Nucleic Acids Research, 1993, Vol. 21, No. 3 719–726

Mitochondrial genes in the colourless alga *Prototheca wickerhamii* resemble plant genes in their exons but fungal genes in their introns

Gabriele Wolff[1,2], Gertraud Burger[2], B.Franz Lang[2] and Ulrich Kück[1,*]
[1]Lehrstuhl für Allgemeine Botanik, Ruhr-Universität Bochum, Postfach 102148, D-4630 Bochum 1, Germany and [2]Département de Biochimie, Université de Montréal, C.P. 6128, Succ. A, Montréal H3C 3J7, Canada

Received August 24, 1992; Revised and Accepted January 6, 1993 EMBL accession nos X68721 and X68722

ABSTRACT

The mitochondrial DNA from the colourless alga *Prototheca wickerhamii* contains two mosaic genes as was revealed from complete sequencing of the circular extranuclear genome. The genes for the large subunit of the ribosomal RNA (LSUrRNA) as well as for subunit I of the cytochrome oxidase (coxI) carry two and three intronic sequences respectively. On the basis of their canonical nucleotide sequences they can be classified as group I introns. Phylogenetic comparisons of the coxI protein sequences allow us to conclude that the *P.wickerhamii* mtDNA is much closer related to higher plant mtDNAs than to those of the chlorophyte alga *C.reinhardtii*. The comparison of the intron sequences revealed several unusual features: (1) The *P.wickerhamii* introns are structurally related to mitochondrial introns from various ascomycetous fungi. (2) Phylogenetic analyses indicate a close relationship between fungal and algal intronic sequences. (3) The *P. wickerhamii* introns are located at positions within the structural genes which can be considered as preferred intron insertion sites in homologous mitochondrial genes from fungi or liverwort. In all cases, the sequences adjacent to the insertion sites are very well conserved over large evolutionary distances. Our finding of highly similar introns in fungi and algae is consistent with the idea that introns have already been present in the bacterial ancestors of present day mitochondria and evolved concomitantly with the organells.

INTRODUCTION

Two types of introns have been identified in mitochondrial genes. Their classification into group I and group II is based on canonical nucleotide sequence motifs and on conserved characteristics of the secondary structure potentially formed by the intron RNA (1, 2, 3). Numerous fungal species contain both groups of mitochondrial introns (reviewed by 2, 4) and the variation in intron number is considerable. The gene for subunit I of the cytochrome oxidase (coxI), for example, contains 16 introns in the filamentous fungus *Podospora anserina* and up to seven introns in various strains of the yeast *Saccharomyces cerevisiae*, but none in the common laboratory strain of *Neurospora crassa* (5, 6, 7, 8). So far, only few introns have been detected in mitochondria from angiosperms and gymnosperms and all belong to group II (9, 10, 11). However, numerous group I and group II introns were identified in mitochondria of the liverwort *Marchantia polymorpha* (12). The variability in the distribution of intron types and numbers is highly suggestive for an unorthodox way of their propagation.

Their particular site of insertion which is mainly located in well conserved regions of the genes is another peculiar feature of mitochondrial introns (13). Several introns are even inserted at identical positions of homologous genes of unrelated species, such as ascomycetous fungi and liverwort (12, 13, 14, 15).

Due to the conserved structure and the occurence of identical insertion sites of mitochondrial introns, it has been speculated that they may already have existed in the progenitor of fungal mitochondria (7). Other considerations have led to the suggestion that introns could have been transferred horizontally between different species (13, 16, 17, 18, 19).

Our current knowledge is too restricted to retrace the pathway of mitochondrial evolution among the main eukaryotic kingdoms on the one hand and to eubacteria on the other hand. Accordingly, we are still unable to define the origin of mitochondrial introns and their distribution throughout eukaryotes. We expect to find the missing information in the largely unexplored group of algae and protozoans. So far, only one green algal mitochondrial genome has been extensively characterized: that of *C.reinhardtii* has been almost completely sequenced (for a review see 20). Partial mitochondrial sequences of the green algae *Scenedesmus obliquus* and *Chlamydomonas smithii* have also been published (21, 22). The mtDNA from *C.reinhardtii* is strikingly unlike the plant or any other known mitochondrial genome from eukaryotes and there is little indication that *C.reinhardtii* and plants share a common mitochondrial ancestor (23, 24).

* To whom correspondence should be addressed

In order to extend our knowledge of the green algal/plant lineage, we recently started to analyze another chlorophyte mtDNA, namely that of the colourless alga *Prototheca wickerhamii*. Comparative analysis of the mitochondrial small ribosomal subunit RNA (mtSSUrRNA) indicates that the mtDNA from *P. wickerhamii* is very different to that of *C. reinhardtii* and considerably related to higher plants (25).

In this paper we will analyse intron sequences detected in the coxI and LSUrRNA genes of *P. wickerhamii* mitochondria which unexpectedly resemble fungal introns. The peculiarity of this result is supported by our phylogenetic analyses which place the *P. wickerhamii* mtDNA at the basis of higher plant mtDNAs and far from that of fungi and *C. reinhardtii*.

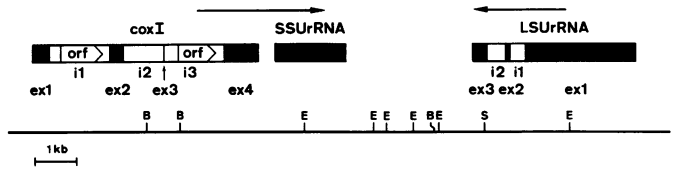

Figure 1. Organization of the coxI and LSUrRNA gene region of *Prototheca wickerhamii* mitochondria. Exons are indicated by black, introns by open bars. The direction of transcription is shown by large arrows. Abbreviations: ex, exon; i, intron; B, *Bam*HI; E, *Eco*RI; S, *Sal*I; coxI, subunit I of cytochrome c oxidase; LSUrRNA & SSUrRNA, large and small subunits of ribosomal RNA.

