J Mol Evol (2001) 53:333–339
DOI: 10.1007/s002390010223

JOURNAL OF MOLECULAR EVOLUTION

Regular Spliceosomal Introns Are Invasive in *Chlamydomonas reinhardtii:* 15 Introns in the Recently Relocated Mitochondrial *cox2* and *cox3* Genes

Kazuo I. Watanabe,[1,*] Takeshi Ohama[2]

[1] Department of Biology, Faculty of Science, Osaka University, 1-1 Machikaneyama-cho, Toyonaka, Osaka 560-004, Japan
[2] Graduate School of Engineering, Department of Environmental Systems Engineering, Kochi University of Technology (KUT), Tosayamada, Kochi 782-8502, Japan

Received: 22 August 2000 / Accepted: 28 February 2001

Abstract. In the unicellular green alga, *Chlamydomonas reinhardtii,* cytochrome oxidase subunit 2 (*cox2*) and 3 (*cox3*) genes are missing from the mitochondrial genome. We isolated and sequenced a BAC clone that carries the whole *cox3* gene and its corresponding cDNA. Almost the entire *cox2* gene and its cDNA were also determined. Comparison of the genomic and the corresponding cDNA sequences revealed that the *cox3* gene contains as many as nine spliceosomal introns and that *cox2* bears six introns. Putative mitochondria targeting signals were predicted at each N terminal of the *cox* genes. These spliceosomal introns were typical GT–AG-type introns, which are very common not only in *Chlamydomonas* nuclear genes but also in diverse eukaryotic taxa. We found no particular distinguishing features in the *cox* introns. Comparative analysis of these genes with the various mitochondrial genes showed that 8 of the 15 introns were interrupting the conserved mature protein coding segments, while the other 7 introns were located in the N-terminal target peptide regions. Phylogenetic analysis of the evolutionary position of *C. reinhardtii* in Chlorophyta was carried out and the existence of the *cox2* and *cox3* genes in the mitochondrial genome was superimposed in the tree. This analysis clearly shows that these *cox* genes were relocated during the evolution of Chlorophyceae. It is apparent that long before the estimated period of relocation of these mitochondrial genes, the cytosol had lost the splicing ability for group II introns. Therefore, at least eight introns located in the mature protein coding region cannot be the direct descendant of group II introns. Here, we conclude that the presence of these introns is due to the invasion of spliceosomal introns, which occurred during the evolution of Chlorophyceae. This finding provides concrete evidence supporting the "intron-late" model, which rests largely on the mobility of spliceosomal introns.

Key words: Group II intron — Signal peptide — Exon shuffling — Proto-splicing site.

Introduction

The 15.8-kilobase (kb)-long mitochondrial genome of the unicellular green alga *Chlamydomonas reinhardtii* is one of the most reduced, encoding genes for only eight proteins, large and small ribosomal RNAs, and two transfer RNAs. In contrast, the typical mitochondrial genome size is ca. 50 kb in most unicellular eukaryote taxa (Gray et al. 1998). The *C. reinhardtii* mitochondrial genome is even missing the representative mitochondrial genome contents, the *cox2* and *cox3* genes (Gray and Boer 1988), as in two other green algal relatives, *Chlamydomonas eugametos* (Denovan-Wright et al. 1998) and *Pedinomonas minor* (Turmel et al. 1999). It was recently shown that in the colorless chlamydomonad algae *Polytomella* spp. and *C. reinhardtii,* COXII is encoded by two independent nuclear genes (*cox2a* and

* *Present address:* Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, TX 78712, USA
Correspondence to: Takeshi Ohama; *email:* ohama@env.kochi-tech.ac.jp

−*2b*) and that COXIII is also in the nuclear genome (Pérez-Martínez et al. 2000, 2001). These analyses were based mainly on cDNA structures or N-terminal peptide sequences. However, the introns in these genes have not been characterized thus far.

Here, we analyzed an unexpected gene structure of the *cox2* and *cox3* genes of *C. reinhardtii*, focusing on the evolutionary origin of the detected introns. These *cox* genes contain a total of 15 spliceosomal introns, which seem to have no particular unique features compared to introns in other genes. This suggests that spliceosomal introns are also highly invasive in nature, similar to organelle group I and II introns. This attribute might be strongly related to the ability of group II introns to invade specific sites of intronless genes with a high efficiency, a process refered to as intron homing (for a review, see Lambowitz and Belfort 1993). Recent evidence suggests that group II introns are the ancestor of spliceosomal introns, i.e., spliceosomal introns and the spliceosome are a type of extremely deformed group II intron (Sharp 1991). This conclusion is based mainly on two facts: first, the core splicing reaction is identical between these two kinds of introns (for reviews, see Saldanha et al. 1993; Weiner 1993); second, experimental data showed trans-activation of group II intron splicing by a small RNA in a spliceosome (U5 snRNA) (Hetzer et al. 1997). Such deformation of group II introns seems to have occurred in the early evolution of eukaryotes, since core features of spliceosomal introns and spliceosomes are essentially identical among diverse eukaryote taxa.

