Hyaloraphidium curvatum: A Linear Mitochondrial Genome, tRNA Editing, and an Evolutionary Link to Lower Fungi

Lise Forget,* Jana Ustinova,† Zhang Wang,* Volker A. R. Huss,† and B. Franz Lang*

*Program in Evolutionary Biology, Département de Biochimie, Canadian Institute for Advanced Research, Université de Montréal, 2900 Boulevard Edouard-Montpetit, Montréal, Québec, Canada; and †Institut für Botanik und Pharmazeutische Biologie, Staudtstr. 5, Erlangen, Germany

We have sequenced the mitochondrial DNA (mtDNA) of *Hyaloraphidium curvatum*, an organism previously classified as a colorless green alga but now recognized as a lower fungus based on molecular data. The 29.97-kbp mitochondrial chromosome is maintained as a monomeric, linear molecule with identical, inverted repeats (1.43 kbp) at both ends, a rare genome architecture in mitochondria. The genome encodes only 14 known mitochondrial proteins, 7 tRNAs, the large subunit rRNA and small subunit rRNA (SSU rRNA), and 3 ORFs. The SSU rRNA is encoded in two gene pieces that are located 8 kbp apart on the mtDNA. Scrambled and fragmented mitochondrial rRNAs are well known from green algae and alveolate protists but are unprecedented in fungi. Protein genes code for apocytochrome *b;* cytochrome oxidase 1, 2, and 3, NADH dehydrogenase 1, 2, 3, 4, 4L, 5, and 6, and ATP synthase 6, 8, and 9 subunits, and several of these genes are organized in operon-like clusters. The set of seven mitochondrially encoded tRNAs is insufficient to recognize all codons that occur in the mitochondrial protein genes. When taking into account the pronounced codon bias, at least 16 nuclear-encoded tRNAs are assumed to be imported into the mitochondria. Three of the seven predicted mitochondria-encoded tRNA sequences carry mispairings in the first three positions of the acceptor stem. This strongly suggests that these tRNAs are edited by a mechanism similar to the one seen in the fungus *Spizellomyces punctatus* and the rhizopod amoeba *Acanthamoeba castellanii*. Our phylogenetic analysis confirms with overwhelming support that *H. curvatum* is a member of the chytridiomycete fungi, specifically related to the Monoblepharidales.

Introduction

Hyaloraphidium curvatum Korschikoff is a rare representative of freshwater nanoplanktons (Korschikoff 1931; Pringsheim 1963) and is the only species within the genus *Hyaloraphidium* that is currently available in axenic culture. It has colorless cells that are short, crescent-like, and nonmotile. The species reproduces by spores that resemble the mature organisms even within mother cells (autospores). These spores are arranged in series of four to eight. Images of *H. curvatum* can be inspected at the Protist Image Database (PID; http://megasun.bch.umontreal.ca/protists/protists.html). Based on the similarities in morphology and reproduction, the entire genus *Hyaloraphidium* has been traditionally classified with algae from the autosporine green algal family Ankistrodesmaceae (syn. Selenastraceae), Chlorophyta (Korschikoff 1931; Komárek and Fott 1983). Questions about the validity of this affiliation emerged in the 1980s because of the absence of photosynthetic pigments (Kiss 1984; Marvan, Komárek, and Comas 1984).

We have recently confirmed by electron-microscopical studies, and by phylogenetic analyses with the nuclear 18S rRNA sequence, that *H. curvatum* indeed does not belong to the green algae but rather to chytridiomycete or zygomycete fungi (Ustinova, Krienitz, and Huss 2000). However, its exact phylogenetic position could not be resolved, as statistical support for its affiliation

to any particular lineage of lower fungi was low. As also documented by others (cf., Nagahama et al. 1995; Jensen et al. 1998; Berbee, Carmean, and Winka 2000; James et al. 2000), nuclear 18S rRNA data fail to resolve phylogenetic relationships within chytridiomycetes, zygomycetes, or even within ascomycetes. Therefore, we reanalyzed the phylogenetic position of *H. curvatum* with alternative sequence data, i.e., a set of concatenated mitochondrial proteins.

Mitochondrial genome sequences have proven to provide valuable information for resolving evolutionary relationships among the various eukaryotic lineages. These data have the potential of exceeding the resolving power of nuclear genes because the evolutionary history of mitochondrial genes can be retraced to a relatively recent (~1.5 Byr), most likely single, endosymbiotic event involving an alpha-Proteobacterium that gave rise to the mitochondrion (for a recent review, see Lang, Gray, and Burger 1999). In addition, a set of five, highly conserved protein-coding genes, *cob, cox1, 2, 3* and *atp6, 9,* is present in essentially all mitochondrial DNAs (mtDNAs) of fungi, animals (only *atp9* is absent), and eukaryotes in general, with the notable exception of only a few most unusual, highly derived protists (for a review, see Gray et al. 1998). These mitochondrial protein sequences, when concatenated, provide a large and information-rich data set for the inference of deep, or difficult to resolve, evolutionary divergences. Still at the present time, the available mtDNA data are biased toward a phylogenetically limited range of eukaryotes, primarily encompassing animals (132, most of which are vertebrates), whereas fungi (5, four of which are ascomycetes), land plants (4), and protists (22) are underrepresented. A few additional, currently unpublished fungal and protist mtDNA sequences are available from us: http://megasun.bch.umontreal.ca/ogmp/projects/sumprog.

html; http://megasun.bch.umontreal.ca/People/lang/FMGP/ progress.html. We will show that despite this bias, the current mitochondrial database contains enough representatives to clearly place *H. curvatum* into one of the lower fungal lineages, with overwhelming support.

In order to determine the evolutionary affiliation of *H. curvatum* within the lower fungi, we have sequenced and analyzed its complete mtDNA and compared it with mtDNAs from selected representatives of major lower fungal lineages (Blastocladiales, Chytridiales, Monoblepharidales, Spizellomycetales, and Zygomycota). In the following study, we will (1) analyze the unusual mitochondrial genome architecture of the *H. curvatum* mtDNA; (2) describe the gene content, discuss the apparent correlation between the strongly reduced set of mitochondrial tRNA genes and codon bias, and the presence of a gene in pieces; (3) provide preliminary evidence for mitochondrial 5′ tRNA editing; and finally (4) present the results of our phylogenetic analyses.

