J Mol Evol (1998) 47:431–440

JOURNAL OF MOLECULAR EVOLUTION

The Mitochondrial Genome of *Chlorogonium elongatum* Inferred from the Complete Sequence

Jürgen Kroymann,* Klaus Zetsche

Institut für Pflanzenphysiologie der Justus-Liebig-Universität Giessen, Heinrich-Buff-Ring 54-62, 35392 Giessen, Germany

Received: 3 November 1997 / Accepted: 12 January 1998

Abstract. The 22,704-bp circular mitochondrial DNA (mtDNA) of the chlamydomonad alga *Chlorogonium elongatum* was completely cloned and sequenced. The genome encodes seven proteins of the respiratory electron transport chain, subunit 1 of the cytochrome oxidase complex (*cox1*), apocytochrome b (*cob*), five subunits of the NADH dehydrogenase complex (*nad1, nad2, nad4, nad5,* and *nad6*), a set of three tRNAs (Q, W, M), and the large (LSU)- and small (SSU)-subunit ribosomal RNAs. Six group-I introns were found, two each in the *cox1, cob,* and *nad5* genes. In each intron an open reading frame (ORF) related to maturases or endonucleases was identified. Both the LSU and the SSU rRNA genes are split into fragments intermingled with each other and with other genes. Although the average A + T content is 62.2%, GC-rich clusters were detected in intergenic regions, in variable domains of the rRNA genes, and in introns and intron-encoded ORFs. A comparison of the genome maps reveals that *C. elongatum* and *Chlamydomonas eugametos* mtDNAs are more closely related to one another than either is to *Chlamydomonas reinhardtii* mtDNA.

Key words: *Chlamydomonas — Chlorogonium elongatum* — Mitochondrial genome — Genome map — Group-I introns — GC-rich elements

Introduction

Among the few green algal mitochondrial DNAs (mtDNAs) investigated so far, the mitochondrial genomes (mt genomes) of two chlamydomonad algae, *Chlamydomonas reinhardtii* and *Chlamydomonas eugametos,* and their interfertile relatives, *Chlamydomonas smithii* and *Chlamydomonas moewusii,* respectively, have been characterized in detail. Mitochondrial genomes from both *C. reinhardtii* (Pratje et al. 1984, 1989; Boer et al. 1985; Vahrenholz et al. 1985; Boer and Gray 1986a, b, 1988a–1989; Ma et al. 1988, 1989a, b, 1990; Michaelis et al. 1990; Bennoun et al. 1991) and *C. eugametos* (Denovan-Wright and Lee 1992, 1993, 1994) encode genes for seven subunits of the mitochondrial respiratory chain (*nad1, −2, −4, −5,* and *−6, cob,* and *cox1*), for three tRNAs (tRNA-M, tRNA-W, and tRNA-Q), and for the large (LSU rRNA)- and small (SSU rRNA)-subunit ribosomal RNAs. The latter are discontinuous, and the LSU and SSU rRNA gene fragments are interspersed with one another and with protein-coding and tRNA genes (Boer and Gray 1988a, b; Denovan-Wright and Lee 1994). Additionally, in *C. reinhardtii* an *rtl* gene coding for a reverse transcriptase-like protein was found (Boer and Gray 1988a). This gene is not present in *C. eugametos* or *C. moewusii* mtDNA (Denovan-Wright and Lee 1993; R.W. Lee, personal communication).

Despite this very similar gene content, the mt genomes from *C. reinhardtii*/*C. smithii* and *C. eugametos*/*C. moewusii* display quite different structural features. The 15.8-kilo base pair (kb) mitochondrial genome from *C. reinhardtii* is linear (Ryan et al. 1978; Michaelis et al. 1990; Harris 1993) and flanked by characteristic inverted

* *Present address:* Max-Planck-Institut für Chemische Ökologie, Talzend promenade 1a, 07745 Jena, Germany
Correspondence to: K. Zetsche; *e-mail:* Klaus.Zetsche@bot3.bio.uni-giessen.de

repeats with 3′ single-stranded, noncomplementary extensions of 39 to 41 nucleotides (Ma et al. 1992; Vahrenholz et al. 1993). Both DNA strands encode genes. With the exception of a 1.075-kb group-I intron present in *C. smithii* mtDNA, the mt genomes from the interfertile species *C. reinhardtii* and *C. smithii* are colinear (Colleaux et al. 1990). In contrast, circular maps have been presented for the colinear 22- and 24-kb mt genomes of *C. moewusii* (Lee et al. 1991) and *C. eugametos* (Denovan-Wright and Lee 1992; Denovan-Wright and Lee 1993). In these species, all genes and gene fragments are encoded on the same strand.

To investigate further the evolution of mt genomes among chlamydomonad algae, we have examined the mtDNA of *Chlorogonium elongatum*, a species distantly related to *C. reinhardtii*/*C. smithii* and *C. eugametos*/*C. moewusii* (Buchheim et al. 1990, 1996; Buchheim and Chapman 1991). We present a physical and functional map of the mt genome from *C. elongatum* inferred from the complete nucleotide sequence.

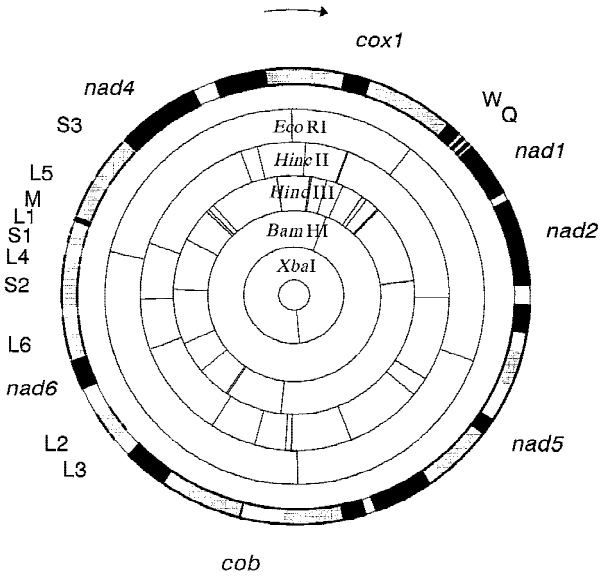

Fig. 1. Physical and functional map of *Chlorogonium elongatum* mtDNA. The circular map results from sequencing of overlapping clones. The five *inner circles* indicate restriction sites for *Hin*dIII, *Hin*cII, *Eco*RI, *Xba*I, and *Bam*HI. Note that *Xba*I and *Bam*HI each cut the mtDNA only once. The *outer circle* shows the position, size, and organization of protein-coding (*nad1, −2, −4, −5,* and *−6, cob, cox1*) and tRNA (M, W, Q) genes as well as the position and approximate size of rRNA gene fragments (S1–S3, L1–L6). The *cox1, nad5,* and *cob* genes each contain two group-I introns. Exons are shown as *black boxes;* intronic ORFs, as *gray-tone* regions. All genes are oriented in a clockwise direction, as indicated by the *arrow.*

Materials and Methods

Strain, Culture Conditions, Preparation of Organellar DNA, Cloning into pUC18, and pBluescript II Vectors

Chlorogonium elongatum Dangeard strain 2e was obtained from the Sammlung von Algenkulturen (Göttingen, Germany). Cell cultures, isolation of organellar (i.e., plastid and mitochondrial) DNA, and cloning procedures have been described previously (Kroymann and Zetsche 1997).