Materials and Methods

Algal Culture Conditions. Chlamydomonas reinhardtii strain CC277 (cw-15) was obtained from The *Chlamydomonas* Genetics Center (Duke University, Durham, NC, USA). Cells were grown at 25°C in TAP medium (Harris 1989) in constant light (200 $\mu E/m^2$ s).

Preparation of DNA and cDNA. Total DNA was extracted following the method described by Harris (1989). Total RNA was extracted using TRIzol reagent (GIBCO BRL) following the manufacturer's instructions. From the total RNA, poly(A) RNA was extracted using oligo(dT)-latex particles, Oligotex-dT30 (Roche). Purified poly(A) RNA was used as the template for complementary DNA (cDNA) synthesis with oligonucleotide primers, using the Ready-To-Go You Prime First Strand Kit (Pharmacia Biotech).

PCR Conditions A buffer kit optimized for amplifications from GC-rich templates, i.e., LA-Taq with GC buffer I (Takara Shuzo, Japan), was used to amplify a part of the *cox3* gene from genomic DNA or cDNA. The programs for the thermal cyclers were 94°C for 5 min, followed by 25–35 cycles at 94°C for 30 s, 48°C for 1 min, and 72°C for 2 min, followed by 72°C for 5 min.

Cloning of the cox2 *and* cox3 *Genes.* A part of the *cox3* cDNA was synthesized with a degenerated oligonucleotide primer (COX3–CRn–349a, AAIAACCAIACIACRTCNACRAARTG; R = A/G, N = G/A/T/C, I = inosine), using extracted poly(A) RNA as the template. From the cDNA, a partial *cox3* fragment, 179 base pairs (bp), was amplified by PCR with nested pairs of degenerated primers (COX3–CRn–281s, TCSGACGGGATHTAYGGNTCNACNTT; COX3–CRn–347a, TAYTGGCAYTTYGTNGAYGTIGTITGGTTITT; COX3–CRn–280s, TCSGACGGGATHTAYGGNTCNAC; COX3–CRn–348a, AAIAACCAIACIACRTCNACRAARTGCCA: S = C/G, H = A/T/C, Y = C/T, N = G/A/T/C, I = inosine).

The 3′-terminal sequence of the *cox3* mRNA was determined by reverse transcription polymerase chain reaction (RT-PCR), using the Ready-To-Go T-Primed First-Strand-Kit (Pharmacia Biotech). For the 5′-terminal part of the *cox3* cDNA sequence, we referred to the expressed sequence tag (EST) database for *C. reinhardtii* (c-9) (Kazusa DNA Research Institute, Japan). To get a clue about the missing *Chlamydomonas cox2* gene, we searched the EST database using the *cox2* sequence of the chlorophyte *Prototheca wickerhamii.* Genomic PCR fragments of the *cox2* gene corresponding to the EST sequences were obtained using two sets of primers based on the EST sequences (CRCOX2–m66s, TGGTTGCTTCCGAGGGCGTCGCACA; CRCOX2–113a, CTCCATGGTGGTGTGTTGGGTCAGC; CRCOX2–107s, AGCATACATCTTCAGCTAGAACGGG; CRCOX2–288a, ACGAGATCCACTTCTTCACGTACTC).

These genomic PCR fragments were cloned into pT7Blue T-Vector (Novagen, USA) and sequenced. The resultant clones were also used as the probe to screen the high-density BAC library filter for *C. reinhardtii* (cw 92) (IncyteGenomics, USA). Probes were labeled using the AlkPhos Direct Labeling system (Amersham Life Science), and signal detection was performed using the CDP-Star chemiluminescent detection reagent (Amersham Life Science) according to the supplier's instructions. Two positive BAC clones for each of the *cox2* and *cox3* probes were detected in the library. Their inserts were subcloned into pT7Blue-Vector or were subjected directly to sequencing.

Phylogenetic Analysis of Chlorophyte Algae. The neighbor-joining (NJ) (Saitou and Nei 1987) tree was constructed using SINCA software version 3.0 (Fujitsu System Engineering, Japan). Bootstrap resampling (500 times) was carried out to quantify the relative support for each of the branches of the phylogenetic tree.

