

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Fungal Genetics and Biology 40 (2003) 215–224

www.elsevier.com/locate/yfgeb

Mapping and characterization of polymorphism in mtDNA of *Cryphonectria parasitica*: evidence of the presence of an optional intron

Emanuela Gobbi,^{a,*} Giuseppe Firrao,^a Anna Carpanelli,^a
Romano Locci,^a and Neal K. Van Alfen^b

^a Dipartimento di Biologia applicata alla Difesa delle Piante, Università di Udine, Udine, Italy

^b Department of Plant Pathology, UC Davis, Davis, CA, USA

Received 8 July 2002; accepted 11 August 2003

Abstract

The mtDNA of the filamentous ascomycete *Cryphonectria parasitica* is large and polymorphic so, to better understand the nature of the polymorphisms within populations, a small collection of Italian strains of the fungus was examined. Known mtDNA polymorphisms were mapped and found to cluster in four regions of the mtDNA molecule, particularly in the RFLP region 2 where five different mtDNA haplotypes out of 13 strains were identified. This region included an area of 8.4 kbp which was entirely sequenced in strain Ep155 showing the presence of two introns. An internal 3.2 kbp portion was sequenced also in six additional strains. Sequence comparison of the *C. parasitica* mitochondrial intronic ORFs revealed similarities to known endonucleases such as those of *Podospora anserina* and *Neurospora crassa*. DNA sequence analysis showed that three polymorphisms of this mtDNA region within this population of 12 strains were due to the optional presence in the *ND5* gene of an intron and of an intervening sequence within the intron. Evidence was also found within this population of mixed mitochondrial types within a single strain. © 2003 Elsevier Inc. All rights reserved.

Keywords: mtDNA; Optional intron; ND5; Polymorphism; Physical map; Interrupted intron; *Cryphonectria parasitica*; Fungi

1. Introduction

The mitochondrial DNA (mtDNA) possesses a genetic role universally conserved in all eukaryotes; nonetheless this genome exhibits remarkable variation in conformation, size, gene organization, arrangement, and expression among different groups of eukaryotes (Gray et al., 1999). Metazoan mtDNAs are characterized by their small, uniform size while plants have large (200 to over 2000 kb in size) and nonuniform mtDNA (Palmer, 1990). Even single individuals may show heterogeneous populations of mtDNA molecules. Among

fungi, despite the similar content of coding information, the mtDNA varies considerably in size between different fungi, like that in plants. The mtDNA varies from a metazoan-like size, 17.6 kbp in *Schizosaccharomyces pombe* (Zimmer et al., 1987) to an almost plant-like size, 176.3 kbp in *Agaricus bitorquis* (Hintz et al., 1985). Moreover size variability is exhibited even within the same genus, for example in *Suillus* the mtDNAs of different species have a 3-fold variation in size (Bruns et al., 1988), and within the same species, as in *S. pombe* where a 1.4-fold variation has been reported (Zimmer et al., 1987).

Much of the length heterogeneity of the mtDNA molecule is the result of variability in content of non-coding regions. These regions contain short tandemly repeated sequences, pseudogenes or other types of duplications of functional genes and other miscellaneous non-transcribed sequences (Palmer and Herbon, 1986;

* Corresponding author. Fax: +39-030-2450064.

E-mail address: micol@pldef.uniud.it (E. Gobbi).

¹ Present address: Biodiversity s.r.l., Via Corfù, 71, 25124 Brescia, Italy.

Small et al., 1989). Additionally, introns, mobile introns, and other transposable elements are responsible for mtDNA heterogeneity (Belcour et al., 1997, Hamari et al., 1999, Salvo et al., 1998). Group I and group II introns are very common in fungal mitochondria: *Aspergillus nidulans* possesses five group I introns (Brown et al., 1985), *Neurospora crassa* nine group I introns (Griffiths et al., 1995), and *Podospira anserina* harbors 33 group I introns and three group II introns (Cummings et al., 1990).

The mtDNA of the filamentous ascomycete *Cryphonectria parasitica* is of interest because of its importance as a plant pathogen. This fungus is one of the most successful pathogens in recorded history since it essentially eliminated its host, the American chestnut trees, *Castanea dentata*, from its native range (Van Alfen, 1988). In Europe this introduced pathogen, while less destructive than in North America, has become a serious disease of the native cultivated and forest chestnut trees (Bisiach et al., 1988). Some strains of the fungus contain cytoplasmically transmissible elements that limit fungal virulence toward chestnuts. In most cases, these disease-causing agents are double-stranded RNA (dsRNA) viruses (Heiniger and Rigling, 1994), but a form of this transmissible hypovirulence has been found to be associated with mtDNA mutations (Monteiro-Vitorello et al., 1995). Considering the potential role of mtDNA mutations in the population biology of this fungus it is important to better understand the nature of these mutations and to learn more about the mtDNA variability of natural populations of the fungus.

The mtDNA of *C. parasitica* is one of the larger in fungi with sizes reported to range from 135 to 158 kbp (Bell et al., 1996; Gobbi et al., 1990). Restriction fragment length polymorphism (RFLP) analysis has revealed significant polymorphism in the mtDNA of *C. parasitica* (Bell et al., 1996; Gobbi and Locci, 1990; Milgroom and Lipari, 1993). In certain geographically limited areas each isolate represented a unique mtDNA RFLP haplotype (Gobbi and Locci, 1994). The mtDNA mutations in *C. parasitica* are poorly understood and yet of importance considering their potential role in the population biology of the fungus. To begin a characterization of the types of mutations present in the mtDNA genome a population of the fungus in Italy was examined for mtDNA polymorphisms and the basis for three of these polymorphisms determined. As part of the process a physical map of the mtDNA was constructed using three restriction enzymes to locate the regions of mtDNA responsible for the polymorphisms. An 8.4 kbp polymorphic region was sequenced in strain Ep155 and a 3.2 kbp subportion of this region was sequenced in six other strains to identify the nature of the mutations in this region. A putative optional intron of the *ND5* gene is reported to be one cause of the mtDNA diversity in this fungus.

2. Materials and methods

2.1. Collection of isolates and culture conditions

Twelve strains of *C. parasitica* used in this experiment were collected from near Grimacco, Italy, in the province of Friuli Venezia Giulia and were designated L₁. The chestnut trees were commercially farmed, but were located within a mixed-stand forest. No attempt had been made to keep other trees and shrubs from invading the plantation although the trees had obviously been planted. Samples were taken in 1989 from each infected tree located along a linear transect of approximately 150 m. Strain EP155 was isolated from Connecticut in the United States and has become the most commonly used virulent strain for laboratory studies of this fungus. The fungal cultures were maintained on potato dextrose agar (PDA). DNA was extracted from cultures grown following the method of Hansen et al. (1985).

