9-1-2004

The Mitochondrial Genome of the Peronosporomycete Saprolegnia Ferax: Organization, Gene Content, and Nucleotide Sequence

W. Scott Grayburn

D.S.S. Hudspeth

M. Gane

M.E.S. Hudspeth

Original Citation
Grayburn, W.S., Hudspeth, D.S.S., Gane, M., and Hudspeth, M.E.S. (2004). The mitochondrial genome of the peronosporomycete Saprolegnia ferax: organization, gene content, and nucleotide sequence. Mycologia 96:981-989.

Mycologia, 96(5), 2004, pp. 981–989.

The mitochondrial genome of *Saprolegnia ferax*: organization, gene content and nucleotide sequence

W. Scott Grayburn
Deborah S. S. Hudspeth
Melody K. Gane
Michael E. S. Hudspeth[1]

Department of Biological Sciences, Northern Illinois University, DeKalb, Illinois 60115

Abstract: The mitochondrial genome of the peronosporomycete water mold *Saprolegnia ferax* has been characterized as a 46 930 bp circle containing an 8618 bp large inverted repeat (LIR). Eighteen reading frames encode identified subunits of respiratory complexes I, III, IV and V; 16 encode polypeptides of small and large mitoribosome subunits; and one encodes a subunit of the *sec*-independent protein translocation pathway. Of four additional putative reading frames three are homologues of those found in the related *Phytophthora infestans* genome. Protein encoding loci in the tightly compacted genome typically are arranged in operon-like clusters including three abutting and two overlapping pairs of reading frames. Translational RNAs include the mitochondrial small and large subunit rRNAs and 25 tRNA species. No tRNAs are encoded to enable translation of any threonine or the arginine CGR codons. The LIR separates the molecule into 19 274 bp large and 10 420 bp small single copy regions, and it encodes intact duplicate copies of four reading frames encoding known proteins, both rRNAs, and five tRNAs. Partial 3′ sequences of three additional reading frames are duplicated at single copy sequence junctions. Active recombination between LIR elements generates two distinctive gene orders and uses the duplicated 3′ sequences to maintain intact copies of the partially duplicated loci.

Key words: mitochondrial genes, mtDNA, Oomycetes, Peronosporomycetes, stramenopiles

INTRODUCTION

Ascertaining the genetic content and mitochondrial organization of lower eukaryotes has been invaluable in providing insight into mitochondrial origins and in demonstrating the diversity of mitochondrial genomes (reviewed in Gray et al 1998, Lang et al 1999). Genetically, variation in the number of identified protein-encoding loci ranges from 3 to 67, tRNAs from 0 to 27 and rRNAs from 0 to 3. The assembled sequence data physically describe the 6–77 kb range of genomes as compact gene-rich circular or linear molecules of high AT content. Despite these variations mtDNAs of recognized eukaryotic lineages are consistent in their overall genetic and physical organization. Thus, mitochondrial genomes of protists are potentially useful tools for resolving evolutionary relationships between organisms whose morphological and biochemical characters often are limited and/or difficult to obtain.

The Peronosporomycetes (Oomycetes) are a class of fungal-like protists included in an assemblage of lower eukaryotes often referred to as stramenopiles (Patterson 1989, Dick 2001). Members of this lineage are unified taxonomically by the presence of tripartite tubular hairs (stramenopiles) at some stage of their life cycle. The diverse collection of organisms also includes autotrophic chromophytes (chlorophyll $a + c$ containing algae such as chrysophytes, fucophytes, xanthophytes and diatoms), additional fungal-like heterotrophs (thraustochytrids, labyrinthulids and hyphochytriomycetes) and the heterotrophic bicosoecids. Complete mitochondrial DNA sequences have been determined for six of these organisms—the chromophytes *Chrysodidymus synuroideus* (Chesnick et al 2000), *Laminaria digitata* (Oudot-LeSeq et al 2002), *Ochromonas danica* (Burger et al 2002) and *Pylaiella littoralis* (Oudot-LeSeq et al 2001); the heterotrophic bicosecid *Cafeteria roenbergensis* (Burger 1999); and the peronosporomycete *Phytophthora infestans* (Paquin et al 1997). Like most protists, mitochondrial genomes of peronosporomycetes typically are compact mtDNAs and encode at least 30 polypeptides of known function.

Two distinct mtDNA organizational patterns have emerged from restriction endonuclease analyses of a variety of peronosporomycete taxa. The more prevalent LIR pattern initially was described for *Achlya ambisexualis* (Hudspeth et al 1983) and subsequently was found in at least some representatives from all other examined orders, with the exception of the

Accepted for publication February 21, 2004.
[1] Corresponding author. E-mail: mykes@niu.edu

Leptomitales. This pattern features a large inverted repeat, ranging from about 8.5 to 28.9 kb (Hudspeth and Hudspeth 1996), which separates the molecule into small and large single copy regions. Intramolecular recombination events between the repeat elements in *Achlya* (Hudspeth et al 1983, Boyd et al 1984) were shown to generate two equimolar orientational genome isomers, each with its own gene order.