Results

Spliceosomal Introns in the cox2 *and* cox3 *Genes*

One of the two positive BAC clones (24p23), which was selected using the labeled probe for the *cox3* gene, was sequenced over 4.0 kb (Genbank AB046570). The resultant sequence contained the *cox3*-like gene. Homology search of the sequence against the amino acid sequence database SWISSPROT using the BLAST program (Altschul et al. 1997) showed that it has significant homology to the mitochondrial *cox3* genes of various eukaryotes (score = 150–172), while similarity values for the bacterial *cox3* homologues are less than 131. Comparison of the genome sequence with the raised cDNA sequences by RT-PCR revealed that nine short introns ranging in size from 133 to 233 nucleotides (nt) were inserted in this gene, the total length of which occupied more than half of the *cox3* gene. [Our genomic sequence data had no conflict with the *cox3* cDNA sequence that was determined independently (Pérez-Martínez et al. 2000)].

Among the *Chlamydomonas* EST data (Asamizu et al. 1999), the cDNA sequences of LC027d01_r (503 bp) and

CM066h07_r (540 bp) were found to posses significant homologies to the *cox2* gene of the chlorophyte *P. wickerhamii.* These two cDNA sequences did not possess any sequence overlap; instead, they probably correspond to the N-terminal half and the remaining C-terminal part of the *cox2* gene (Fig. 1b). To obtain the full genomic *cox2* sequence, we surveyed a BAC library (IncyteGenomycs, USA). Two BAC clones, 16d21 and 40n6, showed apparent signals when the PCR fragment corresponding to EST CM066h07_r was used as a probe. Subsequent sequencing confirmed that those BAC clones actually contained the proper sequence corresponding to CM066h07_r. The genomic sequence corresponding to LC027d01_r was determined by sequencing the PCR fragments obtained using the total DNA as template. The cDNA sequence of LC027d01_r (containing the N-terminal region of the *cox2* sequence of 503 bp) and CM066h07_r (containing the C-terminal region of the *cox2* sequence of 540 bp) corresponded to 1307-bp (Genbank AB046571) and 727-bp (Genbank AB046572) genomic sequences, respectively. These genomic sequences contained five and one short intervening sequences (ranging in size from 79 to 297 nt) compared to the cDNA sequences, respectively.

Conserved sequences for exon–intron boundaries and for the putative internal branch sites of these introns (Fig. 2) showed that all 15 introns resemble the typical U2-type GT–AG intron, which is the most common type of spliceosomal intron, but not the U12 AT–AC-type intron (Sharp and Burge 1997). None of these nine introns showed significant similarity to one another. Furthermore, each intron sequence was subjected to a BLAST search (Altschul et al. 1997), revealing that these sequences have no significant homology to any introns or to any other sequence in the GenBank database (release 117). As expected, no prominent secondary structures within the introns were predicted.

Among the nine introns of the *cox3* gene, the five downstream introns (5 through 9) (Fig. 1a), and also the three downstream introns (4 through 6) of the *cox2* gene (Fig. 1b), were interrupting segments conserved among the mitochondrial *cox* genes, while upstream introns were located in the N-terminal regions that are unique to the *C. reinhardtii cox* genes.

We also analyzed the positive correlation of intron phases (the position of the intron in the codon) between the adjacent introns. Only 3 of 13 neighboring positions among introns [between 3 and 4 (3/4) and 4/5 of *cox2* and 8/9 of *cox3*] had the same phase (Fig. 2). Therefore, exon shuffling was not positively supported by the intron phase analysis (de Souza et al. 1998).

Signal Peptides in the cox2 *and* cox3 *Genes*

The deduced amino acid (aa) sequence of the *C. reinhardtii cox3* gene was aligned with eight homologues from extremely divergent eukaryotes and, in addition, with two from bacteria (Fig. 1a). Of the 382 aa residues, only the 250 C-terminal aa showed high and nearly the same extent of similarity to the various mitochondrial *cox3* genes, for example, 28% to *Prototheca wickerhamil* (green alga), 28% to *Marchantia polymorpha* (land plant), 31% to *Allomyces macrogynus* (fungus), and 29% to *Homo sapiens.* In contrast, the N-terminal region of the *cox* gene was peculiarly long in *C. reinhardtii.* BLAST homology search of the 145 N-terminal aa showed no significant homology to any sequence in the database. Likewise, the *cox2*-like gene matched in a similar fashion; i.e., the 250 C-terminal aa of the gene showed significant homology to various mitochondrial *cox2* genes, whereas no prominent similarity existed for the extremely long N-terminal region (Fig. 1b).