2.2. DNA isolation and analysis

Mitochondrial DNA of *C. parasitica* was isolated by CsCl-bisbenzimidazole isopycnic centrifugation as previously described (Gobbi et al., 1990). The total DNA for RFLP analysis was isolated following the methods reported by Zhang et al. (1994). Nucleic acids used in PCR experiments were extracted as described by Lecellier and Silar (1994). Standard methods for preparations of plasmid DNA, purifications of phages and DNA extraction and RFLP analysis were used (Ausubel et al., 1991). Restriction enzymes (*Kpn*I, *Sac*I, *Eco*RI, *Sal*I, *Xho*I, *Apa*I, *Hind*III, *Mbo*I, *Sau*III, and *Not*I) were purchased from BMB (Germany). For fragments larger than 31 kbp, the length was estimated by isolation and digestion of the fragments with other enzymes or from the hybridization analysis results.

The bands of DNA were visualized by UV-ethidium bromide and digitalized by E.A.S.Y. (Ultra-Violet Products, UK). Cloning procedures were as described in Sambrook et al. (1989).

2.3. Selection of mtDNA RFLP probes

Random clones of about 5 kbp each from a mtDNA plasmid library of *C. parasitica* strain EP155 (N.K. Van Alfen, unpublished) were used to screen *Eco*RI digested DNA from a large number of different strains of *C. parasitica* for detection of RFLPs (data not shown). From the 32 mtDNA clones screened, eight detected RFLPs among the strains tested. These eight clones, labeled probe 4, 5, 7, 13, 17, 26, 38, and 47, were used to screen the isolates from Grimacco, Italy for mtDNA RFLPs following standard protocols (Gobbi et al., 1990).

2.4. Construction of mtDNA physical map

Double reciprocal digestions of the mtDNA of strain EP155 with *Kpn*I, *Sac*I, and *Eco*RI were initially attempted to develop a physical map of the mtDNA. The large size of many of the fragments and the complexity of their restriction patterns did not allow the definitive construction of a map. For this reason single restriction fragments were either cloned or purified from agarose gels using GeneClean (Bio101, USA) and used as probes against mtDNA digested with *Kpn*I, *Sac*I, *Eco*RI alone, or in combination. Hybridizations of the mtDNA with lambda mtDNA clones and with the eight DNA fragments that were previously used as probes to detect mtDNA RFLPs were used to help clarify ambiguous sections of the map. The lambda clones were from strain EP155 and were identified from the total genomic library described by Powell and Van Alfen (1987) by screening the library with mtDNA probes.

2.5. PCR analysis and DNA sequencing

The 19.3 kbp portion of the mtDNA genome to which probes 4 and 13 mapped, i.e. RFLP region 2, was chosen for sequencing. A lambda clone was selected as described above by probing the library with probe 4 and subcloned into pBSKS+ (Stratagene, USA). This region of the mtDNA was sequenced using the primer walking method. The sequence obtained was used to design primers to allow sequencing of the same region from six of the Italian strains. The primer sets used were mt2 (5'-AAATTAATATTCATTAGGGAGG-3') and eg1f4

(5'-CCAATAACTAAACAATACC-3') located at positions 2511 and 5701, respectively, according to the partial sequence of the mtDNA of Ep155.

Each PCR mixture contained 10x buffer (Qiagen, Germany), 150 ng each primer, 200 μ M dNTPs (Amersham-Pharmacia Biotech, USA) and 2U *Taq* Polymerase (Qiagen, Germany) in a standard reaction volume of 50 μ l. Reactions were cycled 25 times at 95 °C (30 s), 50 °C (30 s), and 72 °C (180 s), preceded by 5 min at 95 °C and followed by 7 min at 72 °C. All reactions were performed in a GeneAmp PCR System 2400 (Perkin-Elmer, USA). Two separate total DNA preparations were amplified for each of the six examined strains. The PCR products were purified with spin columns (Qiagen, Germany), cloned with the pGEM-T Easy system by Promega, USA, and sequenced by standard methods of primer walking. The sequences were manually aligned and the sequence data were analyzed using the sequence analysis software package from the Genetics Computer Group of the University of Wisconsin (Devereux et al., 1984) and by the BLAST program (Pearson and Lipman, 1988).

3. Results

3.1. Physical map of mtDNA

The physical map of the mtDNA is shown in Fig. 1 with the cleavage sites for *Eco*RI, *Kpn*I, and *Sac*I identified. The names, sizes of the restriction fragments, and their map positions are provided in Table 1.

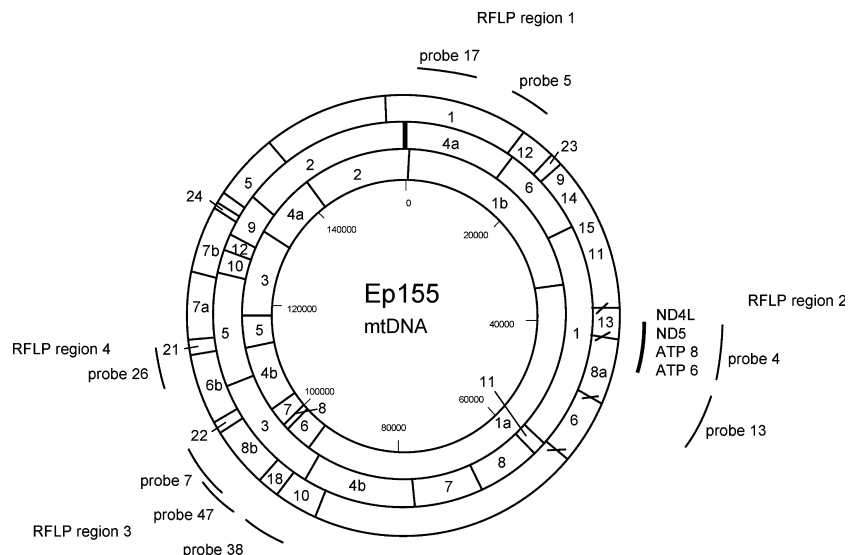


Fig. 1. Physical map of mtDNA of the strain EP155 of *C. parvula*. The restriction enzyme sites and fragments for *Kpn*I, *Sac*I, and *Eco*RI are indicated, respectively, on the inner, central and outer circles. A few *Eco*RI restriction sites are not precisely known so their approximate localizations are indicated by a diagonal line; these sites can be anywhere within the areas of K1a and S1. The positions of the eight probes used to detect mtDNA polymorphisms are shown and the locations of the four polymorphic regions mentioned in the text are indicated. Since probe 17 does not have *Kpn*I, *Sac*I, and *Eco*RI sites, its approximate location between the first *Kpn*I site of the mtDNA and the beginning of probe 5 is shown. The sequenced region is indicated by an outer broken circle and the four genes are reported in order, their nomenclature is explained in the text.