In the less prevalent non-LIR pattern, initially described for *Phytophthora infestans* (Klimczak and Prell 1984) of the order Pythiales, the LIR is conspicuously absent. Non-LIR genomes subsequently have been described only for additional *Phytophthora* species (Forster et al 1987, Shumard-Hudspeth and Hudspeth 1990) and *Apodachlya* and *Leptomitus* of the Leptomitales (McNabb and Klassen 1988, Hudspeth 1992).

In this work we report the complete mtDNA sequence for the peronosporomycete *Saprolegnia ferax,* compare it with the non-LIR pattern *P. infestans,* and examine the consequences of active intramolecular recombination events between LIR elements.

MATERIALS AND METHODS

Culturing of Saprolegnia ferax *and DNA isolation.*—Stock cultures of *Saprolegnia ferax* (ATCC 36051) were obtained from I.B. Heath (York University, Toronto, Canada), maintained on potato-dextrose or cornmeal agar (Difco Laboratories, Troy, Michigan) and stored at 4 C. Mycelia for DNA preparations were propagated as 16 L peptone yeast glucose (PYG; Griffin et al 1974) aerated carboys at ambient temperatures.

Total DNA was prepared from late log-phase mycelia with purified mtDNA obtained as the upper band in bis-benzimide (Hoechst 33258; Calbiochem, La Jolla, California) CsCl gradients (Hudspeth et al 1980, Shumard et al 1986). Plasmid DNAs were isolated by the alkaline lysis method (Birnboim and Doly 1979).

Cloning and DNA sequencing.—Before sequencing of the genome, a preliminary seven-enzyme restriction map was constructed to approximate the limits of the LIR. Initial DNAs for sequencing then were prepared from clones generated from a combination of mapped *Eco*RI or unmapped *Hin*dIII restriction fragments ligated into pUC19 and transformed into *E. coli* JM83 following established procedures (Maniatis et al 1982). Clones were sequenced using sequence-generated primers prepared using either a 392 or PCRmate 391 DNA synthesizer (Applied Biosystems, Foster City, California) or were obtained commercially (MWG Biotech, High Point, North Carolina). Positions of abutting *Eco*RI or *Hin*dIII clones were confirmed by sequencing of mtDNA PCR products generated by crossing of presumptive adjacent restriction sites. Regions of sequence not included in the initial clones were generated by direct sequencing of purified mtDNA using primers derived from sequenced noncontiguous clones. Subsequent PCR products were generated and sequenced to include these regions. DNA sequence data were obtained using a 373 DNA Sequencer (Applied Biosystems, Foster City, California) or a Beckman Coulter CEQ 2000XL automated DNA sequencer (Beckman Coulter, Fullerton, California).

Sequence data were assembled using Sequencher (Gene Codes Corp., Ann Arbor, Michigan). Protein-encoding and rRNA loci were identified using the NCBI BLASTP and BLASTN similarity searches (Altschul et al 1990). tRNAs were identified using tRNAscan-SE (Lowe and Eddy 1997). Protein alignments were performed using the Clustal X program (Thompson et al 1997).

The complete mtDNA sequence has been deposited in GenBank as accession number AY534144.

RESULTS

Physical organization of S. ferax *mtDNA.*—The nucleotide sequence of the *S. ferax* mitochondrial genome assembles and restriction maps as a 46 930 bp circle (FIG. 1) in good agreement with the 14 μm (44.5 kb) circular molecule detected by electron microscopy for a *Saprolegnia* sp. (Clark-Walker and Gleason 1973). Located within the circle is an 8618 bp large inverted repeat (LIR) separating the genome into large (19 274 bp) and small (10 420 bp) single-copy regions. The repeat is similar in size and location to that initially described for the *Achlya* genome (Hudspeth et al 1983) and is a characteristic of the vast majority of peronosporomycete mtDNAs examined to date (Hudspeth and Hudspeth 1996, McNabb and Klassen 1988).

Given the potential for intramolecular recombination events in circular molecules containing inverted repeat sequences, a restriction endonuclease analysis was used to assay for the presence of isomeric forms of the genome. In this assay *Pst*I restriction sites, present within each single-copy region (FIG. 1), but absent within the LIR, were used to detect pairs of fragments resulting from the reorientation of single-copy regions as a consequence of recombination between the LIR elements (Hudspeth et al 1983). The results (not shown) identified two sets of such *Pst*I fragments in apparent equimolar concentration. One pair of 26.7 kb and 14.3 kb fragments, when combined with a 5.9 kb fragment located wholly within the large single-copy region, was consistent with one orientation of the genome. A second pair of 23.7 kb and 17.3 kb fragments, when likewise combined with the 5.9 kb fragment, was consistent with the alternate genome orientation. The presence of equimolar concentrations of both sets of the larger fragment pairs in the mtDNA population was interpreted as evidence for a dual population of orientational iso-

