns in *cis*- and *trans*-arrangements (Chapdelaine and Bonen 1991; Wissinger et al. 1991). Exons a and b (5/a, 5/b) of the *nad*5 gene are separated by a conventional group II intron while the other exons have to be added in *trans* on the RNA level (Knoop et al. 1991). *Cis*-arranged group II introns interrupt the coding regions of the genes for ribosomal protein S3 (S3/a, S3/b; Hunt and Newton 1991) and of the *nad*4 gene (4/b, 4/c; Lamattina and Grienenberger 1991). *Solid pentangles* show sequences involved in interorganellar transfers, while *open pentangles* represent intraorganellar duplication events. Sizes of the promiscuous mitochondrial intron fragments and sequence similarities to the nuclear copies which are in the range of 80% (*nad*1 d/e vs acidic glucanase gene) to 97% (*nad*1 c/d vs *oee*2-A) are indicated. *Question marks* indicate cases where the presence of the mitochondrial intron sequences in the respective other members of the gene families is not yet clear.

A

```
283 gctagctaggctagcgcgcaatggctttctctgataggggcttgcttctt 234

233 taagctccgcccaacctatacaaggtgctgctcgctttctctcagccact 184
                                                    ¦¦¦
                                nuclear oee2-A gene › ACT   3

183 agtggcttatgtagtggtcggcattcctacgtgctcctccttgccccctt 134
    ¦¦¦¦¦¦¦¦¦ ¦¦¦¦¦¦¦ ¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦
  4 AGTGGCTTATATAGTGGTCAGCATTCCTACGTGCTCCTCCTTGCCCCCCT  53

133 acttaagaatccaaaggtcacctttggatgttgtcgtgacgctttggtta  84
    ¦¦¦¦¦¦¦¦¦¦¦ ¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦  ¦¦
 54 ACTTAAGAATCTAAAGGTCACCTTTGGATGTTGTCGTGACGCTTATTTTC 103

 83 cgaaggttaccgcggtctaagtttatggatggattcagtcagctaaactc  34

 33 catcaaactcctcaatcaatatcaacaagatgt 1
```

B

```
    C
U — A
G — C
G — C
A — U
A — U
A — U
C — G
C — G
U — A
A — U
```

Fig. 3. An insertion derived from the mitochondrial intron sequence downstream of *nad*1 exon c in the nuclear *oee*2-A gene of tobacco 772 bp upstream of the start codon. **A** Only three nucleotide transitions are observed in the promiscuous region of 97 nucleotides comparing the tobacco mitochondrial intron sequence (*lowercase letters*, accession No. X72288) with the tobacco *oee*2-A gene (*uppercase letters*, Hua et al. 1991). Numbering is given according to the database entries. **B** The sequence *underlined* in **A** within the transferred region has the potential to fold into a hairpin structure.

genes in tobacco (Hua et al. 1992). The limited sequences available for the other members of the gene family do not allow any conclusion about the presence or absence of the mitochondrial sequence in these genes. Absence or variation of the promiscuous sequence element in the other *oee*2 genes and a correlation with possibly different expression of these genes would raise speculations on the functional significance of the inserted intron fragment analogous to the *Dolichos* situation. In this context an inverted repeat upstream of *oee*2-A (Hua et al. 1991) in the inserted mitochondrial fragment should be noted, which can on the RNA level be folded into a hairpin structure (Fig. 3).

An intramitochondrial duplication might be responsible for a similarity between the 200 bp downstream of *nad*1 exon c including the transferred sequence (Fig. 2) and the second intron of the mitochondrial *nad*4 gene in wheat (Lamattina and Grienenberger 1991). The comparatively low sequence similarity of 70% (alignment not shown) complicates classification of this homology as a sequence duplication, although the absence of additional extensive homologies between the two introns makes their derivation from a common ancestor unlikely. The intraorganellar duplication event in the *nad*1/*nad*4 introns is, like the promiscuous *nad*5/*rps*3 intron sequences, far older than the respective interorganellar sequence duplications.

Mitochondrial Intron Fragments Downstream of Glucanase Genes

Downstream of the coding regions of genes for two types of glucanases, fragments from different introns of the mitochondrial *nad*1 gene are detected. Glucanases (glucan endo-1,3-β-glucosidases) belong to the pathogenesis-related proteins (PR), whose genes are locally and systemically activated upon pathogen attack (Linthorst et al. 1990). Genes for the acidic and basic type of glucanases have been sequenced in tobacco (Linthorst et al. 1990; Sperisen et al. 1991).