MATERIALS AND METHODS

Strains, vectors and gene libraries

Prototheca wickerhamii (strain 263-11) was obtained from the 'Sammlung von Algenkulturen, Gottingen, Germany' and grown as described (25). Lambda vectors EMBL3 and EMBL4

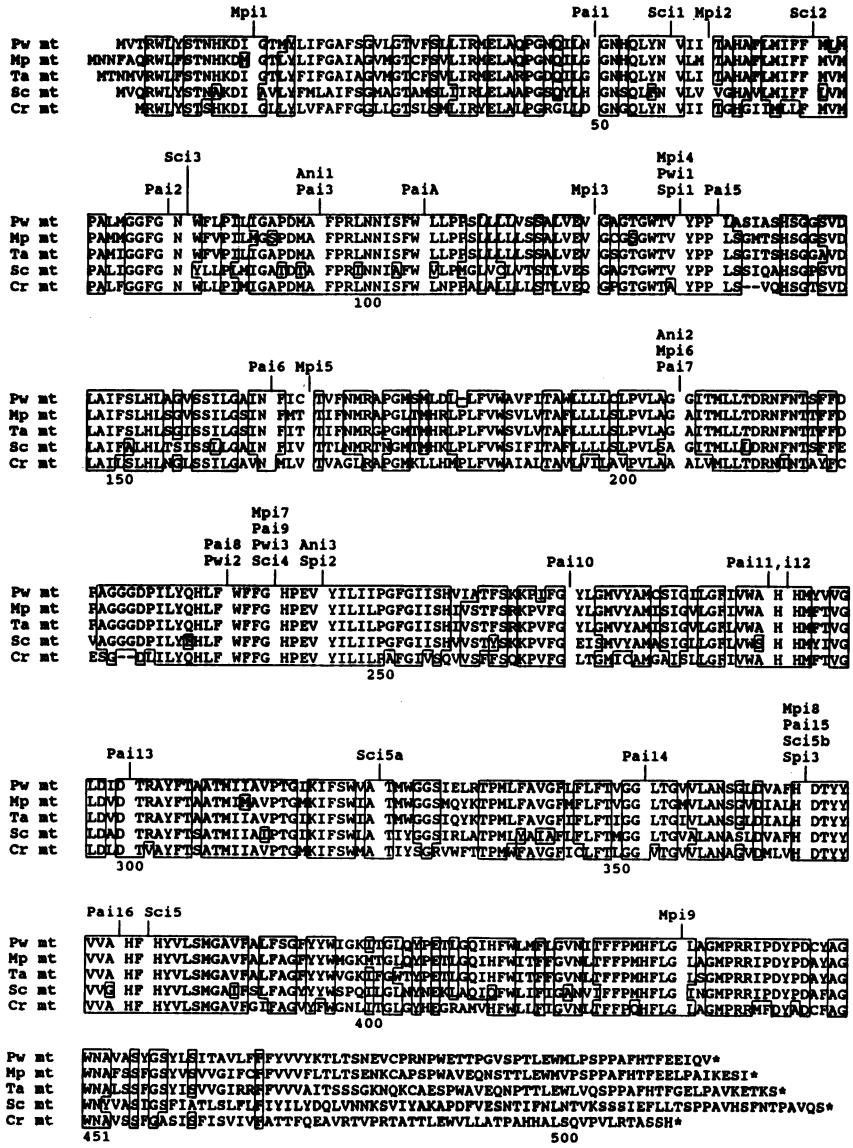

Figure 2. Alignment of the coxI proteins from algae and fungi. The amino acid sequences were deduced from the DNA sequences. Intron insertion sites are indicated. Abbreviations and references: Pw, *Prototheca wickerhamii*; Mp, *Marchantia polymorpha* (12); Ta, *Triticum aestivum* (65); Sc, *Saccharomyces cerevisiae* (66); Cr, *Chlamydomonas reinhardtii* (67); Pa, *Podospora anserina* (7); An, *Aspergillus nidulans* (14); Sp, *Schizosaccharomyces pombe* (68, 13, 69).

(Stratagene) were used for the cloning of partial *Sau*3AI and *Eco*RI digests of extranuclear DNA, enriched by CsCl centrifugation. Random restriction fragment libraries for sequencing were subcloned using vectors M13mp19 (26) and pBluescriptII KS+ (Stratagene).

The recombinant lambda-phage LPw2-13 was isolated from the EMBL4 gene bank by probing with the recombinant plasmid pGW1, which contains a 4.3 kbp *Hind*III fragment of *P.wickerhamii* mtDNA coding for the 3′ part of the coxI gene and the SSUrRNA gene (25). Screening of the EMBL3 gene bank with terminal *Eco*RI fragments from LPw2-13 identified the overlapping clone LPw3-1.1. The recombinant phages LPw2-13 and LPw3-1.1 contain *P. wickerhamii* mtDNA fragments of 13.5 and 9.6 kbp length, respectively. Possible rearrangements of the insert DNA were excluded by hybridizing subfragments of the lambda inserts with corresponding extranuclear DNA fragments (results not shown).

Isolation of nucleic acids, gel electrophoresis, hybridization conditions, oligonucleotide synthesis and standard in vitro recombinant techniques were carried out as described elsewhere (27, 28).

cDNA cloning. Total RNA from *P.wickerhamii* was isolated according to (28), and primed with oligonucleotide 164 (5′ CTCCAGTTAAACCACCTAC) followed by reverse transcription (27). The resulting cDNA was amplified by the polymerase chain reaction using oligonucleotide 164 and oligonucleotide 198 (5′ TGTATGGGCTGTATTTATTAC) as primers. The PCR fragments were cloned into M13mp19 and six independent clones were analysed by DNA sequencing.

DNA sequencing and computer analysis. The dideoxy chain termination method was used (29). High resolution polyacrylamide gels for long range reading were prepared according to (30). Computer analysis included homology searches with FASTA (31), Multalin (32) and programs developed by (32, unpublished data). Phylogenetic trees based on coxI amino acid sequences were established either with (i) a parsimony program (34) or (ii) the neighbour joining program of (35) using a distance matrix calculated with 'protdist' (Felsenstein, unpublished). Phylogenetic trees based on the nucleic acid sequences of introns inserted at the identical position as intron 1 of the LSUrRNA gene from *P.wickerhamii* were calculated with a maximum parsimony program (36), a maximum likelihood and a neighbour joining program (35). The results of the maximum parsimony program were submitted to a bootstrapping procedure (36).

