Theor Appl Genet (2008) 117:459–470
DOI 10.1007/s00122-008-0790-7

ORIGINAL PAPER

The mitochondrial genome of a cytoplasmic male sterile line of perennial ryegrass (*Lolium perenne* L.) contains an integrated linear plasmid-like element

Paul McDermott · Vincent Connolly · Tony A. Kavanagh

Received: 24 January 2008 / Accepted: 2 May 2008 / Published online: 27 May 2008

Abstract The mitochondrial genome of a cytoplasmic male sterile line of perennial ryegrass (*Lolium perenne* L.) was shown to contain a 9.6 kb element, LpCMSi, that is absent in the mitochondrial genome of fertile lines. LpCMSi contains the previously described chimeric gene *orfC9*, and three additional open reading frames (orfs) encoding a unique 45 kDa predicted protein of unknown function, a family B-like DNA polymerase (LpDpo), and a phage-type single subunit RNA polymerase (LpRpo). The latter two proteins shared significant similarity with DNA and RNA polymerases encoded by extrachromosomal linear mitochondrial plasmids of plants and fungi, and also to integrated plasmid-like sequences found in various plant and fungal mitochondrial genomes. Transcripts for both LpDpo and LpRpo were detected by RT-PCR in mitochondria of the CMS line. PCR-based investigations further revealed the presence of LpCMSi-like sequences in fertile *L. perenne* lines that are likely maintained as low-copy number extrachromosomal replicons. The absence of integrated forms of LpCMSi in the mitochondrial genome of fertile lines suggests that LpCMSi integration adjacent to the *atp9* gene may be responsible, directly or indirectly, for the sterility phenotype of the CMS line.

Communicated by Y. Xue.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-008-0790-7) contains supplementary material, which is available to authorized users.

P. McDermott · V. Connolly · T. A. Kavanagh (✉)
Plant Molecular Genetics Laboratory,
Smurfit Institute of Genetics, Trinity College,
Dublin 2, Ireland
e-mail: tkvanagh@mail.tcd.ie

Present Address:
V. Connolly
Teagasc, Oakpark Research Centre, Carlow, Ireland

Introduction

Extrachromosomal plasmids are frequently associated with the mitochondria of filamentous fungi (Griffiths 1995), but only occasionally with mitochondria of higher plants (Brown and Zhang 1995). The most common class of mitochondrial plasmids comprises linear double-stranded DNA molecules of approximately 5–12 kb, the ends of which are defined by terminal inverted repeats (TIRs), and the presence of a terminal protein (TP) covalently attached to the 5′ termini. Transcription and replication of linear plasmids is initiated in the TIRs by, respectively, a plasmid-encoded phage-like single subunit RNA polymerase (Rpo) and a DNA polymerase (Dpo) of the B-type family (Meinhardt et al. 1990). Replication is therefore thought to occur by a protein-primed mechanism similar to that used by the *Bacillus* phage *phi 29* (Salas 1991) in which the TP forms a complex with the plasmid-encoded Dpo to initiate DNA synthesis in the TIRs. The efficiency of replication can be such that plasmid copy number may equal or exceed that of the mitochondrial genome (Brown and Zhang 1995).

In contrast to fungi, extrachromosomal linear plasmids have been identified in the mitochondria of only four species of higher plants. These include the high copy number S1 and S2 plasmids (Paillard et al. 1985; Levings and Sederoff 1983) associated with cytoplasmic male sterile (CMS) lines of *Zea mays* (maize) carrying the S cytoplasm (CMS-S or S-mitotype maize); an 11.6 kb linear plasmid in the mitochondria of *Brassica napus* (oilseed rape) (Palmer et al. 1983; Turpen et al. 1987; Handa et al. 2002); a

10.4 kb plasmid in *Beta vulgaris* subsp. *maritima* (sea beet) (Saumitou-Laprade et al. 1989; GenBank Accession Y10854); and a 9.2 kb plasmid recently identified in the FG21 mitotype of *Daucus carota* (carrot) (Robison and Wolyn 2005). All of these plasmids encode a phage-like Rpo and Dpo except for those of maize, which encode only one or the other polymerase.

Although linear plasmids do not encode a dedicated integrase function, they can integrate into the mitochondrial genome by recombining with short homologous sequences in mtDNA (Brown and Zhang 1995). A recent extensive search of the nucleotide sequence databases for plasmid-like sequences in the mitochondrial genomes of filamentous fungi suggests that most linear plasmids can integrate in this way (Cahan and Kennell 2005). Evidence for the integration of linear plasmids into plant mitochondrial genomes was originally provided by molecular investigations of the maize CMS-S mitotype where integration of the S plasmids is associated with extensive linearization of mtDNA (Schardl et al. 1985). The mitochondrial genomes of *Secale cereale* (rye) (Dohmen and Tudzynski 1994) and the liverwort *Marchantia polymorpha* (Weber et al. 1995) were subsequently shown to contain fragments of a linear plasmid-related Dpo despite the apparent absence of homologous extrachromosomal plasmids in the mitochondria of either species. Additional examples of integrated fragments of linear plasmid Rpo- and/or Dpo-like orfs were identified in the completely sequenced mitochondrial genomes of *Arabidopsis thaliana* (Unseld et al. 1997), *B. vulgaris* (sugar beet) (Kubo et al. 2000; Satoh et al. 2004), *Oryza sativa* (rice) (Notsu et al. 2002), *Z. mays* (Clifton et al. 2004), *Nicotiana tabacum* (tobacco) (Sugiyama et al. 2005) and *Triticum aestivum* (wheat) (Ogihara et al. 2005). Although not yet completely sequenced, the mitochondrial genome of *D. carota* mitotype SW3 has been shown to contain an integrated but truncated copy of the 9.2 kb extrachromosomal linear plasmid originally identified in carrot mitotype FG21 (Robison and Wolyn 2005).

Linear plasmids, by virtue of their ability to integrate into the mitochondrial genome, are potential insertional mutagens. In the fungus *Neurospora*, for example, the integration of either of two linear plasmids, p*Kalilo* (Chan et al. 1991) and p*Maranhar* (Court and Bertrand 1992) results in abnormal physiology and a premature senescence or 'death' phenotype. In plants, episomal and/or integrated linear plasmids are frequently associated with CMS phenotypes. However, evidence from maize (Escote et al. 1985), *Brassica* (Kemble et al. 1986) and sugar beet (*B. vulgaris*) (Satoh et al. 2004) suggests that this association is fortuitous rather than causative.

We previously reported the presence of sequences of apparently non-mitochondrial origin (relative to wild-type mtDNA) located downstream of a chimeric gene, *orfC9*, in the mitochondrial genome of a CMS line of *Lolium perenne* (perennial ryegrass) that could be used as a diagnostic probe for the CMS cytoplasm (Kiang et al. 1993; Kiang and Kavanagh, 1996). Here, we report a more extensive analysis of this region which shows that the CMS line contains an integrated 9.6 kb element, LpCMSi, that shares extensive homology with linear mitochondrial plasmids of fungi and higher plants.

Materials and methods

Plant material

The *L. perenne* male sterile line CMS-St (CMS9B290), a fertile-restored line CMS-Rst and the fertile maintainer line LPSB21 were described in Kiang et al. (1993). The fertile wild *L. perenne* lines LpFIp2, 4, 5 were isolated from Irish pastures. The *L. perenne* line LpFNz, a commercial seed cultivar from New Zealand was confirmed to be infected with the hybrid fungal endophyte *Epichloe typhina* × *Acremonium lollii* by PCR amplification and sequencing of the *pyr4-2* and *tub2-2* genes encoding, respectively, orotidine-5-monophosphate decarboxylase and tubulin was carried out using the primer pair Pyr4F, Pyr4R (Collett et al. 1995), and Tub2F, Tub2R (Schardl et al. 1994), respectively (data not shown).

Isolation of mtDNA

Mitochondrial DNA was purified from 7-day-old etiolated *L. perenne* seedlings. All steps were carried out on ice or at 4°C. Etiolated seedlings (150 g) were homogenized to a paste in 350 ml sterile Extraction Buffer (EB) (0.44 M sucrose, 50 mM Tris–HCl (pH 8.0), 3 mM EDTA (pH 8.0), 0.5% (w/v) bovine serum albumin, 10 mM β-mercaptoethanol (Kiang et al. 1993). A further 350 ml EB was added, followed by vigorous homogenization. The suspension was filtered through four layers of sterilised muslin and centrifuged at 1,000×*g* for 10 min to pellet cell debris and nuclei. The supernatant was centrifuged at 5,000×*g* for 20 min to pellet chloroplasts. The recovered supernatant was then centrifuged at 16,000×*g* for 20 min to pellet mitochondria.