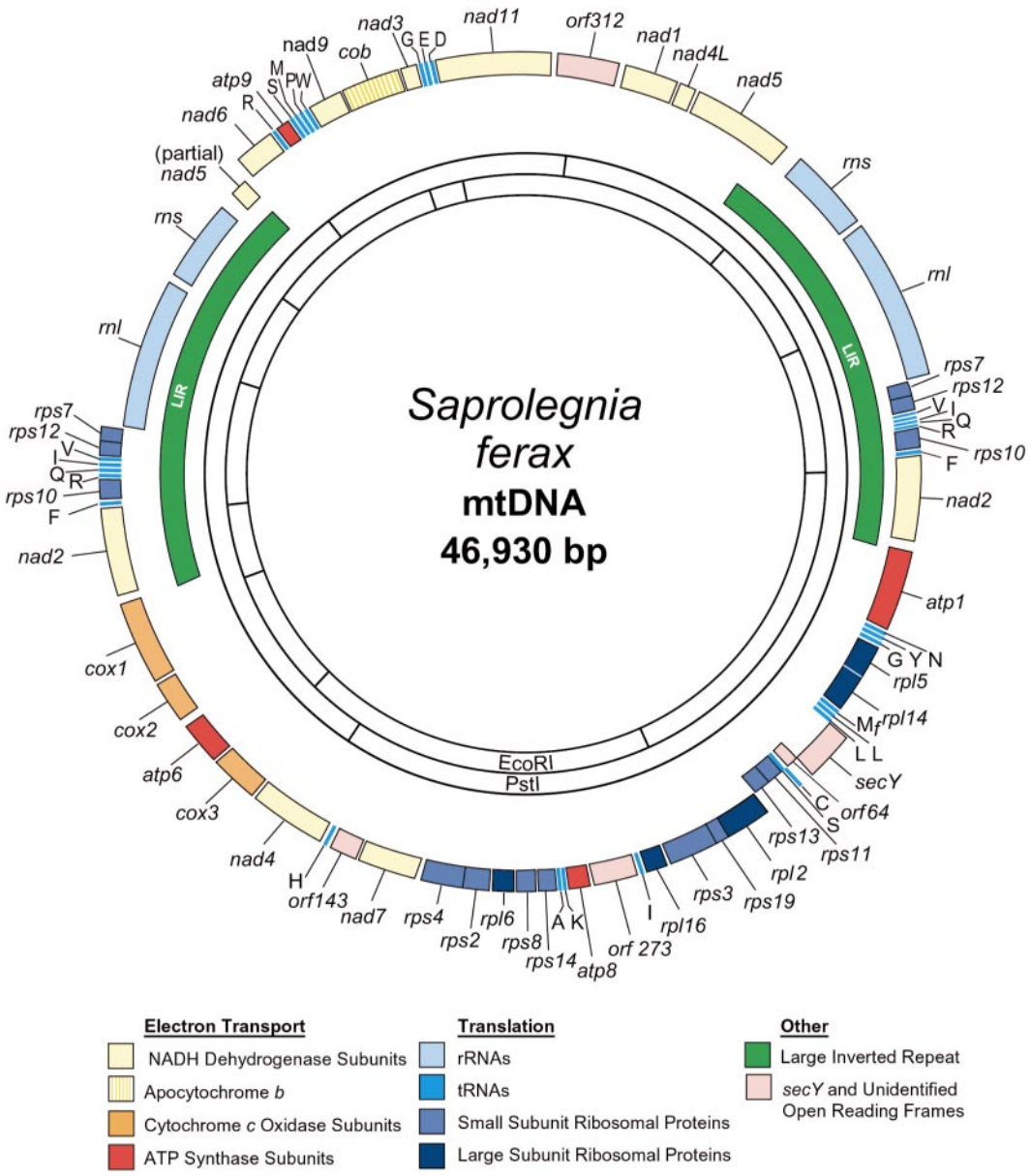

FIG. 1. Gene map of *Saprolegnia ferax* mtDNA. Genes in the outermost circle are transcribed clockwise, whereas those in the adjacent circle are transcribed anticlockwise; the two innermost circles identify locations of *Eco*RI and *Pst*I restriction sites. Intergenic spaces are not to scale.

mers, and as support for an active intramolecular recombination mechanism.

