J Mol Evol (1992) 34:254–258

Journal of Molecular Evolution

Genetic Code and Phylogenetic Origin of Oomycetous Mitochondria

Petr Karlovsky and Berthold Fartmann

Institute of Plant Pathology, University of Goettingen, Grisebachstrasse 6, W-3400 Goettingen, Germany

Summary. We sequenced the 3′-terminal part of the *COX3* gene encoding cytochrome c oxidase subunit 3 from mitochondria of *Phytophthora parasitica* (phylum Oomycota, kingdom Protoctista). Comparison of the sequence with known *COX3* genes revealed that UGG is used as a tryptophan codon in contrast to UGA in the mitochondrial codes of most organisms other than green plants. A very high AT mutation pressure operates on the mitochondrial genome of *Phytophthora,* as revealed by codon usage and by A+T content of noncoding regions, which seems paradoxical because AT pressure causes tryptophan codon reassignment from UGG to UGA in mitochondria of most species. The genetic code and other data suggest that mitochondria of Oomycota share a direct common ancestor with mitochondria of plants and that mitochondria of the ancestor of Planta and Oomycota were acquired in a second endosymbiotic event, which occurred later than the acquisition of mitochondria by other eukaryotes.

Key words: Genetic code — Mitochondria — Oomycetes — AT pressure — Codon usage — Cytochrome oxidase — Phylogeny — Endosymbiotic hypothesis — Polyphyletic mitochondrial origin

Introduction

The phylogenetic origin of mitochondria as eubacterial endosymbionts is generally accepted (Margulis 1981). Mitochondrial genomes of different eukaryotes exhibit an extraordinary diversity, which suggests a polyphyletic origin (Gray 1989). A clear demarcation line can be drawn between mitochondrial genomes of plants and that of other eukaryotes. Plant mitochondrial genomes are very large (up to the size of a eubacterial genome); even the smallest one known is larger than the largest mitochondrial DNA (mtDNA) of a nonplant organism. The open structure of plant mitochondrial genomes with its genes floating in a sea of noncoding sequences strongly contrasts with the economized structure found in mitochondrial genomes of other organisms. Intergenic sequences in plant mtDNA contain recombinogenic repeats that promote frequent genome rearrangements or even the coexistence of various DNA molecules in equilibrium with the master chromosome (Gray 1989; Palmer 1990). In contrast to the frequent recombination and the resulting structural divergence, the rate of sequence change by point mutations is extremely low in plant mitochondria, around 100 times slower than for chloroplast genomes or animal mitochondria (Palmer and Herbon 1988; Palmer 1990). Additional characteristics that distinguish plant mitochondrial genomes from those of other organisms are the presence of genes for the α subunit of F_1-ATPase, a 5S RNA, and a protein homologous to the S13 ribosomal protein of *Escherichia coli*; larger rRNAs; and frequent acquisition of genes from nuclei and chloroplasts, resulting in the formation of mosaic genomes (for a review see Levings and Brown 1989).

The most generally applicable method for inferring long-range phylogenetic relationships within the mitochondrial lineage is the comparison of rRNA sequences. The results of such studies indicate that plant mitochondria, or at least their rRNA genes, originated at a later time in evolution than did the mitochondria or their rRNA genes in other organisms (Gray 1989; Gray et al. 1989).

Another feature relevant to the evolution of the mitochondrial genome is the genetic code. Plant mitochondria retain the universal UGG code for tryp-