Identification of Mitochondrial Clones and PCR

Recombinant plasmids containing mtDNA inserts were identified by clone walking, starting with heterologous *cob* (Schuster and Brennicke 1985) and *cox1* (Kessler and Zetsche 1995) gene probes kindly provided by Axel Brennicke, Ulm, and Ulrike Kessler, Giessen, Germany. A 2.45-kb *Eco*RI fragment encoding the second exon and parts of both introns of the *cox1* gene was missing in the clone banks. This one was amplified via PCR (Mullis and Faloona 1987) and cloned into the pCRII vector using the TA cloning kit (Invitrogen, Leek, The Netherlands). Three fragments were obtained from independent reactions.

Blotting and Hybridization

Restriction fragments were electrophoretically separated on 0.8% agarose gels and blotted onto Hybond N+ membranes (Amersham) (Southern 1975). Hybridizations were performed nonradioactively using the ECL kit RPN 3001 (Amersham).

Sequencing and Data Analysis

DNA sequencing of double-stranded plasmid templates was performed by the dideoxy chain termination method (Sanger et al. 1977) using the T7 sequencing kit (Pharmacia) and universal primers or, when necessary, various synthetic oligonucleotide primers (Genosys Biotechnologies, England). Compressions were resolved with Deaza G/A sequencing mixes (Pharmacia). Sequences were read manually. Computer analyses and similarity searches were performed with DNASIS and PROSIS (Hitachi Software, Yokohama, Japan). Furthermore, the BLAST network service (Altschul et al. 1990) was used. Similarity scores for intronic open reading frames (ORFs) were calculated from pairwise comparisons of sequences aligned with Clustal V (Higgins et al. 1992). Secondary structures were drawn with CARD (Winnepenninckx et al. 1995). The complete sequence of the mtDNA of *C. elongatum* is deposited in the EMBL database (accession Nos. Y07814, Y13643, Y13644).

Results

Gene Content

Seven genes coding for subunits of the mitochondrial electron transport chain were identified on the *Chlorogonium* mt genome (Fig. 1). These specify subunits 1, 2, 4, 5, and 6 of the NADH dehydrogenase complex (*nad1, −2, −4, −5,* and *−6*), apocytochrome b (*cob*), and subunit 1 of the cytochrome oxidase complex (*cox1*). Furthermore, split genes for the SSU and LSU rRNAs (see below) and genes encoding the transfer RNAs for methionine (CAU), tryptophan (CCA), and glutamine

Table 1. Insertion sites of group-I introns[a]

Ce	Introns from other species with identical insertion sites		
*nad5*i1 ID/L	*Allomyces macrogynus nad5*i1 ns/L (U41288)		
*nad5*i2 IB/L	*A. macrogynus nad5*i2, ns/L (U41288)	*Marchantia polymorpha nad5*i1, ns/— (M68929)	*Neurospora crassa nad5*i2 IB4[b]/L (X05115)
*cob*i1 ID/G	*N. crassa cob*i1, ID[b]/G (X06884)	*Podospora anserina cob*i2 ID[b]/G (X55026)	
*cob*i2 ID/L	*Saccharomyces cerevisiae cob*i2, ID[b]/L (J01472)	*A. macrogynus cob*i3 ID[c]/L (U41288)	
*coxl*i1 IB2/L	*P. anserina cox1*i11, IB2[b]/L (Y00403)	*A. macrogynus cox1*i8 ns/L (U41288)	
*cox1*i2 IB/G	*M. polymorpha cox1*i9, IB[d]/— (M68929)	*A. macrogynus cox1*i12 ns/— (U41288)	

[a] Intron subgroups, ORF types, and accession numbers (in parentheses) of group-I introns in mitochondrial genes with insertion sites shared by *C. elongatum (Ce)* and other species. G, an intronic ORF of the GIY–YIG type is encoded; L, the intron encodes an ORF exhibiting a LAGLI–DADG motif or a degenerate version of it; —, no intronic ORF. ns, intron subtype not specified.[b] Data from Michel and Westhof (1990).[c] Data from Kroymann and Zetsche (1997).[d] Data from Ohta et al. (1993).

(UUG) were identified. Remnants of other mitochondrial genes or of an *rtl* gene were not detected.

Genome Structure

The map of *Chlorogonium* mtDNA (Fig. 1) was deduced from the complete nucleotide sequence as obtained from the analysis of overlapping mtDNA fragments. To determine whether a linear form of the mtDNA exists in addition to the circular one, hybridization experiments were performed: both *Bam*HI and *Xba*I have single restriction sites in the mtDNA from *C. elongatum* (Fig. 1). In Southern blots from restricted and electrophoretically separated organellar DNA probed with fragments from different sites of the mtDNA, only a single DNA species of about 23 kb could be detected in these digests (data not shown). Furthermore, no regions homologous to the *C. reinhardtii* inverted repeats (Ma et al. 1992; Vahrenholz et al. 1993) were found. These results are consistent with a circular conformation, as shown in Fig. 1.