Table 1
Restriction site distances and map positions (A), names and sizes of the fragments (B) of EP155 mtDNA digested with *KpnI*, *SacI*, and *EcoRI*

A			B					
Enzyme	Size (kb)	Map position	<i>KpnI</i>		<i>SacI</i>		<i>EcoRI</i>	
			Fragment	kb	Fragment	kb	Fragment	kb
<i>Sac</i>	0.0	0.0						
<i>Kpn</i>	0.6	0.6	K1a	59.9	S1	30.3	E1	17.0
<i>Eco</i>	14.0	14.6	K1b	34.0	S2	23.1	E2	13.5
<i>Sac</i>	0.7	15.3	K2	16.5	S3	16.6	E3	9.5
<i>Eco</i>	3.5	18.8	K3	13.4	S4a	15.3	E4	9.0
<i>Eco</i>	1.3	20.1	K4a	10.4	S4b	15.3	E5	8.8
<i>Sac</i>	7.2	27.3	K4b	10.4	S5	14.8	E6a	8.1
<i>Kpn</i>	7.3	34.6	K5	4.6	S6	12.0	E6b	8.1
<i>Sac</i>	23.0	57.6	K6	4.1	S7	9.0	E7a	7.6
<i>Sac</i>	2.4	60.0	K7	3.7	S8	8.1	E7b	7.6
<i>Sac</i>	8.1	68.1	K8	0.6	S9	5.6	E8a	7.2
<i>Sac</i>	9.0	77.1			S10	3.3	E8b	7.2
<i>Eco</i>	12.3	89.4			S11	2.4	E9	6.9
<i>Sac</i>	3.0	92.4			S12	1.9	E10	4.9
<i>Eco</i>	1.9	94.3					E11	4.6
<i>Kpn</i>	0.2	94.5					E12	4.2
<i>Eco</i>	2.4	96.9					E13	4.0
<i>Kpn</i>	1.7	98.6					E14	3.8
<i>Kpn</i>	0.6	99.2					E15	3.6
<i>Kpn</i>	3.7	102.9					E16	3.2
<i>Eco</i>	1.2	104.1					E17	3.0
<i>Eco</i>	1.4	105.5					E18	2.6
<i>Sac</i>	3.5	109.0					E19	2.4
<i>Kpn</i>	4.3	113.3					E20	1.9
<i>Eco</i>	0.4	113.7					E21	1.6
<i>Eco</i>	1.6	115.3					E22	1.4
<i>Kpn</i>	2.6	117.9					E23	1.3
<i>Eco</i>	5.0	122.9					E24	1.2
<i>Sac</i>	0.9	123.8					E25	0.8
<i>Sac</i>	3.3	127.1						
<i>Sac</i>	1.9	129.0						
<i>Eco</i>	1.5	130.5						
<i>Kpn</i>	0.8	131.3						
<i>Eco</i>	0.4	131.7						
<i>Eco</i>	1.2	132.9						
<i>Sac</i>	1.7	134.6						
<i>Eco</i>	7.0	141.6						
<i>Kpn</i>	0.1	141.7						
<i>Eco</i>	13.6	155.3						
<i>Sac</i>	2.4	157.7						

Fragments smaller than 0.5 kbp are not identified on the map. Fragments K1a and K1b co-migrated in agarose gels and could not be separated by electrophoretic analysis so their existence is hypothesized from data of the hybridization analysis (data not shown). Although it was possible to detect small regions of the mtDNA that showed some homology by hybridization, there was no evidence of significant regions of duplicated mtDNA. Digestion of the mtDNA with *NotI* resulted in a single fragment that migrated slowly as a linear form and was of genomic size. Uncut mtDNA of the same preparation had a faster migration and appeared as a fuzzy band on the gel suggesting that it consisted of super coiled circular molecules (data not shown). A few *EcoRI* sites

could not be precisely mapped but their approximate locations are shown in Fig. 1. A detailed physical map of the sequenced region is shown in Fig. 2.

3.2. Mitochondrial DNA RFLPs

The eight probes for detecting mtDNA RFLPs (N.K. Van Alfen, unpublished) were located on the mtDNA map (Fig. 1) and, with the exception of probes 7 and 47, which partially overlap, were found to be independent. These probes detected ten mtDNA haplotypes among the twelve strains isolated near Grimacco, Italy (Table 2). Probes 5, 17, and 26 did not identify any polymorphism within the Grimacco population, while in