Analysis with the MITOPROT program (Claros 1995, http://www.mips.biochem.mpg.de/proj/medgen/mitop/) predicted a high probability for these N-terminal regions as efficient mitochondria targeting signals (0.6805 and 0.8824 for *cox2* and *cox3*, respectively). At 361 nt downstream from the translational stop codon of the *cox3* gene, a poly(A) additional signal (TGTAA) was detected. In fact, each of the six analyzed cDNA clones had a poly(A) tract 14, 15, or 58 bp downstream from the poly(A) addition signal. Taken altogether, the *C. reinhardtii cox2* and *cox3* genes are most likely typical active genes that translocated from the mitochondrion to the nucleus.

Discussion

The *Chlamydomonas cox2* and *cox3* genes showed almost the same extent of similarity to various mitochondrial genes, apparently not reflecting evolutionary relationships. This is probably because the translocation of these genes from an organellar genome to the nucleus led to a significant change in their sequence evolution. Results of our targeting sequence analyses in silico were confirmed by peptide analyses of their mature products (Pérez-Martínez et al. 2000, 2001).

Group II introns seem to be completely absent from current eukaryote genomes. This is most likely because group II introns that are suddenly relocated into their current nuclei are unable to be spliced out successfully in the cytosol. However, from an evolutionary perspective, spliceosomal introns are supposedly the descendants of group II introns that were once efficiently spliced out into the cytosol of ancient eukaryote cells (Sharp 1991). It follows that the existence of spliceosomal introns even in the relocated organelle genes cannot be concrete evidence for the insertional gain of spliceosomal introns. Therefore, without information on the time period of the gene relocation, the following possibility cannot be excluded: an ancestral organellar gene containing group II