RESULTS

The coxI gene of *Prototheca* contains three group I introns

Isolation of recombinant lambda phage LPw2-13 containing the coxI gene is described in the Material and Methods section. The complete sequencing of the recombinant phage revealed the gene organization of this mitochondrial DNA segment including the coxI gene as shown in Fig.1. The coxI gene of *P.wickerhamii* is 5369 nucleotides long and contains three intervening sequences at amino acid position 129, 235 and 239 splitting the amino acid coding region in four exons of 387, 319, 11, and 828 nucleotides, respectively (EMBL acc. no. X68721). The existence of the mini-exon 3 and the positions of the adjacent intron splice points were verified by cDNA sequencing. The deduced protein is 514 amino acids in length (Fig. 2).

All introns in the coxI gene of *Prototheca* contain the canonical P, Q, R, S sequence elements and can be folded into the conserved potential RNA secondary structure characteristic for group I introns (2, 37). The first and third intron (1451 nt and 1323 nt, respectively) each contain an open reading frame (ORF) for 342 and 277 amino acids in length, respectively. Intron 2 is shorter (1056 nt) and possesses no ORF. As its potential RNA folding is extraordinarily similar to that of the *Podospora anserina* coxI intron 8 (Pa ai8) (7), we analyzed its structure in detail. Fig. 4a shows that the *P.wickerhamii* ai2 carries all postulated helices (P1−P9) in the conserved core sequence and displays the characteristics of subgroup ID introns: a CUA sequence motif separates P6 and P7 and a UGUA motif occurs between P8 and P7′ (7, 15).

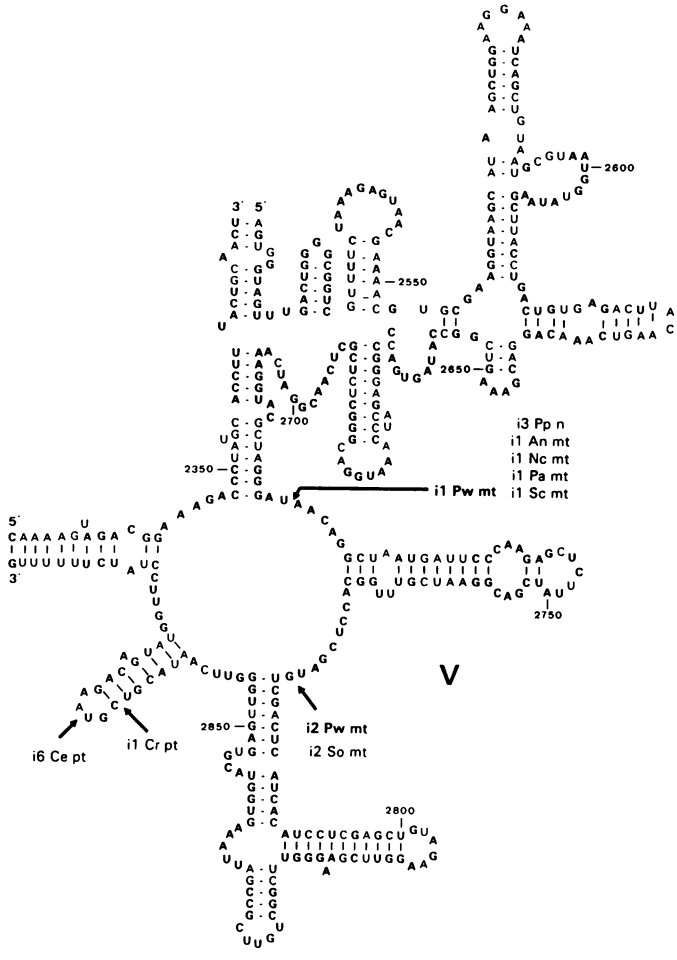

Figure 3. Secondary structure model of domain V of the LSUrRNA from *P.wickerhamii*, identical nucleotides compared to the corresponding sequence from *E.coli* (38) are printed in bold. Several group I intron insertion sites are indicated. The model was constructed according to (70) and the numbering of domains was performed following (71). Abbreviations and references: Ce, *Chlamydomonas eugametos* (72); Cr, *Chlamydomonas reinhardtii* (73); Nc, *Neurospora crassa* (40); Pa, *Podospora anserina* (74); Pp, *Physarum polycephalum* (75, 76); Pw, *Prototheca wickerhamii*; Sc, *Saccharomyces cerevisiae* (42); So, *Scenedesmus obliquus* (21); mt, mitochondrial; n, nuclear; pt, plastid.