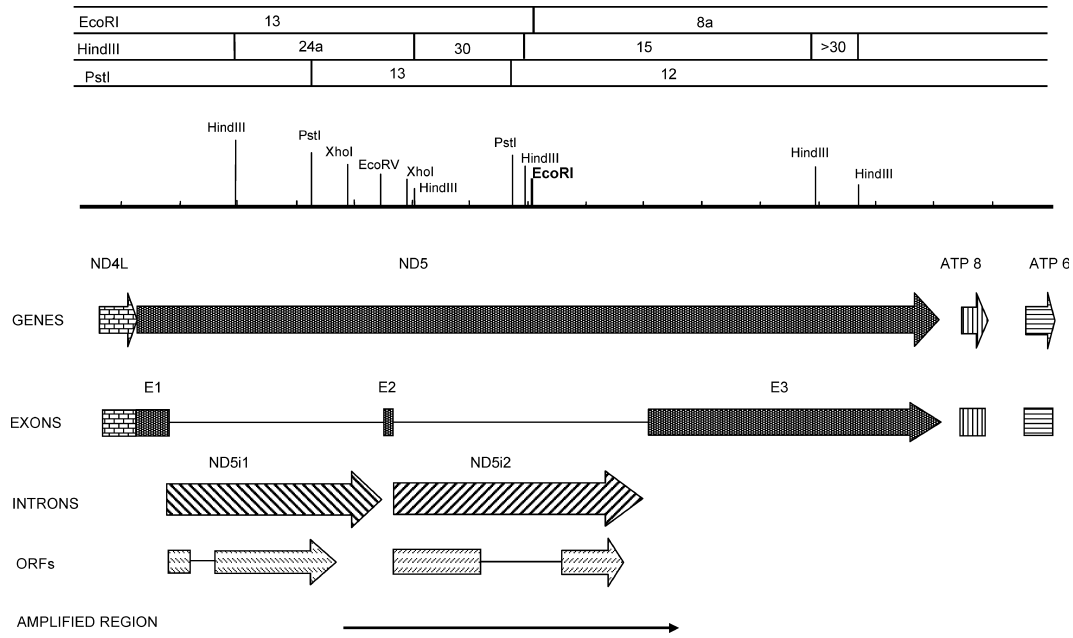


Fig. 2. Structure of the portion of polymorphic region 2 (see Fig. 1) sequenced in EP155. The locations of four genes are shown: ND4L (*NADH dehydrogenase* subunit 4L); ND5 (*NADH dehydrogenase* subunit 5); ATP 8 (*ATP synthase* subunit 8); and ATP 6 (*ATP synthase* subunit 6). In ND5 three exons (E1, E2, E3) and two introns, ND5i1 (intron 1) and ND5i2 (intron 2), are present and intron ORFs with the intervening sequences (black lines) are indicated. In the lowest line the amplified region of this sequence is reported. The restriction map shows in bold the position of the *EcoRI* site located between fragments E8a and E13 of the mtDNA map (Fig. 1). In addition a partial physical map with also *HindIII* and *PstI* is given.

previous studies they detected, respectively, three, two, and three different haplotypes of mtDNA out of 21 strains deriving from a world-wide collection (N.K. Van Alfen, unpublished). The eight probes identify four regions in the mtDNA; the RFLP region 1 is defined by probes 17 and 5, the RFLP region 2 by probes 4 and 13 while the RFLP region 3 contains probes 7, 47 and also probe 38, the RFLP region 4 consists only of probe 26 (Fig. 1). Each RFLP regions measures at least 10 kb. Probes 4 and 13, which mapped in the RFLP region 2 of the mtDNA, identified the largest number of haplotypes in this and other population studies (N.K. Van Alfen,

unpublished). Probe 4 and RFLP region 2 were therefore selected for further investigations.

The RFLP region 2 of the mtDNA (Fig. 1) was partially mapped in the Italian strains and the results suggested that the detected polymorphisms were the result of insertion/deletion events rather than restriction site changes. Five RFLP patterns were found by using probe 4; the RFLP patterns D, A, and H differed in the size of the fragment E8a that was respectively 7.9, 7.2, and 6.8 kbp due to small deletions. The E and F patterns had a deletion causing the loss of the *EcoRI* site, moreover the former also had an insertion of 2.2 kbp in

Table 2
Mitochondrial DNA RFLP haplotypes detected by the eight RFLP clones in 13 strains of *C. parasitica*

Strains	RFLP probes								mtDNA haplotype
	4	5	7	13	17	26	38	47	
I-1	H	A	A	C	A	A	D	A	1
I-2a	A	A	A	D	A	A	A	A	2
I-2b	F	A	B	D	A	A	D	A	3
I-3	E	A	B	C	A	A	D	B	4
I-4	A	A	B	A	A	A	D	B	5
I-6	A	A	B	A	A	A	A	B	6
I-7	F	A	D	D	A	A	E	E	7
I-8	A	A	B	A	A	A	A	B	6
I-9	E	A	A	E	A	A	D	A	8
I-10	A	A	B	A	A	A	A	B	6
I-11	H	A	A	B	A	A	D	A	9
I-12b	H	A	A	C	A	A	D	C	10
EP155	D	A	A	A	A	A	A	A	11

Letters indicate different restriction patterns of mtDNA when hybridized to each probes.

EP8a while the latter another deletion of about 1.6 kbp (data not shown).

3.3. Sequencing of polymorphic region 2

The polymorphic region 2 was sequenced in strain EP155. The sequence of 8379 bp of the mtDNA of this region was deposited in GenBank under Accession No. AF456838. A database search showed that the region has four open reading frames with homology to the following known mtDNA genes: *NADH dehydrogenase* subunit 4L (*ND4L*), *NADH dehydrogenase* subunit 5 (*ND5*), *ATP synthase F0* subunit 8 (*ATPase 8*), and *ATP synthase F0* subunit 6 (*ATPase 6*). In Fig. 2 the predicted gene locations are depicted.

3.4. Structure of the *ND4L*, *ND5*, *ATPase 8*, and *ATPase 6* genes

The entire exon sequence of *ND4L* was contained in the region sequenced. The *C. parasitica ND4L* showed extensive sequence homology with those of *P. anserina* (98.9% amino acidic similarity of the putative coded protein), *N. crassa* (88.8% amino acidic similarity) and *Hypocrea jecorina* (88% amino acidic identity). The *ND4L* termination codon overlapped the ATG start codon of the *ND5* gene as occurs in these genes of *P. anserina* and *N. crassa*. The structure of the *C. parasitica ND4L* gene differed, however, from the gene of these other fungi since, like that of the *ND4L* genes of *Aspergillus nidulans* (Brown et al., 1985) and *H. jecorina* (Chambergo et al., 2002), it does not contain introns in strain EP155.

ND5 had a more complex structure than that of *ND4L* since it was interrupted by 2 group I introns, each containing one ORF. The first intron (*ND5i1*) was a group I intron, 2235 bp long, starting at position 248 of the exon sequence. It displayed a nucleotidic sequence 77.2% identical with that of *P. anserina ND5i1* and a 78% similarity in amino acids. The *ND5i1* ORF was 1625 bp

long and contained two LAGLIDADG dodecapeptide sequence motifs typical of endonucleases (Dalgaard et al., 1997; Jurica and Stoddard, 1999). This ORF sequence was 82% similar to the *ND5* intron 1 protein of *P. anserina* mitochondrion (Genbank accession number S09142) and 56% similar to the *Ophiostoma novo-ulmi* possible RNA splicing or intron mobility element (Genbank accession number CAA87391). The *ND5i1* intron sequence was interrupted at position 218 by a 156 bp sequence not significantly similar to any known sequence found by Blast analysis.