Genetic organization and gene content.—The 66 intronless genetic loci (77 including LIR duplicated sequences) encoded in the genome are shown in FIG. 1 and listed in TABLE I. These loci, representing 92.1% of the genomic sequence, often are arranged in operon-like clusters and are densely packed in the genome. Two pairs of genes overlap (*rps12-rps7, rpl2-rps19*) and three pairs abut each other (*rps2-rps4, rps13-rps11, rps19-rps3*). Intergenic regions average only 47 bp with a range of −23 (overlapping loci) to 341 bp. The overall A+T content of 76.9% is significantly higher in intergenic regions (94.1%) than in protein-encoding (79.2%) sequences. Thirty-nine loci encode polypeptides of which 35 are readily assigned by similarity searches of the GenBank protein database. Eighteen of these are components of the mitochondrial respiratory chain, 16 are subunits of the mitoribosome, and one, *secY,* is a homolog of *E. coli tatC*—a component of the sec-independent protein translocation pathway (Bogsch et al 1998, Wei-

TABLE I. Genetic content of *Saprolegnia ferax* mtDNA

Respiratory chain proteins
  Complex I
    nadl, ***nad2****, nad3, nad4, nad4L, nad5*, nad6, nad7, nad9, nad11*
  Complex III
    cob
  Complex IV
    cox1, cox2, cox3*
  Complex V
    atp1, atp6, atp8, atp9*

Ribosomal proteins
  Small subunit
    rps2, rps3, rps4, ***rps7****, rps8,* ***rps10****, rps11,* ***rps12****, rps13, rps14, rps19*
  Large subunit
    rpl2, rpl5, rpl6, rpl14, rpl16

Import proteins
    secY

Translational RNAs
  Ribosomal RNAs
    rns*,* ***rnl***
  tRNAs
    $trnA_{UGC}$, $trnC_{GCA}$, $trnD_{GUC}$, $trnE_{UUC}$, $\mathbf{trnF_{GAA}}$, $trnG_{UCC}$, $trnG_{GCC}$, $trnH_{GUG}$, $\mathbf{trnI_{GAU}}$, $trnI_{CAU}$, $trnK_{UUU}$, $trnL_{UAG}$, $trnL_{UAA}$, $trnM_{eCAU}$, $trnM_{fCAU}$, $trnN_{GUU}$, $trnP_{UGG}$, $\mathbf{trnQ_{UUG}}$, $trnR_{\mathbf{GCG}}$, $trnR_{UCU}$, $trnS_{UGA}$, $trnS_{GCU}$, $\mathbf{trnV_{UAC}}$, $trnW_{CCA}$, $trnY_{GUA}$

Unassigned open reading frames
    orf64, orf143, orf273, orf312

Bold, encoded in LIR; asterisk, 3′ sequence encoded in LIR.

ner et al 1998). Three of four unassigned reading frames (*orf64, orf143* and *orf273*) share regions of homology (FIG. 2) with ''unique'' *P. infestans* ORFs (Paquin et al 1997, below) of approximately equivalent sizes and the carboxyl region of *rps4* shares homology with *orf100* of *P. infestans*. The remaining 27 loci (exclusive of LIR duplicated sequences) encode RNA components of the mitochondrial translational apparatus. Twenty-five are tRNAs, and the remaining two encode the small (*rns*) and large (*rnl*) subunit rRNAs. The tRNAs include both initiator and elongator *trnM* plus cognate species for *trnG, trnI, trnL, trnR,* and *trnS*. No *trnT* was detected.

Codon-usage frequencies are shown in TABLE 2. Notably absent are uses of the leucine codon CUC, the arginine codons AGG and CGG and the termination codon UGA. A single use of the ACG threonine codon occurs in *orf312*. GUG serves as the initiator codon for *atp1* and *atp6,* with the standard AUG employed for all other polypeptides. Similarly, the termination codon UAG is used only for *rpl2* with UAA used for all other reading frames.

The LIR extends for 8618 bp and encodes 11 intact and three partial loci. Eleven genes thus are duplicated fully in the genome and include the two rRNAs, five tRNAs, three small subunit mitoribosome polypeptides, and *nad2* of respiratory complex I. It was a surprise to discover, however, that each end of the LIR terminates with the partial 3′ sequence of a protein encoding gene. At the small single copy junctions, 319 nucleotides of *nad5* (encoding 106 full codons) are duplicated and both a lysine and a termination codon, completing both the *atp1* and *cox1* reading frames, are duplicated at the large single copy junctions.

DISCUSSION

The impetus for this study was the opportunity to compare the organizational and genetic features of two phylogenetically distant representative peronosporomycete mtDNAs. By selecting *S. ferax* for comparison with the available *P. infestans* genome (Paquin et al 1997) we were able to include well-studied representatives from the two major peronosporomycete subclasses (the Saprolegniomycetidae and the Peronosporomycetidae) as well as to contrast LIR and non-LIR genomes.

The overall genetic content and organization of the two genomes is similar. All 35 assignable reading frames, both rRNAs, and the 25 species of tRNAs are common to both genomes with differences reflected only in the numbers and types of ORFs.

