Theor Appl Genet (1996) 92:308–315

A-S. Kiang · T. A. Kavanagh

Cytoplasmic male sterility (CMS) in *Lolium perenne* L. 2. The mitochondrial genome of a CMS line is rearranged and contains a chimaeric *atp* 9 gene

Received: 5 May 1995 / Accepted: 22 September 1995

Abstract The most striking difference between the mtDNAs of the fertile *L. perenne* line LPSB21 and the male-sterile line CMS9B290, is the presence in the former and the absence in the latter of a 5.6-kb *Hin*dIII fragment. This difference between fertile and sterile lines was the starting point for a detailed molecular analysis of the mitochondrial genome in the region spanning the 5.6-kb *Hin*dIII fragment in fertile *L. perenne* and the corresponding region in CMS9B290. Restriction mapping and Southern-blot analyses indicated that rearrangement of the mitochondrial genome consistent with a deletion/insertion event had occurred in the sterile line. Nucleotide-sequence analysis of the rearranged region in CMS9B290 revealed the presence of (1) a novel chimaeric gene, *orf-C9*, comprising the first six codons of *atp9* fused to a further 118 codons of an unknown sequence and (2) a truncated version of an open reading frame, *orf-L*, originally identified in LPSB21 mtDNA. Northern-blot analysis confirmed the absence of *orf-L* transcripts and the presence of *orf-C9* transcripts in the mtRNA of CMS9B290.

Key words Perennial ryegrass (*Lolium perenne*) · Cytoplasmic male sterility · Mitochondrial DNA · *atp9* · Chimaeric open reading frame

Introduction

Cytoplasmic male sterility (CMS) is a maternally inherited trait in which plants fail to produce functional pollen. It thus provides one of the most efficient strategies for the commercial production of F_1 hybrid seed since self-fertilization of the seed parent is prevented. Male fertility can typically be restored in the progeny of CMS plants by the introduction of nuclear restorer genes (Hanson 1991). CMS can arise spontaneously in natural plant populations but in a number of species, for example ryegrass (*Lolium perenne*), the CMS trait can be induced by making interspecific or even intergeneric crosses. In these cases, CMS is thought to be a consequence of an incompatibility between the novel nucleus-cytoplasm combination.

Two well characterized sources of CMS *L. perenne* have been described: the first was derived from an interspecific cross between *L. perenne* and *L. multiflorum* (Wit 1974) and the second from an intergeneric cross between *Festuca pratensis* as the female parent and *L. perenne* as the pollen parent (Connolly and Wright-Turner 1984). In the latter case, the *L. perenne* nuclear background was subsequently re-introduced into the male-sterile hybrid by repeated backcrossing to the pollen parent (*L. perenne*) over many generations. Nuclear genes conferring restoration of male fertility have also been identified in *L. perenne* (Kiang et al. 1993). However, since ryegrass is a forage crop, there has been little emphasis on the production of seed from F_1 hybrids obtained using the CMS lines. Consequently restorer genes have not been studied in detail.

Molecular analyses of the CMS trait in a range of species (reviewed in Hanson 1991) have tended to focus in the first instance on the identification of restriction-fragment differences between the mitochondrial genomes of fertile and sterile lines that correlate with the male-sterile phenotype. Many such studies have identified extensive rearrangement of the mitochondrial genome in CMS lines (Hanson 1991). More detailed investigations of CMS-associated mtDNA rearrangements indicate that they frequently result in the fortuitous creation of chimaeric transcription units which may also contain novel chimaeric open reading frames (*orfs*).

The translation products of chimaeric orfs have been identified definitively in CMS lines of maize (TURF-13; Dujardin et al. 1983; Levings 1990), *Petunia* (S-PCF protein; Nivison and Hanson 1989), sunflower (ORF522 protein; Kohler et al. 1991; Laver et al. 1991), wheat

Communicated by R. Hagemann

A-S. Kiang · T. A. Kavanagh (✉)
Department of Genetics, Trinity College Dublin, Dublin 2, Ireland

(ORF256 protein; Song and Hedgcoth 1994) and *Brassica* cybrids carrying the Ogura CMS cytoplasm (ORF138 protein; Grelon et al. 1994). By contrast, in rice the CMS phenotype does not appear to involve chimaeric *orfs* but instead correlates with the presence of chimaeric transcripts which potentially encode an additional normal ATP6 protein but which show altered post-transcriptional processing and editing patterns (Iwabuchi et al. 1993).

The most convincing evidence that chimaeric transcription units or the translation products of chimaeric orfs play a causative role in CMS derives from studies of CMS lines showing reversion to fertility and from investigations of the effect of nuclear fertility restorer genes. In maize, fertile revertant plants obtained by in vitro regeneration of a CMS-T cell line were shown to have accumulated small insertions which introduced frameshifts into the T-urf13 ORF (Wise et al. 1987) or deletions which effectively removed the T-urf13 gene from the mitochondrial genome (Rottman et al. 1987). Furthermore in maize (Levings 1990), *Petunia* (Pruitt and Hanson 1991), sunflower (Moneger et al. 1994), wheat (Song and Hedgcoth 1994) and radish carrying the Ogura cytoplasm (Krishnasamy and Makaroff 1994), the introduction of nuclear fertility restoring (Rf) genes dramatically decreases the abundance of the CMS-specific chimaeric transcripts and proteins. In rice the Rf-1 gene appears to promote correct processing and editing of the CMS-specific chimaeric *atp6* transcripts (Iwabuchi et al. 1993).

