

Identification of the entire set of transferred chloroplast DNA sequences in the mitochondrial genome of rice

Mikio Nakazono¹ and Atsushi Hirai^{1, 2}

¹ Graduate Division of Biochemical Regulation, School of Agricultural Sciences, Nagoya University, Nagoya 464-01, Japan
² Laboratory of Radiation Genetic, Faculty of Agriculture, The University of Tokyo, 1-1-1 Yayoi, Bunkyo, Tokyo 113, Japan

Received June 26, 1992 / Accepted August 8, 1992

Summary. The entire set of transferred chloroplast DNA sequences in the mitochondrial genome of rice (Oryza sativa cv. Nipponbare) was identified using clone banks that cover the chloroplast and mitochondrial genomes. The mitochondrial fragments that were homologous to chloroplast DNA were mapped and sequenced. The nucleotide sequences around the termini of integrated chloroplast sequences in the rice mtDNA revealed no common sequences or structures that might enhance the transfer of DNA. Sixteen chloroplast sequences, ranging from 32 bases to 6.8 kb in length, were found to be dispersed throughout the rice mitochondrial genome. The total length of these sequences is equal to approximately 6% (22 kb) of the rice mitochondrial genome and to 19% of the chloroplast genome. The transfer of segments of chloroplast DNA seems to have occurred at different times, both before and after the divergence of rice and maize. The mitochondrial genome appears to have been rearranged after the transfer of chloroplast sequences as a result of recombination at these sequences. The rice mitochondrial DNA contains nine intact tRNA genes and three tRNA pseudogenes derived from the chloroplast genome.

Key words: *Oryza sativa* – Mitochondrial DNA – Chloroplast DNA – Interorganellar DNA transfer – Intramitochondrial recombination

Introduction

The genetic information of higher plants is contained in three separate compartments, the nucleus, the mitochondrion and the chloroplast. The transfer of DNA

Correspondence to: A. Hirai

between subcellular genomes during the evolution of eukaryotic cells has been recognized as a general phenomenon (Stern and Lonsdale 1982; Timmis and Scott 1983; Kemble et al. 1983; Stern and Palmer 1984; Schuster and Brennicke 1987a; Gantt et al. 1991; Nugent and Palmer 1991). Similar sequences present in more than one subcellular genome have been termed "promiscuous DNA" (Ellis 1982). Stern and Lonsdale (1982) were the first to demonstrate promiscuous DNA in a mitochondrial genome. Mitochondrial DNA (mtDNA) of maize was found to contain a 12 kb chloroplast sequence that includes genes for 16 S rRNA, tRNA^{Ile} and tRNA^{Val}. Since then, a number of other chloroplast sequences have been found in the mitochondrial genomes of higher plants (Lonsdale et al. 1983; Stern and Palmer 1984; Schuster and Brennicke 1988). However, the composition of the transferred chloroplast sequences in the mitochondrial genome varies among angiosperms (Stern and Palmer 1984).

The mitochondrial genomes of higher plants are much larger and more complicated than those of other eukaryotic organisms (Newton 1988). It has been suggested that one of the reasons for such complexity might be the presence of many chloroplast sequences in the mtDNA of higher plants (Schuster and Brennicke 1988). Although there are many reports of chloroplast-like sequences in the mtDNA of higher plant, there are still few comprehensive surveys of these sequences. However, Stern and Palmer identified all locations of chloroplastlike sequences in the mtDNA of spinach by Southern hybridization (Stern and Palmer 1986).

Physical maps and overlapping clone banks of chloroplast DNA (ctDNA) and mtDNA from rice (*Oryza* sativa cv. Nipponbare) have been available for some time (Hirai et al. 1985; Iwahashi et al. 1992), and the complete sequence of the chloroplast genome has been determined (Hiratsuka et al. 1989). We have carried out hybridization analysis using two clone banks in order to examine the chloroplast sequences in the mtDNA of rice. Nucleotide sequences of hybridized fragments were determined and compared with the reported chloroplast sequences. Thus, we could determine the exact length

The sequences reported in this paper have been deposited in the DDBJ, EMBL and GenBank Data Library under accession numbers as follows: D13097(1), D13098(3), D13099(4), D13100(7), D13101(8), D13102(9), D13103(10, 11), D13104(12), D13105(13), D13106(14), D13107(15), D13108(16), D13109(2), D13110(2), D13111(2'), D13112(5, 6). Numbers in parentheses correspond with fragment numbers in Table 1

of and the sequences around the termini of each transferred fragment. In this report, we describe the entire set of chloroplast sequences in the rice mitochondrial genome.