Materials and Methods
Fungal Strains, Culture of *H. curvatum*, and mtDNA Isolation

*Rhizophydium*136 was kindly supplied by Dr. Joyce Longcore (University of Maine), *Monoblepharella*15 by Marilyn Mollicone (University of Maine), and *Mortierella verticillata* by Kerry O'Donnell (National Center for Agricultural Utilization Research, Peoria, Ill.). *Hyaloraphidium curvatum* SAG 235-1 was obtained from the Culture Collection of Algae at the University of Göttingen (http://www.gwdg.de/~epsag/ phykologia/catalogue/abc/h.htm). The same strain, designated CCAP 235/1, can be ordered from the Culture Collection of Algae and Protozoa, U.K. (http://www. ife.ac.uk/ccap/). Cells were grown in the recommended medium (*Polytoma* medium) at 15°C without aeration but with light shaking to prevent sedimentation. Cells (25 g wet weight) were harvested after 1 week and broken with glass beads. Total DNA was extracted according to standard procedures (http://megasun.bch. umontreal.ca/People/lang/FMGP/methods/mtDNA. html), and the bulk of high molecular weight polysaccharides was removed by centrifugation (60 min at 120,000*g*). Mitochondrial and nuclear DNA were separated by CsCl-bisbenzimide isopycnic centrifugation, with mtDNA forming a single A+T–rich band.

Cloning and DNA sequencing

Mitochondrial DNA of *H. curvatum* was physically sheared by nebulization (Okpodu et al. 1994), and a size fraction of 1,300–4,000 bp was recovered after agarose gel electrophoresis. The DNA was incubated with a mixture of T7 DNA polymerase and *Escherichia coli* DNA polymerase I (the Klenow fragment) to generate blunt ends and was then cloned into the *Eco*RV cloning site of the phagemid pBFL6 (B. F. Lang, unpublished data). This cloning vector has a size of 1.85 kbp and contains a downsized origin of plasmid replication, an M13 origin of replication, a chloramphenicol resistance marker, and a *lac*Z gene containing a short multicloning-site.

Recombinant plasmids containing mtDNA inserts were identified by colony hybridization using mtDNA as a probe. Clones contained in this random library encompassed the *Hyaloraphidium* mitochondrial genome, including most of the inverted repeats at its linear ends (1.24 of 1.43 kbp).

The mtDNA sequences of *H. curvatum*, *Spizellomyces punctatus*, *Rhizophydium*136, and *Schizophyllum commune* have been deposited in GenBank (accession numbers AF402142, AF404303, AF404306, and AF402141, respectively).

DNA Sequencing and Data Analysis

DNA sequencing was performed on a double-dye Li-Cor 4200L apparatus, using plasmid DNAs as a template, end-labeled primers, and a cycle-sequencing protocol (ThermoSequenase, Amersham).

Sequences were assembled using the XGAP package (Staden, Beal, and Bonfield 1998). Feature annotations were stored in masterfile format that integrates annotations with primary sequences, and this masterfile was regularly synchronized with the growing XGAP sequence database, using tools developed by the Organelle Genome Megasequencing Program (OGMP; http://megasun.bch. umontreal.ca/ogmp/ogmpid.html). The FASTA program (Pearson 1990) was employed for similarity searches in local databases, and the BLAST network service (Altschul et al. 1990) was used for remote searches in GenBank at the National Center for Biotechnology Information. Custom-made batch utilities were used for submitting queries and browsing the output (BBLAST, TBOB, BFASTA, and FOB). A number of additional programs, including multiple sequence file manipulation, preprocessing, conversion, and batch utilities for XGAP, FASTA, and GDE, as well as a masterfile maintenance suite, have been developed by the OGMP. These utilities are described in more detail and are available through the OGMP website.

Phylogenetic Analysis

Multiple protein alignments of the concatenated Cox1, Cox2, Cox3, Cob, Atp6, and Atp9 protein sequences were performed with the CLUSTAL W program (Thompson, Higgins, and Gibson 1994), which was launched from GDE (Genetic Data Environment; Smith et al. 1994). Only unambiguously aligned amino acid positions (a total of 1,305), with at least 10% amino acid identity among the 33 selected species, were used in the phylogenetic analysis. For tree construction, we used either maximum likelihood (PUZZLE 4.02; Strimmer and von Haeseler 1996) or distance approaches (PHYLIP 3.6 a2.1; Felsenstein 2001). A distance table was calculated using the most recent implementation of PROTDIST, which allows a Jin-Nei correction for unequal rates of change at different amino acid positions (alpha-version of this program in PHYLIP release 3.6 a), and the tree topology was inferred using WEIGHBOUR (Bruno, Socci, and Halpern 2000). Bootstrap analysis (1,000 replicates) was performed according to Felsenstein (Felsenstein 1985).

Fig. 1.—Gene and physical map of the *H. curvatum* mitochondrial genome. The linear mtDNA is shown in an open, circular format to facilitate comparisons with circular mapping mitochondrial genomes. Black blocks represent genes and ORFs, which are all transcribed clockwise in this representation; the light-gray block marks the ORF in the cob intron. Names of tRNA genes are indicated by the amino acid (one-letter code) they specify (M_f, initiator *trnM[cau]*). The middle circle shows the size scale, and the inner ring the restriction fragments after digestion with the enzyme *Eco*RI. Three potential operon-like structures (*cox,* rRNA, and *nad* genes) are indicated by underlining of gene names. The linear ends of the mtDNA are marked by filled circles, and the inverted repeats are indicated by arrows. Sequence alignments of the two repeat element families rep1 and rep2 are given at the bottom of the figure; their respective position and directionality (orientation of triangles) is indicated in the map.