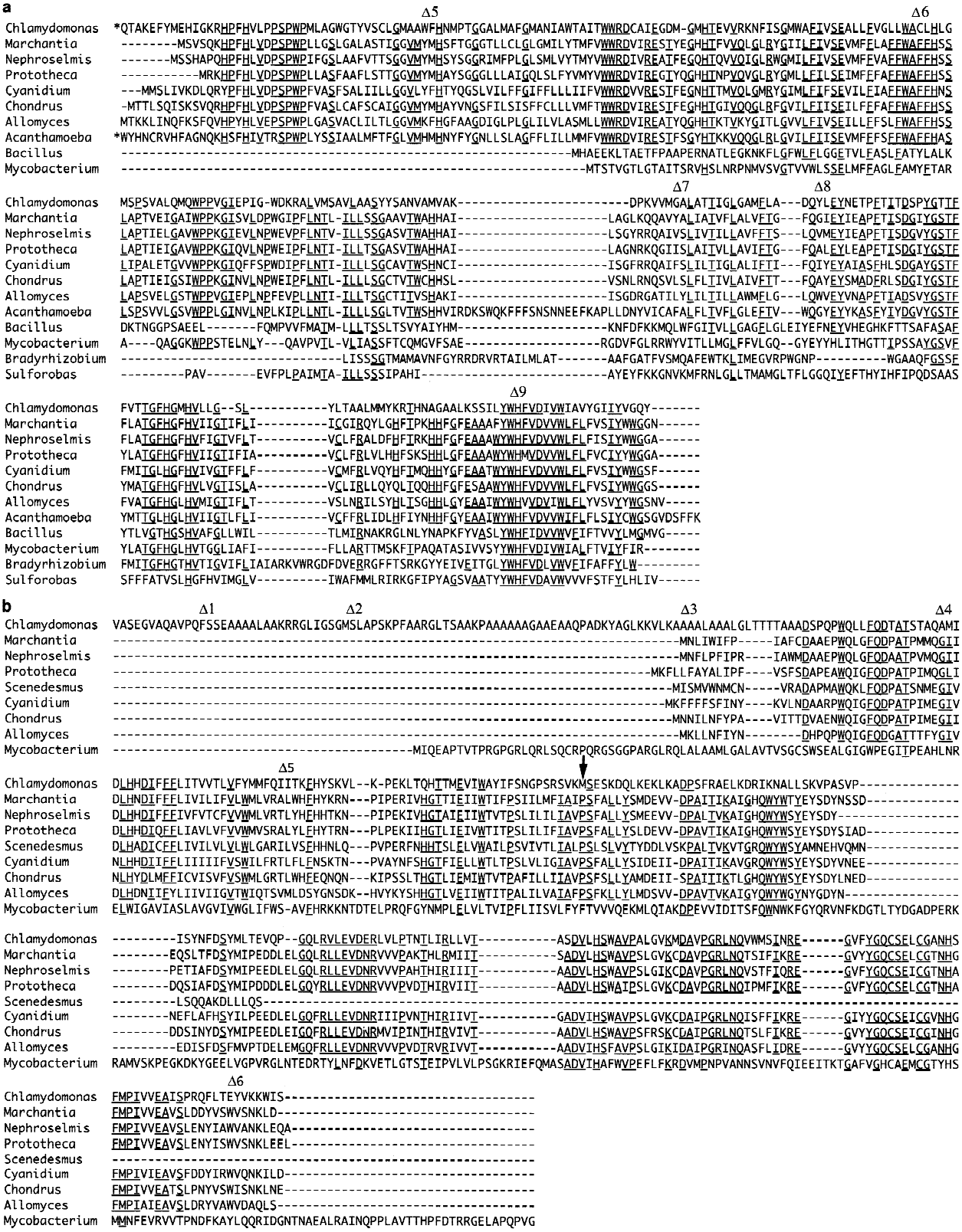

Fig. 1. Alignment of the *Chlamydomonas reinhardtii cox3* **(a)** and *cox2* **(b)** genes with those of various other organisms. The *asterisk* at the beginning of the sequence for *C. reinhardtii* and *Acanthamoeba castellanii* indicates that farther upstream 142 and 29 amino acid (aa) residues were omitted. Conserved amino acids are *underlined.* Sequence data, except for those for *C. reinhardtii*, are from the GenBank sequence database. *Triangles* indicate the position of each intron, and the following *digit* corresponds to the serial number of introns in the gene, ordered with the most N-terminal one first. The boundary of the sequence obtained from LC027d01_r and CM066h07_r is shown by the *arrow* (see the text). Organism names and GenBank accession numbers are as follows: *Allomyces macrogynus* (P80439), *Marchantia polymorpha* (P26858), *Nephroselmis olivacea* (AAF03208), *Prototheca wickerhamii* (Q37620), *Cyanidium caldarium* (P48873), *Chondrus crispus* (P48872), *Acanthamoeba castellanii* (Q37374), *Theileria parva* (S41690), *Bacillus* sp. PS3 (Q03439), and *Mycobacterium leprae* (CAA18701).

	No	5′ exon	intron	(branch)		3′ exon	intron length	phase
cox2	1	CTTCCGAGGGCGTCGCACAG	GTGCGAACAGACCTCCACGG	... CTCAC ... 29 ...		TTTGCATGCTTCGATTCCAG	99	0
	2	GCGCCGTGGCCTGATTGGCA	GTAAGTCCAGCACAACTGGC	... CTGGC ... 27 ...		ACTGGCTATTGCACGTGCAG	79	1
	3	GCCTGAAGAAGGTCCTGAAG	GTGGGGAGGATGCTGGAGCG	... GTCAC ... 25 ...		CTACTATATAACGCGTGCAG	183	0
	4	CGGCCACCTCCACTGCTCAG	GTGCGCCGGGCGTAAAGCTG	... CTGAC ... 29 ...		GACGGCTCTCCCTCCGACAG	297	0
	5	TGTTCTACATGATGTTCCAG	GTGGGGCAGCGCCTGTAATT	... CTGAA ... 42 ...		TTGGGTTGCGCCTACCGCAG	145	0
	6	CCCCGTCAGTTCCTGACCGA	GTGAGTTGCGCATGGGGCCG	... CTAAC ... 25 ...		TCGGTTCTGCTCTATTGCGA	187	2
cox3	1	CCTGACGCGGGCTCCCGCAG	GTGTGTATAATTGGATGTAA	... CTGAC ... 33 ...		CGCTCCTGGATGCTATGCAG	267	1
	2	CTGCAGTCCTCTGCCTTCGG	GTGAGCGTAGGGCATCGCAG	... CTGAA ... 32 ...		GCACATTTTGTGTCCCGCAG	133	2
	3	GCGGCCTGGGCTTCGGCAAG	GTAAGGACTTTACGGACATA	... CTGAC ... 39 ...		CGCGCGTCTCGCCCTCGCAG	139	0
	4	CTGCCCCCTCGCGGCACCCG	GTGAGGAGGCGCTGCAGCGG	... CTAAC ... 30 ...		CCTCCATCTCTGTCTTGCAG	256	2
	5	CACCTACGTGAGCTGCCTGG	GTGAGCAGCAGTTGCAGTGG	... TTCAT ... 98 ...		TCCTTGCCCCCCGCGCGCAG	290	1
	6	CTGCTGTTCGTGGGCCTGCT	GTGAGTACGGCAGCGTGCGC	... CTAAT ... 41 ...		CTTCCGCCTCGCCTCCGCAG	323	2
	7	TGGTGGCCAAGGACCCTAAG	GTGGGGAGGCTGGGGGAAGG	... CTGAC ... 37 ...		TTCCGTCTTGGTGTGCGCAG	160	0
	8	GCCATGTTCTTGGCTGACCA	GTGAGTGCCGAGTCTCTTGA	... GTGAC ... 44 ...		CACCACCTTGCGCGGCGCAG	180	2
	9	AAGAGCTCCATCCTGTACTG	GTGGGTCGCGCTGTACTGCA	... GTGAC ... 57 ...		CCAACCCCATCAAAACACAG	204	2
			G GTRAGT	CTRAY		YAG S		