Following this inserted sequence there was a frame shift in the intron sequence that presumably makes the encoded intron protein nonfunctional. The second exon of *ND5* continued for 75 bp and then was interrupted by a second group I intron (*ND5i2*) at position 324. The *ND5i2* intron was 1641 bp in length and contained a 1359 bp long ORF coding a putative endonuclease protein with two LAGLIDADG motifs (Dalgaard et al., 1997; Jurica and Stoddard, 1999). The *ND5i2* of *C. parasitica* had 81% nucleotide homology to the *ND5* intron 2 of *P. anserina* and 70% identity to the *ND5* intron 1 of *N. crassa* while its ORF showed, respectively, 83 and 76% similarity to the *ND5* intron 2 protein of *P. anserina* and to the *ND5* intron 1 protein of *N. crassa*. The *C. parasitica ND5i2* was interrupted at position 812 by a 759 bp insertion that lacked significant homology to sequences in GenBank, thus its origin is unknown. Nevertheless within the insertion, there was a 46 bp sequence with similarity in sequence to group II introns (Fig. 3), moreover it was 86% identical to a portion of the InC9 insertion (GenBank Accession No. AF218209) in the small subunit ribosomal RNA of *C. parasitica* and 88% identical to a portion of the intron 3 of the same gene (GenBank Accession No. AF029891) that has recently been identified as a group II intron in spite of its LAGLIDADG ORFs (Toor and Zimmerly, 2002).

The secondary structure of *ND5i1* (data not shown) demonstrated the lack of an extensive C domain, indicating that this intron belonged to subgroup IB

Accession number	Organism and gene	Sequence
AF456838	<i>C. parasitica</i> Intervening sequence IS	GGAGAGCCGT ATGCTGAGAA ATTCGCACGT ACGGTTCCGA
AF218209	<i>C. parasitica</i> InC9	GGAGAGCCGT GTGCCGGGAA ACTCGCATGC ACGGTTCCGA
AF029891	<i>C. parasitica</i> mtSrRNA	TTTGAGCCGT ATGCGATGAA AGTCGCACGT ACGGTTCTTA
Z69899	<i>Podospora comata</i> cox1	GGAGAGCCGT ATGCCGTGAA AGTCGCACGT ACGGTTCCGA
M36911	<i>Podospora anserina</i> cox1	GGAGAGCCGT ATGCCGTGAA AGTCGCACGT ACGGTTCCGA
X03085	<i>Podospora anserina</i> α -sen	GGAGAGCCGT ATGCCGTGAA AGTCGCACGT ACGGTTCCGA
X04336	<i>Podospora anserina</i> class II intron	GGAGAGCCGT ATGCAGGGAA ACTTGCACGT ACGGTTCCGA
Z69898	<i>Podospora curvicolla</i> intron	GGAGAGCCGT ATGATGGGAA ACTATCACGT ACGGTTCCGT
M68929	<i>Marchantia polymorpha</i> mtDNA	GGAGAGCCGT ATGCCGGGAG ACTCGCACGT ACGGTTCCGA
U41288	<i>Allomyces macrogynus</i> mtDNA	GGAGAGCCGT ATGCCGTGAA AGTTGCACGT ACGGTTCCGT

Fig. 3. DNA sequence comparison of the 46 bp portion of the intervening sequence of *ND5i2* of *Cryphonectria parasitica* with the consensus sequences of similar group II introns found in fungal mitochondrial DNA.

(Cummings and Domenico, 1988). The *ND5i2* had an extended C domain so it could be defined as a type C intron (data not shown). The entire exon sequence of *C.parasitica ND5* resembled that of *P. anserina*, with 87.6% amino acid similarity and that of *N. crassa*, with 86.8% amino acid similarity.

The *ATPase 8* gene started at position 7384, 369 bp from the end of the *ND5* gene. Its peptide sequence shared a very high similarity with those of *P. anserina* and *N. crassa* (95 and 91%, respectively). Only a partial sequence, 219 bp, of *ATPase 6* gene was obtained, starting at position 8086, but its sequence was again similar to those of *N. crassa* and *P. anserina* DNA sequences, with 88 and 81% identity, respectively.

3.5. Amplifications and sequences of the polymorphic region 2

Mapping of the polymorphic region 2 of the Italian strains identified 7 length polymorphisms within this

Table 3
mtDNA amplification products of strains of *C. parasitica*

Strain	PCR product size (bp)		
	3200	2400	800
Ep 155	+		
I-12b	+§		
I-10	+§		
I-9		+§	+§
I-7			+§
I-4		+	+§
I-3		+§	+

+, fragment originated by PCR; §, fragment cloned and sequenced.

region. Sequencing of this region in strain EP155 indicated that some of the detected polymorphisms could be the result of presence or absence of the intron *ND5i2* including the *EcoRI* site within the *ND5* gene. Two primers mt2 and eg1f4 were therefore designed to anneal respectively upstream and downstream of the *ND5i2*. The oligonucleotide primer mt2 was situated in the 3' end of *ND5i1* and the primer eg1f4 within exon E3 of *ND5* to contain the full length of the intron and to compare the exon–intron junctions of *ND5* of the Italian strains polymorphic in this region.

Using the selected primers, amplification of EP155 and Italian strains I-10 and I-12b yielded a single fragment of approximately 3200 bp, whereas an 800 bp was amplified from strain I-7; from strains I-3, I-4, and I-9 two fragments were amplified that were about 2400 and 800 bp. The amplification products were cloned and sequenced as follows: the largest fragment amplified from strain I-3 (2400 bp), the smallest from I-4 (800 bp) and the single fragments amplified from strains I-7, I-10, and I-12b (Table 3). Their sequences were deposited at GenBank under Accession Nos. AF456839, AF456840, AF456841, AF456844, and AF46845. Each of the two fragments amplified in isolate I-9 was cloned and sequenced. The smaller of these two fragments was named I-9/3 and the larger I-9/4. Their GenBank sequence accession numbers are AF456842 and AF456843, respectively.