Table 1
Gene Content and Introns in Fungal and Animal mtDNAs

Genes	CHYTRIDIOMYCETES			ZYGOMYCETES	ASCOMYCETES		ANIMALS
	H. curvatum	*S. punctatus*	*A. macrogynus*	*R. stolonifer*	*S. cerevisiae*[a]	*S. pombe*	*H. sapiens*
rns, rnl	■	■	■	■	■	■	■
atp6, 8	■	■	■	■	■	■	■
atp9	■	■	■	■	■	■	○
cob	■	■	■	■	■	■	■
cox1-3	■	■	■	■	■	■	■
nad1-6[b]	■	■	■	■	○	○	■
trnA-W	7	8	25	25	25	25	23
rnpB	○	○	○	■	■	■	○
rps3	○	○	■	■	■	■	○
ORFs[c]	3	5	3	4	4	0	0
intronI[d]	1	12	26	9	9	2	0
intronII	0	0	2	0	4	1	0

NOTE.—Filled squares indicate presence, open circle absence of a gene. Gene maps of *S. punctatus, R. stolonifer, H. curvatum,* and *S. pombe* mtDNAs are deposited at http://megasun.bch.umontreal.ca/People/lang/FMGP/.

[a] The reported number of introns applies to the long form version of yeast mtDNA, GenBank accession number SCE011856.

[b] Includes *nad1-4, nad4L,* and *nad5-6*. Note that several ascomycete mtDNA (e.g., *Aspergillus nidulans, Pichia canadensis*) also encode this set of *nad* genes.

[c] Only ORFs that have predicted initiation codons and are greater than 100 amino acids in length are listed.

[d] IntronI and IntronII are introns of group I and group II, respectively.

Results and Discussion

### Unusual Genome Architecture of *H. curvatum* mtDNA

Figure 1 depicts the physical and gene map of the 29.97-kbp *H. curvatum* mtDNA. As predicted from the sequence assembly of randomly cloned DNA fragments, and as independently confirmed by restriction analysis with eight different restriction enzymes (*Bst*XI, *Cla*I, *Eco*RI, *Hae*II, *Hpa*I, *Sac*I, *Sal*I, and *Sca*I) and by pulsed-field electrophoresis of the purified mtDNA (data not shown), the majority of the DNA molecules occur in monomeric, linear configuration, with identical inverted repeats of about 1.43 kbp at both ends. Cloning of the linear DNA ends was less efficient than the central part of the genome, and ~190 bp on both ends were not recovered at all (estimated by restriction analysis with double-stranded DNA fragments as the marker), which might be because of terminally attached proteins or covalently closed terminal hairpin structures (Nosek et al. 1998). Because of the potential presence of such terminal complexes, our estimates of the genome size that are only based on restriction analysis might be imprecise. The known part of the inverted terminal repeat sequence does not code for genes or ORFs, lacks any obvious potential secondary structure, and is not composed of multiple, tandemly arranged telomere-like repeat elements.

It should be noted that linear mtDNAs are more common than usually assumed. Many of the circular mapping mtDNAs are large concatemers, which are likely the product of a rolling-circle type of replication (Maleszka, Skelly, and Clark-Walker 1991; Bendich 1996). These molecules have to be distinguished from linear mtDNAs that occur predominantly as monomers, and these have been identified in several protists and in ascomycete fungi (Dinouel et al. 1993; Nosek et al. 1995; for a recent review, see Nosek et al. 1998). The linear mtDNAs of some fungal species possess cova-lently closed, single-stranded DNA termini (i.e., single-stranded loops connect the two DNA strands, at both ends) and carry long inverted repeats that are similar in size to the ones identified in the *H. curvatum* mtDNA. We have tested by PCR experiments whether the *H. curvatum* mtDNA also has covalently closed, single-stranded DNA termini, but no specific DNA amplification products were obtained. Alternatively, long inverted terminal repeats also occur in numerous linear, mitochondrial plasmids of fungi and plants. These replicate through participation of terminally attached proteins that serve as primers for synthesis of the complementary DNA strand (Sakaguchi 1990), a mechanism that might have also been adopted by the *H. curvatum* mtDNA. The strict conservation of its repeat sequences (no nucleotide difference in a total of 10 kbp of sequence from the terminal repeat regions; homogenous terminal ends of identical length) implies an efficient copy correction mechanism.