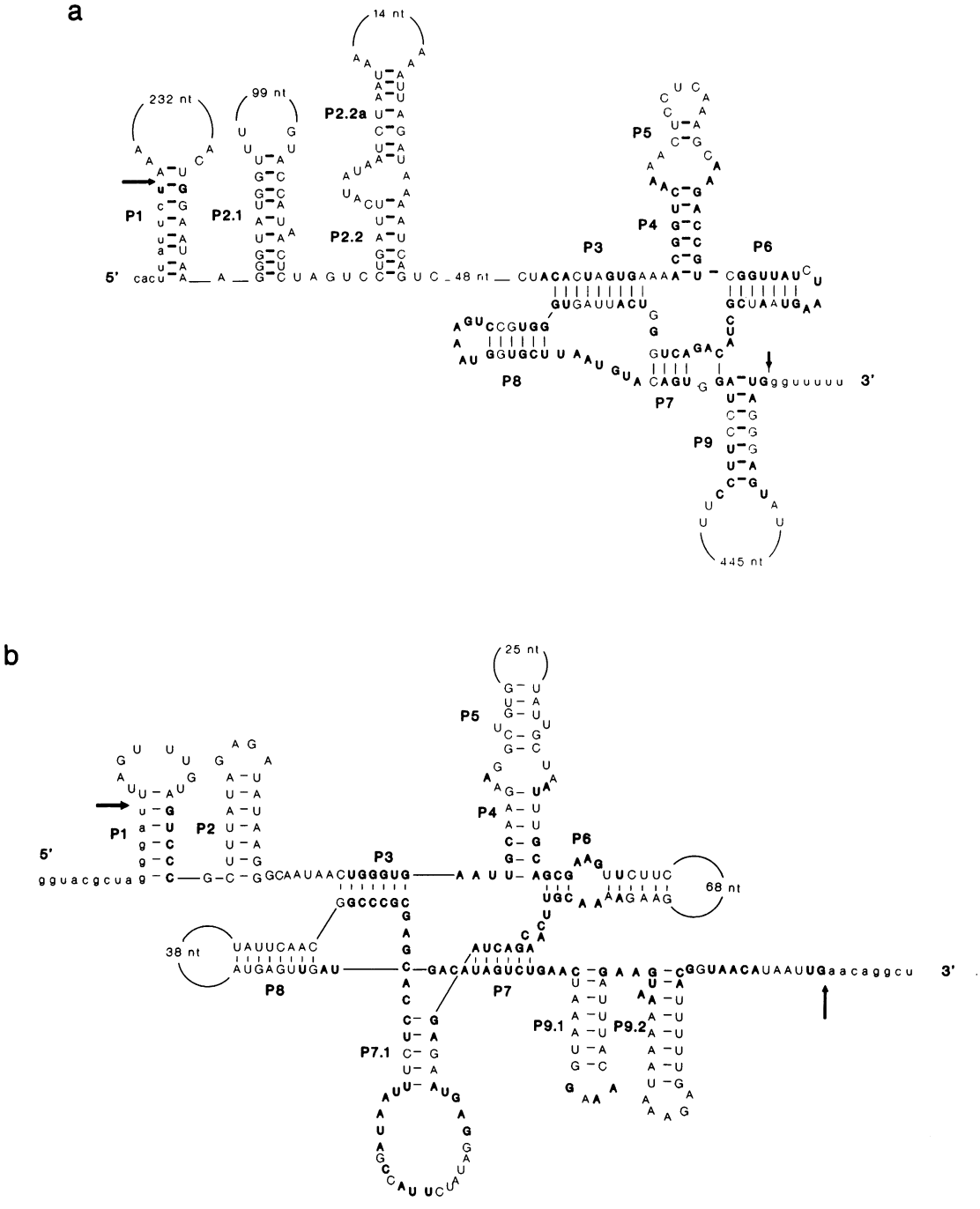

Fig. 4. RNA secondary structure model of **(a)** Pw ai2 and **(b)** Pw ri1. The helix anumeration and structure corresponds to the model and the helix anumeration proposed by Burke et al. (39) and Michel and Westhof (87). Identical nucleotides to Pa ai8 (Fig. 4a) (7) and to Sc ri1 (Fig. 4b) (42) are printed in bold. Exons are indicated in lower case letters, introns in capitals; arrows specify splice points.

Group I introns in the LSUrRNA gene of *P.wickerhamii*

The gene organization of the mitochondrial LSUrRNA identified by DNA sequencing of lambda phage LPw3-1.1 is depicted in Fig. 1. (The complete sequence, EMBL acc. no. X68722, and the secondary structure model of the LSUrRNA will be published elsewhere.) Fig. 3 shows the secondary structure of the fifth domain in which two introns of 375 nt and 500 nt length are inserted at position 2721 and 3096, respectively. The structure of the *P.wickerhamii* domain V is almost identical to the corresponding model of the LSUrRNA from *E.coli* (38) and the introns are located in highly conserved stretches. Both introns display all distinctive features of group I introns (2, 39) and, moreover, are members of subgroup IA which is characterized by the existence of an additional stem-loop in the P7 region (P7.1) (19) (Fig. 4b).

Close phylogenetic relationship between *Prototheca* and plant mitochondria

The phylogenetic distances of the coxI protein from various eukaryotes were calculated with the program of (34) and

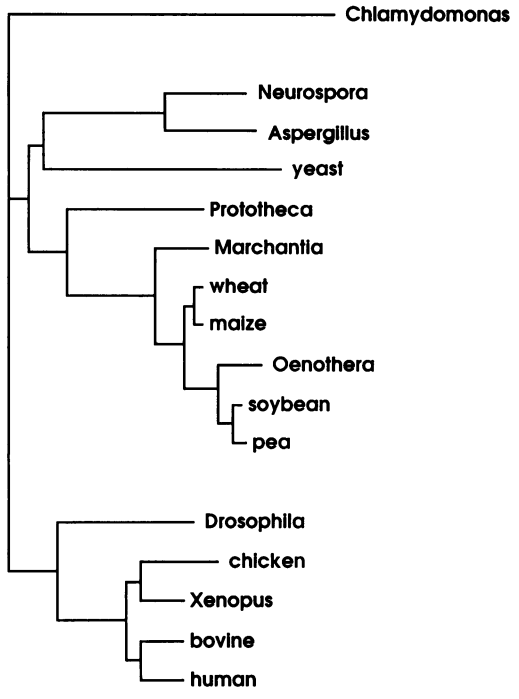

Figure 5. Phylogenetic tree based on the coxI protein from various eukaryotes. The unrooted tree has been calculated with Felsenstein's 'protdist' (unpublished) and neighbour joining programs (35). Essentially the same result was obtained with the program of Hein (34). The length of the horizontal branches represents the relative phylogenetic distance between the species. Residue 4 to 463 of the coxI protein of *P.wickerhamii* (Fig. 2) and the corresponding stretches of the sequences from other species were used. The species are *Chlamydomonas*, *C.reinhardtii* (67); *Neurospora*, *N.crassa* (8); *Aspergillus*, *A.nidulans* (77, 14); yeast, *Saccharomyces cerevisiae* (66); *Prototheca*, *P.wickerhamii*; *Marchantia*, *M.polymorpha* (12); wheat, *Triticium aestivum* (65); maize, *Zea mays* (78); *Oenothera*, *O.berteriana* (79); soybean, *Glycine max* (80); pea, *Pisum sativum* (81); *Drosophila*, *D.melanogaster* (82); chicken, *Gallus gallus* (83); *Xenopus*, *X.laevis* (84); bovine, *Bos tauris* (85); human, *Homo sapiens* (86).