Materials and methods

Hybridization analysis and DNA sequencing. We used the cloned DNAs from the clone banks of rice mtDNA (Iwahashi et al. 1992) and ctDNA (Hirai et al. 1985). A maize clone designated pLSH20 was a gift from Dr. D. Stern, Cornell University (Lonsdale et al. 1983). Digestion with restriction enzymes and Southern hybridization were performed as described previously (Iwahashi et al. 1992). DNA sequencing was conducted using a DNA sequencer (model 373A, Applied Biosystems, Foster City, USA). DNA sequencing data were analyzed with GENETYX Software (SDC, Tokyo, Japan).

Results

The ctDNA sequences in the rice mitochondrial genome

Seventeen cloned fragments covering the entire length of rice ctDNA were allowed to hybridize with the cloned mtDNAs. Under the hybridization conditions used, a 32 bp ctDNA sequence in the 4 kb mtDNA fragment was detected when probed with a 5 kb fragment (ctDNA clone B-7). Rice mtDNA exhibited homology to 15 of the ctDNA fragments. It was necessary to sequence the hybridized fragments of the mtDNA clones to determine

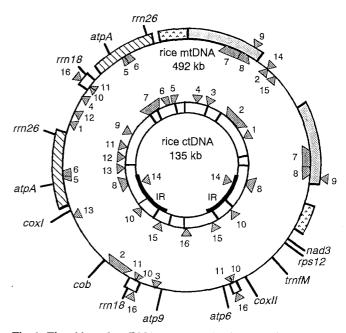


Fig. 1. The chloroplast DNA sequences in the rice mitochondrial genome. The *inner circle* is a physical map of *PstI* fragments of the rice chloroplast genome (Hirai et al. 1985). *Thick lines* indicate the inverted repeat sequence (IR). The *outer circle* is a master circle of the rice mitochondrial genome (Iwahashi et al. 1992). The repeated sequences are indicated by *boxes* outside the master circle. The sequences homologous to ctDNA are indicated by *numbered stippled triangles*. The numbers correspond to those in Table 1

Table 1. Chloroplast DNA sequences in the mitochondrial genome of rice

Fragment no.	Fragment size (bp)	Position in ctDNA ^a	Gene(s)	Sequences near the junctions	
				mt ct	ct mt
1	570	17852- 18425	trnC	ATCAATTCAA ATTAATAAAA	-CCATTAACTA TATTAATTAA
2	~6800	18985 25732	rpoB, rpoC1, rpoC2	TTTATATACC CATTTTCATT	-GAGCCAGGGA CAATTGATAC
2	371	18985- 19356	rpoB	TTTATATACC CATTTTCATT	-ATCACGGAAT GGATACGAAA
3	463	32051- 32533	atpH	CATTCCACTA ATTGCTGCTG	-CTTTTTCCTT CTCCCGTGCT
4	1090	34853- 35963	atpA, trnR	CAAGAAGAAA AAGAGGGGGC	-TCTTTCATTT TGTAAACAAC
5 6	1530 2536	43974– 45573 46932– 49319	trnS, rps4 trnL ^b , trnF, ORF159, ndhK	ATATGGGCTT TCTGCTCAAA	-TCCAATGCTC GCAAGCAAGC
7 8	~ 5500 2094	50595 55954 132158134393	trnV ^b , trnM, atpE, atpB, rbcL rpl23, rpl2 ^b , trnH, rps19	AATGATGCAG GAGCAATACC	-CAGATCGTAT ACTAGTATTC
9	457	64009- 64458	trnW, trnP	TTTTTTGATC CGACATAACA	-ACCCAACCTA ACGTATAAAG
10 11	32 48	90916– 90947 68198– 68245	Spacer region (ORF85- <i>trnV</i>) <i>clpP</i>	TGATTGGTCG AGCCCGGAGG	-GCTTCTTCAT TTCTCTAAAC
12	821	71837- 72654	petB ^b	AGTTGAATGC TGAGATTTTT	-GCCCTTTCTA TGAAGCAATG
13	358	77238- 77609	rpl14	ATAGTCTCTA CTACTAGTAT	-AGGGTCTGAG TGAGATGGAT
14	32	131904-131935	trnI	TCACTCGCTA AAGCATCCAT	-AAAGCGCCCA CCAACCTAGA
15	86	99203- 99288	trnN	TTCGAATTCT GAATGAATCA	- TACTGAGGAA GAACGGACTT
16	57	109079-109131	Spacer region $(ndhE - ndhG)$	AACAGGTACG ATTCACTAGA	-GGTATTTGTG AAAAGAGAAA