Introns and Intron-Encoded ORFs

Six introns displaying the structural features of group-I introns (Burke 1988; Cech 1988; Michel and Westhof 1990) were found, two each in the *cox1, nad5,* and *cob* genes. Both introns in the *cob* gene were placed in subgroup ID (Kroymann and Zetsche 1997), as well as the first intron in the *nad5* gene. The first intron in the *cox1* gene belongs to subgroup IB2. The others are most similar to introns from subgroup IB but could not be specified unambigously. The intron insertion sites are shared by various group-I introns from other species (Table 1). In several cases, these introns belong to the same or a closely related structural subgroup as the respective *Chlorogonium* introns: *Saccharomyces cerevisiae cob* intron 2, as well as *Neurospora crassa cob* intron 1 and *Podospora anserina cob* intron 2, which are inserted at the same site as *Chlorogonium cob* introns 1 and 2, respectively, belongs to subgroup ID. *N. crassa nad5* intron 2, which shares its insertion site with *Chlorogonium nad5* intron 2, was placed in subgroup IB4, and *P. anserina* cox1 intron 11, inserted at the same position as *Chlorogonium cox1* intron 1, belongs to subgroup IB2 (Michel and Westhof 1990). Furthermore, *Marchantia polymorpha cox1* intron 9, which shares its insertion site with *Chlorogonium cox1* intron 2, was classified as a group IB intron (Ohta et al. 1993).

Each *Chlorogonium* intron contains a single ORF potentially encoding proteins related to RNA maturases or endonucleases. Two of these (*cob*i1 and *cox1*i2) are of the GIY–YIG type (Burger and Werner 1985; Michel and Cummings 1985; Kroymann and Zetsche 1997); the others exhibit highly degenerate LAGLI–DADG motifs (Dujon 1980; Lazowska et al. 1980). In general, the intronic ORFs of both subtypes appear to be only distantly related to one another. Only the ORFs in *Chlorogonium cob* intron 2 and *nad5* intron 2 show an identity of approx. 25% and a similarity of approx. 45%; in all other cases, pairwise comparisons reveal identity values well below 25%. BLAST (Altschul et al. 1990) searches were performed with all intron-encoded ORFs; in all cases, the highest similarity scores were obtained with ORFs encoded by introns inserted at various locations of mitochondrial genes.

Regarding the ORFs that are encoded by the introns listed in Table 1, they belong to the same family as the respective *Chlorogonium* ORFs. Relatively high degrees of similarity were observed between the ORFs encoded by the *Chlorogonium cob* genes and their fungal counterparts, as well as between the second intronic ORF in the *Chlorogonium* and the sole intronic ORF in the *Chlamydomonas smithii cob* genes (cf. Kroymann and

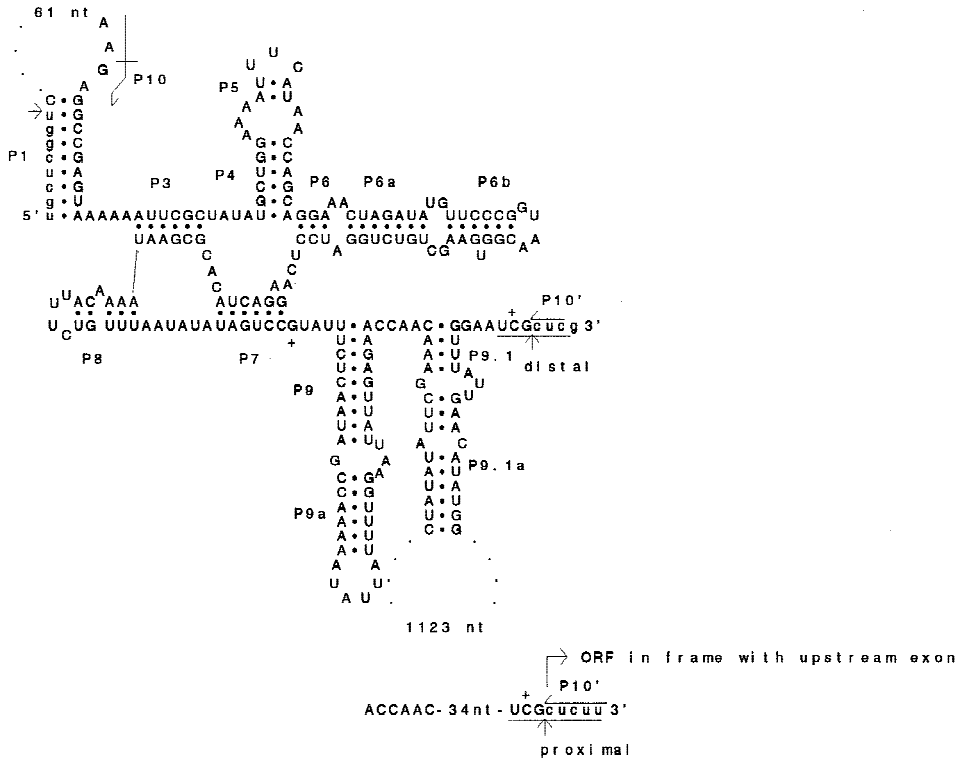

Fig. 2. Potential secondary structures of the second intron in the *C. elongatum cox1* gene drawn according to Burke et al. (1987). Intron sequences are shown in *capital letters;* exon sequences, in *lowercase.* Secondary structure elements are indicated; the P9.0 interaction is marked by *crosses.* The putative proximal 3′ splice site as well as the intron-encoded ORF lies within the loop of P9.1a. The alternate secondary structures are identical up to the sequence ''ACCAAC.'' Elements P9.1 and P9.1a are omitted when the proximal site is used for splicing. In this case, the P10–P10′ interaction consists of 5 bp (P10, 5′AAGAG3′; P10′, 3′uucuc5′). Alternatively, i.e., when the distal splice site is used, only 3 bp are involved (P10, 5′GAG3′; P10′, 3′cuc5′).

Zetsche 1997). In addition, alignment of the ORFs encoded in *nad5* intron 2 of the fungus *Allomyces macrogynus* (Paquin and Lang 1996) and *Chlorogonium nad5* intron 2 reveals that they share 104 identical residues in 285 positions, and a further 68 amino acids result from conservative substitutions (36% identity/60% similarity). The ORF potentially encoded by the *Neurospora crassa* intron (Nelson and Macino 1987) inserted at the same position shows no significant similarity to either the *Chlorogonium* or the *Allomyces* intronic ORF.

In the other cases, no significant similarity was detected. One intron (*cox1*i2) putatively possesses two alternate 3′ splice sites (Cummings et al. 1988). The use of the proximal site would fuse the intron-encoded ORF in phase to the preceding exon; alternatively, splicing at the distal site would result in excision of the intron from the precursor RNA (Fig. 2).

rRNA Genes

The rRNA genes were identified according to their similarity at the primary and secondary structure level with their counterparts from *C. reinhardtii* (Boer and Gray 1988b) and *C. eugametos* (Denovan-Wright and Lee 1994). The exact 5′ and 3′ termini were not determined. The SSU rRNA gene is split into (at least) three pieces, S1, S2, and S3; the LSU rRNA gene consists of (at least) six fragments, L1 to L6, numbered with respect to their order in continuous rRNA genes. The breakpoints lie within regions corresponding to variable domains of the rRNAs. Altogether, the SSU and LSU rDNA fragments span a region of about 5.1 kb on the mt genome from *C. elongatum.* They are intermingled with each other and with the *nad6* and tRNA-M genes in the following order: L3, L2, *nad6,* L6, S2, L4, S1, L1, M, L5, S3.