Fig. 2. Alignment of 15 spliceosomal introns in the *cox2* and *cox3* genes of *Chlamydomonas reinhardtii* with the intron phase analysis. Two *vertical lines* indicate the 5′ and 3′ splice junctions, respectively. The putative branch site, and the distance between the branch site and the 3′ splice site, is also shown. In the bottom line, a representative sequence of the U2-type GT–AG intron (Sharp and Burge 1997) is indicated: R indicates A or G, Y indicates C or T, and S indicates C or G. The last column lists the phase of the intron: phase 0, the intron is inserted between two codons; phase 1, the intron is inserted between the first and the second nucleotide of the codon; phase 2, the intron is inserted between the second and the last nucleotide of the codon.

introns was relocated to the nucleus in the primitive eukaryote cell. These group II introns were subsequently deformed during evolution into the currently existing spliceosomal introns.

To deduce the evolutionary period when the relocation of the *Chlamydomonas cox2* and *cox3* genes occurred, phylogenetic relationships of representative algae belonging to Chlorophyta were analyzed based on the *cox1* sequences, followed by superimposition of the existence of the *cox2* and *cox3* genes in the mitochondrial genome onto the tree (Fig. 3). This clearly shows that relocation of the *Chlamydomonas cox2* and *cox3* genes occurred during the evolution of Chlorophyceae, after the branch of Volvocales and Chlorococcales. The suggested period is modern enough to reject the scenario described above.

This led us to conclude that at least five downstream introns in the *cox3* gene and three downstream introns in the *cox2* gene are highly likely due to invasion of spliceosomal introns. It has been suggested that in *Scenedesmus obliquus* the *cox2* gene is also split into two genes, and probably the N-terminal coding region (*cox2a*) is retained in the mitochondrial genome, whereas the C terminal-containing region (*cox2b*) has been transferred to the nucleus (Pérez-Martínez et al. 2001). Taking the order of splitting *cox2* gene transfer as the same as that for *S. obliquus* in the chlamydomonad algae *C. reinhardtii* and *Polytomella* spp. simply leads to the conclusion that intron density in the *cox2b* is probably higher than in *cox2a.* However, in the two chlamydomonad algae, the intron density is strongly biased toward the *cox2a* over *cox2b.* This suggests that the gene transfer of *cox2a* preceded that of *cox2b* in the ancestor of these chlamydomonad algae, not following the case of *S. obliquus.* Besides our investigation, there are several reported instances of mitochondrial genes that relocated quite recently to the nucleus, e.g., the *cox2* genes in various species of beans (Nugent and Palmer 1991) and ribosomal protein S11 gene in rice (Kadowaki et al. 1996). However, in these cases, introns were detected only at the junction between the transit peptide and the mature protein coding region. Based on these observations, the evolutional composition of these genes has been proposed such that the targeting presequence region-bearing introns were added to the relocated organellar gene as a result of exon shuffling. This is probably due to the fact that in these flowering plants, chromosome recombination is very active, whereas intron invasion (after relocation of the reported genes) is less active than in *Chlamydomonas.* Spliceosomal introns in the chloroplast-targeted acetolactate synthase (ALS) genes of *C. reinhardtii* (Genbank acc. no. AF047459) and *Volvox carteri* (AF04490) are instances that clearly show intron invasion into organelle genes (besides those in this investigation), consistently showing the high activity of intron invasion in them (Funke et al. 1999).