The molecular basis of CMS in perennial ryegrasses has not been extensively investigated. Molecular analyses of the CMS *L. perenne* lines developed by Wit (1974) revealed rearrangements of the mitochondrial *atp6* and *coxI* loci that correlate with the CMS phenotype (Rouwendal et al. 1992) but the precise molecular consequences of the rearrangements have yet to be elucidated.

We are currently investigating the molecular basis of the CMS trait in the *L. perenne* male-sterile lines developed by Connolly and Wright-Turner (1984). In a previous report (Kiang et al. 1994) we described how a Southern-blot analysis of total DNA from the male-sterile line CMS9B290 and several early *L. perenne* backcross generation hybrids revealed that, contrary to expectation, the mitochondrial genome of CMS9B290 was not that of *F. pratensis*, the maternal parent used in the original intergeneric cross (Connolly and Wright-Turner 1984), but was instead almost identical to that of the male parent, *L. perenne*. Indeed a comparative Southern-blot analysis of purified mtDNA from CMS9B290 and from fertile *L. perenne*, employing a range of restriction enzymes and probes for the mitochondrial genes *coxI*, *coxII*, *atpA* and *cob*, showed identical hybridization profiles (Kiang and Kavanagh, unpublished). However, close inspection of ethidium bromide-stained gels revealed a small number of restriction-fragment differences between the mtDNAs of fertile and sterile lines suggesting that limited rearrangement of the progenitor fertile *L. perenne* mitochondrial genome had occurred in the CMS line. The single most striking difference was the absence in CMS9B290 mtDNA of a 5.6-kb *Hin*dIII fragment that was present in the mtDNA of fertile *L. perenne* (Kiang et al. 1993).

In the present report, we describe the structure, nucleotide sequence and expression of the mtDNA region spanning the 5.6-kb *Hin*dIII fragment in fertile *L. perenne* and the corresponding rearranged region in CMS9B290.

Materials and methods

Plant material

The source of commercial varieties of *F. pratensis* and *L. perenne* and material from the original CMS *L. perenne* breeding program have been described elsewhere (Connolly and Wright-Turner 1984; Kiang et al. 1993). The CMS line used in the present analysis was CMS9B290, a ninth-generation backcross hybrid. The fertile *L. perenne* line used in these studies was the *L. perenne* maintainer line LPSB21 which was used as the backcross parent to CMS9B290 for many generations. Consequently CMS9B290 and LPSB21 were deemed to be a near-isogenic pair.

DNA isolation and molecular cloning

Mitochondria, mtDNA and total cellular DNA were isolated using modifications of standard protocols as described in Kiang et al. (1993, 1994). DNA was cloned into the plasmid vector pUC19 using standard protocols (Maniatis et al. 1982). Restriction fragments of mtDNA, electrophoresed in 0.8% agarose gels, were purified prior to cloning by electroelution into dialysis membranes. The recovered DNA was extracted twice with buffer-saturated phenol and the aqueous layer rendered free of phenol and other contaminants by spin-dialysis through Sepharose-CL6B mini-columns (Murphy and Kavanagh 1988).

Isolation of mitochondrial RNA

Intact mitochondria were prepared as described by Kiang et al. (1993) and resuspended in 1 mlof ice-cold TE (10mM Tris-HCl, 1 mM EDTA pH 8.0). Five volumes of GT buffer [50% (w/v) guanidinium isothiocyanate, 100 mM Tris-HCl pH 7.5, 1% beta-mercaptoethanol] were added to the mitochondrial suspension which was then adjusted to 0.5% sodium sarkosyl. The mixture was clarified by centrifugation at 5000 g for 10 min at room temperature and the supernatant layered on top of a 9-ml cushion of CsCl/EDTA solution [96% (w/v) CsCl in 10 mM EDTA and 0.1% (v/v) diethylprocarbonate]. Mitochondrial RNA was pelted by centrifugation in a Sorvall SW28 rotor at 26000 rpm for 24 h. The RNA pellet was washed in 70% ethanol, dried at room temperature and dissolved in 150 µl of TE. The solution was adjusted to 0.6M sodium acetate, pH 5.2, and the RNA reprecipitated by the addition of 3 vol of ethanol. The RNA was stored as a precipitate at − 20 °C.