Gene Content and Gene Organization

Table 1 lists the 23 genes and 3 ORFs (>80 amino acids, none of them similar to known proteins) residing in the *H. curvatum* mtDNA and compares them with those of selected fungal and animal mitochondrial genomes. A single, 961-bp-long group I intron is present which is inserted in the *cob* gene. It is a homolog of the second cob intron of *Allomyces macrogynus* with which it shares its insertion site and highly similar intronic ORFs (an omega-type ORF with potential endonuclease-maturase activities [Dujon et al. 1986; Paquin and Lang 1996]). Identified protein genes in *H. curvatum* mtDNA exclusively code for components involved in respiration and oxidative phosphorylation, namely, subunits of NADH:ubiquinone oxidoreductase (respiratory complex I), ubiquinone:cytochrome *c* oxidoreductase (complex III), cytochrome oxidase (complex IV), and ATP syn-
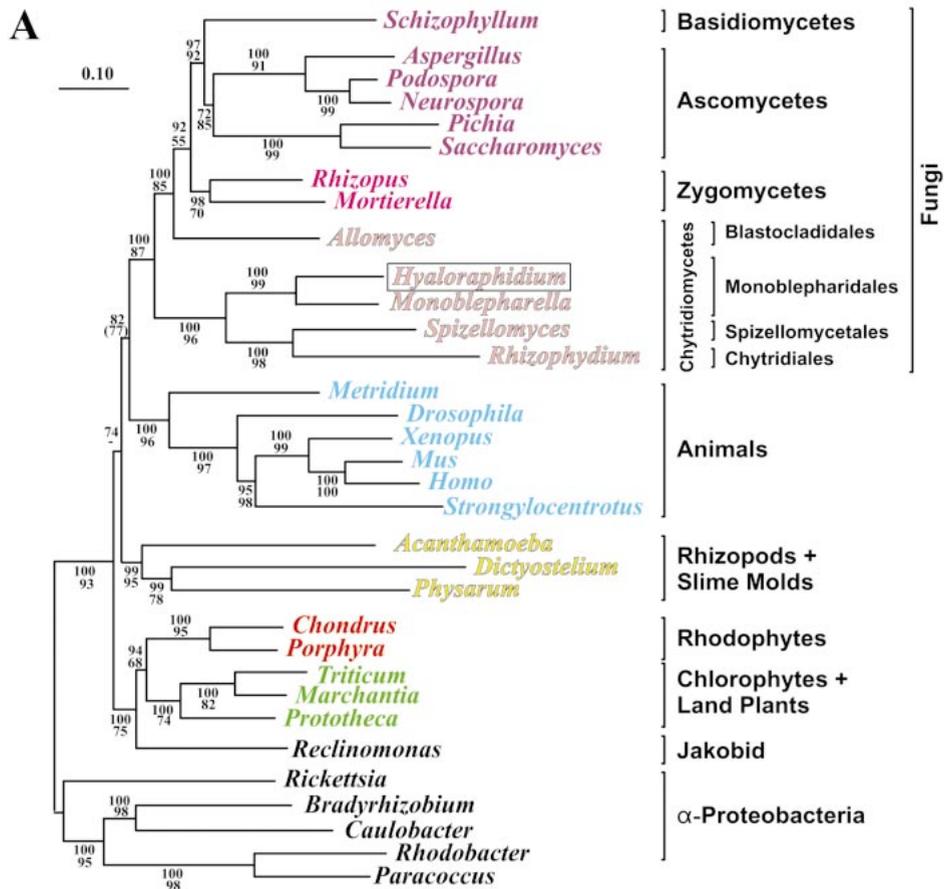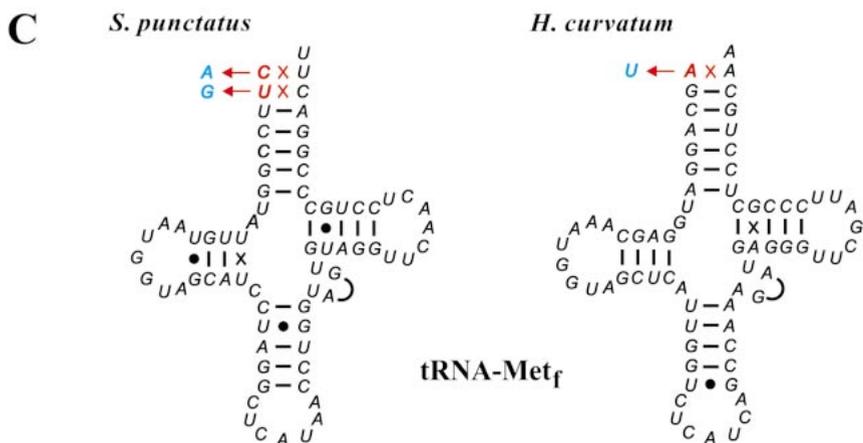
FIG. 2.—*A,* phylogenetic analysis. The phylogenetic tree was inferred from the concatenated protein sequences of Cox1, Cox2, Cox3, Cob, Atp6 and Atp9 (Atp9 of animals was not included because it is nuclear encoded). Only unambiguously aligned portions of these protein sequences were used in the analysis. The tree shown was inferred employing the most recent implementation of (1) PROTDIST (beta test-version; Felsenstein 2001), which calculates a distance table using a Jin-Nei correction for unequal rates of change at different amino acid positions; and (2) WEIGHBOR (Bruno, Socci, and Halpern 2000), which is a likelihood-based weighted neighbor-joining method. The variation coefficient was 0.67. Bootstrap support (%, upper number) was calculated from 1,000 replicates using PARBOOT parallel bootstrapping (Felsenstein 1985). The scale bar (0.10) denotes mean number of substitutions per site. The same tree topology was obtained with TREE-PUZZLE (see lower number for support values), except for the unresolved divergence point of the animal-fungal lineages and the placement of the rhizopod-slime

thase (complex V). These genes belong to the standard repertoire of fungal mtDNAs. Note that *rps3* and *rnpB* (RNA subunit of RNase P) are not encoded in the *H. curvatum* or any other known chytridiomycete mtDNA (table 1).

Hyaloraphidium curvatum mitochondrial genes are less densely packed than usual (including fungi such as *Schizosaccharomyces pombe* and *A. macrogynus*; Gray et al. 1998), covering only 67% of the entire sequence and leaving intergenic regions from as few as seven up to several hundred base pairs. The largest noncoding region of 2.5 kbp includes the terminal repeat adjacent to *orf182*. Like the terminal repeat, the empty stretch between the repeat region and *orf182* appears to be non-coding, contains no obvious potential secondary structure, and has a G+C content similar to that of coding regions (49.4%). Genes are encoded on only one DNA strand, and three operon-like gene clusters have been recognized (fig. 1), a ribosomal RNA cluster, a cytochrome oxidase, and an NADH dehydrogenase gene cluster (the respective gene names are underlined). The *H. curvatum* gene clustering is secondary because the gene order within these clusters is different from comparable eubacterial or protist mitochondrial operons (Lang et al. 1997). It is likely driven by similar principles as in eubacteria (concerted transcriptional or translational control [or both] of genes that function in common biological processes; Dandekar et al. 1998; Niehrs and Pollet 1999). Among other organisms featuring secondary mitochondrial gene clustering, *Chrysodidymus synuroideus* exhibits the most evident examples involving riboprotein genes (see also Wolff et al. 1994 for the green alga *Prototheca wickerhamii*; Chesnick et al. 2000).

The A+T content of *H. curvatum* mtDNA is atypically low relative to other fungal mtDNAs, both in coding (61%) and in intergenic regions (57%). The more G+C–rich fungal mtDNAs are characterized by intergenic regions with numerous G+C–rich, folded repeat elements, such as DHEs (double-hairpin elements) in *Allomyces* species (Paquin, Laforest, and Lang 2000) or Pst-palindromes in *Neurospora crassa* (Yin, Heckman, and RajBhandary 1981). Although we did find a few repeats in the *H. curvatum* mtDNA (fig. 1), they are not particularly G+C–rich and do not occur frequently enough to explain the unusual, overall high G+C content of intergenic regions. We think that the high overall G+C content in the *H. curvatum* mtDNA results from differences in its mutational or repair mechanisms (or both). It will be interesting to see whether a similar G+C bias is found in mitochondrial genomes of species closely related to *H. curvatum*, and whether mutational changes include frequent A to T transversions, as in many A+T–rich mtDNAs.

The repeats in the *H. curvatum* mtDNA fall into two groups. The shorter motif (rep1; fig. 1) might be involved in transcription initiation based on its orientation and distribution in the genome. The rep2 motif is located in four out of six instances immediately downstream of protein-coding genes (fig. 1); its potential function, if any, might be in RNA 3'-end processing or transcription termination.