Genome organization, with the exception of the LIR, also is strikingly similar, with the genes organized as compact blocks of loci. In *S. ferax* these blocks are suggestive of three major polycistronic transcripts. A 20-loci clockwise transcript (FIG. 1) originates at *nad6* in the small single copy region and terminates with *rnl* in the LIR, while a 22-loci anticlockwise transcript originates at *rps13* in the large single copy region and terminates following *rps7* in the LIR. Finally, a 33-loci clockwise transcript extends from *secY* in the large single copy region to *rps7* in the other LIR. This latter transcript would overlap the anticlockwise transcript at its origin and duplicates transcription for nine of the LIR encoded loci.

Gene-order conservation.—Comparison of gene order and potential transcriptional units in the *S. ferax* and *P. infestans* genomes revealed two major blocks of colinear loci. With the inclusion of a newly identified *orf68* in the *P. infestans* genome (see below) a colin-

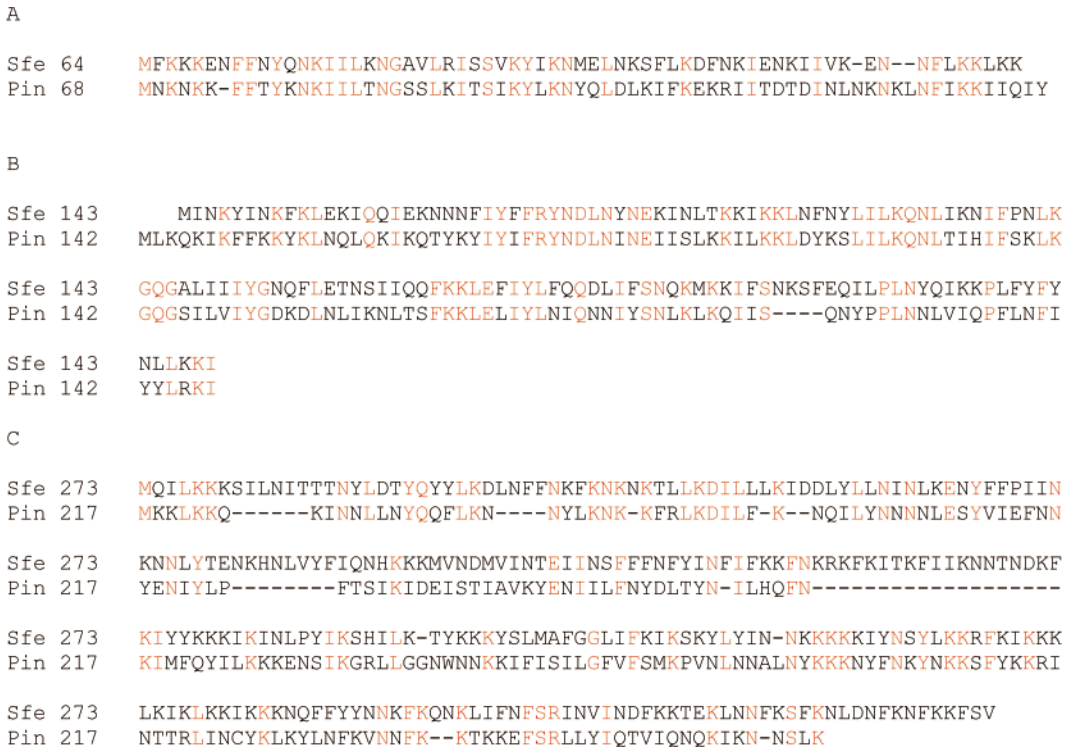

FIG. 2. Comparisons of deduced amino acid sequences of *S. ferax* ORFs with their *P. infestans* counterparts. A. *S. ferax* ORF64 and *P. infestans* ORF68. B. *S. ferax* ORF143 and *P. infestans* ORF142. C. *S. ferax* ORF273 and *P. infestans* ORF217. Red identifies identical residues. Sfe, *S. ferax*; Pin, *P. infestans*; dashes indicate spaces introduced to facilitate alignments.

ear set of 26 loci extending from $trnL_{UAG}$ through *rps4* (FIG. 1) is apparent. Retained within this block are the remnants of the conserved prokaryotic ribosomal protein linkages including *rps13-rps11, rpl2-rps19-rps3-rpl16* and *rps14-rps8-rpl6,* all previously noted for *P. infestans* (Lang et al 1999). The presumptive stramenopile linkage of *rps8-rpl6-rps2-rps4* (Chesnick et al 2000) is similarly present. A second major block includes the eight LIR genes from *nad2* through *rps7*.

The retention of large colinear blocks of loci in the Peronosporomycetes, even in distantly related taxa, is expected. This is based on the necessity of either retaining or acquiring promoter sites for the translocated loci. Due to the paucity of intergenic target sequences for translocation events that avoid disruption of other loci, the rate of viable translocations is expected to be low and the retention of genetic linkages to be high.