GC-Rich Elements

The average A + T content of the *C. elongatum* mt genome is 62.2%. However, many sequence stretches displaying an unusually high G + C content were found in intergenic regions, in introns and intron-encoded ORFs, and in regions corresponding to variable domains of the rRNAs. These ''GC-rich elements'' vary in size, from a few bases up to ca. 150 bp. Characteristic is the succession of repeated G or C residues, in some cases up to 8mers, but more often in blocks of 3mers, 4mers, or 5mers (Fig. 3). These may be part of shorter or longer palindromic or hairpin loop sequences, and might be able to form secondary structures. Based on sequence simi-

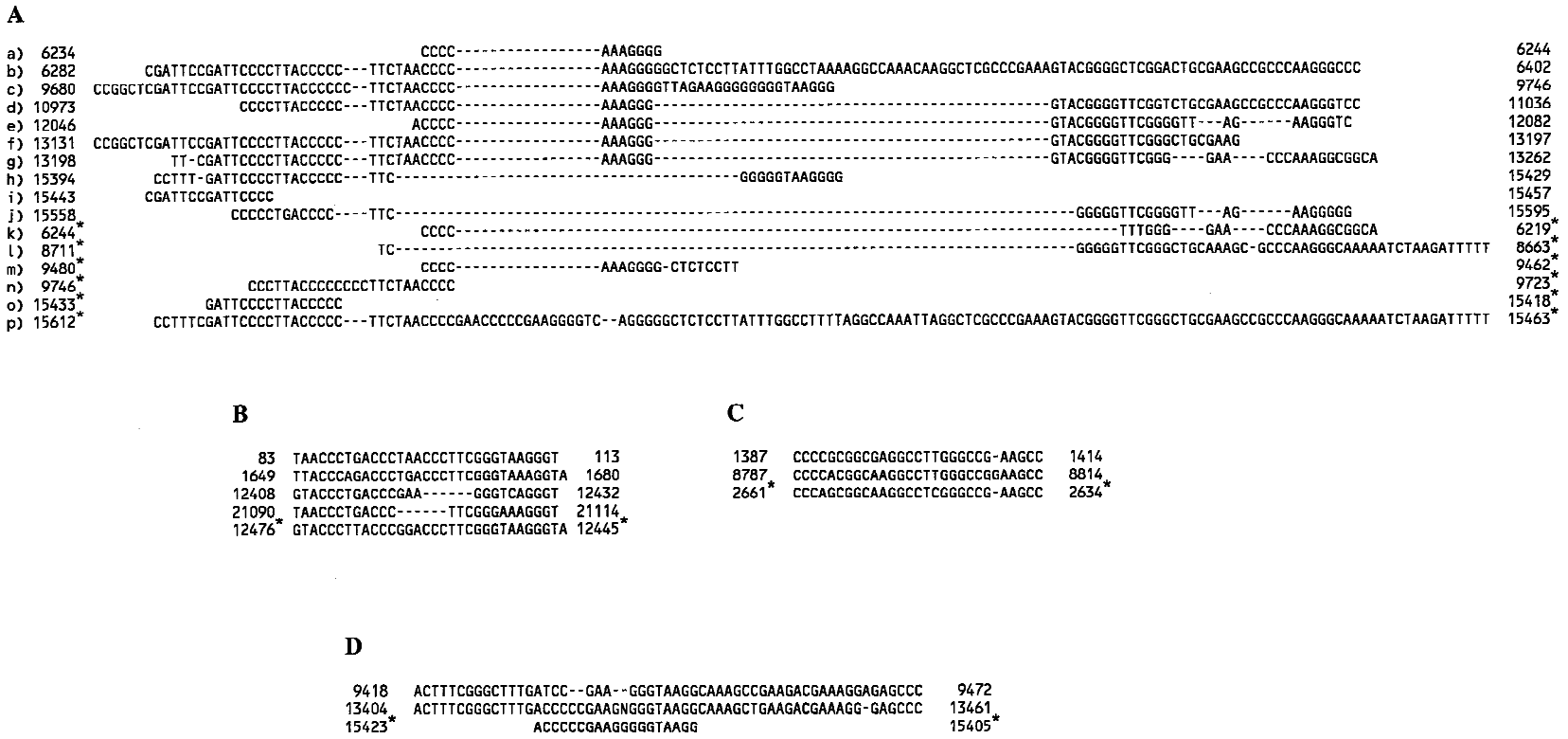

Fig. 3. GC-rich elements. Based on sequence similarities, GC-rich elements were divided into four groups (**A, B, C,** and **D**). Characteristic are blocks of repeated G and C residues. The elements of each group share portions of sequence identity. To optimize sequence identity, gaps (—) were introduced. Due to shorter or longer stretches with a palindromic or hairpin loop character, some of the elements are shown in both directions (elements shown in a counterclockwise direction are marked with an *asterisk*). *Numbers* indicate the first and last bases of the elements and refer to their position within the complete genomic sequence. Position 1 corresponds to the *Hin*cII site following the *nad2* gene (see Fig. 1).

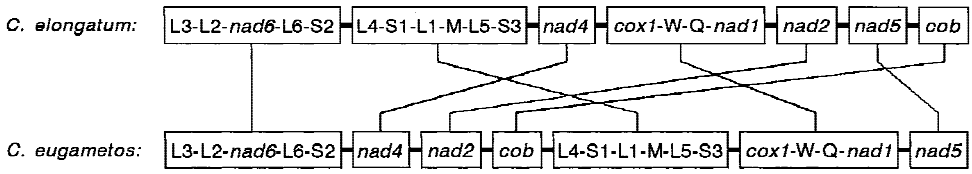

Fig. 4. Schematic representation of the *C. elongatum* and *C. eugametos* mitochondrial gene orders. To faciliate comparison, both genomes were linearized at the rRNA gene fragment L3. *Boxed regions* indicate genes and gene fragments in identical succession in both genomes.

larities, these elements were divided into four groups (Kroymann 1997).