A Reduced Set of Mitochondrial tRNAs and Pronounced Codon Bias

Similar to the situation in the chytridiomycete *S. punctatus* (fig. 2B; Laforest, Roewer, and Lang 1997), and contrary to other fungi, *H. curvatum* has a reduced set of only seven mitochondrially encoded tRNAs which is insufficient to recognize all codons present in the mitochondrial protein-coding genes. Therefore, the remainder of its tRNAs are likely nuclear encoded and imported into the mitochondria. Unlike *S. punctatus* that

---

←

mold clade. It should be noted that the depicted tree further confirms divergences much deeper than fungi, e.g., the previously reported phylogenetic association of rhizopods and slime molds based on mitochondrial data (Burger et al. 2000). The support for this affiliation is much higher in this analysis, likely because of the inclusion of more protein sequences and species. Another recent analysis with concatenated nuclear protein sequences has come to the same conclusion, with robust support (Baldauf et al. 2000). For the first time, there is also a demonstration for the divergence of *Reclinomonas americana* prior to the red algal-green plant lineage. Finally, the recent availability of more data in the alpha-Proteobacteria has enabled us to resolve a branching position of *Rickettsia prowazekii* (Andersson et al. 1998) closest to mitochondria, still placed at a long distance from the mitochondrial radiation. Species (Genbank accession numbers): *Homo*, *Homo sapiens* (J01415); *Mus*, *Mus musculus* (J01420); *Xenopus*, *Xenopus laevis* (M10217); *Drosophila*, *Drosophila yakuba* (X03240); *Strongylocentrotus*, *Strongylocentrotus purpuratus* (X12631); *Metridium*, *Metridium senile* (AF000023); *Allomyces*, *A. macrogynus* (U41288); *Hyaloraphidium*, *H. curvatum* (this study; AF402142) *Saccharomyces*, *Saccharomyces cerevisiae* (AJ011856); *Pichia*, *Pichia canadensis* (*Hansenula wingei*) (D31785); *Aspergillus*, *Aspergillus* (*Emericella*) *nidulans* (X00790, X15441, X06960, J01387, X01507, M35967); *Neurospora*, *N. crassa* (X18150, K00825, A2855, K01181); *Podospora*, *Podospora anserina* (X55026); *Chondrus*, *Chondrus crispus* (Z47547); *Porphyra*, *Porphyra purpurea* (AF114794); *Schizophyllum*, *S. commune* (this study; AF402141); *Spizellomyces*, *S. punctatus* (this study; AF404303); *Rhizophydium*, *Rhizophydium*136 (this study; AF404306); *Triticum*, *Triticum aestivum* (Y00417, X01108, P15953 (protein), P007747 (protein), M24084, X54621); *Marchantia*, *Marchantia polymorpha* (M68929); *Prototheca*, *P. wickerhamii* (U2970); *Rickettsia*, *R. prowazekii* (AJ235270–73); *Bradyrhizobium*, *Bradyrhizobium japonicum* (X54800, U33883, AAF78800, AAF78802, AAF78811, P51131); *Rhodobacter*, *Rhodobacter sphaeroides* (X62645, M57680, C45164 (protein), X56157, Y12313); *Paracoccus*, *Paracoccus denitrificans* (X05829, M17522, X05934, X05828); and *Caulobacter*, *Caulobacter crescentus* (sequences kindly provided by the *Caulobacter* project at http://www.tigr.org/tdb/mdb/mdbinprogress.html). Protein sequences of *Monoblepharella*, *Monoblepharella*15; *Rhizopus*, *Rhizopus stolonifer*; and *Mortierella*, *M. verticillata* can be downloaded from http://megasun.bch.umontreal.ca/People/lang/FMGP/proteins/. *B*, Partially shared, reduced sets of tRNA genes in *H. curvatum* and *S. punctatus*. tRNA species are identified by their corresponding amino acid (one-letter code) and anticodon sequence. Filled black or red rectangles, gene present; filled red rectangles, tRNA editing demonstrated or predicted; open circles, gene absent. It should be noted that *S. punctatus* has a leucine tRNA decoding UAG stop codons, which is not present in *H. curvatum*, as it uses the standard translation code. *C*, Similar tRNA editing in *H. curvatum* and *S. punctatus* mtDNAs. The initiator methionine tRNA of *S. punctatus* has been shown to be edited (Laforest et al. 1997), and we predict editing in the *H. curvatum* tRNA because of a terminal, unpaired A-A in the acceptor stem that would otherwise abolish its function. Mispaired nucleotides are marked red, and nucleotide replacements are marked blue.

Table 2
Codon Usage in Genes and ORFs of *H. curvatum* mtDNA[a]

a.a./codon[b]	Genes	ORFs	a.a./codon	Genes	ORFs	a.a./codon	Genes	ORFs	a.a./codon	Genes	ORFs
F TTT	103	26	S TCT	59	38	**Y TAT**	82	35	C TGT	1	3
F TTC	192	22	S TCC	55	20	**Y TAC**	101	13	C TGC	12	2
L TTA	56	29	S TCA	—	—	* TAA	13	3	* TGA	—	—
L TTG	76	23	S TCG	14	6	* TAG	1	1	**W TGG**	87	22
L CTT	187	28	**P CCT**	86	30	H CAT	25	21	R CGT	56	23
L CTC	175	26	**P CCC**	20	13	H CAC	66	9	R CGC	12	21
L CTA	86	15	**P CCA**	50	13	Q CAA	53	22	R CGA	26	9
L CTG	114	10	**P CCG**	11	4	Q CAG	20	18	R CGG	1	2
I ATT	187	26	T ACT	108	20	N AAT	37	17	S AGT	61	21
I ATC	92	16	T ACC	76	15	N AAC	72	17	S AGC	55	7
I ATA	—	—	T ACA	13	7	K AAA	45	33	R AGA	—	1
M ATG ...	182	24	T ACG	17	8	K AAG	9	24	R AGG	—	—
V GTT	82	28	A GCT	209	36	**D GAT**	42	45	G GGT	84	22
V GTC	84	7	A GCC	96	9	**D GAC**	49	19	G GGC	13	5
V GTA	72	6	A GCA	—	4	**E GAA**	51	15	G GGA	172	13
V GTG	56	10	A GCG	—	4	**E GAG**	27	30	G GGG	32	13

[a] The correlation of amino acids (a.a.) and codons reflects the standard translation code, as the analysis of multiple protein alignments did not reveal any reassignments of stop codons (e.g., TAG, leucine; *S. punctatus*) or other deviations from the standard code (e.g., CTN, threonine, ATA, methionine; yeast mitochondria). Numbers indicate the total numbers of codons in either identified protein coding genes or ORFs (including both free-standing and intronic). **Bold**, codons recognized by mitochondrially encoded tRNAs.