A Mitochondrial Intron Fragment Downstream of an Acidic Glucanase Gene

A sequence derived from the maturase-related reading frame in the intron between exons d and e of the mito-

chondrial *nad*1 gene is located 280 bp downstream of the stop codon of an acidic glucanase gene in tobacco (Linthorst et al. 1990). The 140-bp intron fragment is inserted in inverted orientation and shares 80% identical nucleotides with its mitochondrial counterpart (Fig. 4). The 28 sequence differences (23 single-base substitutions and five insertions/deletions) within the nuclear copy increase the AT content of the transferred sequence by 10%.

Sequence comparison of the nuclear and mitochondrial copies in tobacco to the mitochondrial sequences in the three other species suggests that only one tobacco-specific point mutation has occurred in the common ancestral mitochondrial sequence before the sequence transfer, a C-to-A exchange in position 140 (Fig. 4). After the interorganellar sequence transfer a triplet inversion of a GGA (Gly) to TCC (Ser) codon has occurred in the mitochondrial tobacco maturase sequence. Such an uncommon sequence variation in a coding region that would require three independent point mutations — if no other mechanisms are invoked — has also been observed in exon a of the *nad*5 gene in *Oenothera*. (See alignment in Knoop et al. 1991.)

The reading frame of the maturase-related polypeptide in the investigated tobacco mitochondrial intron sequence is conserved in comparison to the respective sequences of *Vicia faba* (Wahleithner et al. 1990), *Oenothera* (Wissinger et al. 1991b), and wheat (Chapdelaine and Bonen 1991). The predominance of nucleotide exchanges in the mitochondrial reading frames in third codon positions supports the assumption of a functional importance of the maturase-related reading frame in yet another plant species.

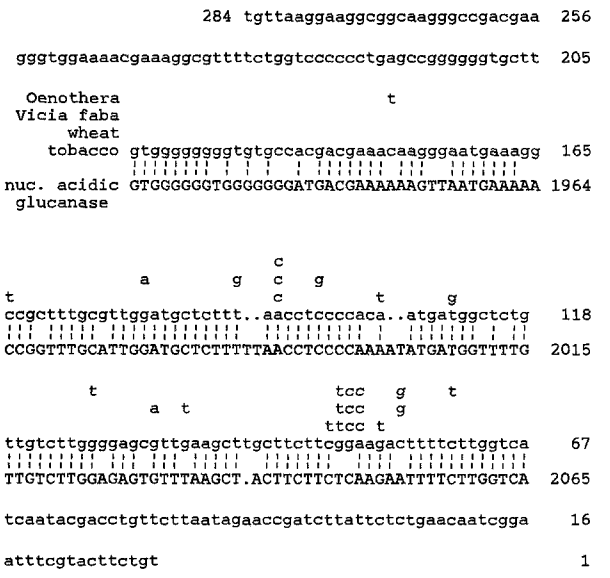

Fig. 4. An intron fragment from the mitochondrial maturase-related open reading (mat-r) frame between exons d and e of the *nad*1 gene is found inserted 280 bp downstream of the termination codon of an acidic glucanase gene in tobacco. The mitochondrial tobacco sequence in *nad*1 intron d/e (*lowercase letters*, accession No. X72289) is aligned with the acidic glucanase gene (*uppercase letters*, Linthorst et al. 1990). *Numbering* is according to the database entries. Nucleotide differences in the mitochondrial sequences of *Oenothera*, *Vicia faba*, and wheat are indicated only within the promiscuous sequence block.

Differential Sequence Drift of a Mitochondrial Intron Sequence Inserted in Two Basic Glucanase Genes

In tobacco basic glucanases are encoded by a multigene family with at least four members (Sperisen et al. 1991). A 75-bp fragment derived from the (*trans*-splicing) intron sequence downstream of *nad*1 exon a is present approximately 900 bp downstream of the coding regions of (at least) two members of this gene family. The alignment of the tobacco mitochondrial intron sequence with the two genes GL-A and GL-B (Sperisen et al. 1991) reveals differential sequence drift in the two copies in the nucleus (Fig. 5), indicating insertion of the intron sequence before duplication of the glucanase genes. A third basic glucanase gene (Linthorst et al. 1990) with 99.7% sequence identity to GL-A might be a third representative of the gene family. This sequence displays three additional nucleotide insertions in the promiscuous intron sequence (alignment not shown). No sequence data of other members of the basic glucanase gene family in tobacco are as yet available.

While GL-A and GL-B display eight and 11 point deviations from the mitochondrial copy, only one of these, a G-to-A transition in position 384 (Fig. 5), is in common. The transfer of the promiscuous intron sequence into the ancestral glucanase gene has thus presumably occurred shortly before the nuclear gene duplication event.

An additional point mutation within the promiscuous sequence in the corresponding *nad*1 intron sequence of *Oenothera* (Wissinger et al. 1991b) suggests the sequence transfer to have occurred after speciation of these dicots.