Discussion

SSU and LSU rRNA, *nad1, −2, −4, −5,* and *−6, cob,* and *cox1* genes were found to be encoded in *C. elongatum* mtDNA as well as in *C. reinhardtii*/*C. smithii* and *C. eugametos*/*C. moewusii* mtDNAs. Only three genes specify tRNAs, i.e., for methionine, tryptophan, and glutamine. Apart from the *rtl* gene, an identical set of mitochondrially encoded genes is shared by *C. elongatum, C. eugametos*/*C. moewusii,* and *C. reinhardtii*/*C. smithii.* This suggests that already their common ancestor possessed a highly reduced coding capacity, compared to other green algae, e.g., *Tetraselmis subcordiformis* (Kessler and Zetsche 1995) and *Prototheca wickerhamii* (Wolff et al. 1994).

Another feature typical of chlamydomonad mt genomes is the presence of fragmented SSU and LSU rRNA genes. Some of the rRNA fragment breakpoints are shared between *C. reinhardtii* (Boer and Gray 1988b) and *C. eugametos* (Denovan-Wright and Lee 1994); others are not. The simplest explanation for this is that common breakpoints were present in the progenitor of both species, while unique ones resulted from rearrangements after the separation of lineages (Denovan-Wright et al. 1996).

The breakpoints of mitochondrial rRNA gene fragments from *C. elongatum* appear to correspond to those from *C. eugametos.* Furthermore, the fragment order is identical, with the following exception: in *C. eugametos* the rDNA fragments are dispersed in two distinct clusters separated by several kilobase pairs (Denovan-Wright and Lee 1993), whereas in *C. elongatum* they are located in a single region of about 5.1 kb.

Several features are specifically shared by the mt genomes from *C. elongatum* and *C. eugametos/C. moewusii* (Lee et al. 1991; Denovan-Wright and Lee 1992, 1993, 1994), i.e., a circular genome map, the absence of an *rtl* gene, and apparently identical breakpoints of the rRNA gene fragments. A circular genome structure was also reported for the mtDNA from *Chlamydomonas pitschmannii,* a species very closely related to *C. eugametos*/*C. moewusii* (Boudreau and Turmel 1995). In *C. elongatum* and *C. eugametos*/*C. moewusii,* all mitochondrial genes are encoded by the same DNA strand. Furthermore, in three regions, ranging from L3 to S2, from L4 to S3, and from *cox1* to *nad1,* the gene order is exactly the same (Fig. 4). In contrast, *nad5, nad4,* and *cob* are transcribed from one DNA strand of the *C. reinhardtii* mt genome, whereas the remaining genes are encoded by the complementary strand (Michaelis et al. 1990; Harris 1993). Except for two rDNA modules (Denovan-Wright et al. 1996), the gene order in *C. reinhardtii* mtDNA completely differs from both the *C. eugametos* and the *C. elongatum* mt genome order. Therefore, we conclude that *C. elongatum* and *C. eugametos*/*C. moewusii* are more closely related to each other than either of these species is to *C. reinhardtii*/*C. smithii*. This is consistent with recently published results of a phylogenetic analysis based on nuclear and plastid rRNA gene sequences (Buchheim et al. 1996). Furthermore, our results strengthen the view that the genus *Chlamydomonas* is a polyphyletic assemblage of only distantly related species.

At least three transpositions are required to transform the *C. elongatum* genome order into the *C. eugametos* order, or vice versa. However, this could be achieved in various ways, and current data concerning mt genomes in the *C. elongatum*/*C. eugametos* lineage of chlamydomonads do not allow one to determine the succession of events that led to the different genome arrangements in these species. Notably, transpositions retaining the direction of genes are obviously preferred to rearrangements resulting in a ''head-to-head'' or ''tail-to-tail'' order of genes: starting from a gene order as in *C. elongatum* or *C. eugametos,* three—independent—transpositions lead in only one of eight imaginable cases to a gene order in which all genes are arranged ''head to tail.'' Provided that both directions of a gene (with respect to the others) could be realized with equal opportunity following a transpositional event, the probability of such a case is 1/8, or 12.5%.

Thus, the *Chlorogonium* and *C. eugametos* mt genome organization, i.e., with all genes and gene fragments arranged in a head-to-tail order, might be due to properties of the mechanism(s) responsible for the arrangements, but could also reflect some special properties of the genome itself, e.g., the presence of a single promoter, from which all genes in *C. elongatum* and *C. eugametos* are transcribed. As yet, such an element has not been identified, in either *C. elongatum* or *C. eugametos,* although mitochondrial mRNAs were detected in

C. elongatum (Kroymann et al. 1995; our unpublished results).

So far, the character of the mechanism(s) underlying the rearrangements remains speculative. A potential role could be assigned to the GC-rich elements, which are present repetitively on the *Chlorogonium* mt genome (see below). However, an interaction of repetitive elements within a single circular mtDNA molecule does not seem to be an appropriate explanation: a recombination between two direct repeats gives rise to two subgenomic circles, each comprising only a part of the genome, whereas a recombination between two indirect repeats leads to an inversion of the intervening sequence.

Therefore, an interaction between two distinct mtDNA molecules might be envisaged: an event resembling unequal crossover and involving GC-rich elements located at different positions, which could lead to the transfer of an additional gene copy to one mtDNA molecule accompanied by a deletion in the other. In a transient stage, both copies would be active. If later the ''original'' gene copy is inactivated, a genome rearrangement results. This implies the presence of a pseudogene in the mt genome for some period of time. However, remnants of such a gene could not be detected in *Chlorogonium* mtDNA.

Furthermore, the proposed model suggests that single genes are transferred, i.e., *nad4, nad2, nad5,* and/or *cob,* rather than continuous stretches of several genes, i.e., L3 to S2, L4 to S3, and *cox1* to *nad1*. In the latter case it would be difficult to explain why these stretches are kept intact, i.e., why all genes and gene fragments should be inactivated in one copy, but none in the other.