Felsenstein's 'PROTDIST' plus 'NEIGHBOR' (Felsenstein, unpublished, 35) using 16 sequences from species including monocots, dicots, mosses, green algae, fungi and animals. Both programs gave essentially the same result (Fig. 5). Obviously, the tree is consistent with phylogenies based on nuclear SSUrRNA sequences: animal and fungal sequences each form a cluster, the higher plants form a separate branch divided into monocots and dicots, *M.polymorpha* branches close to higher plants and *Prototheca* near the root of plant mitochondria. Only *Chlamydomonas reinhardtii* which has been chosen as an outgroup is unreliably positionned in the presented tree as the bootstrap values are lower than 80% between the connecting intermediate modes of the three eukaryotic kingdoms of fungi, plant, and animals. The recurrent difficulty to position mitochondrial sequences from *C.reinhardtii* in phylogenetic trees is due to the extreme branch length of *C.reinhardtii* and is expected to disappear with the availability of more data from green algal species.

Phylogeny of mitochondrial introns in domain V of LSUrRNA

To date, five species are known which contain a group I intron at an identical position of the mtLSUrRNA (Fig. 3): the four ascomycetes *P.anserina* (intron ri2), *N.crassa* (intron ri1), *A.nidulans* (intron ri1), *S.cerevisiae* (intron ri1) and *P.wickerhamii* (intron ri1) (refernces are given in the legend of Fig. 3).

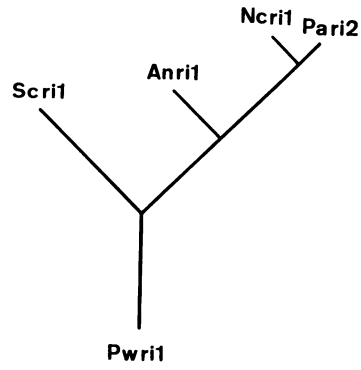

Figure 6. Phylogenetic tree based on mitochondrial LSUrRNA introns (ri) from fungi and algae which correspond to Pw ri1. The unrooted tree was calculated using the parsimony program of (36). The length of the branches indicates the relative distance between the species. Abbreviations see legend of Fig. 3.

A phylogenetic analysis was performed with five corresponding algal and fungal intron nucleotide sequences. Fig. 6 shows the phylogeny of these LSUrRNA introns, constructed with a maximum parsimony program (36). The tree is robust as indicated by its bootstrap values (all above 93%). Consistent results were obtained by using informative positions varying in numbers between 453 and 98 or by applying further algorithms like a maximum likelihood approach and a neighbour-joining method (35). Two main aspects of the presented tree can be noted. i) It is congruent with the traditional fungal taxonomy and with phylogenetic trees of fungi based on nuclear SSUrRNA data: the two pyrenomycetes (*P.anserina* and *N.crassa*) branch very closely together and their distance to *A.nidulans*, a representative of the plectomycetes, is shorter than that to *S.cerevisiae*, belonging to the endomycetes. ii) The position of *P.wickerhamii* is unexpectedly closer to *S.cerevisiae* than to all other ascomycetes, whereas the distance between *S.cerevisiae* and *P.wickerhamii* lies in the same range as that between the two ascomycetes, *S.cerevisiae* and *A.nidulans*.

Two further introns were included in our comparisons: the second LSUrRNA intron from *P.wickerhamii*, which resembles Pw ri1 in sequence and secondary structure, and ri2 from *Scenedesmus obliquus* mitochondria, which is located at an identical position as ri2 from *P.wickerhamii* (21). Focusing exclusively on the positions of *P.wickerhamii* ri1, ri2 and So ri2 in a phylogenetic tree, the distance between the two homologous introns So ri2 and Pw ri2 is shorter than that between ri1 and ri2 of *P.wickerhamii* (data not shown). This suggests that ri2 originated from the transposition of ri1, but that this event dates back to a time before *P.wickerhamii* and *S.obliquus* diverged.

Conserved intron insertion positions in the coxI and LSUrRNA genes of *P.wickerhamii*

The availability of sequence data from two mosaic genes of *P.wickerhamii* mitochondria allows for the first time a comparison of algal intron positions with those from fungi and liverwort. The insertion sites from *P.wickerhamii*, ascomycetous fungi and *Marchantia* coxI introns are summarized in a coxI protein alignment (Fig. 2). As already noted, introns tend to reside in regions corresponding to highly conserved protein segments (e.g., 40, 13; for a review, see 41). The same helds true for the three *P.wickerhamii* introns. These are, however, even inserted at positions identical to those of corresponding fungal and liverwort intervening sequences.