^a Numbered according to Hiratsuka et al. 1989

^b Intron-containing genes

mt, mitochondrion; ct, chloroplast

the exact length of the transferred chloroplast sequences. All the hybridizing regions were localized by mapping, and the entire nucleotide sequences of these fragments and the regions around their ends were determined, except in two cases. Two regions, containing rpoB-rpoC1-rpoC2 and rbcL-atpB-atpE-trnM-trnV, which are each more than 5 kb long, were compared in terms of their restriction maps with corresponding ctDNAs and were sequenced only around the borders between the chloroplast-specific and mitochondrial sequences.

We found 16 chloroplast fragments in rice mtDNA ranging from 32 bp to about 6.8 kb in length. As shown in Fig. 1, the transfer of ctDNA occurred from widely separated regions of ctDNA to widely separated sites in the mtDNA. Transferred chloroplast sequences are summarized in Table 1. The results show that about 6% (22 kb) of rice mtDNA, excluding repeated sequences, is made up of chloroplast sequences. Thus, about 19% of the rice ctDNA, omitting one of the inverted repeats, must have been transferred to rice mtDNA. We did not find any common sequences or structures around the termini of transferred chloroplast sequences that might explain why these particular sequences should have been inserted.

Rearranged ctDNA sequences in rice mtDNA

 $rps19 - trnH - rpl2/\psi rpl23 - rbcL - atpB - atpE - trnM - trnV$. It was previously reported that the mitochondrial genome of rice (*Oryza sativa* var. Labelle) contains a rearranged cluster of chloroplast genes, comprising *rpl2* plus $\psi rpl23 - rbcL - atpB - atpE - trnM - trnV$ (Moon et al. 1988). As shown in Fig. 2, we also detected

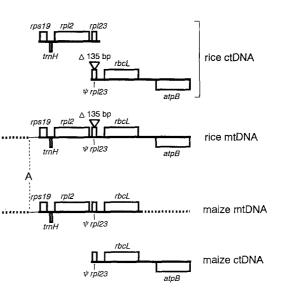
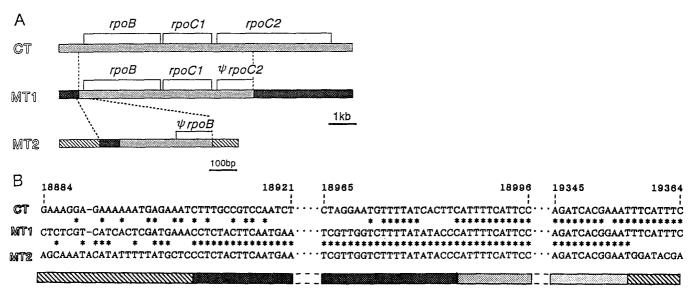
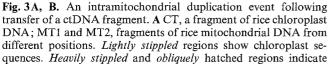


Fig. 2. Comparison between chloroplast sequences in the mitochondrial genomes of rice and maize. *Thick lines* indicate chloroplast sequences. Regions of high homology between rice and maize mitochondrion-specific sequences at the junctions with chloroplast sequences are shown by *thin lines*. The sequence upstream from position A in maize mitochondrial DNA was not available at the time of writing

 $rpl2/\psi rpl23 - rbcL - atpB - atpE - trnM - trnV$ (fragments 7 and 8 in Table 1) in the mitochondrial genome of our cultivar. In the previous report, the terminus on the rpl2 side was not reported. However, this terminus was found to be located within rps19 in our analysis. We compared rice mitochondrion-specific sequences at the border with available sequence from maize (Iams et al. 1985), and we found 93% homology over 84 bases. The result indicates that the transfer of this region oc-