The mitochondrial pellet was resuspended in 9 ml DNAase I buffer (0.44 M Sucrose, 50 mM Tris–HCl (pH 8.0), 10 mM $MgCl_2$). DNAase I (1 ml at 10 mg/ml) was added to the mitochondrial suspension and digestion was allowed to proceed on ice for 45–60 min. The digestion was terminated by the addition of 0.5 M EDTA (pH 8.0) to a final concentration of 25 mM. Mitochondria were re-pelleted at 16,000×*g* for 10 min. The pellet was resuspended in 20 ml of Wash Buffer (WB) (0.6 M Sucrose, 25 mM EDTA (pH 8.0), 50 mM Tris–HCl (pH 8.0). Once again the pellet was

resuspended in 20 ml WB and the mitochondria re-pelleted at 16,000×*g* for 10 min. Mitochondria were lysed by resuspension in 2 ml NTE buffer (10 mM NaCl, 10 mM Tris–HCl (pH 8.0), 1 mM EDTA (pH 8.0)), followed by the addition of SDS to a final concentration of 1% (w/v) and incubation at 37°C for 30 min. Mitochondrial nucleic acids were purified by three phenol extractions followed by two phenol–chloroform extractions. Nucleic acids were precipitated by the addition of two volumes of ice-cold absolute ethanol and recovered by centrifugation at 16,000×*g* for 10 min. The pellet was air-dried and resuspended in 1× TE.

DNA and RNA analyses

Total tissue DNA was isolated from 14 day old *L. perenne* seedling leaf tissue as previously described (Kiang et al. 1994). Samples (10 µg), digested with restriction enzymes, were electrophoresed in a 0.8% (w/v) agarose gel, blotted onto Hybond N (Amersham) membrane and hybridized with a ^{32}P-labeled probe (Kiang et al. 1993). Total RNA was extracted from leaf tissue by homogenization in extraction buffer containing 100 mM Tris, 50 mM EDTA, 50 mM NaCl, 2% SDS, 10 mM β-mercaptoethanol. The mixture was emulsified with an equal volume of buffer-saturated phenol and the phases separated by centrifugation. The aqueous phase was extracted twice more with phenol and once with a phenol:chloroform (50:50) mixture. The aqueous phase was adjusted to 4 M lithium chloride from which total RNA was precipitated at −20°C. The recovered RNA precipitate was washed with 70% ethanol and resuspended in 1× TE (pH 8.0) (Kiang and Kavanagh 1996).

PCR and RT-PCR

PCR amplifications using *L. perenne* DNA (50–100 ng) were performed using Taq DNA polymerase (Invitrogen), and primer pairs are outlined in Table 1 and in Fig. 6, and the following cycle parameters: denaturation at 94°C for 5 min followed by 25 cycles of 1 min at 94°C, 1 min at 60°C, 1 min at 72°C and a final extension at 72°C for 5 min. All PCR products were sequenced directly with the primers used in their amplification. Oligonucleotide primer sequences are listed in Table 1. The location of the various primer sequences is illustrated in Fig. 1b.

For RT-PCR, total RNA (1–4 µg) from *L. perenne* leaf tissue pre-treated with RNAase-free DNAase I was added to 100–200 pmol of random primers (La Roche) and denatured at 70°C for 10 min, rapidly cooled on ice and used in the synthesis of first strand cDNA. The reaction mixture (50 µl) was incubated at 20–25°C for 10 min, preheated to 42°C for 2 min before adding 200 units of Superscript II reverse transcriptase (Invitrogen). Following incubation at 42°C for 60 min, the reaction was stopped by heating to 70°C for 15 min. RT-PCR (35 cycles) was performed by using 1 µl of first strand cDNA with primer pairs as detailed in Table 1 and Fig. 5. PCR product sizes up to 1.5 kb were produced using 2.5 units Platinum Taq DNA polymerase (Invitrogen) and the conditions described above. RT-PCR products were sequenced directly with the primers used in their amplification.

Table 1 Primer sequences used to amplify LpCMSi fragments

Fragment (size bp)	Forward (f) and reverse (r) primers (5′→ 3′)	
D1 (738)	f	CAATCAGTATACAGAAGGATAAGTAC
	r	GTACTCGTCATTGTGCTTAGTAAAC
D2 (1,295)	f	CACCAATAGTCTATTTTCATAAC
	r	GTCCGTATAATAGCAGTCATC
D3 (703)	f	GATGACTGCTATTATACGGACAC
	r	GGAAAGGTATAAGATGAAATCTGATG
D4 (555)	f	GTTATTGCTAAATTCATTATATGGTAG
	r	GTATACGTGACCTCATCTATATG
DF1 (549)	f	CTAGAACATCTATAGTGACAGATG
	r	GTACTTATCCTTCTGTATACTGATTG
DF2 (1,115)	f	CTCACTGAACTCATGATAAATGTTC
	r	GTACTCGTCATTGTGCTTAGTAAAC
RP1 (843)	f	CAGATGTTTGATTACTTAGCGAAG
	r	GATAATTTCTACCACGGAAATCAAG
RP2 (1,022)	f	CTTGATTTCCGTGGTAGAAATTATC
	r	CTGATGGATGAAGTTCGCAAAG
R3 (241)	f	CGTTCAGAAAGCCATGTTTGAACAAG
	r	CAATCCAACTGATCTTGGGTTAACTC
R4 (502)	f	CAATCGGTCTGCATGATCACTAC
	r	GAAACTGAGCGTAATCAGCATCAC
R5 (808)	f	CTTTGCGAACTTCATCCATCA
	r	CACTTTTGAATGAACACTATACTCG
R6 (1,482)	f	GAGTTACTAGAGTCGACTCTAATGTTG
	r	CTTCGCTAAGTAATCAAACATCTG
R7 (836)	f	CTTCTGGGATACACTTTTAGCTAGTG
	r	CAAACATCTGTGTCCCTATAGCATAAG
JL (156)	f	CGAAATATCTCTCTGTCTTGTTCC
	r	GCACCTTTTCTTTGGAATTGTCATTCC
JR (268)	f	CAACATTAGAGTCGACTCTAGTAACTA
	r	GAGTTACCAATCCGTGTCGATAATAAG
DPN1 (1,474)	f	CTCAGAAGTCCTGATACTGTCAC
	r	CAACACAATAGGCTAAGGTTATCCAC
N1 (678)	f	GATCAAGATAGTGCATCCAATGATGG
	r	GATATCCTTCAACCACTGGTGATTC
DP1 (637)	f	CACCAATAGTCTATTTTCATAAC
	r	GATGGATATAATGAATTCACATCATAGAC
DP2 (684)	f	GTCTATGATGTGAATTCATTATATCCATC
	r	GTCCGTATAATAGCAGTCATC
D5 (2,022)	f	CTCAGAAGTCCTGATACTGTCAC
	r	CTAGTTTGTCGTAACATGTTCAACG

DNA sequencing, BLAST searches and phylogenetic analysis

Sequencing was performed using fluorescent dideoxy terminators (Applied Biosystem) and analyzed with the ABI 373A sequencer. Sequence homology searches were performed by querying the EMBL and GenBank databases using the BLAST algorithm (Altschul et al. 1997). Translated nucleotide sequences were aligned using ClustalW (Thompson et al. 1994). Phylogenetic analysis and tree construction was performed using the maximum parsimony method employed by the Molecular Evolutionary Genetics Analysis (MEGA) Version 2.1 software package (http://www.megasoftware.net/) (Kumar et al. 2001).