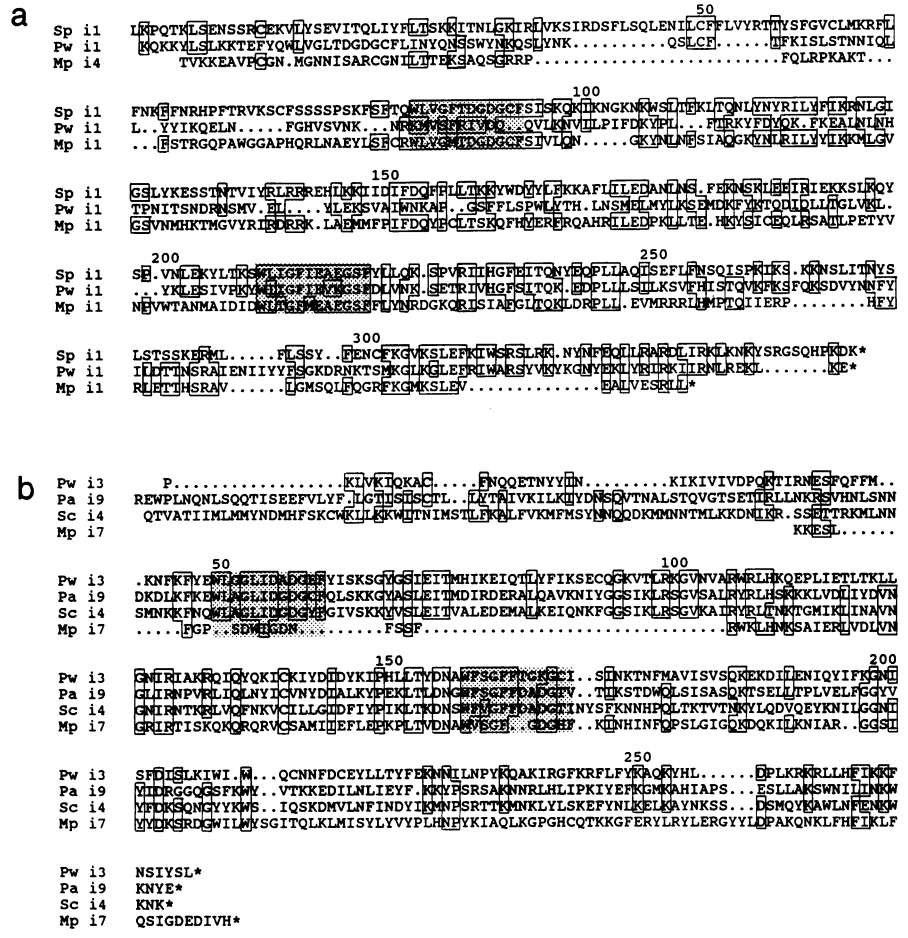

Figure 7. Alignment of homologous group I intronic open reading frames from **(a)** Mp ai4, Sp ai1 and Pw ai1 and **(b)** Mp ai7, Pw ai3, Pa ai9 and Sc ai4. Conserved dodecapeptide regions are dotted, additional regions of similarity are indicated by boxes. For abbreviations and references see legend of Fig. 2.

All LSUrRNA introns of *P.wickerhamii* are located in domain V (Fig. 3) which also seems to be the preferred target for introns in many other species. As much as five mitochondrial introns are known today which are inserted at the same position as intron 1 in *P.wickerhamii*. The sequence adjacent to the insertion site is very well conserved over large evolutionary distances such as between fungi, algae and bacteria.

Mitochondrial introns from *Prototheca* and fungi are highly similar in sequence and secondary structure

As mentioned above, the potential RNA folding of Pw ai2 extraordinarily resembles that of Pa ai8, both members of intron subgroup ID (7). In fact, the structures of both introns are nearly superimposable. At the nucleotide sequence level, the similarity exceeds the conserved P, Q, R, S (corresponding to P4, P7) motifs and includes the additional helical elements P1 to P9 (39). A sequence identity of about 80 % has been found between both introns in the secondary structure core (Fig. 4a).

The comparison of the *P.wickerhamii* ri1 and ri2 with *S.cerevisiae* ri1 (42) which belong to group IA (Fig. 4b) reveal similar results. Regions of sequence similarity exceed beyond the very conserved core into the helices P4, P6, P8 and P9.1 as well as into the loops of P7.1 and P9.2.

ORFs in *P.wickerhamii* introns resemble their counterparts in ascomycetous fungi and the liverwort *Marchantia polymorpha*

Group I intronic ORFs from different organisms usually display only limited similarities, which are mainly restricted to two conserved dodecapeptide motifs (37, 2, 1). However, the intronic ORFs from Pw ai1, Mp ai4 and Sp ai1 on the one hand and those from Pw ai3, Mp ai7, Sc ai4 and Pa ai9 on the other hand, show significant sequence identities, despite the large evolutionary distance between fungal and algal mitochondria (Fig. 7a,b). For example, Pw ai1 ORF shows a 28 % identity over a stretch of 229 amino acids with the Mp ai4 ORF. Similarly, Pw ai3 ORF has an identity of 30 % over a stretch of 260 amino acids with Sc ai4 ORF. Possible implications will be discussed below.

DISCUSSION

The phylogenetic analysis based on coxI sequences implies a close relationship between plant and *P.wickerhamii* mitochondria

At the level of nuclear sequences, the specific link between the higher plants and unicellular eukaryotes is well established. Chlorophytes are the acknowledged progenitors of embryophytes

(mosses, ferns, gymnosperms and angiosperms). So far, however, this link does not seem to apply to the mitochondrial genome and available data do not suggest that green algal and plant mtDNAs share a common ancestor. Phylogenetic trees based on mitochondrial SSUrRNA and LSUrRNA sequences including *Chlamydomonas reinhardtii* led to the assumption that the mitochondrial lineage of embryophytes and chlorophytes is incoherent (24, 23, 43).

Our phylogenetic analysis including *P. wickerhamii* shows that this green algal mtDNA does cluster with corresponding sequences from lower and higher plants. It appears that *P.wickerhamii* has accumulated far less mutations in its mtDNA compared to higher plants than *Chlamydomonas*, thus providing us with a well suited model organism to study the plant-algal mitochondrial lineage for the first time. Therefore we propose a recent common ancestor for mitochondria of *P.wickerhamii* and plants.

Were group I introns already present in the bacterial ancestor of mitochondria?

The phylogeny based on mitochondrial LSUrRNA introns inserted at an identical position as Pw ril coincides strikingly with trees based on nuclear sequences and with phylogenies using morphological data of the corresponding taxa. The phylogenetic tree indicates that the taxa diverged at the same time as did the mitochondrial LSUrRNA introns (Fig. 6). Therefore we assume that this particular intron is ancient and propagated via mendelian inheritance rather than via horizontal, interspecific gene transfer. As a consequence it might have been present in the putative progenitors of mitochondria.

A more complete record of sequencing data specifically in the algal, protozoan and prokaryotic groups would be necessary to trace the evolutionary origin of this intron type back to the alpha-purple bacteria, the bona fide eubacterial ancestors of mitochondria and the suspected common origin of these introns.

Even if introns from different species are located at identical sites within genes, there may or may not be an intronic ORF present. In cases where ORFs are present, their position in the RNA secondary structure may differ and the similarity of their deduced amino acid sequence is usually very low, i.e., restricted to dodecapetide motifs (44, 45). Therefore it has been proposed that intronic ORFs may be capable of their own evolution, independent of the intervening sequence in which they reside (44, 46, 19). However, this view is contradictory with the results described here. For example, we noticed consistency for those four coxI introns, from fungi, algae and liverwort which are inserted at amino acid position 239 (referring to the anummeration in *P.wickerhamii*, see fig. 2). Not only are these introns inserted at identical positions of the coxI gene, but in addition their ORFs are quite similar (Fig. 7b). Concerning Pw ail and its ORF a similar situation can be observed: the ORFs of the corresponding introns from *M.polymorpha* and *S.pombe* (Fig. 2) show significant similarities to the algal intronic ORF (Fig. 7a). In both cases the intronic ORFs and their introns seem to be intimately linked and appear to be propagated together.