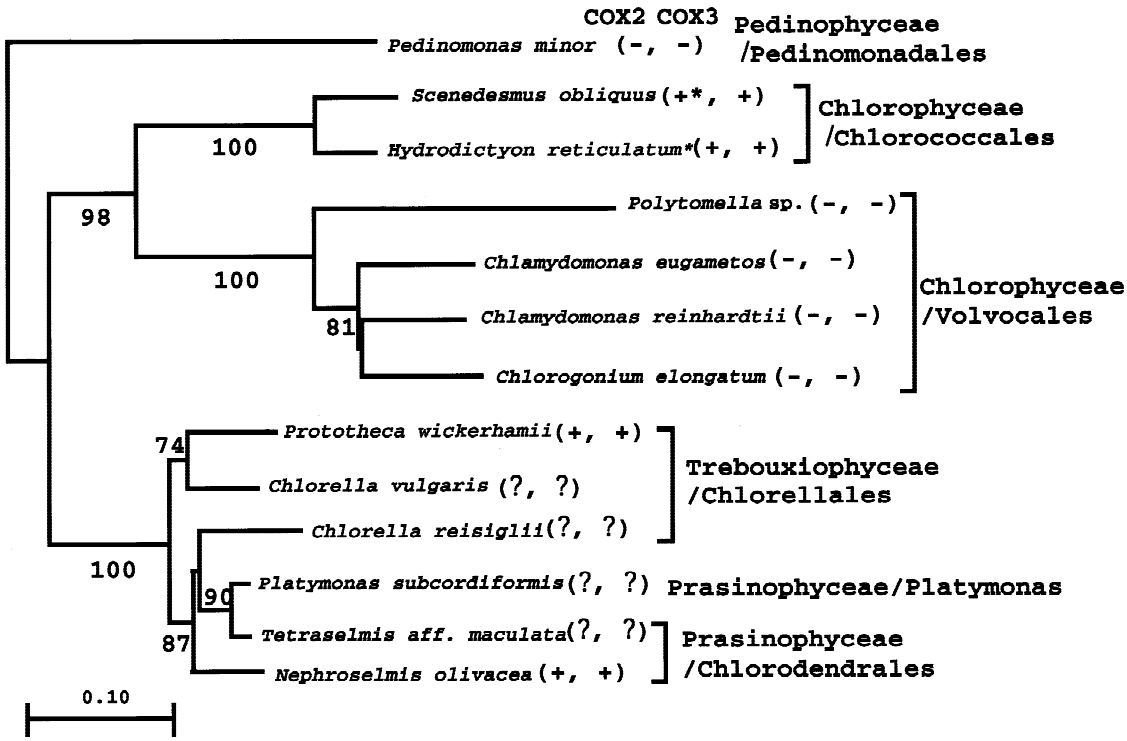

Fig. 3. Unrooted phylogenetic tree based on the amino acid sequences of the *cox1* genes of chlorophyte algae. The tree, obtained by the neighbor-joining method, is shown with bootstrap values (those less than 50% are omitted). The existence of the *cox2* and *cox3* genes in each species of mitochondria is also indicated: +, present; −, absent. +*: A *cox2a* gene that encodes a protein corresponding to the amino-terminal half of a typical one-polypeptide COXII is detected (Pérez-Martínez et al. 2001). *Hydrodictyon reticulatum**: Ohama et al. (unpublished results).

We closely analyzed the exon–intron boundaries and also any possibly informative secondary structures of the introns in the *cox* genes to specify the process for intron invasion, but no particular traits were detected. This led us to speculate that the extremely high invasive ability of spliceosomal introns in *Chlamydomonas* is due to the spliceosome or some other enzymes in the cytosol.

A particular type of noncanonical reverse splicing utilizing a previously excised intron (i.e., insertion of the spliced-out intron into another mRNA) might generate a mRNA which is invaded by an intron, which could then be reverse transcribed into DNA, followed by recombination of the cDNA fragment into the genome. The reported insertion of a spliced-out group II intron into an unrelated mRNA supports this model (Augustin et al. 1990). A non-mutually exclusive and alternative hypothesis is that excised intron RNA is directly inserted into an intronless gene, after which the intronic RNA region is changed into DNA by reverse transcriptase in the fashion of group II intron invasion (Yang et al. 1996). Not only genes of organellar origin, but also nuclear genes must have been invaded severely in *C. reinhardtii.* This scenario is the most probable reason for the high density of spliceosomal introns in *Chlamydomonas* nuclear genes (Logsdon et al. 1998).