Results

Nucleotide sequence analysis of a putative integrated linear plasmid in the mitochondrial genome of perennial ryegrass

An RFLP-based analysis of the mitochondrial genome of a CMS line of *L. perenne* (CMS-St) identified a novel 9.6 kb region, designated LpCMSi, located 89 bp downstream of the *atp9* gene which was absent in fertile lines and was therefore diagnostic of the CMS cytoplasm (Kiang et al. 1993; Kiang and Kavanagh 1996). The LpCMSi region and approximately 3.6 kb of flanking mtDNA is contained in three contiguous restriction fragments (from left-to-right, relative to *atp9*): 6.0 kb *Hin*dIII, 4.5 kb *Hin*dIII/*Bam*HI, 2.9 kb *Bam*HI/*Hin*dIII (Fig. 1), which were cloned in the vector pUC19 to give, respectively, pCMS60, pCMS45 and pCMS29 (Kiang and Kavanagh 1996). These were used in the present study as hybridization probes and as LpCMSi sequencing templates.

LpCMSi possesses an overall GC content of 36%, lying within the 30–39% range reported for extrachromosomal mitochondrial linear plasmids in plants, but significantly lower than that reported for the main mitochondrial genome (43–45%) (Handa et al. 2002; Clifton et al. 2004). The leftward junction (JL) between LpCMSi and mtDNA consists of an 'AT-rich' motif containing a partial 6-bp repeat (5′ ATAATATATA ATATAATA 3′), while the rightward junction, JR, comprises an inverted repeat (5′ ACTTCAATCGTAAGTTGATTCAAGT 3′) capable of forming a stem loop structure (Fig. 1). Translation of the LpCMSi nucleotide sequence using the universal genetic code identified 4 reading frames each with methionine (ATG) as the translational initiation codon (Fig. 1a). The two largest reading frames, *orf1* and *orf2* are located on opposite strands at either end of LpCMSi: *orf1*, located near JL has a coding capacity of 1,002 amino acids (molecular mass 111 kDa), while *orf2* located near JR encodes 1,285 amino acids (molecular mass 141 kDa). The smaller *orf3*, located immediately downstream of *orf1* on the same strand, encodes 364 amino acids (molecular mass 45 kDa). In addition, a chimeric reading frame, *orfC9*, encoding 167 amino acids (molecular mass 12 kDa) was identified upstream of *orf1* adjacent to JL. This reading frame appears to have arisen fortuitously as a result of the fusion of sequences derived from a partial duplication of the nearby *atp9* gene (located just outside JL), comprising the promoter region and first 6 codons of *atp9,* with adjacent LpCMSi sequences (Kiang and Kavanagh 1996).

The predicted translation products of *orfs* 1–3 were submitted as query sequences to the GenBank/EMBL/DDBJ databases with a view to identifying homologs. ORF1 was found to be significantly similar to members of the protein-primed family B DNA polymerases (Dpos) encoded by bacteriophages, eukaryotic viruses and linear plasmids (Jung et al. 1987) and is, hereafter, referred to as LpDpo. LpDpo showed the greatest similarity to predicted Dpos encoded by linear mitochondrial plasmids identified in plants, sharing for example 32% amino acid identity with the putative Dpo encoded by the *Z. mays* S1 plasmid (Paillard et al. 1985), 31% identity with ORF1 of a linear plasmid from *B. vulgaris* (Acc. No. Y10854; unpublished), 29% identity with ORF of the *B. napus* plasmid described by Handa et al. (2002), and 27% identity with ORF2 of the linear plasmid recently described from *D. carota* mitotype FG21 (Robison and Wolyn 2005). LpDpo shared lower levels of similarity with the putative Dpos encoded by fungal linear mitochondrial plasmids, for example, 20% identity with ORF1 of the *Claviceps purpurea* plasmid pCLK1 (Oeser and Tudzynski 1989).

Conserved motifs characteristic of the N-terminal 3′–5′ exonuclease (Exo) domain and the C-terminal polymerization (Pol) domain, conserved in all proof-reading Dpos were identified in LpDpo, and in highly homologous sequences amplified by PCR from FIp4, a fertile wild *L. perenne* accession isolated from Irish grasslands and FNz, a commercial *L. perenne* cultivar from New Zealand infected with the fungal endophyte *Epichloe typhina* x *Acremonium lollii* (Figs. 1b, 2). Moreover, the Exo domain of LpDpo, corresponding to amino acid residues 246–486, contains the highly conserved catalytic motifs *Exo*I, *Exo*II and *Exo*III (Bernad et al. 1989), while the Pol domain, located between amino acid residues 532–920, contains the highly conserved motifs *Pol*I, *Pol*IIa, *Pol*IIb and *Pol*III (Bernad et al. 1987) (Fig. 2).

A BLAST search of the predicted ORF2 translation product revealed extensive similarity with the single subunit RNA polymerases encoded by bacteriophages and mitochondrial plasmids. The level of sequence similarity was highest between ORF2 (referred to hereafter as LpRpo) and RNA polymerases (Rpos) encoded by plant mitochon-

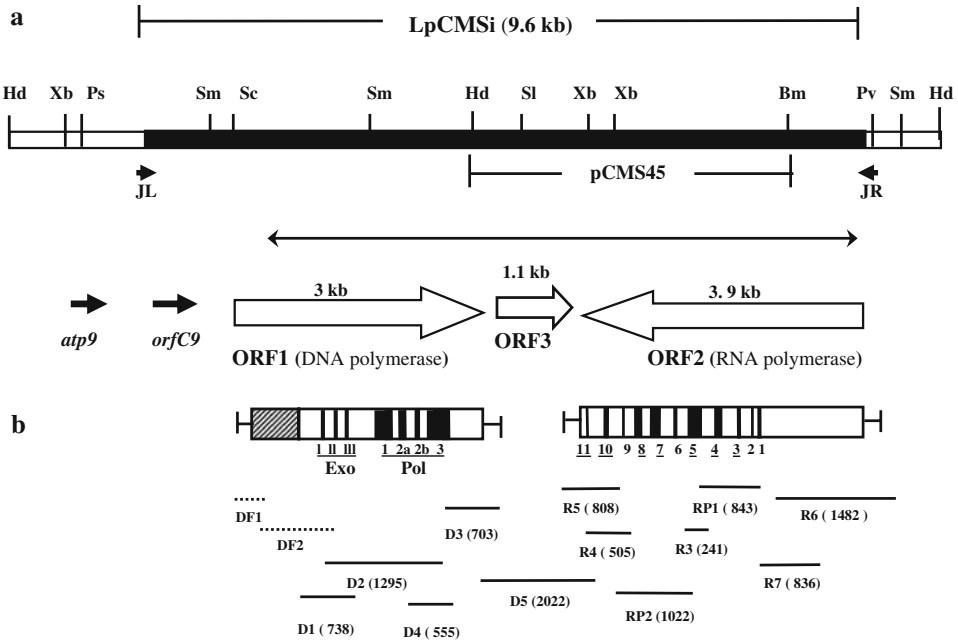

Fig. 1 Physical map of LpCMSi and the flanking mtDNA regions. **a** LpCMSi and associated flanking mtDNA are represented by solid and open rectangles, respectively. *Arrow heads* indicate the leftward (*JL*) and rightward junctions (*JR*) between LpCMSi and mtDNA. Restriction sites are as follows: *Bm Bam*HI, *Hd Hind*III, *Pv Pvu*II, *Ps Pst*I, *Sl Sal*I, *Sc Sac*I, *Xb Xba*I. The location of *atp9* and the chimeric gene *orfC9* is indicated by *solid arrows*. The LpCMSi reading frames are represented by *open horizontal arrows*. The region corresponding to pCMS45 was used as a diagnostic probe for LpRpo sequences in male sterile and restored lines of *L. perenne*. The long double-headed arrow indicates the total extent of LpCMSi-homologous sequences amplified in the fertile *L. perenne* lines FNz and FIp4. **b** Schematic representation of the domain structure of the Dpo and Rpo encoded by ORF1 and ORF2, respectively. The conserved domains are indicated by *solid boxes*. *Horizontal lines* denote the individual LpCMSi-homologous DNA sequences amplified by PCR from fertile lines (FNz and FIp4), with fragment names and sizes (bp) shown below the lines. The *dotted lines* indicate PCR fragments that were amplified from the CMS line but not from FNz or FIp4