Offprint requests to: P. Karlovsky

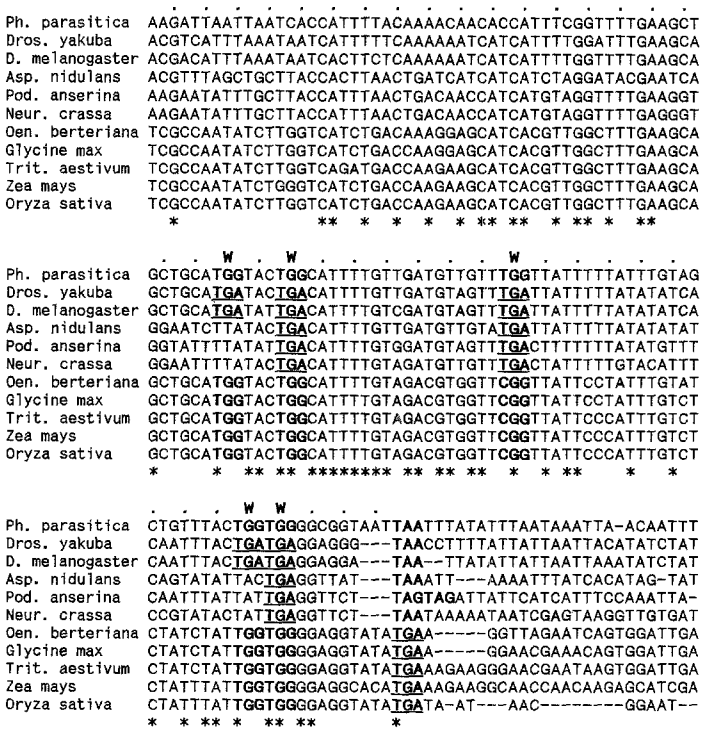

Fig. 1. Multiple alignment of 3′ parts of mitochondrial genes encoding cytochrome c oxidase subunit 3 and their downstream regions. Reading frame (·) and five conservative tryptophan residues (W) are indicated above the sequences. Tryptophan codons and stop codons of the cytochrome oxidase genes are in bold type, TGA codons are underlined. Note that the CGG codons are edited to UGG in plant mitochondrial mRNA. Asterisks indicate places with identical bases in all sequences. Full names of organisms and accession numbers of the respective sequences in GenBank and EMBL libraries are *Aspergillus nidulans* (J01390), *Drosophila melanogaster* (M37275), *Drosophila yakuba* (complete mitochondrial genome sequence compiled under DRYMTCG), *Glycine max* (X15131), *Neurospora crassa* (J01430), *Oenothera berteriana* (X04764), *Oryza sativa* (X17040), *Podospora anserina* (X14734), *Triticum aestivum* (X15944), and *Zea mays* (X12728).

tophan (Trp), whereas most nonplant mitochondria use UGA, as well as other nonstandard amino acid codons (reviewed by Jukes and Osawa 1990).

In this paper, similarities between the genetic code in mitochondria of the oomycetous plant pathogen *Phytophthora parasitica* and that of green plants is interpreted with respect to endosymbiotic theory and implications for the phylogeny of oomycetes.

Materials and Methods

A cloned fragment from mtDNA of *P. parasitica* DSM 1829 (denoted Hind-ARSI), selected for its ability to drive autonomous plasmid replication in yeast (Mühlbauer 1988), was partially digested with DNase I in the presence of Mn^{2+} (Campbell and Jackson 1980). Fragments 150–240 bp long were isolated on an agarose gel, repaired by incubation with Klenow fragment of *E. coli* DNA polymerase I and all four deoxyribonucleotide triphosphates, and cloned into the *Sma*I site of pARSEQ. The cloning vector pARSEQ is a derivative of pUC19, which carries a yeast marker gene *URA3* and a polylinker from pSP64CS (Eckert 1987). This polylinker contains one *Sma*I site flanked on both sides by *Tth*111I recognition sequences with different central bases. After the cleavage of the constructs with *Tth*111I, each end of the cloned fragment was separately labeled using Klenow fragment and a ^{32}P-deoxyribonucleotide triphosphate complementary to the respective protruding 5′ nucleotide. Labeled fragments were sequenced using a modified chemical degradation protocol. After comparing the sequence assembled from overlapping fragments with GenBank and EMBL libraries, a coding sequence for the carboxyterminal part of subunit 3 of cytochrome c oxidase was identified. The second sequence used in this study (part of the coding sequence for the α subunit of the F_1-ATPase, Acc. No. X14351) was found by searching computer libraries for *Phytophthora* mitochondrial sequences.

Results and Discussion

Genetic Code in Mitochondria of Phytophthora parasitica

In the course of analysis of ARS (autonomously replicating sequences) from mitochondria of *P. parasitica,* we sequenced a 3′ part of the gene for cytochrome c oxidase subunit 3 (*COX3*). Figure 1 shows an alignment of this sequence with homologous sequences from other organisms and the position of five conservative Trp residues. In *Phytophthora,* all of these Trp are encoded by UGG codons. By comparing the aligned sequences we found no difference between the *Phytophthora* mitochondrial codon assignment and the universal code.