Why do plant mitochondria lack group I introns ?

Group I introns reside in a large number of genomes from various sources such as phages T4 and SPO1 (47, 48, 49), cyanobacteria (50, 51), plastids (52, 53), mitochondria and nuclei (*Tetrahymena* (54); *Physarum* (55)). In higher plant mitochondria, however, not a single group I intron has so far been detected. Instead, a small number of group II introns, often involved in trans-splicing processes, are present in these organelles (10, 11, 56, 57). A variable intron distribution was recently discussed for the coxII and the nad4 genes in plant mitochondria (58, 59). Surprisingly, several group I as well as group II introns were found in the mtDNA of the liverwort *Marchantia polymorpha*, which are located in protein and RNA coding genes (12).

Intron loss was attributed to a reverse transcriptase activity potentially encoded in group II intronic ORFs and certain, freestanding genes present in higher plant mitochondria (60, 61, 62, 63, 64, 28). It was proposed that, the putative reverse transcriptase (e.g. encoded by group II intronic ORFs) could produce a cDNA copy of an intron-less RNA, which, if followed by a homologous recombination event involving the cDNA and the genomic DNA, would promote intron loss (46).

This process was suggested to account for the unbalanced numbers of group I versus group II introns in mitochondria from various species (46). In our view, however, this process would rather cause the concommitant loss of group I and group II introns. The unexpected finding of very few and exclusively group II introns in plant mitochondria may be explained by their requirement for trans-splicing. All introns may have been eliminated in this organelle by the above mentioned mechanism, except those which are implicated in trans-splicing. Trans-splicing introns are expected to be very resistant to intron-loss, because the correct replacement of genomic copies of trans-spliced genes cannot be achieved by simple homologous recombination and should therefore be extremely rare. Whatever the mechanism of the assumed intron loss is in the chlorophyte mitochondrial lineage, that massive loss should have appeared after the bifurcation of lower and higher plants.

ACKNOWLEDGEMENTS

We thank M.Luber and I.Plante for excellent technical assistance. GW was a part-time stipend of the Deutscher Akademischer Austauschdienst (Bonn-Bad Godesberg, Germany). This work is supported by grants from the Deutsche Forschungsgemeinschaft (Bonn-Bad Godesberg, Germany), FCAR (Canada), CIAR (Canada) and MRC (Canada).

REFERENCES

1. Waring RB, Davies RW, Scazzocchio C, Brown TA (1982) Proc Natl. Acad Sci USA 79, 6332–6336.
2. Michel F, Jacquier A, Dujon B (1982) Biochimie 64, 867–881.
3. Michel F, Dujon B (1983) EMBO Journal 2, 33–38.
4. Lambowitz AM (1989) Cell 56, 323–326.
5. Jacquier A, Dujon B (1983) Mol Gen Genet 192, 487–499.
6. Karsch T, Kück U, Esser K (1987) 6743–6744.
7. Cummings DJ, Michel F, McNally KL (1989) Curr Genet 16, 381–406.
8. Burger G, Scriven C, Machleidt W, Werner S (1982) EMBO J 1, 1385–1391.
9. Lonsdale DM (1989) In The Biochemistry of Plants, A comprehensive Treatise, ed A. Marcus. Vol 15, 229–295, New York, Academic Press.
10. Chapdelaine Y, Bonen L (1991) Cell 65, 465–472.
11. Wissinger B, Schuster W, Brennicke A (1991) Cell 65, 473–482.
12. Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T, Ohyama K (1992) J Mol Biol 223, 1–7.
13. Lang BF (1984) EMBO J 3, 2129–2136.
14. Waring RB, Brown TA, Scazzocchio C, Davies RW (1984) EMBO J 3, 2121–2128.