While the intron downstream of *nad*1 exon a is well conserved, the tobacco sequence carries a 43-bp insertion with a — for group II intron sequences rather unusual — AT content of 88% (Fig. 5). In spite of the high degree of sequence conservation in plant mitochondria the promiscuous sequence cannot be identified in the corresponding wheat sequence downstream of *nad*1 exon a (Chapdelaine and Bonen 1991), possibly indicating a monocot-specific (additional) recombination event in this *trans*-splicing intron (Fig. 5).

Discussion

In a plant cell with three different genomes nucleic acids can theoretically be transferred in six different directions. While foreign sequences derived from the respective other organelles have been identified in both mitochondria and nuclei of some plant species, no examples for nuclear or mitochondrial sequences are re-

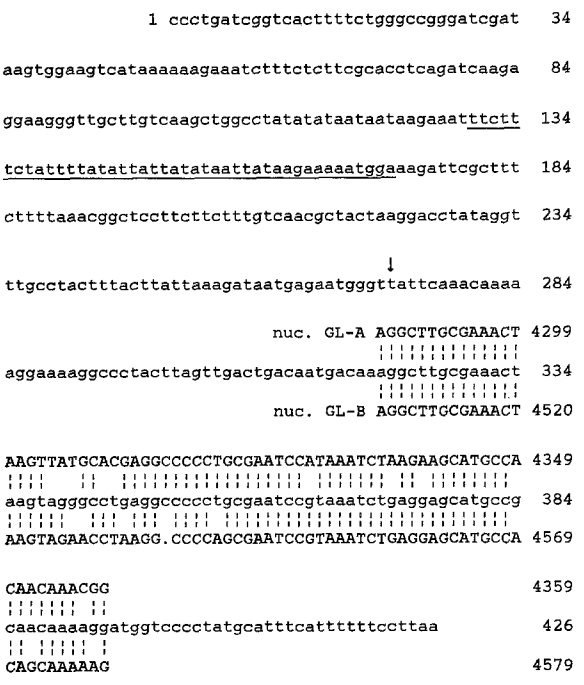

Fig. 5. An insertion of mitochondrial origin approximately 900 bp downstream of the termination codons in two basic glucanase genes. The alignment of the tobacco mitochondrial intron sequence downstream of *nad*1 exon a (accession No. X72290, *lowercase letters*) suggests that the intron fragment was inserted before the glucanase gene duplication that gave rise to GL-A and GL-B (*uppercase letters*, Sperisen et al. 1991). *Numbering* is as in the database entries. The tobacco intron sequence displays an AT-rich insertion with respect to the *Oenothera* sequence (*underlined*). The *vertical arrow* represents the end of homology downstream of *nad*1 exon a between wheat and the dicots.

ported in the completely sequenced chloroplast genomes of several plant species. This status quo could reflect different directional permeabilities of the organellar membrane systems for nucleic acids. Alternatively, the capacity of chloroplast genomes to integrate and maintain "foreign" sequences may limit the genetic propagation of such a priori unnecessary genetic information. The descriptive analysis of foreign DNA sequences in mitochondria and nuclei has resulted in indirect, yet compelling, evidence that both DNA and RNA could be the physical carrier of the sequence information during the transfer process.

The endosymbiont hypothesis postulates a massive transfer of functional genes into the nucleus and their adaptation to the nuclear requirements (Brennicke et al. 1993). The identification of organellar intron sequences in the nuclear genome, however, clearly demonstrates that no selection mechanism for functionally expressable sequences is operative. Nevertheless, their presence in close proximity to nuclear genes might influence the expression of the respective genes (for which the intron sequence upstream of the differentially regulated lectin gene in *Dolichos* is presently the strongest candidate). Whether the other intron sequences inserted in the proximity of nuclear reading frames reported in this paper contribute to the regulation of their expression is an open question. It might be of interest in this respect that all four intron fragments in the nuclear genome are inserted in members of gene families. Except for the lectin gene example, however, it is as yet unclear whether only some or all members of these families carry the mitochondrial invaders.

An evolutionary timing of the sequence transfer events must necessarily rely on nuclear sequence data since mitochondrial sequences in the plant kingdom are too conserved for comparisons over short evolutionary distances (Palmer and Herbon 1988). The identification of promiscuous mitochondrial sequences in members of nuclear gene families and a careful comparison in closely related species will allow one to determine more precisely the time scale of the sequence transfers.

The mitochondrial sequences identified here in the nuclear genome by database searching of available sequences suggests that numerous sequences of mitochondrial origin are scattered throughout the nuclear genome (at least in plants). This expectation is supported by the observation that at least 75% of a plasmidlike mitochondrial DNA is present in the nuclear genome of a rice variety (Fukuchi et al. 1991).