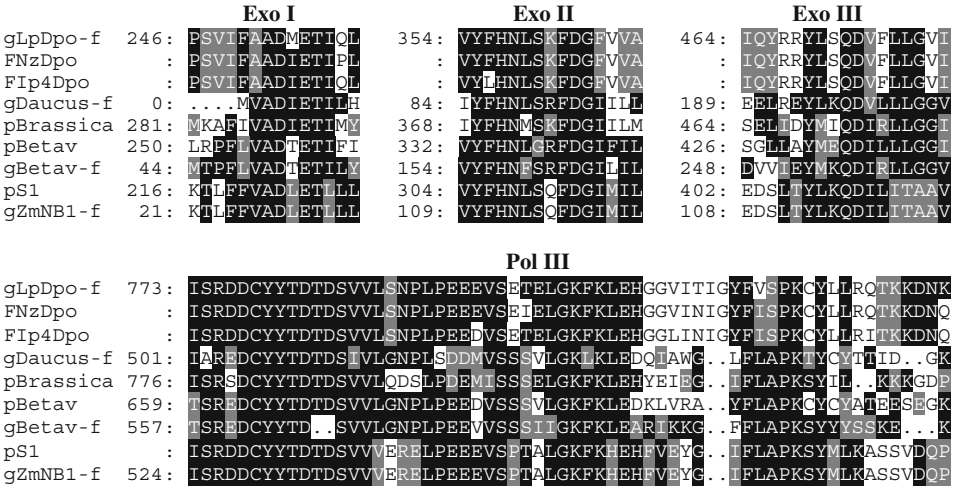

Fig. 2 Multiple alignment of conserved motifs shared by LpDpo and Dpo-like proteins of linear plasmid origin. Amino acid sequence motifs characteristic of the exonuclease (*Exo*) and polymerase (*Pol*) domains of family B-type DNA polymerases were identified in plant mitochondrial genomes and episomal linear plasmids. Some of the genomic (*g*) sequences consisted of reading frame fragments; most, however, were full-length (*-f*). The *L. perenne* sequences LpDpo (gLpDpo-f), FNzDpo and FIp4Dpo were aligned with homologous sequences in the mtDNA of *B. vulgaris* Z34298 [gBetav-f]; *D. carota* AY521591 [gDaucus-f]; *Z. mays* (NB) AY506529 [gZmNB1-f (*orf734*)]; and Dpo sequences encoded by linear plasmids (p) of *B. napus* AB073400 [pBrassica]; *Z. mays* (S) X02451 [pS1]; *B. vulgaris* Y10854 [pBetav]. Amino acid positions are indicated for Dpo reading frames that contained each of the seven conserved domains. However, only the *Exo*I–III domains and the *Pol*III domain are shown in the alignment

drial linear plasmids, ranging from 24.8% identity with ORF1 of the *Z. mays* S2 plasmid (Levings and Sederoff 1983) to 22.0% identity with ORF6 of the *B. napus* 11.6 kb linear plasmid (Handa et al. 2002). The amino acid sequence of LpRpo and highly homologous sequences from the fertile *L. perenne* lines FIp4 and FNz (see below) were compared with Rpo sequences from linear plasmid-like elements of fungi and other plant species (Fig. 3). The single subunit Rpo family is characterized by eleven conserved motifs (Fig. 1b). Motifs I and II are poorly conserved between different subfamily groupings. However, motifs III–XI are highly conserved between diverse species (McAllistair and Raskin 1993) and of these, seven are easily recognisable in LpRpo (Fig. 3).

The predicted 40 kDa polypeptide encoded by *orf3* was novel, showing no significant similarity to any proteins in the public sequence databases. Such proteins are commonly encoded by linear plasmids and may play a role in their maintenance and/or transmission, as suggested by Handa et al. (2002).

Phylogenetic relationships of Dpo- and Rpo-like sequences in plants and fungi

The maximum parsimony (MP) method described by Kumar et al. (2001) was used to determine phylogenetic relationships among Dpo-like (Fig. 4a) and Rpo-like (Fig. 4b) polypeptides encoded by extrachromosomal and integrated linear plasmid-like sequences of plants and fungi. The Dpo encoded by the *Bacillus* phage *phi29* (Salas 1991) and the nucleus-encoded mitochondrial Rpo from *A. thaliana* (Hedtke et al. 1997) (At-nRpo) were used as outgroups. The resulting phylogenetic trees show that the three *L. perenne* sequences form a discrete monophyletic subgroup within the plant mitochondrial Dpo-like and Rpo-like sequences and further indicate that in species which contain both integrated and extrachromosomal plasmid-like sequences (e.g., *Z. mays* and *B. vulgaris*), the former are likely to have been derived from the latter. Significantly, for each tree, the plant sequences occupy a discrete clade well-separated from similar sequences encoded by linear mitochondrial plasmids of ascomycete (pClK1 and pBgh) and basidiomycete (pEM and pLMP1) fungi. This clear separation of fungal and plant sequences was confirmed by a second more extensive phylogenetic analysis that included 12 fungal linear plasmid sequences (data not shown).

Transcripts of LpDpo and LpRpo are detectable in *L. perenne* lines containing the CMS cytoplasm

Unlike *atp9* and *orfC9*, transcripts of which were readily detectable by northern blot analysis of total RNA of the *L. perenne* CMS line (data not shown; Kiang and Kavanagh 1996), transcripts of LpDpo and LpRpo were detected only when RT-PCR was employed (Fig. 5; Table 1). Thus, PCR products of 637 bp (DP1) and 684 bp (DP2) representing 44% of the LpDpo reading frame, and products of 1,022 bp (RP2) and 843 bp (RP1) representing 48% of the LpRpo reading frame, were amplified from first-strand cDNA prepared from total RNA of the CMS line (St) and an *L. perenne* line restored to fertility (Rst), but which also contained the CMS cytoplasm. A 678 bp product (N1) representing 49% of orf3 was also amplified by RT-PCR.

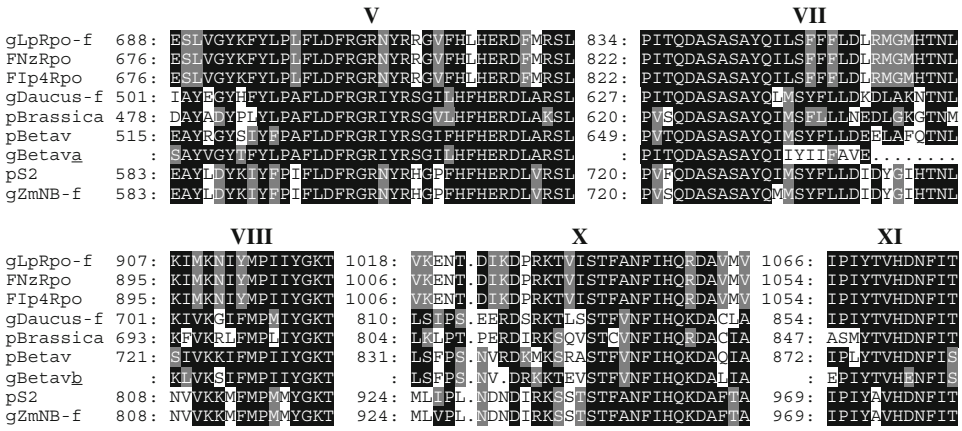

Fig. 3 Multiple alignment of conserved motifs shared by LpRpo and Rpo-like proteins of linear plasmid origin. Conserved amino acid sequence motifs characteristic of single-subunit Rpo domains were identified in plant mitochondrial genomes and episomal linear plasmids. The *L. perenne* sequences LpRpo (gLpRpo-f), FNzRpo and FIp4Rpo were aligned with homologous sequences identified in the mtDNA of *Daucus carota* AY521591 [gDaucus-f]; *B. vulgaris* Z34298 [gBetava (463–2,152), gBetavb (2,139–3,235)]; *Z. mays* (NB) AY506529 [gZmNB-f (*orf1159*)]; and Rpo sequences encoded by linear plasmids (p) of *B. napus* AB073400 [pBrassica]; *B. vulgaris* Y10854 [pBetav]; *Z. mays* (S) J01426 [pS2]. Amino acid positions are indicated for Rpo reading frames that contained each of the seven conserved Pol domains. However, only the Pol domains V, VII, VIII, X and XI are shown in the alignment