Unassigned reading frames.—Deduced amino acid sequences of the four unassigned *Saprolegnia* ORFs were compared with the five free-standing ORFs reported for *P. infestans* (Paquin et al 1997). Two of the ORFs lie in the largest positionally conserved gene block with *orf143* located between *nad7* and $trnH_{GUG}$, and *orf273* between $trnI_{CAU}$ and *atp8*. Each ORF shares obvious regions of amino acid similarity with its *P. infestans* equivalent. The *orf143/orf142* pair (FIG. 2b) encodes the more highly conserved deduced polypeptide with an amino acid identity of about 45%. The *orf273/orf217* pair (FIG. 2c) is significantly less conserved but still clearly retains three regions of conservation—one centrally located and the others near the termini.

No homologue to *Saprolegnia orf64* was identified previously in the *P. infestans* genome. However, because this ORF is included in a conserved 26 loci gene block (see above), and, because the corresponding block in the *P. infestans* genome contains an intergenic region of sufficient length to encode a comparable reading frame, we re-evaluated this region of the *P. infestans* genome. A short open reading frame, designated *orf68,* subsequently was located between *secY* and $trnC_{GCA}$ on the complementary strand as in *S. ferax,* but in *P. infestans* this ORF shares a 12-codon overlap with the carboxyl region of *secY*. The alignment of deduced amino acid sequences from this ORF pair revealed the conserved *P. infestans* polypeptide (48% identity) shown in FIG. 2a.

Searches of the protein databases, using short regions of identity/similarity derived from each of the

TABLE II. Codon Usage

		A	B	C	D	E			A	B	C	D	E
UUU	F	92	97	94	93	765	UCU	S	38	38	33	37	226
UUC	F	8	3	6	7	55	UCC	S	<1	2	2	1	6
UUA	L	93	92	86	92	1096	UCA	S	30	30	21	29	177
UUG	L	1	1	5	2	19	UCG	S	2	3	13	3	19
CUU	L	4	5	6	5	55	CCU	P	67	56	53	64	178
CUC	L	—	—	—	—	—	CCC	P	2	—	5	2	5
CUA	L	2	1	2	2	23	CCA	P	28	44	32	32	88
CUG	L	—	—	<1	<1	1	CCG	P	3	—	11	3	8
AUU	I	64	54	48	59	706	ACU	T	72	68	61	70	316
AUC	I	4	2	3	4	43	ACC	T	4	4	11	5	22
AUA	I	32	43	49	37	448	ACA	T	23	28	26	25	111
AUG	M	100	100	100	100	203	ACG	T	—	—	3	<1	1
GUU	V	50	49	47	50	210	GCU	A	56	58	50	56	186
GUC	V	<1	3	11	1	5	GCC	A	5	—	—	4	13
GUA	V	47	47	37	47	198	GCA	A	35	40	30	36	120
GUG	V	2	1	5	2	10	GCG	A	4	2	20	4	15
UAU	Y	92	96	89	93	433	UGU	C	95	95	86	94	81
UAC	Y	8	4	11	7	35	UGC	C	5	5	14	6	5
UAA	–	94	94	100	95	37	UGA	–	—	—	—	—	—
UAG	–	6	6	—	5	2	UGG	W	100	100	100	100	106
CAU	H	80	86	90	82	143	CGU	R	33	26	42	31	76
CAC	H	20	14	10	18	31	CGC	R	<1	3	—	2	4
CAA	Q	96	98	97	97	283	CGA	R	6	14	33	10	26
CAG	Q	4	2	3	3	9	CGG	R	—	—	—	—	—
AAU	N	90	96	93	93	614	AGU	S	29	26	27	30	169
AAC	N	10	4	7	7	47	AGC	S	1	1	4	1	8
AAA	K	99	99	98	99	900	AGA	R	61	57	25	57	143
AAG	K	1	1	2	1	11	AGG	R	—	—	—	—	—
GAU	D	95	92	92	94	271	GGU	G	76	63	53	72	345
GAC	D	5	8	8	6	17	GGC	G	1	3	13	2	9
GAA	E	97	94	92	96	311	GGA	G	19	34	27	23	108
GAG	E	3	6	8	4	13	GGG	G	3	<1	7	3	14

A = % of that codon used for respiratory proteins
B = % of that codon used for ribosomal proteins
C = % of that codon used for ORFs
D = % of that codon used in all proteins (including ORFs)
E = Total number of codons used

three ORF pairs noted above, failed to identify likely mitochondrial homologues. However, their positional and relative amino acid conservation in two distinct subclass peronosporomycete lineages strongly argues for their validity as functional, albeit highly derived, mitochondrial proteins.