The only known mitochondrial protein coding sequence of *Phytophthora* in addition to *COX3* is a part of the gene for the α subunit of F_1-ATPase (Förster et al. 1989, Acc. No. X14351). There is no Trp codon in this sequence. A comparison with homologous sequences revealed no difference from the universal code (data not shown). Green plants were the only organisms previously known to use the universal code in their mitochondria (Jukes and Osawa 1990).

One of the five conservative Trp residues in the carboxyterminal end of the *COX3* product is coded for by CGG in plant mitochondria (Fig. 1). This codon is edited to UGG in mRNA sequences (reviewed by Cattaneo 1990). Another difference between *Phytophthora* and plant *COX3* genes is the

Table 1. Codons for the most frequently used amino acids in sequenced parts of *COX3* and of the gene encoding the α subunit of the F_1-ATPase from mitochondrial DNA of *Phytophthora*

Leucine		Alanine		Glycine		Valine		Isoleucine	
TTA	13	GCT	13	GGT	11	GTT	8	ATT	11
TTG	–	GCC	–	GGC	1	GTC	–	ATC	–
CTT	–	GCA	3	GGA	1	GTA	4	ATA	–
CTC	–	GCG	1	GGG	–	GTG	–		
CTA	–								
CTG	–								

use of the stop codons UAA (*Phytophthora*) and UGA (plants). We do not know whether the UGA triplet would still function in *Phytophthora* mtDNA properly as a translation termination signal. However, we predict that because of the extremely high AT pressure (see the next section), most if not all UGA stops have been converted to UAA.

Some species are known to use both UGG and UGA Trp codons in their mitochondria (Anderson et al. 1982; Roe et al. 1985). Because these mitochondria evolved under a high AT mutational pressure, only a small fraction of Trp codons is represented by UGG. Regarding this, we examined the *COX3* genes used in the alignment in Fig. 1. We counted 50 Trp codons in *COX3* genes of two *Drosophila* and three fungal species, 49 were UGA and 1 was UGG (in the gene of *Podospora anserina*). We predict that no UGA Trp codon will be found in mtDNA of *Phytophthora* because at least some of the UGG codons found in *COX3* should have already been converted to UGA under a strong AT pressure if it were possible. More than 90% of codons allowed to contain A or T in the third position are actually NNA or NNT (see the next section). Therefore, we predict that the CCA Trp anticodon (and not UCA) will be found in mitochondria of *Phytophthora*. The definitive answer can be only obtained by analyzing the respective $tRNA^{Trp}$ sequences.

AT Pressure and Codon Usage

AT pressure is the cause of the UGA capture by Trp in fungal and animal mitochondria and in some bacteria with a very high A/T content (Osawa and Jukes 1989; Osawa et al. 1990). The presence of AT or GC pressure can be ascertained by calculating the G/C content of noncoding regions and by analyzing codon usage, especially with respect to the frequency of A/T versus G/C in the third codon position. Excluding methionine and Trp, which possess only one codon each, codons used in both of the *Phytophthora* mitochondrial protein coding sequences mentioned above have A or T in their third position 92.4% of the time. This is an indication of very strong AT pressure. Codons for the five most frequently used amino acids are shown in Table 1. Two hierarchically ordered preferences governing the codon usage are evident.

1) Preference for A and T over G and C in all codon positions. (Only one leucine codon consisting exclusively of A and T nucleotides is used out of six possible codons. The A/T-richest codons available are used in all other cases.)
2) Preference for T over A. (In all four amino acids where two codons share the same high A/T content, the codon with T in the third position is strongly preferred.)

In the first noncoding region of *Phytophthora* mtDNA, the beginning of which is shown in Fig. 1, 83 out of 103 bases following the stop codon of *COX3* are A or T. Two other sequenced noncoding fragments with ARS activities, located at different places in the mitochondrial genome, also have a high A/T content: ARSII, with a length of 414 bp and 87.7% A+T, and ARSIII, with a length of 368 bp and 78.8% A+T (Karlovsky and Fartmann, unpublished).