[b] Cognate amino acid in one-letter code; asterisk, stop codon.

decodes UAG stop codons as leucine by a specific mitochondrial *trnL(cua)* (fig. 2*B*), the standard genetic code is used in translation of *H. curvatum* mitochondrial-encoded proteins (table 2).

The third codon positions of *H. curvatum* mitochondrial protein-coding genes are less biased toward A and T (table 2), compared with mitochondrial genes of other fungi (Gray et al. 1998). Furthermore, there is complete lack of ATA, TCA, and AGG codons, and AGA, GCA, and GCG are lacking in identified protein-coding genes and are rarely used in ORFs. It is conceivable that this bias indicates either that the respective tRNAs are not imported from the cytosol or that the imported tRNAs are not optimized to recognize these codons efficiently (or both), and thus are avoided in the highly expressed, identified genes.

Given that all chytridiomycete fungi other than the Blastocladiales have strongly reduced sets of mitochondrially encoded tRNAs and pronounced codon bias, they could serve as attractive models to study the concerted evolution of tRNA specificity and codon bias. To pinpoint common evolutionary trends in codon bias, it will be particularly interesting to analyze sister taxa of *H. curvatum*, such as *Monoblepharella*15 (see fig. 2*A*) and *Harpochytrium* species (M. J. Laforest, personal communication).

The Small Subunit rRNA is Encoded in Two Pieces

The mitochondrial rRNA sequences have predicted sizes of 2,841 nt (large subunit) and 1,477 nt (small subunit) and can be folded into conventional eubacteria-like secondary structures. Whereas the large subunit rRNA is encoded in a single gene, the SSU rRNA has to be assembled from two genes (fig. 1), *rns_a* (481 bp, coding for the 5′ portion) and *rns_b* (996 bp, the 3′ part of the SSU rRNA). The break point of the SSU rRNA molecule is in a highly variable region in fungal mitochondrial sequences, corresponding to nucleotides 590–649 of the *E. coli* RNA molecule. The two *H. curvatum* mitochondrial SSU rRNA gene fragments are encoded on the same DNA strand but in inverted order (*rns_b* upstream of *rns_a*), and are separated from each other by a 3.2-kbp region that codes for *rnl* and two tRNAs.

Fragmented and scrambled ribosomal genes have been first described in ciliates (e.g., *Tetrahymena pyriformis*; Schnare et al. 1986). They occur frequently in mtDNAs of green algae such as Chlamydomonas species, *Pedinomonas minor*, and *Scenedesmus obliquus* (e.g., Boer and Gray 1988; Turmel et al. 1999; Nedelcu et al. 2000) and in alveolates such as *Plasmodium falciparum* (Gillespie et al. 1999) and *Theileria parva* (Kairo et al. 1994) but have not been seen in fungal mtDNAs. In most of those cases, it has been shown that the break points occur in structurally little conserved, predominantly single-stranded loop regions of the rRNA molecules, that splicing of the rRNA does not occur, and that the rRNA pieces have the potential to correctly assemble by base pairing. We assume that the *H. curvatum* mitochondrial SSU rRNA pieces are also combined by intermolecular base pairing because we find no indication for a groupI or groupII transspliced intron structure.

A Second Instance of 5′ tRNA Editing in Lower Fungi

RNA editing has been shown to alter the sequences of various nuclear and organellar transcripts in a wide range of organisms, and it is usually essential for expression because it restores the sequence of the functional gene product. 5′ tRNA editing was first discovered in the rhizopod amoeba *Acanthamoeba castellanii* (Lonergan and Gray 1993) and later in the chytridiomycete fungus *S. punctatus* (Laforest, Roewer, and Lang

1997). It is characterized by the replacement of one to three mismatched nucleotides at the 5′ end of the acceptor stem, thereby reconstituting canonical Watson-Crick base pairing. The same type of tRNA editing very likely occurs in three of the seven *H. curvatum* mitochondrial tRNAs (fig. 2*B*). For example, the gene sequence of tRNA-Met$_f$ of *H. curvatum* displays a non-canonical A-A mispairing at the 5′ end of its acceptor stem (fig. 2*C*). This would interfere both with 5′-end processing of its tRNA precursor and with its translational function. A comprehensive evolutionary comparative analysis of 5′ tRNA editing in several lower fungi, including *H. curvatum*, will be published elsewhere (M. J. Laforest, personal communication). For the structures of all *H. curvatum* mitochondrial tRNA species, see http://megasun.bch.umontreal.ca/People/lang/species/hyalo/hyalo.html.

Hyaloraphidium curvatum is a Member of the Monoblepharidales Fungi

Fungal phylogeny has been largely studied based on single, nuclear genes, such as the SSU rRNA, tubulins, RPB1 (encoding the largest subunit of RNA polymerase II), etc. (e.g., Bruns et al. 1992; Liu, Whelen, and Hall 1999; James et al. 2000; Keeling, Luker, and Palmer 2000). However, these data sets are problematic because they either contain insufficient phylogenetic signal (as is evidently the case with rRNA sequences; e.g., Bruns et al. 1992; Nagahama et al. 1995; Jensen et al. 1998; Berbee, Carmean, and Winka 2000; James et al. 2000), or the taxon sampling is currently poor (RBP1). In addition, these nuclear phylogenies are often plagued by inequality in evolutionary rates (in particular, tubulins, in the higher fungal lineages), which casts doubt on the validity of phylogenetic inferences. Attempts to overcome these problems have been made by using multiple, concatenated nuclear protein sequences. This approach did not significantly improve the resolution of the fungal tree (Baldauf et al. 2000).