The *rtl* gene present in *C. reinhardtii* (and *C. smithii*) mtDNA(s) is suggested to be of foreign origin, due to several characteristics that distinguish this gene from the other protein-coding genes (Boer and Gray 1988a). An *rtl* gene is not encoded in the mtDNAs of either *C. elongatum* or *C. eugametos*/*C. moewusii* (Denovan-Wright and Lee 1993; R.W. Lee, personal communication). Furthermore, no remnants of such a sequence could be detected in the *C. elongatum* mt genome. Therefore, it seems likely that this gene was introduced into the mtDNA in a chlamydomonad lineage leading to *C. reinhardtii*/*C. smithii* after separation from a chlamydomonad lineage leading to *C. elongatum* and *C. eugametos*/*C. moewusii*. Similarly, no remnants of the inverted repeats characteristic of the linear *C. reinhardtii* mt genome (Ma et al. 1992; Vahrenholz et al. 1993) were found in *C. elongatum*. Hence, the same explanation might apply to them. Vahrenholz et al. (1993) proposed a hypothesis explaining the functions of both the inverted repeats and the *rtl* gene in *C. reinhardtii* mtDNA: an internal 86-bp repeat of the two outermost sequences serves as a matrice for the generation of an RNA molecule, which binds complementarily to the 3′ single-stranded extensions and primes DNA synthesis during replication. The 3′ single-stranded extensions themselves are then generated by a reverse transcriptase using the RNA molecule as a template. Alternately, the ends of the mtDNA molecule could be generated by an interaction between the internal and the two terminal repeats. The latter model would require a site-specific nuclease, rather than a reverse transcriptase.

A linear mt genome flanked by inverted repeats has also been reported for the green alga *Pandorina morum* (Moore and Coleman 1989). In a phylogenetic analysis based on molecular, i.e., nuclear-encoded SSU and LSU rRNA sequences, and organismal data, this multicellular species turned out to be more closely related to *C. reinhardtii* and *C. smithii* than to *C. eugametos* and *C. moewusii* (Buchheim et al. 1994). Therefore, it seems possible that an *rtl* gene might be encoded in this mt genome, too.

Differences in the mt-genome sizes of chlamydomonad algae result mainly from varying numbers of introns. In *C. elongatum,* the six group-I introns cover approximately 8.1 kb, i.e., more than one-third of the total size of 22.7 kb. Group-I introns have also been detected in the *cob* gene from *C. smithii* (Colleaux et al. 1990) and in the protein-coding (Denovan-Wright and Lee 1993) and rRNA (Denovan-Wright and Lee 1994) genes from *C. eugametos*. The inverted repeats, together with an internal repeat of the 5′ and 3′ terminal sequences, contribute ca. 1.2 kb of the *C. reinhardtii* mt genome; noncoding regions additionally cover approximately 1.5 kb. Short repeated sequences, characterized by the consensus motif CTCGG ($N_{4–14}$) CCGAG, were detected in these spacer regions flanking transcribed coding sequences. Because of their location, a potential function in gene regulation was suggested (Boer and Gray 1991).

In *C. elongatum,* noncoding intergenic regions occupy about 2.4 kb of the mt genome. Here, numerous GC-rich elements were found. Since such sequences are also present in introns, intronic ORFs, and rDNA sequences corresponding to variable domains of the rRNAs, it seems unlikely that they are involved in gene expression or in the formation of higher-order structures of the mtDNA. Rather, their distribution suggests that they either are or were mobile within the mt genome. Similar putatively mobile elements have been found in various fungal mtDNAs, e.g., in *Saccharomyces cerevisiae, Neurospora crassa, Kluyveromyces lactis,* and *Aspergillus nidulans* (summarized by Clark-Walker 1992), as well as in *Allomyces macrogynus* (Paquin and Lang 1996). However, it is not clear whether these elements and those from *C. elongatum* share a common origin, since no particular sequence homology was observed.

The GIY–YIG ORF in *Chlorogonium cob* intron 1 was the first to be detected in a mt genome except those from fungi. It was found to be most similar to the GIY–YIG ORFs encoded by *Podospora anderina cob* intron 2 (Cummings et al. 1989) and by *Neurospora crassa cob*

intron 1 (Collins et al. 1988); the respective introns are inserted at identical sites within the *cob* genes of these species and belong to the same subgroup (ID). Therefore, a common ancestry of these introns was proposed (Kroymann and Zetsche 1997). A second GIY–YIG ORF is encoded in *Chlorogonium cox1* intron 2, which belongs to subgroup IB. This ORF shows only a low degree of similarity to the other one of the same type in *Chlorogonium* (20% identity/41% similarity). Hence lateral transfer of these GIY–YIG ORFs within the *Chlorogonium* mt genome is very unlikely. *Marchantia polymorpha cox1* intron 9 and *Allomyces macrogynus cox1* intron 12, which share identical insertion sites with *Chlorogonium cox1* intron 2, lack intronic ORFs.

The *Chlorogonium* introns encoding ORFs with a degenerate LAGLI–DADG motif belong to two subgroups, and the ORFs themselves are not closely related to one another. Therefore, lateral transfer of these introns within the genome can be ruled out in this case, also. Many introns from other organisms, which are inserted at identical positions as these *Chlorogonium* introns, belong to the same or a closely related structural subgroup and contain ORFs with similar properties, i.e., a degenerate LAGLI–DADG motif. Therefore, a common ancestry of these introns seems possible, although a comparably high degree of similarity was observed only for the LAGLI–DADG ORF encoded by the *Chlorogonium cob* gene and its counterparts (Kroymann and Zetsche 1997), as well as for those in *Chlorogonium nad5* intron 2 and *A. macrogynus nad5* intron 2.

Common ancestry of introns (and/or their ORFs) can result either from vertical transmission from a common ancestor organism or from lateral transfer between only distantly related taxa. A high degree of similarity of both introns and ORFs should be expected when lateral transfer between taxa, i.e., a fungus and a green alga, does not date back too far in the past. In contrast, little similarity should be observable in two cases: (1) when introns in different taxa are vertically inherited from a common ancestor, and the respective taxa evolve independently from one another for a long period of time, and (2) when a lateral transfer dates far back in the past.

Notably, Ohta et al. (1993) drew the conclusion that vertical transmission was the simplest explanation for the presence of cognate introns in liverwort and fungal mt genomes. This would imply that introns are ancestral within the green algal/plant lineage of evolution. The detection of introns from a variety of green algal mitochondria, i.e., those of *Chlamydomonas smithii* (Colleaux et al. 1990), *Scenedesmus obliquus* (Kück et al. 1990), *Prototheca wickerhamii* (Wolff et al. 1993), and *Chlamydomonas eugametos* (Denovan-Wright and Lee 1993), and now, from *Chlorogonium*, supports this hypothesis. However, further data, on both chlamydomonad mt genomes and those from other green algae, are required to decide which possibility explains the presence of the *Chlorogonium* group-I introns and their ORFs.