Fragments with ARS activity are special selected sequences that may not be representative of the whole mitochondrial genome. Digestion with restriction enzymes whose target sequences consist exclusively of A and T nucleotides can be used to test the distribution of A/T-rich sequences in the whole genome. Förster et al. (1987) found that digestion of *Phytophthora* mtDNA (40–45 kb) with *Dra*I (recognition sequence TTTAAA) produced numerous small fragments, all of them less than 1.3 kb. Essentially the same result was obtained with *Asn*I (target sequence ATTAAT; de Cock and Karlovsky, unpublished), strongly suggesting that sequences with a very high A/T content are scattered throughout the whole mitochondrial genome, perhaps encompassing all intergenic spacers.

Phylogeny of Oomycetous Mitochondria

The genetic code is the most conservative evolutionary feature for inferring evolutionary relationships that can be exploited by analyzing DNA sequences. Many time-coordinated mutational changes are required to achieve a codon reassignment without killing the cell through the accumulation of de-

```
Ph. parasitica  RLINHHFTKQ HHFGFEAAAW YWHFVDVVWL FLFVAVYWWG GN
-------------------------------------------------------------
Dros. yakuba    .HL.N..S.N ..... ..... ..YITI.... .-  32
D. melanogaster .HL.N..S.N ..... ..... ..YITI.... .-  32

Oen. berteriana .QYLG.L..E ..V..... ..... ....SI.... .I  32
Glycine max     .QYLG.L..E ..V..... ..... ....SI.... .I  32
Zea mays        .QYLG.L..K ..V..... ..... .P..SI.... .T  31
Oryza sativa    .QYLG.L..K ..V..... ..... .P..SI.... .I  31
Trit. aestivum  .QYLGQM..K ..V..... ..... .P..SI.... .I  30

Asp. nidulans   ..AAY.L.DH ..L.Y.SGIL ..... ..YMS..Y.. Y-  25
Pod. anserina   .MFAY.L.DN ..V...GGIL ..... ..Y.SI.Y.. S-  25
Neur. crassa    .MFAY.L.DN ..V...GGIL ..... ..YIS..Y.. S-  25
-------------------------------------------------------------
```

Fig. 2. Comparison of carboxyterminal parts of cytochrome c oxidase subunit 3 proteins as deduced from mitochondrial DNA sequences (see Fig. 1). Amino acids identical with those at homologous positions of the *Phytophthora parasitica* protein are indicated by dots; nonidentical residues are abbreviated using the standard one-letter code. The number of positions with identical residues as compared with *Phytophthora* (total of 42 positions) is indicated to the right of the sequences.

fective proteins (Osawa and Jukes 1989), including changes in the codon in question at relevant positions in all genes encoding essential proteins, mutations in tRNA genes, and in some cases changes in the specificity of aminoacyl-tRNA synthetases and release factors.

Alterations in the genetic code are probably recent evolutionary events (Osawa et al. 1990). Capture of the stop codon UGA by Trp in mitochondrial codes must have occurred after the separation of mitochondria of nonplant organisms from an ancestor shared with plant mitochondria (Jukes and Osawa 1990). The multistage process involved in codon reassignment probably did not occur two successive times in oomycetous mitochondria: first by the capture of UGA by Trp and then by reassigning the stop codon again. In this case, the second change would have been contraselected by AT pressure. Evidently, oomycetous mitochondria developed from the same ancestor as did plant mitochondria after its separation from the nonplant lineage. Analysis of rRNA sequences and other specific features of plant mitochondria support the origin of these mitochondria in a second endosymbiotic event. The presence of the gene for the α subunit of the F_1-ATPase in mitochondria of *Phytophthora* (Förster et al. 1989) and our results concerning the genetic code suggest that the second endosymbiotic event gave rise to an endosymbiotic lineage that diverged into plant and oomycetous mitochondria and possibly also into mitochondria of heterokont algae. There was not enough time for the AT pressure in the evolution of oomycetes to cause the codon reassignment.