The sequences of the region of the *ND5* gene containing intron *ND5i2* were identical in the strains I-10, I-12b, and EP155 (Fig. 4). This region was thus not polymorphic between these strains. The *ND5i2* intron was present in I-3 and I-9/4 but these strains lacked the

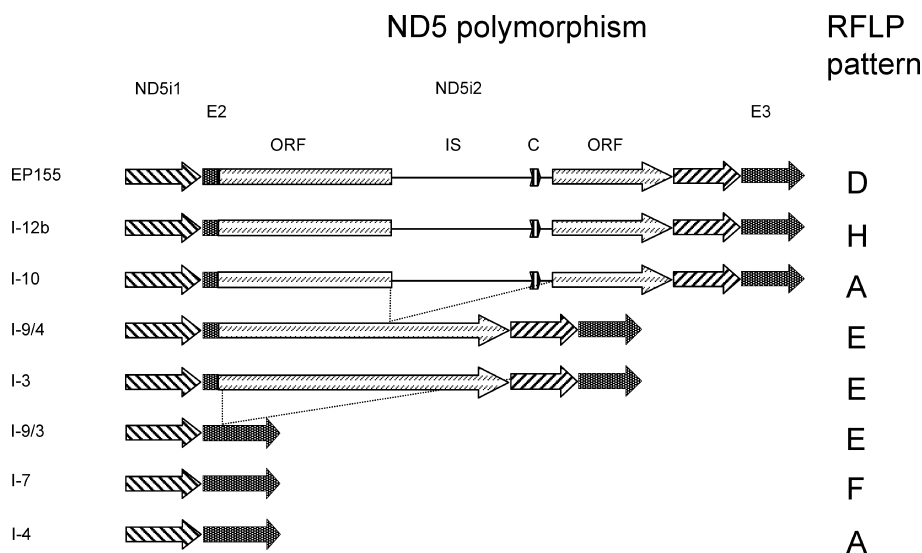


Fig. 4. Map of a portion of polymorphic region 2 of the mtDNA in seven strains of *C. parasitica* (see Fig. 2). The region sequenced in these strains included part of the intron *ND5i1* and all of intron *ND5i2* of the *ND5* gene. E2 and E3 indicate the locations of the *ND5* exons 2 and 3; ORF indicates the *ND5i2* intron ORF; IS is the intervening sequence and C is the group II intron consensus sequence. The mtDNA haplotypes identified by Southern analysis for each strain are reported.

759 bp insertion in the intron. In strains I-4, I-7, and I-9/3 the *ND5* gene did not contain intron *ND5i2*. The polymorphisms detectable within this region of the *ND5* gene were thus based also on the presence or absence of an intron and an insertion within that intron.

The six Italian strains that were sequenced divided into four haplotype groups as determined in hybridization experiments using probe 4 while sequencing part of this region revealed only three polymorphic possibilities (Fig. 4). For instance, the sequence of the mt2/eg1f4 amplified fragment was identical in the strains I-10 and I-12b, although their *EcoRI* patterns revealed by hybridization with probe 4 were different. It is assumed that this is due to polymorphisms present in the *EcoRI* fragments E8a or E13 upstream or downstream from the sequenced region. The portion of the fragment E8a that was not amplified and sequenced in the six Italian strains did not contain any other introns, while the portion not sequenced of the fragment E13 included the *ND5i1* ORF and its intervening sequence (Fig. 2). The differences between the RFLP pattern D, A and H possibly resided in that portion of the E13 fragment. The intervening sequences within the *ND5i2* intron contained an *EcoRI* site whose loss is responsible for the RFLP pattern E and F where only one *EcoRI* fragment was detected by Southern analysis using probe 4. The fragment of the F pattern had another 1.6 kbp deletion corresponding exactly to the lack of the *ND5i2*. The pair of strains I-4 and I-10 as well as I-9/3 and I-9/4 shared the same haplotypes A and E, respectively, but their sequences were quite different.

4. Discussions

4.1. Mapping of mtDNA

The circular mitochondrial molecule of *C. parasitica* is one of the larger reported in fungi. Its size and thus large number of restriction fragments made mapping it a challenge by the conventional method of comparison of double digested fragments. Two additional approaches, hybridization analysis and mapping of restriction sites within individual fragments, were used to complete the map. About 40% of the mtDNA proved to be difficult to clone, as estimated by those clones under-represented in the lambda genomic library. The difficulties in cloning all of the mtDNA genome did not allow the assembling of overlapping clones covering the entire genome, thus it was not demonstrated with certainty that the mtDNA was circular. Nonetheless, both the electrophoretic mobility of the uncut mtDNA genome and the hybridization data indicate a circular form. A physical map of *C. parasitica* was published by Bell et al. (1996) but it is difficult to make direct comparisons with the map

published here since only *EcoRI* was used in common between the two studies and only fragment E1 (Table 1) was located on their map. Nevertheless in an attempt to contribute to the previous map, a detailed physical map of the sequenced region 2 with the restriction fragments previously used is reported. The order and denomination of some restriction fragments did not correspond to those of Bell et al. (1996) as they estimated the molecular size of the restriction fragments by electrophoretic analysis, a technique not as precise as sequencing. This work therefore extends and rectify the previous map.

The mtDNA of *C. parasitica* has been described in previous studies as being polymorphic (Gobbi and Locci, 1990; Milgroom and Lipari, 1993), but previous studies have used isolated mtDNA as probes and so there was no information provided regarding the nature of the polymorphisms in the mtDNA. One of these studies, however, indicated that in one small population each isolate contained a unique mtDNA haplotype (Gobbi and Locci, 1994). These studies suggested that the mtDNA was variable but they lacked information regarding the nature of the variable regions. Mapping mtDNA clones found to detect RFLPs in different strains of *C. parasitica* showed that there were four general regions of the mtDNA detectable as polymorphic. Using these probes to detect mtDNA polymorphisms within a small population of strains isolated near Grimacco, Italy, showed that there were at least 10 different mtDNA haplotypes among these 12 strains.

4.2. Characterization of a polymorphic region

One of the polymorphic regions of the mtDNA, region 2 (Fig. 1) was selected for sequencing and found to be a region that is dense in genes (Fig. 2). The density of ORFs within the polymorphic region selected for sequencing was not expected, particularly when the bulk of the genome must not be functional. Although there is no direct evidence that the four genes identified in this region are functional, there is nothing in the sequence data that would suggest otherwise. The lack of hybridization of the probes of these genes with other regions of the mtDNA suggests that the ORFs are not repeated elsewhere. Three of these genes, *AT-Pase* subunits 6, 8, and *ND5*, were mapped by hybridization by Bell et al. (1996). These genes were found to be clustered as on the map presented here, although their relative positions differed from our map (Fig. 1). Since in the earlier study the genes were mapped individually by hybridization, the gene order given here, obtained by sequencing the region, is probably the correct one.

Within region 2 of the mtDNA two different group I introns were found in strain EP155. To compare the part of region 2 that contains these introns in the Italian

strains, primers were designed for sequencing of these regions from strains with RFLPs mapped to this area of the mtDNA genome. Sequence analysis of six of the Italian strains indicated that intron *ND5i1* was present in all of the strains, as in EP155, but that intron *ND5i2* was optional in some of the strains (Fig. 4). The optional nature of this intron between strains suggests that it is mobile, a supposition that is supported by the ORF found in the intron with the dodecapeptide motifs typical of mobile endonucleases (Lambowitz and Belfort, 1993). The presence or absence of this intron is obviously responsible for an RFLP difference between some of these strains, e.g., between strains I-3 and I-4.