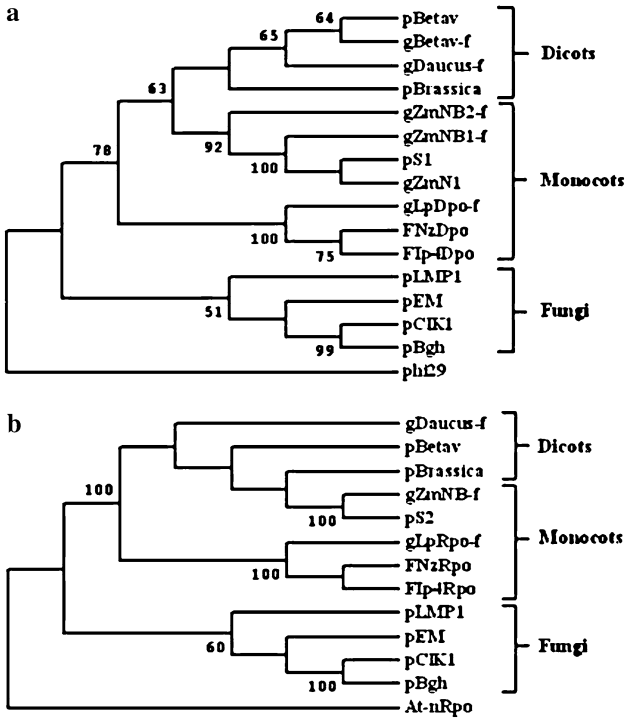

Fig. 4 Phylogenetic relationships of Dpo and Rpo proteins of plants and fungi. Full-length Dpo (**a**) and Rpo (**b**) reading frames encoded by linear plasmids (*p*) and mitochondrial genomes (*g*) (see legends to Figs. 2 and 3) were aligned using ClustalW. Linear plasmid sequences from two fungal species: *C. purpurea* P22373 [pCIK1] and *Blumeria graminis* AY189817 [pBgh] were included in the alignments. The protein parsimony algorithm in Mega-Align v 2.1 (http://www.megasoftware.net/) was then used to construct an unrooted tree taking into account both similar and identical residues. The Dpo of phage *phi29* P03680 [phi29] and the nucleus-encoded single subunit mitochondrial Rpo of *Arabidopsis* AAG51592 [At-nRpo] were used as outgroups. Numbers at the nodes are the bootstrap values that indicate the percentage of trees, in 1,000 replicates, that support a given branch. Only bootstrap values significantly larger than 50% are shown

Moreover, a 1.5 kb product (DPN1) was amplified using a forward primer hybridizing near the 3′ end of LpDpo and a reverse primer hybridizing near the 3′ end of *orf3*, indicating that both reading frames are co-transcribed as part of the same mRNA transcript (Fig. 5). Finally, all of the RT-PCR products were sequenced to investigate editing of LpCMSi transcripts. However, when the nucleotide sequence of RT-PCR products was compared with the corresponding genomic sequence no C to T transitions were identified, suggesting that LpCMSi transcripts are not edited (data not shown).

Evidence for extrachromosomal LpCMSi-like sequences in fertile lines of *L. perenne*

The fertile lines FNz and FIp4, two additional Irish grassland accessions (FIp2 and FIp5), and the CMS maintainer line LPSB21 (Ft) (Kiang and Kavanagh, 1993) were screened by Southern hybridization for the presence of LpCMSi-like sequences. Whereas a single band of the expected size (7.4 kb) was detected in *HindIII*-digested DNA samples from the St and Rst lines (both of which contain the CMS cytoplasm), following hybridization with the LpRpo-specific probe pCMS45, no hybridization signal was observed in DNA samples from fertile lines (Fig. 6). In addition, a Southern blot analysis of proteinase K-treated, intact (i.e., uncleaved) total DNA samples from these lines using the LpDpo and LpRpo probes failed to reveal any hybridizing bands (data not shown). This suggests that neither integrated nor extrachromosomal copies of LpCMSi-like sequences are present in the mitochondria of fertile lines, at least not at levels quantitatively similar to the copy number of LpCMSi in the St and Rst lines.

To facilitate the detection of potentially very low-copy number LpCMSi-like sequences, the same fertile *L. perenne* lines were screened by PCR using primer pairs designed to anneal with the most conserved sequences in LpDpo and LpRpo (Fig. 6). A set of overlapping fragments corresponding to an entire Rpo orf, and all but the N-terminal 134 codons (nucleotides 403–3,005) of a Dpo orf, was amplified from total DNA of FIp4 and FNz, and their nucleotide sequences determined by direct sequencing of PCR products. When compared with LpDpo, the nucleotide sequence of FNzDpo and FIp4Dpo differed at 30 and 70 positions, respectively, resulting in 20 (2.3%) and 43 (5%) amino acid substitutions, respectively, in the corresponding orfs (Supplementary Table 1). Thus, FNzDpo is more closely related to LpDpo than is FIp4Dpo. This relationship was also evident in sequence comparisons between the Rpo orfs: compared with LpRpo, FNzRpo and FIp4Rpo contained 70 and 144 nucleotide substitutions, respectively, resulting in 50 (4%) and 105 (8%) amino acid substitutions (Supplementary Table 2). Surprisingly, alignment of the Rpo nucleotide sequences revealed a 36 bp deletion in FNzRpo and FIp4Rpo (relative to LpRpo) in the region located between the RI and RII domains (Fig. 7). In LpRpo, this region contains 8 copies of an 18 bp repeated sequence motif 5′-(G/T)(A/C)TA(C/T)TC (T/A)AAA(A/T)CAAGA (A/T)-3′, while in FNzRpo and FIp4Rpo the two central repeats (4 and 5) are missing. Since 3 copies of the repeat flank each end of the deleted region, the deletion/insertion event is likely to have been mediated by homologous recombination. Remarkably, none of the nucleotide substitutions identified in the FNz and FIp4 Dpo-like and Rpo-like sequences, nor the 36 bp deletion in FNzRpo and FIp4Rpo, disrupt the corresponding reading frames by creating either premature stop codons or frameshifts.

Finally, in order to assess whether the Rpo-like and Dpo-like sequences identified in the FNz and FIp4 lines might, like LpCMSi, be integrated at the same location in the *L. perenne* mitochondrial genome, PCR amplifications