In contrast, mitochondrial protein data appear to be suited to resolve deep evolutionary divergences in the fungi, with high support (Paquin et al. 1997; Burger et al. 1999). This is most likely because of more equal evolutionary rates throughout most taxa and the availability of several highly conserved protein sequences, with reasonably broad taxon sampling. When inferring phylogenetic trees with six concatenated mitochondrial proteins (Cox1, 2, 3, Cob, and Atp6, 9), *H. curvatum* is placed as a sister species to *Monoblepharella* sp., in the vicinity of the Spizellomycetales (*S. punctatus*) and the Chytridiales (*Rhizophydium*136; fig. 2*A*). This topology is highly supported both in a distance (100% bootstrap) and a maximum likelihood approach (99% bootstrap), and further confirms that the chytridiomycetes are a paraphyletic group as postulated before (Paquin et al. 1997). When using single, or combinations of two, protein sequences from this data set, there is inadequate support for the varying tree topologies obtained (results not shown). Evidently, as in the case of single nuclear gene sequences (see earlier), there is not sufficient phylogenetic signal in single mitochondrial gene sequences to resolve global fungal phylogenies.

LITERATURE CITED

ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS, and D. J. LIPMAN. 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.

ANDERSSON, S. G. E., A. ZOMORODIPOUR, J. O. ANDERSSON, T. SICHERITZ-PONTÉN, C. M. ALSMARK, A. K. NÄSLUND, A.-S. ERIKSSON, H. H. WINKLER, and C. G. KURLAND. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature **396**:133–140.

BALDAUF, S. L., A. J. ROGER, I. WENK-SIEFERT, and W. F. DOOLITTLE. 2000. A kingdom-level phylogeny of eukaryotes based on combined protein data. Science **290**:972–977.

BENDICH, A. J. 1996. Structural analysis of mitochondrial DNA molecules from fungi and plants using moving pictures and pulsed-field gel electrophoresis. J. Mol. Biol. **255**:564–588.

BERBEE, M. L., D. A. CARMEAN, and K. WINKA. 2000. Ribosomal DNA and resolution of branching order among the Ascomycota: how many nucleotides are enough? Mol. Phylogenet. Evol. **17**:337–344.

BOER, P. H., and M. W. GRAY. 1988. Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. Cell **55**:399–411.

BRUNO, W. J., N. D. SOCCI, and A. L. HALPERN. 2000. Weighted neighbor joining: a likelihood-based approach to distance-based phylogeny reconstruction. Mol. Biol. Evol. **17**: 189–197.

BRUNS, T. D., R. VILGALYS, S. M. BARNS et al. (11 co-authors). 1992. Evolutionary relationships within the fungi: analyses of nuclear small subunit rRNA sequences. Mol. Phylogenet. Evol. **1**:231–241.

BURGER, G., D. SAINT-LOUIS, M. W. GRAY, and B. F. LANG. 1999. Complete sequence of the mitochondrial DNA of the red alga *Porphyra purpurea*: cyanobacterial introns and shared ancestry of red and green algae. Plant Cell **11**:1675–1694.

BURGER, G., Y. ZHU, T. G. LITTLEJOHN, S. J. GREENWOOD, M. N. SCHNARE, B. F. LANG, and M. W. GRAY. 2000. Complete sequence of the mitochondrial genome of *Tetrahymena pyriformis* and comparison with *Paramecium aurelia* mitochondrial DNA. J. Mol. Biol. **297**:365–380.

CHESNICK, J. M., M. GOFF, J. GRAHAM, C. OCAMPO, B. F. LANG, E. SEIF, and G. BURGER. 2000. The mitochondrial genome of the stramenopile alga *Chrysodidymus synuroideus*. Complete sequence, gene content and genome organization. Nucleic Acids Res. **28**:2512–2518.

DANDEKAR, T., B. SNEL, M. HUYNEN, and P. BORK. 1998. Conservation of gene order: a fingerprint of proteins that physically interact. Trends Biochem. Sci. **23**:324–328.

DINOUEL, N., R. DRISSI, I. MIYAKAWA, F. SOR, S. ROUSSET, and H. FUKUHARA. 1993. Linear mitochondrial DNAs of yeasts: closed-loop structure of the termini and possible linear-circular conversion mechanisms. Mol. Cell. Biol. **13**: 2315–2323.

DUJON, B., L. COLLEAUX, A. JACQUIER, F. MICHEL, and C. MONTEILHET. 1986. Mitochondrial introns as mobile genetic elements: the role of intron-encoded proteins. Basic Life Sci. **40**:5–27.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.

FELSENSTEIN, J. 2001. PHYLIP (phylogeny inference package). Version 3.6 a. Distributed by the author, University of Washington, Seattle.

GILLESPIE, D. E., N. A. SALAZAR, D. H. REHKOPF, and J. E. FEAGIN. 1999. The fragmented mitochondrial ribosomal RNAs of *Plasmodium falciparum* have short A tails. Nucleic Acids Res. **27**:2416–2422.

GRAY, M. W., B. F. LANG, R. CEDERGREN et al. (15 co-authors). 1998. Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res. **26**:865–878.

JAMES, T. Y., D. PORTER, C. A. LEANDER, R. VILGALYS, and J. E. LONGCORE. 2000. Molecular phylogenetics of the Chytridiomycota supports the utility of ultrastructural data in chytrid systematics. Can. J. Bot. **78**:336–350.

JENSEN, A. B., A. GARGAS, J. EILENBERG, and S. ROSENDAHL. 1998. Relationships of the insect-pathogenic order Entomophthorales (Zygomycota, Fungi) based on phylogenetic analyses of nuclear small subunit ribosomal DNA sequences (SSU rDNA). Fungal Genet. Biol. **24**:325–334.

KAIRO, A., A. H. FAIRLAMB, E. GOBRIGHT, and V. NENE. 1994. A 7.1 kb linear DNA molecule of *Theileria parva* has scrambled rDNA sequences and open reading frames for mitochondrially encoded proteins. EMBO J. **13**:898–905.

KEELING, P. J., M. A. LUKER, and J. D. PALMER. 2000. Evidence from beta-tubulin phylogeny that microsporidia evolved from within the fungi. Mol. Biol. Evol. **17**:23–31.

KISS, K. T. 1984. Special problems in studying phytoplankton associations (occurrence and ecological evaluation of *Hyaloraphidium contortum*). Acta Bot. Hung. **30**:269–276.