Note Added in Proof

Meanwhile, the complete sequence of the *C. eugametos* mtDNA has been published. In this genome an additional copy for $tRNA^{met}$ is present that may be a pseudogene. Furthermore, in total nine group-I introns were detected. For further information, please refer to: Denovan-Wright EM, Nedelcu AM, Lee RW (1998) Complete sequence of the mitochondrial DNA of *Chlamydomonas eugametos.* Plant Mol Biol 36:285–295.

Acknowledgments. We thank M.W. Gray and M.N. Schnare for careful correction of the rRNA secondary structure models and for helpful comments, B. Paquin for detecting of the proximal 3′ intron splice site and helpful discussion on introns and intronic ORFs, and K. Valentin for critical reading of the manuscript. This work was supported by grants from the Deutsche Forschungsgemeinschaft (Ze 71/24-2/3).

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Bennoun P, Delosme M, Kück U (1991) Mitochondrial genetics of *Chlamydomonas reinhardtii:* resistance mutations marking the cytochrome b gene. Genetics 127:335–343

Boer PH, Gray MW (1986a) Nucleotide sequence of a protein coding region in *Chlamydomonas reinhardtii* mitochondrial DNA. Nucleic Acids Res 14:7506–7507

Boer PH, Gray MW (1986b) The *URF 5* gene of *Chlamydomonas reinhardtii* mitochondria: DNA sequence and mode of transcription. EMBO J 5:21–28

Boer PH, Gray MW (1988a) Genes encoding a subunit of respiratory NADH dehydrogenase (*ND1*) and a reverse transcriptase-like protein (*RTL*) are linked to ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. EMBO J 7:3501–3508

Boer PH, Gray MW (1988b) Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. Cell 55:399–411

Boer PH, Gray MW (1988c) Transfer RNA genes and the genetic code in *Chlamydomonas reinhardtii* mitochondria. Curr Genet 14:583–590

Boer PH, Gray MW (1989) Nucleotide sequence of a region encoding subunit 6 of NADH dehydrogenase (*ND6*) and $tRNA^{Trp}$ in *Chlamydomonas reinhardtii* mitochondrial DNA. Nucleic Acids Res 17: 3993

Boer PH, Gray MW (1991) Short dispersed repeats located in spacer regions of *Chlamydomonas reinhardtii* mitochondrial DNA. Curr Genet 19:309–312

Boer PH, Bonen L, Lee RW, Gray MW (1985) Genes for respiratory chain proteins and ribosomal RNAs are present on a 16-kilobase-pair DNA species from *Chlamydomonas reinhardtii* mitochondria. Proc Natl Acad Sci USA 82:3340–3344

Boudreau E, Turmel M (1995) Gene rearrangements in *Chlamydomonas* chloroplast DNAs are accounted for by inversions and by the expansion/contraction of the inverted repeat. Plant Mol Biol 27: 351–364

Buchheim MA, Chapman RL (1991) Phylogeny of the colonial green flagellates: a study of 18S and 26S rRNA sequence data. BioSystems 25:85–100

Buchheim MA, Turmel M, Zimmer EA, Chapman RL (1990) Phylog-

eny of *Chlamydomonas* (Chlorophyta) based on cladistic analysis of nuclear 18S rRNA sequence data. J Phycol 26:689–699

Buchheim MA, McAuley MA, Zimmer EA, Theriot EC, Chapman RL (1994) Multiple origins of colonial green flagellates from unicells: evidence from molecular and organismal characters. Mol Phylogenet Evol 3:322–343

Buchheim MA, Lemieux C, Otis C, Gutell RR, Chapman RL, Turmel M (1996) Phylogeny of the Chlamydomonadales (Chlorophyceae): a comparison of ribosomal RNA gene sequences from the nucleus and the chloroplast. Mol Phylogenet Evol 5:391–402

Burger G, Werner S (1985) The mitochondrial URF1 gene in *Neurospora crassa* has an intron that contains a novel type of URF. J Mol Biol 245:522–537

Burke JM (1988) Molecular genetics of group-I introns: RNA structure and protein factors required for splicing—a review. Gene 73:273–294

Burke JM, Belfort M, Cech TR, Davis RW, Schweyen RJ, Shub DA, Szostak JW, Tabak HF (1987) Structural conventions for group I introns. Nucleic Acids Res 15:7217–7221

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

Clark-Walker GD (1992) Evolution of mitochondrial genomes in fungi. In: Wolstenholme DR, Jeon KW (eds) Mitochondrial genomes. International review of cytology 141. Academic Press, New York, pp 89–127

Colleaux L, Michel-Wolwertz MR, Matagne RF, Dujon B (1990) The apocytochrome b gene of *Chlamydomonas smithii* contains a mobile intron related to both *Saccharomyces* and *Neurospora* introns. Mol Gen Genet 223:288–296

Collins RA, Reynolds CA, Olive J (1988) The self-splicing intron in the *Neurospora* apocytochrome b gene contains a long reading frame in-frame with the upstream exon. Nucleic Acids Res 16: 1125–1134

Cummings DJ, Domenico JM, Michel F (1988) DNA sequence and organization of the mitochondrial *ND1* gene from *Podospora anserina:* analysis of alternate splice sites. Curr Genet 14:253–264

Cummings DJ, Michel F, McNally KL (1989) DNA sequence analysis of the apocytochrome b gene of *Podospora anserina:* a new family of intronic open reading frames. Curr Genet 16:407–418

Denovan-Wright EM, Lee RW (1992) Comparative analysis of the mitochondrial genomes of *Chlamydomonas eugametos* and *Chlamydomonas moewusii.* Curr Genet 21:197–202

Denovan-Wright EM, Lee RW (1993) *Chlamydomonas eugametos* mitochondrial genome. In: O'Brien SJ (ed) Genetic maps. Locus maps of complex genomes, 6th ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 2.170–2.171

Denovan-Wright EM, Lee RW (1994) Comparative structure and genomic organization of the discontinuous ribosomal RNA genes of *Chlamydomonas eugametos* and *Chlamydomonas reinhardtii.* J Mol Biol 241:298–311

Denovan-Wright EM, Sankoff D, Spencer DF, Lee RF (1996) Evolution of fragmented mitochondrial ribosomal RNA genes in *Chlamydomonas.* J Mol Evol 42:382–391

Dujon B (1980) Sequence of the intron and flanking exons of the mitochondrial 21 S rRNA gene of yeast strains having different alleles at the omega and rib-1 loci. Cell 20:185–197

Harris EH (1993) *Chlamydomonas reinhardtii* mitochondrial genome. In: O'Brien SJ (ed) Genetic maps. Locus maps of complex genomes, 6th ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 2.168–2.169

Higgins DG, Bleasby AJ, Fuchs R (1992) Clustal V: improved software for multiple sequence alignment. Comput Appl Biosci 8:189–191

Kessler U, Zetsche K (1995) Physical map and gene organization of the mitochondrial genome from the unicellular green alga *Platymonas* (*Tetraselmis*) *subcordiformis* (Prasinophyceae). Plant Mol Biol 29: 1081–1086

Kroymann J (1997) Das Mitochondriengenom der einzelligen Grünalge *Chlorogonium elongatum.* Neue Erkenntnisse zur Evolution der Mitochondriengenome in Chlamydomonadaceen, PhD thesis, Universität Giessen. Tectum Verlag, Marburg, Germany.