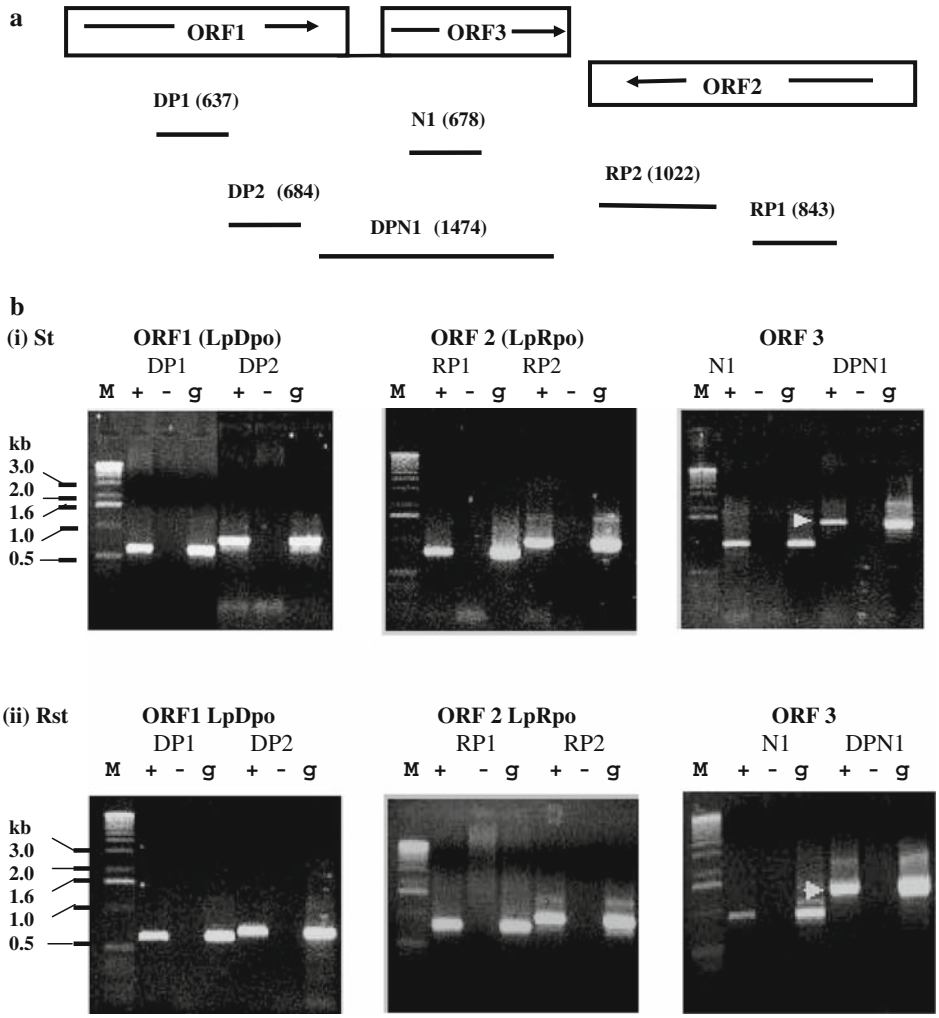

Fig. 5 Detection of LpCMSi transcripts by RT-PCR. **a** Location and size (bp) of the RT-PCR products amplified from ORFs 1, 2 and 3 in LpCMSi. **b** RT-PCR products were amplified from total RNA of (1) the CMS line (St) and (2) a restored line (*Rst*). DNAase I-treated total RNA samples were either used directly as the template in PCR reactions (negative control (−)), or were reverse transcribed (+) prior to amplification by PCR. Amplification from total genomic DNA (g) was used as a positive control. The white arrow head indicates the 1.5 kb DPN1 co-transcript that spans ORF1-ORF3. *M* molecular weight size marker; sizes shown in kb

were performed using the primer pairs JR and JL which span the right and left junctions, respectively, between LpCMSi and mtDNA (Table 1; Fig 1a). PCR products of the expected sizes (156 and 268 bp, respectively) were amplified only in DNA samples from the CMS-St line (Fig. 6c) suggesting that the Rpo-like and Dpo-like sequences identified in the FNz and FIp4 lines are not integrated at the CMSi locus, and therefore, might be of extrachromosomal origin.

Discussion

LpCMSi represents the first documented case in a monocot species of an integrated linear plasmid-like element containing both an intact *Dpo* and an intact *Rpo* orf. Remarkably, despite extensive sequence divergence, the arrangement of the orfs in LpCMSi is similar to that in the *B. napus* 11.6 kb extrachromosomal linear plasmid in which the polymerase orfs are located on opposite strands and transcribed towards each other. This inverted tail-to-tail organization of the polymerase orfs is also found in most linear mitochondrial plasmids of filamentous fungi, for example, the *CLK1* plasmid of *Claviceps purpurea* (Oeser and Tudzynski 1989) and *Maranhar* and *Kalilo* plasmids of *Neurospora* spp. Court and Bertrand (1992) ; Chan et al. 1991). In contrast, the *Rpo* and *Dpo* orfs of the extrachromosomal linear plasmids of *B. vulgaris* subsp. *maritima* (Saumitou-Laprade et al. 1989) and *D. carota* (Robison and Wolyn 2005) are located on the same DNA strand in a head-to-tail arrangement.

With the exception of *L. perenne* (this report), *D. carota* (Robison and Wolyn 2005), and certain *Z. mays* accessions (Garcia et al. 1988), the majority of Dpo- and Rpo-like sequences found in plant mtDNA consist of orf fragments, the products of putative deletion or insertion events that occurred following plasmid integration. The completely sequenced mitochondrial genomes of *B. vulgaris* (Kubo et al. 2000; Satoh et al. 2004), *O. sativa* (Notsu et al. 2002) and *Z. mays* (Clifton et al. 2004) contain both Dpo and Rpo orf fragments. However, the mitochondrial genome of *T. aestivum* (Ogihara et al. 2005) contains only Rpo-like

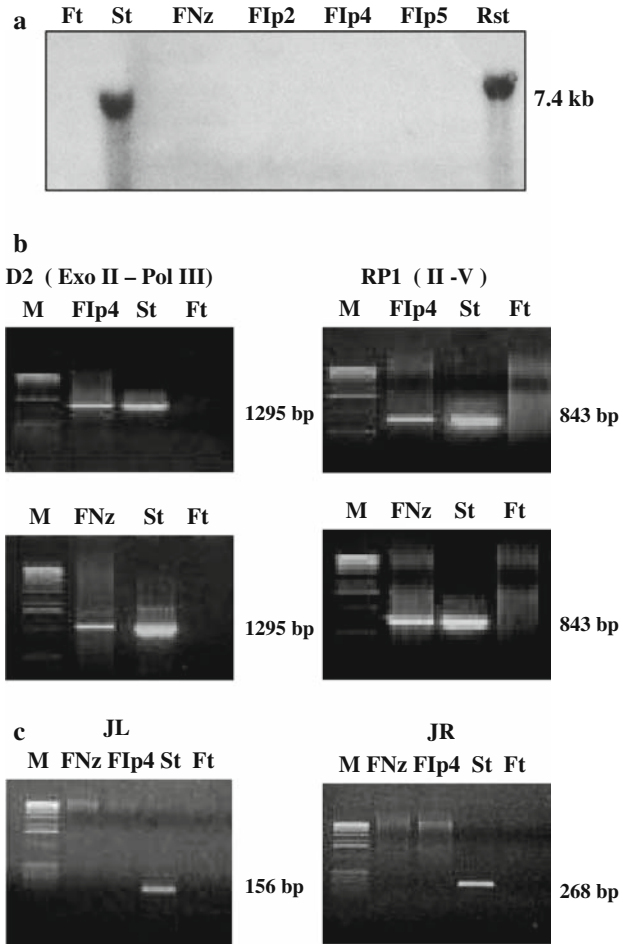

Fig. 6 Detection of CMSi-like sequences in fertile lines of *L. perenne*. **a** Southern blot analysis of *Hin*dIII-digested total DNA from the CMS line (*St*), the fertile lines Ft, FNz, FIp2, FIp4, FIp5, and a restored line (*Rst*). The hybridization probe was ^{32}P-labeled pCMS45. **b** The PCR products D2 and RP1 which are derived from LpCMSi ORF1 and ORF2, respectively, were amplified from St and from the fertile lines FNz and FIp4. **c** The PCR products JL and JR, which span the junctions between CMSi and mtDNA, were amplified from St but not from FNz or FIp4. *M* molecular weight size marker (1 kb ladder). The sizes of the major PCR fragments are indicated in bp

sequences while that of *A. thaliana* (Unseld et al. 1997), *N. tabacum* (Sugiyama et al. 2005) and *Marchantia* (Oda et al. 1992) contain only Dpo-like sequences. Although not yet fully sequenced, *S. cereale* mtDNA appears to lack Rpo-like sequences but contains a truncated Dpo orf encompassing the *Pol*IIa to *Pol*III domains (Dohmen and Tudzynski 1994).

Surprisingly, neither the presence nor the absence of plasmid-like Dpo and/or Rpo sequences in plant mtDNA correlates with the presence/absence of the corresponding extrachromosomal plasmids. Linear mitochondrial plasmids have not, for example, been reported in *L. perenne*, *A. thaliana*, *T. aestivum*, *O. sativa*, *S. cereale*, *N. tabacum*, or *Marchantia,* and remarkably, despite the common occurrence of a high copy number linear plasmid in *Brassica* mitochondria, the sequenced *B. napus* mitochondrial genome contains neither Rpo- nor Dpo-like orf fragments (Handa 2003). Similarly, the *D. carota* mitotype FG21 contains extrachromosomal but not integrated copies of a 9.2 kb linear plasmid, while the SW3 mitotype contains integrated but apparently not extrachromosomal copies of the same plasmid (Robison and Wolyn 2005). In sugar beet (*B. vulgaris*), the completely sequenced mitochondrial genome of a fertile line (369 kb) (Kubo et al. 2000) and a CMS line (501 kb) (Satoh et al. 2004) contained an identical 6,253 bp linear plasmid-like region encoding disrupted Dpo and Rpo orfs. However, the nucleotide sequence of this region was only 43% homologous to that of the 10.4 kb linear plasmid identified in *B. vulgaris* subsp. *maritima* (Acc No.Y10854), from which it was most likely derived. Similarly, a search of the completely sequenced *Z. mays* (NB) genome revealed the presence of two incomplete Dpo orfs, each apparently due to independent S1-like plasmid integration events well-separated over evolutionary time: one, *orf734*, lacks the first 200 codons but shares 99.5% amino acid identity with the extant S1 plasmid Dpo; the other, *orf911*, is in contrast almost full-length but shares only 53% amino acid identity with S1 Dpo (McDermott and Kavanagh, unpublished). These observations suggest that in contrast to the mitochondria of filamentous fungi, plant mitochondria possess efficient mechanisms that either inhibit the integration of linear plasmids or disable or eliminate occasionally integrated copies by mutation. The deletion/insertion of 18 bp repeat units in the *L. perenne* Rpo sequences may be evidence of just such a process.