15. Cummings DJ, Michel F, McNally KL (1989) Curr Genet 16, 407–418.
16. Dujon B, Colleaux L, Jacquier A, Michel F, Monteilhet C (1986) In Extrachromosomal Elements of lower eukaryotes, eds Wickner RB, Hinnenbusch A, Lambowitz AM, Gunsalus C, Hollaender A, pp. 5–27, New York Plenum Press.
17. Michel F, Dujon B (1986) Cell 46,323.
18. Wolf K, Del Guidice L (1988) Adv Genet 25, 185–308.
19. Cummings DJ, Domenico JM (1988) J Mol Biol 204, 815–839.
20. Michaelis G, Vahrenholz C, Pratje E (1990) Mol Gen Genet 223, 211–216.
21. Kück U, Godehardt I, Schmidt U (1990) Nucl Acids Res 18, 2691- 2697.
22. Colleaux L, Michel-Wolwertz M-R, Matagne RF, Dujon B (1990) Mol. Gen Genet 223, 288–296.
23. Cedergren R, Gray MW, Abel Y, Sankoff D (1988) J Mol Evol 28, 98–112.
24. Gray MW (1988) Biochem Cell Biol 66, 325–348.
25. Wolff G, Kück U (1990) Curr Genet 17, 347–351.
26. Yanish-Perron C, Vieira J, Messing J (1985) Gene 33, 103–119.
27. Sambrook J, Fritsch EF, Maniatis T (1989) Cold Spring Harbor Laboratory Press, New York.
28. Kück U (1989) Mol Gen Genet 218, 257–266.
29. Sanger F, Nicklen S, Coulson AR (1977) Proc Natl Acad Sci USA 74, 5463–5467.
30. Lang BF, Burger G (1990) Anal Biochem 188, 176–180.
31. Pearson WA, Lipman DJ (1988) Proc Natl Acad Sci USA 85, 2444 2448.
32. Corpet F (1988) Nucl Acids Res 16, 10881–10890.
33. Lang BF, Burger G (1986) Nucl Acids Res 14,455–465.
34. Hein J (1990) Meth Enzymol 183, 626–644.
35. Felsenstein J (1990) PHYLIP Manual Version 3.3, Herbarium, Unversity of California, Berkeley.
36. Sankoff D, Abel Y, Cedergren R Gray MW (1988) In Classification and Related Methods of Data Analysis, ed Bock HH , pp. 385–394, North Holland, Amsterdam.
37. Lazowska J, Claisse M, Gargouri A, Kotylak Z, Spyridakis A, Slonimski PP (1989) J Mol Biol 205, 275–289.
38. Brosius J, Dull TJ, Noller HF (1980) Nature 290, 457–465.
39. Burke JM, Belfort M, Cech TR, Davies WR, Schweyen RJ, Shub DA, Szostak JW, Tabak HF (1987) Nucl Acids Res 15, 7217–7221.
40. Burke JM, RajBhandary UL (1982) Cell 31, 509–520.
41. Burger G (1985) In Achievements and Perspectives of Mitochondrial Research, vol II Biogenesis, eds Quagliariello E, Slater EC, Saccone C, Kroon AM, pp. 305–316, Elsevier Science Publishers, Amsterdam..
42. Dujon B (1980) Cell 20: 185–197.
43. Gray MW, Cedergren R, Abel Y, Sankoff D (1989) Proc Natl Acad Sci USA 86, 2267–2271.
44. Mota EM, Collins RA (1988) Nature 332, 654–656.
45. Belfort M (1991) Cell 64, 9–11.
46. Dujon B (1989) Gene 82, 91–114.
47. Chu FK, Maley GF, Maley F, Belfort M (1984) Proc Natl Acad Sci. USA 81, 3049–3053.
48. Chu FK, Maley GF, West DK, Belfort M, Maley F (1986) Cell 45, 157–166.
49. Goodrich-Blair H, Scarlato V, Gott JM, Xu M-Q, Shub DA (1990). Cell 63, 417–424.
50. Xu M-Q, Scott DK, Goodrich-Blair H, Nierzwicki-Bauer SA, Shub DA (1990) Science 250, 1566–1570.
51. Kuhsel MG, Strickland R, Palmer JD (1990) Science 250, 1570–1573.
52. Erickson JM, Rahire M, Rochaix JD (1984) EMBO J 3, 2753–2762.
53. Steinmetz A, Gubbins EJ, Bogorad L (1982) Nucl Acids Res 10, 3027–3037.
54. Kan NC, Gall JG (1982) Nucl Acids Res 10, 2809–2822.
55. Muscarella DE, Vogt VM (1989) Cell 56, 443–454.
56. Knoop V, Schuster W, Wissinger B, Brennicke A (1991) EMBO J 10, 3483–3493.
57. Binder S, Marchfelder A, Brennicke A, Wissinger B (1992) J Biol Chem 267: 7615–7623.
58. Gass D A, Makaroff C A, Palmer J D (1992) Curr Genet 21: 423–430.
59. Lippok B, Brennicke A, Wissinger B (1992) Mol Gen Genet 232: 322–327.
60. Schuster W, Brennicke A (1987) EMBO J 6, 2857–2863.
61. Wahleithner J A, Macfarlane J L, Wolstenholme D R (1990) Proc Natl Acad Sci USA 87: 548–552.
62. Nugent JM, Palmer JD (1991) Cell 66, 473–481.
63. Osiewacz HD, Esser K (1984) Curr Genet 8, 299–305.
64. Michel F, Lang F (1985) Nature 316, 641–642.
65. Bonen L, Boer PH, McIntosh JE, Gray MW (1987) Nucl Acids Res 15, 6734.
66. Bonitz SG, Coruzzi G, Thalenfeld BE, Tzagoloff A, Macino G (1980) J Biol Chem 255, 11927–11941.
67. Kück U, Neuhaus H (1986) Appl Microbiol Biotechnol 23, 462–469.
68. Lang BF, Ahne F, Distler S, Trinkl H, Kaudewitz F, Wolf K (1983) in Mitochondria 1983, eds Schweyen RJ, Wolf K, Kaudewitz F, pp. 313–330, Berlin, New York, Walter deGruyter & Co.
69. Trinkl H, Wolf K (1986) Gene 45, 289–297.
70. Noller HF (1984) Ann Rev Biochem 53, 119–162.
71. Lang BF, Cedergren R, Gray MW (1987) Eur J Biochem 169, 527–537.
72. Turmel M, Boulanger J, Schnare MN, Gray MW, Lemieux C (1991) J. Mol Biol 218, 293–311.
73. Allet B, Rochaix J-D (1979) Cell 18: 55–60.
74. Cummings DJ, Domenico JM, Nelson J (1989c) J Mol Evol 28, 242–255.
75. Nomiyama H, Kuhara S, Kukita T, Otsuka T, Sakaki Y (1981a) Nucl Acids Res 9: 5507–5520.
76. Nomiyama H, Sakaki Y, Takagi Y (1981b) Proc Natl Acad Sci USA 78: 1376–1380.
77. Netzker R, Köchel HG, Basak N, Küntzel H (1982) Nucl Acids Res 10, 4783–4794.
78. Isaac PG, Jones VP, Leaver CL (1985) EMBO J 4, 1617–1623.
79. Hiesel R, Schobel W, Schuster W, Brennicke A (1987) EMBO J 6: 29–34.
80. Grabau EA (1986) Plant Mol Biol 7, 377–384.
81. Kemmerer EC, Kao T, Deng G, Wu R (1989) Plant Mol Biol 13, 121–124.
82. DeBruijn MHL (1983) Nature 304, 234–241.
83. Desjardin P, Morais R (1990) J Mol Biol 212, 599–63.
84. Roe BA, Ma DP, Wilson RK, Wong JFH (1985) J Biol Chem 260, 9759–9774.
85. Anderson S, DeBruijn MHL, Coulson AR, Eperon IC, Sanger F, Young IG (1982) J Mol Biol 156, 683–717.
86. Anderson S, Bankier AT, Barrell BG, DeBruijn MHL, Coulson AR, Drouin J, Eperon JC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Nature 290, 457–465.
87. Michel F, Westhof E (1990) J Mol Biol 216: 585–610.