KOMÁREK, J., and B. FOTT. 1983. Das Phytoplankton des Süsswassers 7,1. Chlorophyceae (Grünalgen). Ordnung: Chlorococcales. Schweizerbart, Stuttgart.

KORSCHIKOFF, A. A. 1931. Notizen über einige neue apochlorotische Algen. Archiv Protistenkd. **74**:249–258.

LAFOREST, M. J., I. ROEWER, and B. F. LANG. 1997. Mitochondrial tRNAs in the lower fungus *Spizellomyces punctatus*: tRNA editing and UAG 'stop' codons recognized as leucine. Nucleic Acids Res. **25**:626–632.

LANG, B. F., G. BURGER, C. J. O'KELLY, R. J. CEDERGREN, G. B. GOLDING, C. LEMIEUX, D. SANKOFF, M. TURMEL, and M. W. GRAY. 1997. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature **387**: 493–497.

LANG, B. F., M. W. GRAY, and G. BURGER. 1999. Mitochondrial genome evolution and the origin of eukaryotes. Annu. Rev. Genet. **33**:351–397.

LIU, Y. J., S. WHELEN, and B. D. HALL. 1999. Phylogenetic relationships among ascomycetes: evidence from an RNA polymerase II subunit. Mol. Biol. Evol. **16**:1799–1808.

LONERGAN, K. M., and M. W. GRAY. 1993. Editing of transfer RNAs in *Acanthamoeba castellanii* mitochondria. Science **259**:812–816.

MALESZKA, R., P. J. SKELLY, and G. D. CLARK-WALKER. 1991. Rolling circle replication of DNA in yeast mitochondria. EMBO J. **10**:3923–3929.

MARVAN, P., J. KOMÁREK, and A. COMAS. 1984. Weighting and scaling of features in numerical evaluation of coccal green algae (genera of the Selenastraceae). Algol. Stud. **37**: 363–399.

NAGAHAMA, T., H. SATO, M. SHIMAZU, and J. SUGIYAMA. 1995. Phylogenetic divergence of the Entomophthoralean fungi: evidence from nuclear 18S ribosomal RNA gene sequence. Mycologia **87**:203–209.

NEDELCU, A. M., R. W. LEE, C. LEMIEUX, M. W. GRAY, and G. BURGER. 2000. The complete mitochondrial DNA sequence of *Scenedesmus obliquus* reflects an intermediate stage in the evolution of the green algal mitochondrial genome. Genome Res. **10**:819–831.

NIEHRS, C., and N. POLLET. 1999. Synexpression groups in eukaryotes. Nature **402**:483–487.

NOSEK, J., N. DINOUEL, L. KOVAC, and H. FUKUHARA. 1995. Linear mitochondrial DNAs from yeasts: telomeres with large tandem repetitions. Mol. Gen. Genet. **247**:61–72.

NOSEK, J., L. TOMASKA, H. FUKUHARA, Y. SUYAMA, and L. KOVAC. 1998. Linear mitochondrial genomes: 30 years down the line. Trends Genet. **14**:184–188.

OKPODU, C. M., D. ROBERTSON, W. F. BOSS, R. K. TOGASAKI, and S. J. SURZYCKI. 1994. Rapid isolation of nuclei from carrot suspension culture cells using a BioNebulizer. Biotechniques **16**:154–159.

PAQUIN, B., M. J. LAFOREST, L. FORGET, I. ROEWER, Z. WANG, J. LONGCORE, and B. F. LANG. 1997. The fungal mitochondrial genome project: evolution of fungal mitochondrial genomes and their gene expression. Curr. Genet. **31**:380–395.

PAQUIN, B., M. J. LAFOREST, and B. F. LANG. 2000. Double-hairpin elements in the mitochondrial DNA of *Allomyces*: evidence for mobility. Mol. Biol. Evol. **17**:1760–1768.

PAQUIN, B., and B. F. LANG. 1996. The mitochondrial DNA of *Allomyces macrogynus*: the complete genomic sequence from an ancestral fungus. J. Mol. Biol. **255**:688–701.

PEARSON, W. R. 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol. **183**:63–98.

PRINGSHEIM, E. G. 1963. Farblose Algen. G. Fischer Verlag, Stuttgart.

SAKAGUCHI, K. 1990. Invertrons, a class of structurally and functionally related genetic elements that includes linear DNA plasmids, transposable elements, and genomes of adeno-type viruses. Microbiol. Rev. **54**:66–74.

SCHNARE, M. N., T. Y. HEINONEN, P. G. YOUNG, and M. W. GRAY. 1986. A discontinuous small subunit ribosomal RNA in *Tetrahymena pyriformis* mitochondria. J. Biol. Chem. **261**:5187–5193.

SMITH, S. W., R. OVERBEEK, C. R. WOESE, W. GILBERT, and P. M. GILLEVET. 1994. The genetic data environment, an expandable GUI for multiple sequence analysis. Comput. Appl. Biosci. **10**:671–675.

STADEN, R., K. F. BEAL, and J. K. BONFIELD. 1998. The Staden package. Methods Mol. Biol. **132**:115–130.

STRIMMER, K., and A. VON HAESELER. 1996. Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. Mol. Biol. Evol. **13**:964–969.

THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

TURMEL, M., C. LEMIEUX, G. BURGER, B. F. LANG, C. OTIS, I.

PLANTE, and M. W. GRAY. 1999. The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*: two radically different evolutionary patterns within green algae. Plant Cell **11**:1717–1729.

USTINOVA, I., L. KRIENITZ, and V. A. R. HUSS. 2000. *Hyaloraphidium curvatum* is not a green alga, but a lower fungus; *Amoebidium parasiticum* is not a fungus, but a member of the DRIPs. Protist **151**:253–262.

WOLFF, G., I. PLANTE, B. F. LANG, U. KÜCK, and G. BURGER. 1994. Complete sequence of the mitochondrial DNA of the chlorophyte alga *Prototheca wickerhamii*. J. Mol. Biol. **237**: 75–86.

YIN, S., J. HECKMAN, and U. L. RAJBHANDARY. 1981. Highly conserved GC-rich palindromic DNA sequences flank tRNA genes in *Neurospora crassa* mitochondria. Cell **26**: 325–332.