Within intron *ND5i2* a small insertion of unknown origin was detected in strains EP155, I-12b and I-10. The presence of an identical insertion within this intron in strains from both North America (EP155) and Italy suggests that it is not of recent origin. There may be functional consequences related to the presence of this insertion. A similar insertion was found in intron *ND5i1* in strain EP155. Since the region containing the insertion within this intron was not sequenced in the other strains, it is not known how common it is. Presumably the presence of the intervening sequences within the ORF of the introns generates a non-functional intronic protein, most likely causing the intron to lose its ability to move. Strains I-9/4 and I-3 do not contain this insertion within the ORF of the intron so they may be mobile. In the strains containing the insertion sequences within the introns, such as EP155, there is no evidence that the functions of these mitochondrial genes are impaired. EP155 is the standard virulent laboratory strain of *C. parasitica*, chosen in part for its consistent growth in culture. The nature of these intervening sequences and their biological role are not known, however they result very interesting for the high (about 90%) identity of short stretches of their DNA with group II intron sequences as well as their similarity to the plasmid-like element InC9 of *C. parasitica*, causing a mitochondrial form of hypovirulence, or to the α -sen of *P. anserina*, causing senescence of the mycelium (Cummings et al., 1985).

Amplification of mtDNA sequences using the primer pairs resulted in two different PCR fragments being amplified in strains I-9, I-3, and I-4. It is possible that these were the result of cross contamination, but the consistency with which they were amplified only in these strains makes this unlikely. Sequencing of the two different fragments in I-9 showed that the difference between the fragments was the presence or absence of the *ND5i2* intron within the amplified DNA. Strains I-3, I-4, and I-9 may therefore contain two or more different types of mtDNA, one mtDNA type containing the *ND5i2* intron and another lacking it. PCR amplification using this primer pair in strains EP155, I-10, and I-12 yielded a single fragment, suggesting that the mtDNA molecules of these strains are homogenous with respect to the presence of the

ND5i2 intron. These data suggest that the mtDNA of *C. parasitica* may exist as a population of different mtDNA types within a single strain. This complexity of mtDNA types, however, was not obvious in the haplotype analysis using hybridization probes. There was not complete agreement between the five haplotypes obtained from the RFLP analysis and the three polymorphisms detected by sequencing of the six Italian strains (Fig. 4). The polymorphisms detected by sequencing were entirely within the *ND5* gene and the primer pair used amplified less than 8379 bp of the portion of region 2 that was sequenced. There are clearly more polymorphic areas within the region of mtDNA hybridizing to probe 4, i.e., 11.8 kbp, than those detected by the amplification of 3.2 kbp using the selected primer pairs.

Three possible length polymorphisms were detected due to insertions/deletion events involving an intron and insertions within that intron. The *ND5i1* intron, which was only partially sequenced in the Italian strains, contains insertion sequences within the intron just as *ND5i2* does. We do not have evidence that such strains exist, but presumably introns without the insertions preceded those with insertions.

The purpose of this study was to begin to characterize the basis of the mtDNA haplotype diversity that has been observed in field populations of *C. parasitica*. The strains compared were all isolated from a line of chestnut trees in a small geographic area. A group of mtDNA probes detected 10 RFLP haplotypes among the 12 isolates. A region containing two of these probes was sequenced and then an area representing about a third of this sequence from six of the Italian strains was compared. The polymorphic differences identified in the approximately 3200 bp sequenced region were all the result of length polymorphisms associated with a group I intron. While the polymorphism identified in this region could only explain three of the 10 total mtDNA haplotypes within this small population, the results demonstrate for the first time that insertion/deletion events associated with introns were responsible for some of the polymorphisms detected in this fungus. Given the size of the mtDNA, it was not expected that the polymorphic region selected for analysis would be so dense in presumably functional genes. It remains to be seen if the other polymorphic regions detected by the mtDNA probes are also associated with coding regions of the mtDNA.

Acknowledgments

We thank Dr. Jacqueline Rogers for revising the English text. The work of E.G. in Davis was supported by the fellowship "Short Mobility 2001" of the National Research Council (CNR) of Italy.