Neither Southern hybridization with the diagnostic probe CMS45, nor PCR using primers that amplify the CMSi-mtDNA junction fragments (JL and JR), detected LpCMSi-like sequences in fertile *L. perenne* lines. However, PCR based on primers annealing within the LpRpo and LpDpo reading frames revealed the presence of LpCMSi-like sequences in FNz and FIp4. Moreover, amplification with primers annealing to the 3′ ends of the Rpo and Dpo orfs yielded a 2.2 kb product containing the *orf3* region, indicating that the Rpo and Dpo reading frames in FNz and FIp4 are contained on the same replicon (data not shown). Taken together, these data suggest that FNz and FIp4 mitochondria contain LpCMSi-like elements that are maintained either as very low-copy number extrachromosomal replicons, as has been reported for certain linear mitochondrial plasmids in fungi (Robison and Horgen 1998), or as integrated elements on low-copy number (i.e., substoichiometric) subgenomic mtDNAs (Small et al. 1987). It may be significant that the CMS-St line was generated as part of a hybrid ryegrass breeding programme from an interspecies cross between *Festuca pratensis* (as the female parent) and *L. perenne* (Connolly and Wright-Turner 1984), that unexpectedly resulted in inheritance of the paternal mitochon-

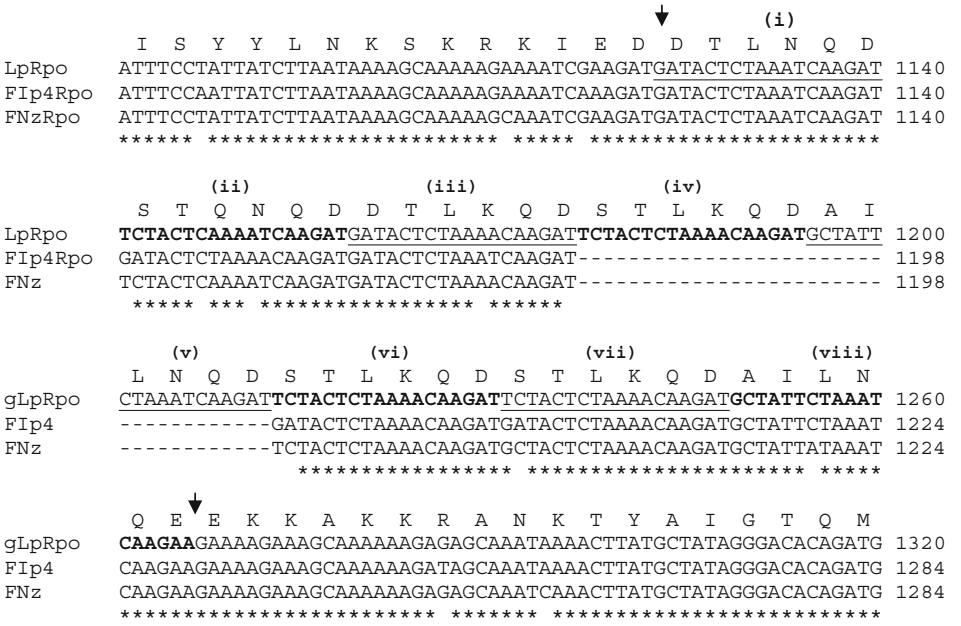

Fig. 7 Multiple alignment of the 18 bp repeat region in LpRpo, FIp4Rpo and FNzRpo The beginning and end of the repeat region is indicated by *arrows*. Alternating repeat units (i)–(viii) are highlighted either by *underlining or in bold*. The predicted amino acid sequence of LpRpo is shown using the *single letter* format above the nucleotide sequence. The *dashed line* indicates nucleotides missing in the FIp4Rpo and FNzRpo sequences. Identical nucleotides in all three sequences are indicated by *asterisks*

drial genome in the resulting hybrid (Kiang et al. 1994). Wide crosses are known to promote illegitimate and asymmetric recombination events that can lead to rearrangements in mtDNA and dramatic changes in the stoichiometry of subgenomic fragments, so-called substoichiometric shifting (SSS) (Abdelnoor et al. 2003). Thus, in the case of LpCMSi, the *Festuca* x *Lolium* cross may have promoted either the original linear plasmid integration event or, alternatively through SSS, dramatically increased the copy number of a subgenomic mtDNA replicon in which LpCMSi already resided.

Although phylogenetic analyses do not appear to support a fungal origin for the LpCMSi sequences, horizontal gene transfer between *L. perenne* (or *F. pratensis*) and a fungus remains a possibility. Fungi of the genera *Epichloe*, *Neotyphodium* and *Acremonium* are common plasmid-containing mutualistic symbionts of grasses (especially of *Festuca* and *Lolium* spp), which during the asexual phase of their lifecycle grow in close association with developing floral organs including ovules, a strategy that ensures their transmission via seeds (Schardl et al. 2004). It is not inconceivable that such a close association between fungal mycelia and plant reproductive organs might also facilitate rare horizontal transmission of genetic information between the symbiont and its host. Horizontal transfer between fungi and higher plants is not without precedent. The best documented case involves a mitochondrial homing group I intron of fungal origin which invaded the *coxI* gene of flowering plant mitochondria on more than 1,000 occasions during angiosperm evolution (Vaughn et al. 1995; Cho et al. 1998). The LpCMSi orfs, like those of other linear plasmids, exhibit a significantly lower % GC value, closer to that of fungal mtDNA than the average for plant mtDNA. Nevertheless, LpCMSi sequences grouped with plant and not fungal plasmids. This discrete grouping might, however, simply reflect differences in mutation rates between plant (Cho et al. 2004; Mower et al. 2007) and fungal mitochondria following an evolutionarily distant horizontal transfer event, and does not rule out the possibility that fungal plasmids may be capable of replicating in plant mitochondria or that horizontal transfer might have played a role in the acquisition of LpCMSi by *L. perenne*. An extensive PCR-based survey of *Lolium* and *Festuca* species and their fungal symbionts might shed more light on the origins of LpCMSi.

Note added in proof Nucleotide sequences for the *L. perenne* lines CMS-St, FIp4 and FNz can be retrieved from Genbank under the following accession numbers: AM998372, AM998373 and AM998374, respectively.

References

Abdelnoor RV, Yule R, Elo A, Christensen AC, Meyer-Gauen G, Mackenzie SA (2003) Substoichiometric shifting in the plant mitochondrial genome is influenced by a gene homologous to *MutS*. Proc Natl Acad Sci USA 100:5968–5973

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Bernad A, Blanco L, Lazaro JM, Martin G, Salas M (1989) A conserved 3′–5′ exonuclease active site in prokaryotic and eukaryotic DNA polymerases. Cell 59:219–228

Bernad A, Zabollos A, Salas M, Blanco L (1987) Structural and functional relationships between prokaryotic and eukaryotic DNA polymerases. EMBO J 6:4219–4225

Brown GG, Zhang M (1995) Mitochondrial plasmids: DNA and RNA. In: Levings CS III, Vasil IK (eds) The molecular biology of plant mitochondria. Kluwer, Dordrecht, pp 61–91

Cahan PC, Kennell JC (2005) Identification and distribution of sequences having similarity to mitochondrial plasmids in mitochondrial genomes of filamentous fungi. Mol Gen Genomics 273:462–473

Chan BS, Court DA, Vierula PJ, Bertrand H (1991) The *kalilo* linear senescence-inducing plasmid of *Neurospora* is an invertron and encodes DNA and RNA polymerases. Curr Genet 20:225–237