Through whole-genome sequencing projects for various organisms, it is becoming clearer that intron density is diverse among organisms without apparent evolutionary relationships (Logsdon et al. 1998). This phenomenon might be explained as follows: the number of introns increases rapidly following the abrupt revival of the activity necessary for intron invasion in a lineage, which lasts for a period of time, while a decrease in the number of introns occurs gradually through reverse transcription of the processed mRNA and integration of the product into the genome. Such a switching on and off of the activity occurred independently in multiple lineages.

Acknowledgments. K.W. was supported by Research Fellowships of the Japanese Society for Promotion of Science for Young Scientists. Special thanks go to Drs. S. Tabata (Kazusa DNA Research Center) and H. Fukuda (Kyoto University) for the EST sequencing data on *C. reinhardtii* and, also, to Dr. M. Ehara for her assistance in drawing the figures.

References

Altschul SF, Madden TL, Zhang AAJ, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Asamizu E, Nakamura Y, Sato S, Fukuzawa H, Tabata S (1999) A large scale structural analysis of cDNAs in a unicellular green alga, *Chlamydomonas reinhardtii.* I. Generation of 3433 nonredundant expressed sequence tags. DNA Res 6:369–373

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

Denovan-Wright EM, Nedelcu AM, Lee RW (1998) Complete sequence of the mitochondrial DNA of *Chlamydomonas eugametos*. Plant Mol Biol 36:285–295

de Souza SJ, Long M, Klein RJ, Roy S, Lin S, Gilbert W (1998) Toward a resolution of the introns early/late debate: Only phase zero introns are correlated with the structure of ancient proteins. Proc Natl Acad Sci USA 95:5094–5099

Funke RP, Kovar JL, Logsdon JM, Corrette-Bennett JC, Straus DR, Weeks DP (1999) Nucleus-encoded, plastid-targeted acetolactate synthase genes in two closely related chlorophytes, *Chlamydomonas reihardtii* and *Volvox carteri:* Phylogenetic origins and recent insertion of introns. Mol Gen Genet 262:12–21

Gray MW, Boer PH (1988) Organization and expression of algal (*Chlamydomonas reinhardtii*) mitochondrial DNA. Philos Trans R Soc Lond B Biol Sci 319:135–147

Gray MW, Lang BF, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Brossard N, Delage E, Littlejohn TG, Plante I, Rioux P, Saint LD, Zhu Y, Burger G (1998) Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res 26:865–878

Harris EH (1989) The Chlamydomonas sourcebook: A comprehensive guide to biology and laboratory use. Academic Press, San Diego

Hetzer M, Wurzer G, Schweyen RJ, Mueller MW (1997) Transactivation of group II intron splicing by nuclear U5 snRNA Nature 386: 417–420

Kadowaki K, Kubo N, Ozawa K, Hirai A (1996) Targeting presequence acquisition after mitochondrial gene transfer to the nucleus occurs by duplication of existing targeting signals. EMBO J 15:6652–6661

Lambowitz AM, Belfort M (1993) Introns as mobile genetic elements. Annu Rev Biochem 62:587–622

Logsdon JM, Stoltzfus A, Doolittle WF (1998) Molecular evolution: Recent cases of spliceosomal intron gain? Curr Biol 8:R560–563

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

Pérez-Martínez X, Vazquez-Acevedo M, Tolkunova E, Funes S, Claros MG, Davidson E, King MP, Gonzalez-Halphen D (2000) Unusual location of a mitochondrial gene. Subunit iii of cytochrome c oxidase is encoded in the nucleus of chlamydomonad algae. J Biol Chem. 275:30144–30152

Pérez-Martínez X, Antaramian A, Vazquez-Acevedo M, Funes S, Tolkunova E, d'Alayer J, Claros MG, Davidson E, King MP, Gonzalez-Halphen D (2001) Subunit II of cytochrome c oxidase in chlamydomonad algae is a heterodimer encoded by two independent nuclear genes. J Biol Chem 276:11302–11309

Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Saldanha R, Mohr G, Belfort M, Lambowitz AM (1993) Group I and group II introns. FASEB J 7:15–24

Sharp PA (1991) "Five easy pieces" Science 254:663

Sharp PA, Burge CB (1997) Classification of introns: U2-type or U12-type. Cell 91:875–879

Turmel M, Lemieux C, Burger G, Lang BF, Otis C, Plante I, Gray MW (1999) The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*. Two radically different evolutionary patterns within green algae. Plant Cell 11:1717–1730

Weiner, AM (1993) mRNA splicing and autocatalytic introns: Distant cousins or the products of chemical determinism? Cell 72:161–164

Yang J, Zimmerly S, Perlman PS, Lambowitz AM (1996) Efficient integration of an intron RNA into double-stranded DNA by reverse splicing. Nature 381:332–335