MT1- and MT2-specific sequences, respectively. **B** Nucleotide sequences of junction regions. Numbers in CT are taken from Hiratsuka et al. (1989). The shading in the schematic representation under MT2 corresponds to that at MT2 in **A**

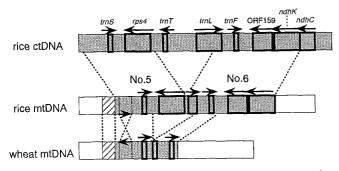


Fig. 4. Schematic comparison of chloroplast sequences between rice and wheat mtDNAs. *Stippled* areas show chloroplast sequences. *Obliquely hatched* areas indicate regions of mitochondrion-specific sequence showing high homology between rice and wheat. Open areas also show mitochondrion-specific sequences. Coding regions are indicated by *heavily outlined boxes*. Nos. 5 and 6 refer to numbers in Table 1

curred before the divergence of rice and maize. By contrast, rice $\psi rpl23$ in ctDNA and mtDNA had a 135 bp deletion. A maize mitochondrial clone, pLSH20, containing $\psi rpl23$ was sequenced, and we found that neither maize gene had a deletion, as illustrated in Fig. 2 (Bowman et al. 1988). These facts reveal that this region was transferred from ctDNA after the divergence of rice and maize.

rpoB-rpoC1-rpoC2. We found that the rice mtDNA contained a sequence homologous to the chloroplast genes for RNA polymerase, namely, rpoB, rpoC1 and rpoC2 (Fig. 3, fragments 2 and 2' in Table 1). One end of this region was located 229 bases upstream from rpoB, and the other end was located within the coding region of rpoC2. The length of this region was about 6.8 kb. This sequence was the largest sequence homologous to ctDNA that we found in the rice mitochondrial genome. A sequence of 450 nucleotides around the border on the rpoB side have been duplicated and transferred to another region about 200 kb distant in the master circle of the rice mitochondrial genome (Fig. 3). In this fragment, 371 nucleotides are chloroplast-specific and 79 nucleotides are mtDNA-specific sequences.

trnS-rps4/3'trnL-trnF-ORF159-ndhK. Rice mtDNA contained a 1530 bp chloroplast sequence that includes trnS and rps4 (fragment 5, Table 1). Furthermore, a 2536 bp sequence, which contains the 3' exon of trnL(3'trnL), trnF, ORF159 and ndhK (fragment 6, Table 1), was found upstream of rps4, as shown in Fig. 4. The two fragments are 1359 bp apart in rice ctDNA (Hiratsuka et al. 1989). The sequence containing trnT and the 5' exon of trnL was absent from rice mtDNA. Since direct repeats of TGAA were found to flank the deleted sequence, we estimate that the deletion occurred via recombination at the direct repeats. It has been reported that chloroplast-like trnS and trnF are also encoded by wheat mtDNA (Joyce and Gray 1989). The rice mtDNA-specific sequence on the trnS side exhibited 90.1% homology over a region of 166 bp with the sequence reported in wheat mtDNA. The spacer regions separating trnL and trnF were more strictly conserved between rice and wheat mtDNAs (95%) than between rice mtDNA and ctDNA (55%).

Spacer region of ORF85-trnV(GAC)/clpP. The spacer sequence (32 bp; fragment 10) between ORF85 and trnV(GAC) was found to be combined with part (48 bp; fragment 11) of clpP in rice mtDNA as a result of apparent recombination at a 7 bp sequence (TTTCTCG) that is common to the two regions. These regions are 22671 bases apart in the ctDNA.

Chloroplast-like tRNA genes

Rice mtDNA contained nine intact tRNA genes and three tRNA pseudogenes derived from the ctDNA (Table 1). Intact tRNA genes displayed 97.3–100% identity in terms of sequence with the corresponding tRNA genes in rice chloroplasts. The rice mtDNA included all the chloroplast-like genes for tRNAs, that have been found by others to be transcribed in the mitochondria of wheat and potato (Joyce and Gray 1989; Maréchal-Drouard et al. 1990). We also found that the chloroplast-like genes trnR(UCU) and trnP(UGG) were intact in the rice mtDNA. The former gene has not been reported in the mitochondrial genome of any other plant, to the best of our knowledge.