The transfer of intron sequences to the nucleus in plants is not restricted to small fragments as reported here. A larger fragment comprising exon b of the *nad*1 gene together with surrounding intron sequences and mitochondrial sequences of other origins is inserted in a nuclear polyubiquitin gene in only one of several ecotypes in *Arabidopsis* (Sun and Callis 1993). Part of a mitochondrial reading frame (*rps3*/*rpl*16) is found inserted in a nuclear intron in wheat (V.K., unpublished observations). Whether the identified transfer events of intron fragments, gene fragments, or functional genes follow the same pathway in plant cells is unclear. Extension of the currently available data set will facilitate deduction of common features of the identified sequence transfers, if there are any.

Note Added at Proof In the meantime, functional analysis in transgenic plants has shown that the intron fragment inserted upstream of the lectin gene in *Dolichos* is actually relevant for tissue-specific expression (M. Etzler, personal communication).

Acknowledgments. We are grateful to Drs. Marilynn Etzler and Christiane Gatz for the gifts of nucleic acids and to Hermann Bock and Wolfgang Schuster for communication of results prior to publication. Work in the authors' laboratory was supported by grants from the Deutsche Forschungsgemeinschaft and the Bundesministerium für Forschung und Technologie.

References

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

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 Letters 325:140–145

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

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

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395

Ecke W, Schmitz U, Michaelis G (1990) The mitochondrial *nad*5 gene of sugar beet (*Beta vulgaris*) encoding a subunit of the respiratory NADH dehydrogenase. Curr Genet 18:133–139

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

Gray MW (1992) The endosymbiont hypothesis revisited. In: Wolstenholme DR, Jeon KW (eds) Mitochondrial genomes. Academic Press, San Diego, pp 233–357

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

Harada JJ, Spadoro-Tank J, Maxwell JC, Schnell DJ, Etzler ME (1990) Two lectin genes differentially expressed in *Dolichos biflorus* differ primarily by a 116-base pair sequence in their 5′ flanking regions. J Biol Chem 265:4997–5001

Hua S, Dube SK, Barnett NM, Kung S-d (1991) Nucleotide sequence of gene *oee*2-A and its cDNA encoding 23 kDa polypeptide of the oxygen-evolving complex of photosystem II in tobacco. Plant Mol Biol 17:551–553

Hua S-B, Dube SK, Barnett NM, Kung S-d (1992) Photosystem II 23 kDa polypeptide of oxygen-evolving complex is encoded by a multigene family in tobacco. Plant Mol Biol 18:997–999

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

Knoop V, Brennicke A (1991) A mitochondrial intron sequence in the 5′-flanking region of a plant nuclear lectin gene. Curr Genet 20:423–425

Knoop V, Schuster W, Wissinger B, Brennicke A (1991) *Trans* splicing integrates an exon of 22 nucleotides into the *nad*5 mRNA in higher plant mitochondria. EMBO J 10:3483–3493

Lamattina L, Grienenberger JM (1991) RNA editing of the transcript coding for subunit 4 of NADH dehydrogenase in wheat mitochondria: uneven distribution of the editing sites among the four exons. Nucleic Acids Res 19:3275–3282

Linthorst HJM, Melchers LS, Mayer A, van Roekel JSC, Cornelissen BJC, Bol JF (1990) Analysis of gene families encoding acid and basic β-1,3-glucanases of tobacco. Proc Natl Acad Sci U S A 87:8756–8760

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

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

Palmer JD, Herbon LA (1988) Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J Mol Evol 28:87–97

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci U S A 85:2444–2448

Pereira de Souza A, Jubier M-F, Delcher E, Lancelin D, Lejeune B (1991) A *trans*-splicing model for the expression of the tripartite *nad*5 gene in wheat and maize mitochondria. Plant Cell 3:1363–1378

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

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

Sperisen C, Ryals J, Meins F (1991) Comparison of cloned genes provides evidence for intergenomic exchange of DNA in the evolution of a tobacco glucan endo-1,3-β-glucosidase gene family. Proc Natl Acad Sci U S A 88:1820–1824

Stern DB, Bang AG, Thompson WF (1986) The watermelon mitochondrial URF-1 gene: Evidence for a complex structure. Curr Genet 10:857–869

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

Wissinger B, Hiesel R, Schuster W, Brennicke A (1988) The NADH-dehydrogenase subunit 5 gene in *Oenothera* mitochondria contains two introns and is co-transcribed with the 5S rRNA gene. Mol Gen Genet 212:56–65

Wissinger B, Hiesel R, Schobel W, Unseld M, Brennicke A, Schuster W (1991a) Duplicated sequence elements and their function in plant mitochondria. Z Naturforsch [C] 46c:709–716

Wissinger B, Schuster W, Brennicke A (1991b) *Trans* splicing in Oenothera mitochondria: *nad*1 mRNAs are edited in exon and *trans*-splicing group II intron sequences. Cell 65:473–482