An alternative hypothesis is that an oomycetous ancestor, living in close parasitic contact with a plant, captured a new mitochondrion from the host and subsequently lost its own. The difficulty with this hypothesis is the assumption that all *Phytophthora* species with mitochondrial genomes closely related to *P. parasitica* developed from one plant parasite species. However, primitive marine species, which are phylogenetically older than species that parasitize land plants, have restriction patterns of mtDNA similar to that of *P. parasitica* (de Cock and Bahnweg 1989). More molecular data on different *Phytophthora* and other oomycetes species are required before the hypothesis can be definitively disproved.

After the separation of the oomycetous and plant mitochondrial lineages, AT mutation pressure lead to the nonrandom nucleotide composition of noncoding sequences and to a biased codon usage in oomycetous mitochondrial genomes. Genomic economization formed mtDNA into a small molecule resembling fungal and animal mitochondrial genomes. The AT pressure and genomic economization are connected with the replication and repair of DNA and with the competition for limited supply of energy. The common denominator of these factors could be the absorbtive nutritional mode, which was the main evolutionary force governing the divergence of oomycetes from plants. In any case, these factors have not operated upon the mitochondrial genome of oomycetes long enough to reassign the Trp codon.

Do other features of oomycetous mtDNA confirm their close relationship to plant mitochondria? Unfortunately, molecular analysis of oomycetous mitochondria is in the early stages. Alignment of the protein sequence of the fragment translated from *COX3* of *P. parasitica* with other cytochrome oxidase sequences (Fig. 2) indicates that oomycetous mitochondria are not more related to mitochondria of Fungi than to that of Metazoa and Planta. However, extensive sequencing is required before a quantitative phylogenetic conclusion can be drawn from such sequence comparisons.

The phylogeny of oomycetes is a controversial issue even at the level of kingdoms. Traditionally, these species were grouped in the class Oomycetes within Phycomycetes or "lower fungi" within the kingdom Fungi (Barr 1983). In the modern systems, they were given the status of phylum (Oomycota) and were first placed in Fungi (Whittaker 1969) only to be later moved into Protoctista (Margulis 1981; Margulis and Schwartz 1988). Protoctista is a polyphyletic kingdom of mostly unicellular eukaryotes defined by exclusion (neither plants nor animals nor fungi). It comprises such different groups as protozoa, algae, and slime molds. To avoid this artificial grouping, the suggestion was made to raise major protoctist phyla to kingdoms (Leedale 1974). However, even these efforts did not lead to a satisfying placement of oomycetes; e.g., according to the 13-kingdom system of Leedale, oomycetes are grouped with brown algae and diatoms (Margulis 1981).

The phylogenetic relationship between oomycetes and heterokont algae (especially chrysophytes and xanthophytes) is, however, a long established concept supported by numerous ultrastructural and biochemical characters (reviewed in Barr 1983). For example, oomycetes contain cellulose in their cell wall and synthesize lysine through diaminopimelic acid-like plants and not through aminoadipic acid-like fungi (LeJohn 1971). Another relevant phylogenetic feature was established after a long and controversial debate (reviewed by Shaw 1983); the oomycetous life cycle has karyogamy immediately following meiosis, and their vegetative phase is diploid. The first molecular data concerning oomycetous phylogeny emerged recently; a study of small-subunit rRNA sequences demonstrated a very close relationship between *Achlya* and chrysophytes (Gunderson et al. 1987). The analysis of the coding sequences for two enzymes of the Trp biosynthetic pathway resulted in phylogenetic trees with *Phytophthora* sequences located outside the clusters of seven fungal genera, which include representatives of ascomycetes, basidiomycetes, and zygomycetes (Karlovsky and Prell 1991). A comparison of actin sequences demonstrated homology of *Phytophthora* actin sequences to those of *Arabidopsis, Hydra,* and *Dictyostelium* rather than to those of *Aspergillus* and *Saccharomyces* (Unkles et al. 1991).

The use of the universal genetic code in the mitochondria of *Phytophthora* confirms the close phylogenetic relationship between these organelles in oomycetes and plants. Thus, the genomes of organelles provide important clues for the study of evolution. This should not surprise us when we bear in mind that genomic mosaicism is a basic feature of eukaryotic organisms and that it possesses an evolutionary role, which is valued even higher than that of Darwinian selection by some researchers (Mann 1991).

Acknowledgments. We are grateful to H.H. Prell for consultation regarding the manuscript and to Margaret Nemec for improving the English. Part of the work was done at Zentrales Isotopenlaboratorium für biologische und medizinische Forschung der Universität Göttingen. The work was partially funded by a grant from Deutsche Forschungsgemeinschaft.