Discussion

We identified the entire set of transferred chloroplast sequences that are present in the rice mitochondrial genome. The sequences homologous to regions of the ctDNA varied from 32 bp to 6.8 kb in length and were widely distributed in the mtDNA. The total length of the chloroplast sequences accounted for about 6% of the rice mtDNA and, therefore, they do not contribute very much to the large size of the mtDNA. Evidence for four recombination events with mtDNA was found in a 22 kb chloroplast sequence. Thus, chloroplast sequences in mtDNA seem to contribute to the complexity of the mitochondrial genome.

It is of interest that the borders of the integrated chloroplast sequences in rice mtDNA exhibited no common sequences or structures that might explain their integration. It is possible that these integration events were the results of nonhomologous or random recombination (Schuster and Brennicke 1988: Manna and Brennicke 1986). Alternatively, the transferred chloroplast sequences that had been integrated via homologous recombination may have been rearranged, which would prevent us from locating any characteristic sequences. Some sequences, if not all, appear not to have been transferred via RNA. We confirmed the sequence from *rbcL* to *atpB* that had been integrated into the mtDNA. However, the two genes are transcribed in different directions in chloroplasts and there is a non-transcribed region between the two genes. This evidence rules out the possibility of RNA-mediated transfer of the chloroplast sequence. Moreover, as shown in Table 1, we identified four intron-containing chloroplast genes in the mtDNA. At least parts of the introns of all four genes were found in the mtDNA.

Intramitochondrial recombination, involving the integrated chloroplast gene for 23S rRNA has been shown to have occurred in the mitochondrial genome of *Oenothera* (Schuster and Brennicke 1987b). We also found evidence that chloroplast sequences must have been rearranged in the rice mtDNA after their transfer to mitochondria.

The sequence including rps19, trnH and rpl2 was linked to $\psi rpl23 - rbcL - atpB - atpE - trnM - trnV$ in rice mtDNA. It seems likely that homologous recombination occurred between rpl23, upstream of rpl2, and ψ rpl23, downstream of rbcL. Such potential recombination was discussed by Moon et al. (1988). In the present study, the rice mtDNA-specific sequence on the rps19 side was found to be homologous to that of maize (Iams et al. 1985). Thus, this region must have been transferred to the mtDNA before the divergence of rice and maize. However, the region corresponding to ψ rpl23 must have been transferred after the divergence of rice and maize, as explained in the Results. Then the two regions must have become connected as a result of at least two recombination events, one in rice and the other in maize. This sequence must be a 'hot spot' for recombination in mitochondria.

As shown in Fig. 3, chloroplast fragments containing genes for RNA polymerase were also transferred to rice mtDNA. Then part of the fragment plus a 79 bp mitochondrion-specific sequence at the border was duplicated and transferred to another region of the rice mitochondrial genome. It is clear that this duplication event occurred by intramitochondrial rearrangement after the transfer of fragments of ctDNA.

The rice mitochondrion-specific sequence on the trnS side of fragment No. 5 was homologous to that of wheat mtDNA, as shown in Fig. 4. The spacer sequence of trnL and trnF was highly conserved between rice and wheat mtDNAs, but this spacer sequence in rice mtDNA exhibited little homology to rice ctDNA. It is suggested that the trnS - rps4/3' trnL - trnF - ORF159 - ndhK region of ctDNA was transferred to mtDNA before the divergence of rice and wheat. In a comparison of the rice sequence with that of wheat, however, we see that the regions that include rps4 and ORF159-ndhK are absent from wheat mtDNA. It is likely that a sequence inversion event occurred at a site 300 bases upstream of wheat trnS. Both termini are associated with incomplete inverted repeat sequences (TTtCaTTCCacTTTCC/ GGAAAtgGGAAaGgAA). These differences between rice and wheat mtDNAs strongly suggest that intramitochondrial rearrangement events involving transferred ctDNA occurred after the divergence of the two species.

In general, homologous recombination and rearrangement occur much more frequently in plant mtDNA than in ctDNA and animal mtDNA (Palmer and Herbon 1988). We found that some of the transferred chloroplast sequences must have been involved in the rearrangement of the rice mitochondrial genome. It was shown previously that DNA was transferred sequentially from the chloroplast to the mitochondrion during evolution (Nugent and Palmer 1988). We confirmed this phenomenon by showing that some fragments were transferred before the divergence of rice and maize and others were transferred after this divergence. Among these sequences, some may have been gradually lost by intramitochondrial rearrangement after the transfer of larger fragments of ctDNA. These results demonstrate that mtDNA is very dynamic and flexible.