Of the three *P. infestans* ORFs without obvious *S. ferax* homologues, *orf100* is a candidate for a carboxyl extension of *rps4*. Similar to other protists and prokaryotes, this locus in *S. ferax* encodes a polypeptide with a carboxyl terminus about 75 amino acids longer than that reported for *P. infestans*. The proximity of *orf100* five nucleotides downstream of *rps4* in *P. infestans* prompted the inclusion of a concatenated *rps4/orf100* sequence in an alignment with deduced *rps4* polypeptides from the ''jakobid'' protist *Reclinomonas americana* (Lang et al 1997), the brown alga *Pylaiella littoralis* (Oudot-LeSeq et al 2001) and the peronosporomycete *Saprolegnia* (FIG. 3). While it is apparent that the carboxyl region of *rps4* is far less conserved than the remainder of the polypeptide among these representative taxa, *orf100* and the *rps4* carboxyl region of *S. ferax* clearly are homologous. It remains to be determined if the *rps4/orf100* combination of *P. infestans* encodes a true frame shift, represents a genomically fragmented *rps4* or is a sequencing artifact.

Assuming *orf100* encodes the carboxyl region of

FIG. 3. Comparisons of deduced *rps4* amino acid sequences identifying *P. infestans orf100* as the carboxyl terminus of *rps4*. *P. infestans orf100* is underlined as an extension of *rps4*; red identifies identical residues; Sfe is *S. ferax*; Pin, *P. infestans*; Pli, *Pylaiella littoralis* (brown alga); Ram, *Reclinomonas Americana* (protest); dashes indicate spaces introduced to facilitate alignments.

rps4, the only unassigned ORFs not shared between the mtDNAs are *orf312* of *S. ferax* and *orf32* and *orf79* of *P. infestans*. Of these the presence of *orf32* has been reported in *P. megasperma* (Lang and Forget 1993) and has been identified subsequently in several additional nonsaprolegnealean lineages of the Peronosporomycetes (unpublished results). We similarly have identified an *orf82* as an *orf79* equivalent in a different *P. megasperma* isolate (unpublished results). Thus, only *orf312* currently remains without precedent in other peronosporomycete mtDNAs.

Codon usage and tRNAs.—Twenty-five species of tRNAs with appropriate clover-leaf structures and unambiguous anticodons are encoded in *S. ferax* mtDNA, with five of these duplicated in the LIR. This set of tRNAs is identical to that encoded in the *P. infestans* genome (Paquin et al 1997) and is insufficient to support mitochondrial protein synthesis. Notably absent are the tRNAs required for translation of arginine CGR and all threonine codons. The two-codon arginine AGR family is accounted for by $trnR_{UCU}$ (only AGA codons are used in *S. ferax*), but the use of $trnR_{GCG}$ in the four-codon arginine family in lieu of the more typical $trnR_{ACG}$ with its wobble ''A'' deaminated to ''I'' (Pfitzinger et al 1990) limits translation to CGY codons. Thus, as in *P. infestans,* it is necessary to postulate the import of cytosolic tRNAs enabling translation of arginine CGR and all threonine codons. Similarly, as originally observed in *E. coli* (Muramatsu et al 1988) and inferred for *P. infestans* and other protists and plants (Gray et al 1998), it is assumed that the wobble ''C'' of $trnI_{CAU}$ is post-transcriptionally modified to lysidine to enable translation of the isoleucine AUA codon.

LIR recombination retains intact loci.—The primary architectural difference between *S. ferax* and *P. infestans* mtDNAs is the presence of the LIR. Size variation in LIR genomes has been well documented among peronosporomycete taxa and has been shown to be the major contributor to genome size variation (Hudspeth and Hudspeth 1996, McNabb and Klassen 1988). *S. ferax* contains the shortest described LIR, and the 8618 bp are shown here to encode intact copies of four polypeptides, five tRNAs and both rRNAs. It was a surprise to find partial protein-encoding sequences, rather than intergenic sequences, at the repeat termini. We had anticipated that, because an active recombination mechanism between repeat elements has been inferred for *S. ferax* and other members of the Saprolegniales (Hudspeth et al 1983, Boyd et al 1984), the LIR would terminate with intergenic sequences and thereby ensure intact encoded loci for both mtDNA orientational isomers.

The LIR ends encode 3′ termini for three polypeptides but still maintain available intact copies for the three loci. At the small single copy junctions, the LIR terminates with 106 codons of *nad5*. When the genome assumes the orientational isomer presented in FIG. 1, the right-handed repeat element generates an intact *nad5* locus while the left-handed element retains only the carboxyl-encoding fragment. In the alternate configuration it is the left-handed repeat that forms the intact locus. A different approach to retaining intact loci is used for the large single copy

junctions. Here the LIR termini encode two in-frame codons—lysine and termination—that are common to the flanking loci. These codons provide translational termination for both *atp1* and *cox1* in both mtDNA orientations. Thus, the extension of coding sequences into the LIR still provides intact *nad5, atp1* and *cox1* loci for either mtDNA isomer.

ACKNOWLEDGMENTS

We thank Barbara Ball for assistance with the figures. This study was supported by the Northern Illinois University Plant Molecular Biology Center and NSF grants DEB-9806785 and DEB-0213076.