Cho Y, Mower JP, Qiu YL, Palmer JD (2004) Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. Proc Natl Acad Sci USA 101:17741–17746

Cho Y, Qiu YL, Kuhlman P, Palmer JD (1998) Explosive invasion of plant mitochondria by a group I intron. Proc Natl Acad Sci USA 95:14003–14005

Clifton SW, Minx P, Fauron CM, Gibson M, Allen JO, Sun H, Thompson M, Barbazuk WB, Kanuganti S, Tayloe C, Meyer L, Wilson RK, Newton KJ (2004) Sequence and comparative analysis of the maize NB mitochondrial genome. Plant Physiol 136:3486–3503

Collett MA, Bradshaw R, Scott DB (1995) A mutalistic fungal symbiont of perennial ryegrass contains two different *pyr4* genes, both expressing orotidine-5′-monophosphate decarboxylase. Gene 158:31–39

Connolly V, Wright-Turner R (1984) Induction of cytoplasmic male sterility into ryegrass (*Lolium perenne* L.). Theor Appl Genet 68:449–453

Court DA, Bertrand H (1992) Genetic organization and structural features of *maranhar*, a senescence-inducing linear mitochondrial plasmid of *Neurospora crassa*. Curr Genet 22:385–397

Dohmen G, Tudzynski PA (1994) DNA-polymerase-related reading frame (pol-r) in the mtDNA of *Secale cereale*. Curr Genet 25:59–65

Escote LJ, Gabay-Laughnan SJ, Laughnan JR (1985) Cytoplasmic reversion to fertility in *cms*-S need not involve loss of linear mitochondrial plasmids. Plasmid 14:264–267

Garcia JA, Kolacz K, Studnicka GM, Gilmore-Herbert M (1988) Sequence of integrated S-1 homologous DNA in the normal maize mitochondrial genome. Nucleic Acids Res 16:4169–4170

Griffiths AJF (1995) Natural plasmids in filamentous fungi. Microbiol Rev 59:673–685

Handa H, Itani K, Sato H (2002) Structural features and expression analysis of a linear mitochondrial plasmid in rapeseed (*Brassica napus* L.). Mol Genet Genomics 267:797–805

Handa H (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. Nucl Acids Res 31:5907–5916

Hedtke B, Boerner T, Weihe A (1997) Mitochondrial and chloroplast phage type RNA polymerases in *Arabidopsis*. Science 277:809–811

Jung G, Leavitt MC, Hsieh J-C, Ito J (1987) Bacteriophage PRD1 DNA polymerase: evolution of DNA polymerases. Proc Natl Acad Sci USA 84:8287–8291

Kemble RJ, Carlson JE, Erickson LR, Serynk JL, Thompson DJ (1986) The *Brassica* mitochondrial DNA plasmid and large RNAs are not exclusively associated with cytoplasmic male sterility. Mol Gen Genet 205:183–185

Kiang AS, Connolly V, McConnell DJ, Kavanagh TA (1993) Cytoplasmic male sterility (CMS) in *Lolium perenne* L. 1. Development of a diagnostic probe for the male-sterile cytoplasm. Theor Appl Genet 86:781–787

Kiang AS, Connolly V, McConnell DJ, Kavanagh TA (1994) Paternal inheritance of mitochondria and chloroplasts in *Festuca pratensis–Lolium perenne* intergeneric hybrids. Theor Appl Genet 87:681–688

Kiang AS, Kavanagh TA (1996) Cytoplasmic male sterility (CMS) in *Lolium perenne* L.2. The mitochondrial genome of a CMS line is rearranged and contains a chimeric *atp9* gene. Theor Appl Genet 92:308–315

Kubo T, Nishizawa S, Sugawara A, Itchoda N, Estiati A, Mikami T (2000) The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNA(Cys)(GCA). Nucleic Acids Res 28:2571–2576

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17:1244–1245

Levings CS III, Sederoff RR (1983) Nucleotide sequence of the S-2 mitochondrial DNA from the S cytoplasm of maize. Proc Natl Acad Sci USA 80:4055–4059

Meinhardt F, Kempken F, Kamper J, Esser K (1990) Linear plasmids among eukaryotes: fundamentals and application. Curr Genet 17:89–95

McAllistair WT, Raskin CA (1993) The phage RNA polymerases are related to DNA polymerases and reverse transcriptases. Mol Microbiol 10:1–6

Mower JP, Touzet P, Gummow JS, Delph LF, Palmer JD (2007) Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. BMC Evol Biol 7:135–149

Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, Hirai A, Kadowaki K (2002) The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. Mol Genet Genomics 268:434–445

Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T, Ohyama K (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. J Mol Biol 223:1–7

Oeser B, Tudzynski P (1989) The linear mitochondrial plasmid pClK1 of the phytopathogenic fungus *Claviceps purpurea* may code for a DNA polymerase and an RNA polymerase. Mol Gen Genet 217:132–140

Ogihara Y, Yamazaki Y, Murai K, Kanno A, Terachi T, Shiina T, Miyashita N, Nasuda S, Nakamura C, Mori N, Takumi S, Murata M, Futo S, Tsunewaki K (2005) Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. Nucl Acids Res 33:6235–6250

Paillard M, Sederoff RR, Levings CS III (1985) Nucleotide sequence of the S-1 mitochondrial DNA from the cytoplasm of maize. EMBO J 4:1125–1128

Palmer JD, Shields CR, Cohen DB, Orton TJ (1983) An unusual mitochondrial DNA plasmid in the genus *Brassica*. Nature 301:725–728

Robison MM, Horgen PA (1998) Widespread distribution of low copy number variants of mitochondrial plasmid plasmid pEM in the genus *Agaricus*. Fungal Genetics Biol 26:62–70

Robison MM, Wolyn DJ (2005) A mitochondrial plasmid and plasmid-like RNA and DNA polymerases encoded within the mitochondrial genome of carrot (*Daucus carota* L). Curr Genet 47:57–66

Salas M (1991) Protein priming of DNA replication. Annu Rev Biochem 60:39–71

Satoh M, Kubo T, Nishizawa S, Estiati A, Itchoda N, Mikami T (2004) The cytoplasmic male-sterile type and normal type mitochondrial genomes of sugar beet share the same complement of genes of known function but differ in the content of expressed ORFs. Mol Genet Genomics 272:247–256

Saumitou-Laprade P, Pannenbecker G, Maggouta F, Jean R, Michaelis G (1989) A linear 10.4 kb plasmid in the mitochondria of *Beta maritima*. Curr Genet 16:181–186

Schardl CL, Leuchtmann A, Tsai H, Collett MA, Watt DM, Scott DB (1994) Origin of a fungal symbiont of perennial ryegrass by inter-

specific hybridisation of a mutualist with the ryegrass choke pathogen *Epichloe typhina*. Genetics 136:1307–1317

Schardl CL, Leuchtmann A, Spiering MJ (2004) Symbioses of grasses with seedborne fungal endophytes. Ann Rev Plant Biol 55:315–340

Schardl CL, Pring DR, Lonsdale DM (1985) Mitochondrial DNA rearrangements associated with fertile revertants of S-type male-sterile maize. Cell 43:361–368

Small I, Isaac P, Leaver C (1987) Stoichiometric differences in DNA molecules containing the *atpA* gene suggest mechanisms for the generation of mitochondrial genome diversity in maize. EMBO J 6:865–869

Sugiyama Y, Watase Y, Nagase M, Makita N, Yagura S, Hirai A, Sugiura M (2005) The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. Mol Genet Genomics 272:603–615

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighing, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Turpen T, Garger SJ, Marks MD, Grill LK (1987) Molecular cloning and physical characterization of a *Brassica* linear mitochondrial plasmid. Mol Gen Genet 209:227–233

Unseld M, Marienfeld JR, Brandt P, Brennicke A (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366, 924 nucleotides. Nat Genet 15:57–61

Vaughn JC, Mason MT, Sper-Whitis GL, Kuhlman P, Palmer JD (1995) Fungal origin by horizontal transfer of a plant mitochondrial group I intron in the chimeric *coxI* gene of *Peperomia*. J Mol Evol 41:563–572

Weber B, Borner T, Weihe A (1995) Remnants of a DNA polymerase gene in the mitochondrial DNA of *Marchantia polymorpha*. Curr Genet 27:488–490