Acknowledgements. We thank Dr. D. Stern, Cornell University, for providing the plasmid clone pLSH20. This research was supported by grants-in-aid from the Ministry of Education, Science and Culture and the Ministry of Agriculture, Forestry and Fisheries of Japan.

References

- Bowman CM, Barker RF, Dyer TA (1988) In wheat ctDNA, segments of ribosomal protein genes are dispersed repeats, probably conserved by nonreciprocal recombination. Curr Genet 14:127–136
- Ellis J (1982) Promiscuous DNA chloroplast genes inside plant mitochondria. Nature 299:678–679
- Gantt JS, Baldauf SL, 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
- Hirai A, Ishibashi T, Morikami A, Iwatsuki N, Shinozaki K, Sugiura M (1985) Rice chloroplast DNA: a physical map and the location of the genes for the large subunit of ribulose-1,5bisphosphate carboxylase and the 32KD photosystem II reaction center protein. Theor Appl Genet 70:117–122
- Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun C, Meng B, 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
- Iams KP, Heckman JE, Sinclair JH (1985) Sequence of histidyl tRNA, present as a chloroplast insert in mtDNA of Zea mays. Plant Mol Biol 4:225–232
- Iwahashi M, Nakazono M, Kanno A, Sugino K, Ishibashi T, Hirai A (1992) Genetic and physical maps and a clone bank of mitochondrial DNA from rice. Theor Appl Genet 84:275–279
- Joyce PBM, Gray MW (1989) Chloroplast-like transfer RNA genes expressed in wheat mitochondria. Nucleic Acids Res 17:5461– 5476
- Kemble RJ, Mans RJ, Gabay-Laughnan S, Laughnan JR (1983) Sequence homologous to episomal mitochondrial DNAs in the maize nuclear genome. Nature 304:744-747
- Lonsdale DM, Hodge TP, Howe CJ, Stern DB (1983) Maize mitochondrial DNA contains a sequence homologous to the ribulose-1,5-bisphosphate carboxylase large subunit gene of chloroplast DNA. Cell 34:1007–1014
- Manna E, Brennicke A (1986) Site-specific circularisation at an intragenic sequence in *Oenothera* mitochondria. Mol Gen Genet 203:377–381
- Maréchal-Drouard L, Guillemaut P, Cosset A, Arbogast M, Weber F, Weil J-H, Dietrich A (1990) Transfer RNAs of potato (Solanum tuberosum) mitochondria have different genetic origins. Nucleic Acids Res 18:3689–3696
- Moon E, Kao T-H, Wu R (1988) Rice mitochondrial genome contains a rearranged chloroplast gene cluster. Mol Gen Genet 213:247-253

- Newton KJ (1988) Plant mitochondrial genomes: organization, expression and variation. Annu Rev Plant Physiol Plant Mol Biol 39:503-532
- Nugent JM, Palmer JD (1988) Location, identity, amount and serial entry of chloroplast DNA sequences in crucifer mitochondrial DNAs. Curr Genet 14:501-509
- 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
- Schuster W, Brennicke A (1987a) 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 (1987b) Plastid DNA in the mitochondrial genome of *Oenothera*: Intra- and interorganellar rearrangements involving part of the plastid ribosomal cistron. Mol Gen Genet 210:44-51

- Schuster W, Brennicke A (1988) Interorganellar sequence transfer: Plant mitochondrial DNA is nuclear, is plastid, is mitochondrial. Plant Science 54:1-10
- Stern DB, Lonsdale DM (1982) Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. Nature 299:698-702
- Stern DB, Palmer JD (1984) Extensive and widespread homologies between mitochondrial DNA and chloroplast DNA in plant. Proc Natl Acad Sci USA 81:1946–1950
- Stern DB, Palmer JD (1986) Tripartite mitochondrial genome of spinach: physical structure, mitochondrial gene mapping, and locations of transposed chloroplast DNA sequences. Nucleic Acids Res 14:5651–5666
- Timmis JN, Scott NS (1983) Spinach nuclear and chloroplast DNAs have homologous sequences. Nature 305:65–67

Communicated by R.G. Herrmann