LITERATURE CITED

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

Bogsch EG, Sargent F, Stanley NR, Berks BC, Robinson C, Palmer T. 1998. An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. J Biol Chem 273:18003–18006.

Birnboim HC, Doly J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523.

Boyd DA, Hobman TC, Gruenke SA, Klassen GR. 1984. Evolutionary stability of mitochondrial DNA organization in *Achlya*. Can J Biochem Cell Biol 62:571–576.

Burger G. 1999. GenBank accession No. AF193903.

———, Lang BF, Gray MW. 2000. GenBank accession No. AF287134.

Chesnick JM, Goff M, Graham J, Ocampo C, Lang BF, Seif E, Burger G. 2000. The mitochondrial genome of the stramenopile alga *Chrysodidymus synuroideus*. Complete sequence, gene content and genome organization. Nucleic Acids Res 28:2512–2518.

Clark-Walker GD, Gleason FH. 1973: Circular DNA from the water mold *Saprolegnia*. Arch. Mikrobiol 92:209–216.

Dick MW. 2001. *Straminipilous Fungi*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

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

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

Griffin DH, Timberlake WE, Cheney JC. 1974. Regulation of macromolecular synthesis, colony development, and specific growth rate of *Achlya bisexualis* during balanced growth. J Gen Microbiol 80:381–388.

Hudspeth MES. 1992. The fungal mitochondrial genome—a broader perspective. In Arora DK, Elander RP, Mukerji KG, eds. Handbook of applied mycology. Fungal biotechnology. Vol. 4. New York: Marcel Dekker Inc. p 213–241.

———, Hudspeth DSS. 1996. Mitochondrial genomes of the zoosporic fungi. In Dayal R, ed. Advances in zoosporic fungi. New Delhi: MD Publications Pvt Ltd. p 173–199.

———, Shumard DS, Bradford CJR, Grossman LI. 1983. Organization of *Achlya* mtDNA: a population with two orientations and a large inverted repeat containing the rRNA genes. Proc Natl Acad Sci USA 80:142–146.

———, ———, Tatti KM, Grossman LI. 1980. Rapid purification of yeast mitochondrial DNA in high yield. Biochim Biophys Acta 61:221–228.

Klimczak J, Prell HH. 1984. Isolation and characterization of mitochondrial DNA of the oomycetous fungus *Phytophthora infestans*. Curr Genet 8:323–326.

Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Gray MW. 1997. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature 387:493–497.

———, Forget L. 1993. The mitochondrial genome of *Phytophthora infestans*. In O'Brien SJ, ed. Genetic Maps. Locus maps of complex genomes. Book 3 Lower Eukaryotes. 6th ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. p 3.133.

———, Gray MW, Burger G. 1999. Mitochondrial genome evolution and the origin of eukaryotes. Ann Rev Genet 33:351–397.

Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964.

Maniatis T, Fritsch EF, Sambrook J. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

McNabb SA, Klassen GR. 1988. Uniformity of mitochondrial complexity in oomycetes and the evolution of the inverted repeat. Expt Mycol 12:233–242.

Muramatsu T, Nishikawa K, Nemoto F, Kuchino Y, Nishimura S, Miyazawa T, Yokoyama S. 1988. Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. Nature 336:179–181.

Patterson DJ. 1989. Stramenopiles: chromophytes from a protistan perspective. In Green JP, Leadbeater BSC, Diver WC, eds. The chromophyte algae: problems and perspectives. Oxford, UK: Clarendon Press. p 357–379.

Oudot-LeSecq M-P, Fontaine JM, Rousvoal S, Kloareg B, Loiseaux-DeGoer S. 2001. The complete sequence of a brown algal mitochondrial genome, the ectocarpale *Pylaiella littoralis* (L.) Kjellm. J Mol Evol 53:80–88.

———, Kloareg B, Loiseaux-DeGoer S. 2002. The mitochondrial genome of the brown alga *Laminaria digitata*: a comparative analysis. Eur J Phycol 37:163–172.

Paquin B, LaForest M-J, Forget L, Roewer I, Wang Z, Longcore J, Lang BF. 1997. The fungal mitochondrial genome project: evolution of fungal mitochondrial genomes and their gene expression. Curr Genet 31:380–395.

Pfitzinger H, Weil JH, Pillay DTN, Guillemaut P. 1990. Codon recognition in plant chloroplasts. Plant Mol Biol 14:805–814.

Shumard-Hudspeth DS, Hudspeth MES. 1990. Genic rearrangements in *Phytophthora* mitochondrial DNA. Curr Genet 17:413–417.

Shumard DS, Grossman LI, Hudspeth MES. 1986. *Achlya* mitochondrial DNA: gene localization and analysis of inverted repeats. Mol Gen Genet 202:16–23.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882.

Weiner JH, Bilous PT, Shaw GM, Lubitz SP, Frost L, Thomas GH, Cole JA, Turner RJ. 1998. A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. Cell 93:93–101.