References

Anderson S, de Bruijn MHL, Coulson AR, Eperon IC, Sanger F, Young IG (1982) Complete sequence of bovine mitochondrial DNA conserved feature of the mammalian mitochondrial genome. J Mol Biol 156:683–717

Barr DJS (1983) The zoosporic grouping of plant pathogens. In: Buczacki ST (ed) Zoosporic plant pathogens. Academic Press, London, pp 43–83

Campbell VW, Jackson DA (1980) The effect of divalent cations on the mode of action of DNase I. J Biol Chem 255:3726–3735

Cattaneo R (1990) Messenger RNA editing and the genetic code. Experientia 46:1142–1148

de Cock AWAM, Bahnweg G (1989) Molecular taxonomy of *Phytophthora*. In: *Phytophthora*—An International Symposium, Trinity College, Dublin, p 22

Eckert R (1987) New vectors for rapid sequencing of DNA fragments by chemical degradation. Gene 51:247–254

Förster H, Kinscherf TG, Leong SA, Maxwell DP (1987) Molecular analysis of the mitochondrial genome of *Phytophthora*. Curr Genet 12:215–218

Förster H, Kinscherf TG, Leong S (1989) F_1-ATPase α subunit in the mitochondrial genome of the Pythiaceae. Nucleic Acids Res 17:7109

Gray MW (1989) Origin and evolution of mitochondrial DNA. Annu Rev Cell Biol 5:25–40

Gray MW, Cedergren R, Abel Y, Sankoff D (1989) On the evolutionary origin of the plant mitochondrion and its genome. Proc Natl Acad Sci USA 86:2267–2271

Gunderson JH, Elwood H, Ingold A, Kindle K, Sogin ML (1987) Phylogenetic relationship between chlorophytes, chrysophytes and oomycetes. Proc Natl Acad Sci USA 84:5823–5827

Jukes TH, Osawa S (1990) The genetic code in mitochondria and chloroplasts. Experientia 46:1117–1126

Karlovsky P, Prell HH (1991) The TRP1 gene of *Phytophthora parasitica* encoding indole-3-glycerol-phosphate synthase N-(5′-phosphoribosyl)-anthranilate isomerase: structure and evolutionary distance from homologous fungal genes. Gene, in press

Leedale G (1974) How many are the kingdoms of organisms? Taxon 23:261–270

LeJohn HB (1971) Enzyme regulation, lysine pathways and cell wall structures as indicators of major lines of evolution of fungi. Nature 231:164–168

Levings CS III, Brown GG (1989) Molecular biology of plant mitochondria. Cell 56:171–179

Mann C (1991) Lynn Margulis: Science's unruly earth mother. Science 252:378–381

Margulis L (1981) Symbiosis in cell evolution. Freeman, San Francisco

Margulis L, Schwartz KV (1988) Five kingdoms. Freeman, San Francisco

Mühlbauer E (1988) Molekulargenetische Untersuchungen an *Phytophthora* spec. PhD dissertation, University of Göttingen, Germany, pp 193–200

Osawa S, Jukes TH (1989) Codon reassignment (codon capture) in evolution. J Mol Evol 28:271–278

Osawa S, Muto A, Jukes TH, Ohama T (1990) Evolutionary changes in the genetic code. Proc R Soc Lond B 241:19–28

Palmer JD (1990) Contrasting modes and tempos of genome evolution in land plant organelles. Trends Genet 6:115–120

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

Roe BA, Ma D-P, Wilson RK, Wong JF-H (1985) The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. J Biol Chem 260:9759–9774

Shaw DS (1983) The perenosporales—a fungal geneticist's nightmare. In: Buczacki ST (ed) Zoosporic plant pathogens. Academic Press, London, pp 85–121

Unkles SE, Moon RP, Hawkins AR, Duncan JM, Kinghorn JR (1991) Actin in the oomycetous fungus *Phytophthora infestans* is the product of several genes. Gene 100:105–112

Whittaker RH (1969) New concepts of kingdoms of organisms. Science 163:150–160

Received June 10, 1991/Revised September 16, 1991