Kroymann J, Zetsche K (1997) The apocytochrome b gene in *Chlorogonium elongatum* (*Chlamydomonadaceae*): an intronic GIY-YIG ORF in green algal mitochondria. Curr Genet 31:414–418

Kroymann J, Schneider W, Zetsche K (1995) Opposite regulation of the copy number and the expression of plastid and mitochondrial genes by light and acetate in the green flagellate *Chlorogonium.* Plant Physiol 108:1641–1646

Kück U, Godehardt I, Schmidt U (1990) A self-splicing group-II intron in the mitochondrial large subunit rRNA (LSUrRNA) gene of the eukaryotic alga *Scenedesmus obliquus.* Nucleic Acids Res 18: 2691–2697

Lazowska J, Jacq C, Slonimski PP (1980) Sequence of introns and flanking exons in wild-type and *box3* mutants of cytochrome *b* reveals an interlaced splicing protein coded by an intron. Cell 22: 333–348

Lee RW, Dumas C, Lemieux C, Turmel M (1991) Cloning and characterization of the *Chlamydomonas moewusii* mitochondrial genome. Mol Gen Genet 231:53–58

Ma DP, Yang YW, Hasnain SE (1988) Nucleotide sequence of *Chlamydomonas reinhardtii* mitochondrial genes coding for subunit 6 of NADH dehydrogenase and $tRNA^{Trp}$. Nucleic Acids Res 16:11373–11373

Ma DP, Yang YW, Hasnain SE (1989a) Nucleotide sequence of *Chlamydomonas reinhardtii* mitochondrial genes coding for $tRNA^{Gln}$ (UUG) and $tRNA^{Met}$ (CAU). Nucleic Acids Res 17:1256

Ma DP, Yang YW, King YT, Hasnain SE (1989b) Nucleotide sequence of cloned *nad4* (*urf4*) gene from *Chlamydomonas reinhardtii* mitochondrial DNA. Gene 85:363–370

Ma DP, Yang YW, King YT, Hasnain SE (1990) The mitochondrial apocytochrome b gene from *Chlamydomonas reinhardtii.* Plant Mol Biol 15:357–359

Ma DP, King YT, Kim Y, Luckett WS, Boyle JA, Chang YF (1992) Amplification and characterization of an inverted repeat from the *Chlamydomonas reinhardtii* mitochondrial genome. Gene 119: 253–257

Michaelis G, Vahrenholz C, Pratje E (1990) Mitochondrial DNA of *Chlamydomonas reinhardtii*: the gene for apocytochrome b and the complete functional map of the 15.8 kb DNA. Mol Gen Genet 223:211–216

Michel F, Cummings DJ (1985) Analysis of class I introns in a mitochondrial plasmid associated with senescence of *Podospora anserina* reveals extraordinary resemblance to the *Tetrahymena* ribosomal intron. Curr Genet 10:69–79

Michel F, Westhof E (1990) Modelling of the three-dimensional architecture of group-I catalytic introns based on comparative sequence analysis. J Mol Biol 216:585–610

Moore LJ, Coleman AW (1989) The linear 20 kb mitochondrial genome of *Pandorina morum* (Volvocaceae, Chlorophyta). Plant Mol Biol 13:459–465

Mullis KB, Faloona FA (1987) Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Methods Enzymol 155: 335–350

Nelson MA, Macino G (1987) Structure and expression of the overlapping *ND4L* and *ND5* genes of *Neurospora crassa* mitochondria. Mol Gen Genet 206:307–317

Ohta E, Oda K, Yamato K, Nakamura Y, Takemura M, Nozato N, Akashi K, Ohyama K, Michel F (1993) Group I introns in the liverwort mitochondrial genome: the gene coding for subunit 1 of cytochrome oxidase shares five intron positions with its fungal counterparts. Nucleic Acids Res 21:1297–1305

Paquin B, Lang BF (1996) The mitochondrial DNA of *Allomyces macrogynus*: the complete genomic sequence from an ancestral fungus. J Mol Biol 255:688–701

Pratje E, Schnierer S, Dujon B (1984) Mitochondrial DNA of *Chlam-*

ydomonas reinhardtii: the DNA sequence of a region showing homology with mammalian *URF2*. Curr Genet 9:75–82

Pratje E, Vahrenholz C, Buhler S, Michaelis G (1989) Mitochondrial DNA of *Chlamydomonas reinhardtii*: the *ND4* gene encoding a subunit of NADH dehydrogenase. Curr Genet 16:61–64

Ryan R, Grant D, Chiang KS, Swift H (1978) Isolation and characterization of mitochondrial DNA from *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA 75:3268–3272

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

Schuster W, Brennicke A (1985) TGA-termination codon in the apocytochrome b gene from *Oenothera* mitochondria. Curr Genet 9: 157–163

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98: 503–517

Vahrenholz C, Pratje E, Michaelis G, Dujon B (1985) Mitochondrial DNA of *Chlamydomonas reinhardtii*: Sequence and arrangement of *URF5* and the gene for cytochrome oxidase subunit I. Mol Gen Genet 201:213–224

Vahrenholz C, Riemen G, Pratje E, Dujon B, Michaelis G (1993) Mitochondrial DNA of *Chlamydomonas reinhardtii*: the structure of the ends of the linear 15.8-kb genome suggests mechanisms for DNA replication. Curr Genet 24:241–247

Winnepenninckx B, Van de Peer Y, Backeljau T, De Wachter R (1995) CARD: a drawing tool for nucleic acid secondary structures. Biotechniques 18:1060–1063

Wolff G, Burger G, Lang BF, Kück U (1993) Mitochondrial genes in the colourless alga *Prototheca wickerhamii* resemble plant genes in their exons but fungal genes in their introns. Nucleic Acids Res 21:719–726

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