References

- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Sideman, J.G., Smith, J.A., Stroh, K., 1991. Current Protocols in Molecular Biology. Greene/Wiley-Intersciences, New York.
- Belcour, L., Rossignol, M., Koll, F., Sellem, C.H., Oldani, C., 1997. Plasticity of the mitochondrial genome in *Podospora*. Polymorphism for 15 optional sequences: group-I, group-II introns, intronic ORFs and an intergenic region. *Curr. Genet.* 31, 308–317.
- Bell, J.A., Monteiro-Vitorello, C.B., Hausner, G., Fulbright, D.W., Bertrand, H., 1996. Physical and genetic map of the mitochondrial genome of *Cryphonectria parasitica* EP155. *Curr. Genet.* 30, 34–43.
- Bisiach, M., De Martino, A., Gobbi, E., Intrepido, M., Vegetti, G., 1988. Studies on chestnut blight: activity report. *Riv. Pat. Veg.*, S. IV, 24, 3–13.
- Brown, T.A., Waring, R.B., Scazzocchio, C., Davies, R.W., 1985. The *Aspergillus nidulans* mitochondrial genome. *Curr. Genet.* 9, 113–117.
- Bruns, T.D., Palmer, J.D., Schumard, D.S., Grossman, L.I., Hudspeth, M.E.S., 1988. Mitochondrial DNAs of *Suillus*: three fold size change in molecules that share a common gene order. *Curr. Genet.* 13, 49–56.
- Chambergo, F.S., Bonaccorsi, E.D., Ferreira, A.J.S., Ramos, A.S.P., Ribamar Ferreira Jr., J., Abrahão-Neto, J., Simon Farah, J.P., El-Dorry, H., 2002. Elucidation of the metabolic fate of glucose in the filamentous fungus *Trichoderma reesei* using expressed sequence tag (EST) analysis and cDNA microarrays. *J. Biol. Chem.* 277, 13983–13988.
- Cummings, D.J., MacNeil, I.A., Domenico, J., Matsuura, E.T., 1985. Excision-amplification of mitochondrial DNA during senescence in *Podospora anserina*. DNA sequence analysis of three unique 'plasmids'. *J. Mol. Biol.* 185, 659–680.
- Cummings, D.J., Domenico, J.M., 1988. Sequence analysis of mitochondrial DNA from *Podospora anserina*. Pervasiveness of a class I intron in three separate genes. *J. Mol. Biol.* 204, 815–839.
- Cummings, D.J., Michel, F., Domenico, J.M., McNally, K.L., 1990. DNA sequence analysis of the mitochondrial *ND4L-ND5* gene complex from *Podospora anserina*. Duplication of the *ND4L* gene within its intron. *J. Mol. Biol.* 212, 269–286.
- Dalgaard, J.Z., Klar, A.J., Moser, M.J., Holley, W.R., Chatterjee, A., Saira Mian, I., 1997. Statistical modeling and analysis of the LAGLIDADG family of site-specific endonucleases and identification of an intein that encodes a site-specific endonuclease of the HNH family. *Nucl. Acids Res.* 25, 4626–4638.
- Devereux, J., Haerberli, P., Smithies, O., 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* 12, 387–395.
- Gobbi, E., Locci, R., 1990. Investigation on strains of *Cryphonectria parasitica*, agent of chestnut blight, isolated in Italy. In: Proceedings of the 8th Congress of the Mediterranean Phytopathological Union, Actes Editions, Rabat, Maroc, pp. 67–69.
- Gobbi, E., Locci, R., 1994. Caratterizzazione intraspecifica di ceppi di *Cryphonectria parasitica* mediante RFLP del DNA mitocondriale. In: Atti del convegno Moderni indirizzi diagnostici in patologia vegetale, Torino 28–29 Ottobre, p. 3.
- Gobbi, E., Wang, Y., Martin, R.M., Powell, W.A., Van Alfen, N.K., 1990. Mitochondrial DNA of *Cryphonectria parasitica*: lack of migration between vegetatively compatible strains. *Mol. Plant Microbe Interact.* 3, 66–71.
- Gray, M.W., Burger, G., Lang, B.F., 1999. Mitochondrial evolution. *Science* 283, 1476–1481.
- Griffiths, A.J., Collins, R.A., Nargang, F.E. 1995. Mitochondrial genetics of *Neurospora*. In: Kuck, U. (Ed.) *The Mycota*, vol. II, Springer, Berlin, pp. 93–105.
- Hamari, Z., Pfeiffer, I., Ferenczy, L., Kevei, F., 1999. Interpretation of variability of mitochondrial genomes in the species *Aspergillus carbonarius*. *Antonie van Leeuwenhoek* 75, 225–231.
- Hansen, D.R., Van Alfen, N.K., Gillies, K., Powell, W.A., 1985. Naked dsRNA associated with hypovirulence of *Endothia parasitica* is packaged in fungal vesicles. *J. Gen. Virol.* 66, 2605–2614.
- Heiniger, U., Rigling, D., 1994. Biological control of chestnut blight in Europe. *Annu. Rev. Phytopathol.* 32, 581–599.
- Hintz, W.E., Mohan, M., Anderson, J.B., Horgen, P.A., 1985. The mitochondrial DNAs of *Agaricus*: heterogeneity in *A. bitorquis* and homogeneity in *A. brunnescens*. *Curr. Genet.* 9, 127–132.
- Jurica, M.S., Stoddard, B.L., 1999. Homing endonucleases: structure, function and evolution. *Cell. Mol. Life Sci.* 55, 1304–1326.
- Lambowitz, A.M., Belfort, M., 1993. Introns as mobile genetic elements. *Annu. Rev. Biochem.* 62, 587–622.
- Lecellier, G., Silar, P., 1994. Rapid methods for nucleic acids extraction from Petri dish-grown mycelia. *Curr. Genet.* 25, 122–123.
- Milgroom, M.G., Lipari, S.E., 1993. Maternal inheritance and diversity of mitochondrial DNA in the chestnut blight fungus, *Cryphonectria parasitica*. *Phytopathology* 83, 563–567.
- Monteiro-Vitorello, C.B., Bell, J.A., Fulbright, D.W., Bertrand, H., 1995. A cytoplasmically transmissible hypovirulence phenotype associated with mitochondrial DNA mutations in the chestnut blight fungus *Cryphonectria parasitica*. *Proc. Natl. Acad. Sci. USA* 92, 5935–5939.
- Palmer, J.D., Herbon, L.A., 1986. Tricircular mitochondrial genomes of *Brassica* and *Raphanus*: reversal of repeat configurations by inversion. *Nucl. Acid Res.* 14, 9755–9764.
- Palmer, J.D., 1990. Contrasting modes and tempos of genome evolution in land plant organelles. *Trends Genet.* 6, 115–120.
- Pearson, W.R., Lipman, D.J., 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* 85, 2444–2448.
- Powell, W.A., Van Alfen, N.K., 1987. Two non homologous viruses of *Cryphonectria (Endothia) parasitica* reduce accumulation of specific virulence-associated polypeptides. *J. Gen. Bacteriol.* 169, 5324–5326.
- Salvo, J.L., Rodeghier, B., Rubin, A., Troischt, T., 1998. Optional introns in mitochondrial DNA of *Podospora anserina* are the primary source of observed size polymorphisms. *Fungal Genet. Biol.* 23, 162–168.
- Sambrook, J., Fritsch, E.F., Maniatis, T.A., 1989. *Molecular Cloning: A Laboratory Manual*, second ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Small, R., Suffolk, C.J., Leaver, 1989. Evolution of plant mitochondrial genomes via substoichiometric intermediates. *Cell* 58, 69–76.
- Toor, N., Zimmerly, S., 2002. Identification of a family of group II introns encoding LAGLIDADG ORFs typical of group I introns. *RNA* 8, 1373–1377.
- Van Alfen, N., 1988. Viruses of *Endothia parasitica*. *Mycol. Ser.* 7, 371–386.
- Zhang, L., Villalon, D., Sun, Y., Kazmierczak, P., Van Alfen, N.K., 1994. Virus-associated down-regulation of the gene encoding cryparin, an abundant cell-surface protein from the chestnut blight fungus, *Cryphonectria parasitica*. *Gene* 139, 59–64.
- Zimmer, M., Welser, F., Oraler, G., Wolf, K., 1987. Distribution of mitochondrial introns in the species *Schizosaccharomyces pombe* and the origin of the group II intron in the gene encoding *apocytochrome b*. *Curr. Genet.* 12, 329–336.