Southern- and Northern-blot analyses

Southern-bolt analyses including the random oligomer labelling of probes were carried out as described by Kiang et al. (1993). For Northern-blot analysis, total mtRNA was electrophoresed through 1.25% formaldehyde-agarose gels in 1 × MOPS running buffer [10 × MOPS buffer is 0.2M 3-(N-morpholino)-propanesulphonic acid, 0.05 M sodium acetate pH 7.0, 0.01 M EDTA] as described in Maniatis et al. (1982) and blotted onto Hybond-N (Amersham Inc.). Hybridisation of Northern blots to random oligomer-labelled probes

was carried out at 42 °C in 50% formamide, 300 mM NaCl, 10 mM PIPES buffer (piperazine-N,N'-bis 2-ethanesulfonic acid), pH 6.8, 1 M Na_2-EDTA, 0.1% SDS, 1 × Denhardt's solution (Maniatis et al. 1982). Hybridised filters were washed to a stringency of 0.3 × SSC at 65°C for 15 min. Probes were stripped off filters by immersion in 0.1% SDS for 5 min at 100 °C.

DNA sequencing

Double-stranded DNA was prepared for sequencing as described by Murphy and Kavanagh (1988) and sequenced using a sequenase kit (United States Biochemical Corporation).

Results

Restriction mapping of a mtDNA rearrangement in CMS *L. perenne*

The previously reported 5.6-kb *Hin*dIII fragment difference observed when restriction digests of mtDNA from the fertile *L. perenne* maintainer line LPSB21 and the CMS line 9B290 are compared (Kiang et al. 1993) was the starting point for a detailed analysis of structural rearrangements in the CMS mitochondrial genome. For this reason, and because of its apparent diagnostic value in distinguishing between the sterile and fertile cytoplasms (Kiang et al. 1993), the 5.6-kb *Hin*dIII fragment was gel-purified from mtDNA restriction digests and cloned into pUC19 to give pLMF56.

When used as a probe, pLMF56 hybridised to *Hin*dIII fragments of 7.4 kb and 6.0 kb in CMS9B290 mtDNA and to a 5.8-kb fragment in fertile LPSB21 mtDNA (in addition to the 5.6-kb fragment from which the probe was itself derived; Fig. 1) suggesting that rearrangement, rather than extensive deletion of sequences sharing homology with pLMF56, had occured in CMS9B290. To investigate this possibility and the transcriptional consequences of rearrangements in this region of the CMS mitochondrial genome, the 7.4-kb *Hin*dIII fragment from CMS9B290 mtDNA was cloned to give pCMS74. *Bam*HI digestion of pCMS74 generated two fragments of 2.9kb and 4.5 kb which were subcloned to give pCMS29 and pCMS45 respectively (Fig. 2). These sub-clones were used as probes in a Southern-blot analysis of single and double *Bam*HI and *Hin*dIII digests of CMS9B290 mtDNA in order to extend the mapped region to the left and right of the 7.4-kb *Hin*dIII fragment. pCMS29 and pCMS45 hybridised to *Bam*HI fragments of 7.8 kb and 13.3 kb respectively in CMS9B290 mtDNA (Fig.1), thus extending the mapped region to 21.1 kb (Fig. 2). pCMS29 also hybridised to the 5.6-kb and 8.5-kb *Hin*dIII fragments in fertile LPSB21 mtDNA that were originally detected by probing blots with pLMF56 (Fig. 1). Surprisingly, however, pCMS45 failed to detect any sequences in fertile LPSB21 mtDNA (Fig.1). Consequently pCMS45 proved to be a useful diagnostic probe specific for the CMS cytoplasm (Kiang et al. 1993).

Restriction mapping of homologous regions in fertile *L. perenne* mtDNA

In order to investigate the structure of the corresponding region in fertile LPSB21 mtDNA (i.e. sharing homol-

Fig 1a–c Southern-blot analysis of mtDNA from fertile and CMS *L. perenne*. **a** *Hin*dIII-digested mtDNA hybridised with the pLMF56 probe. *Lane 1*, LPSB21 mtDNA; *Lane 2*, CMS9B290 mtDNA. **b** CMS9B290 mtDNA digested with *Bam*HI (*Lane 1*), *Bam*HI and *Hin*dIII (*Lane 2*), *Hin*dIII (*Lane 3*) and fertile LPSB21 mtDNA digested with *Hin*dIII (*Lane 4*) hybridised with the pCMS29 probe. **c** Same as in **b** except that the probe was pCMS45

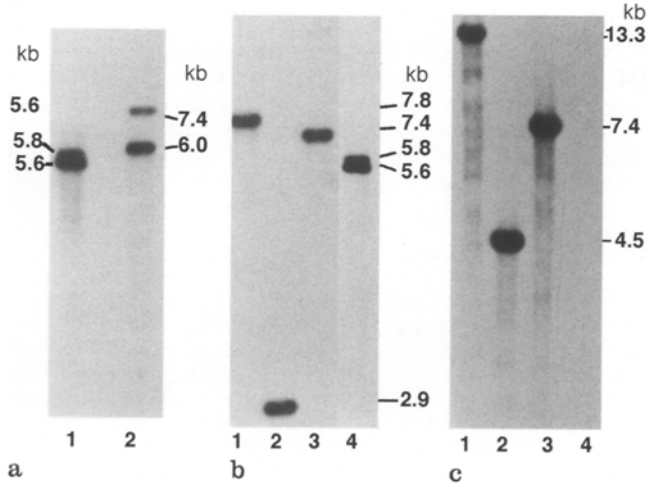
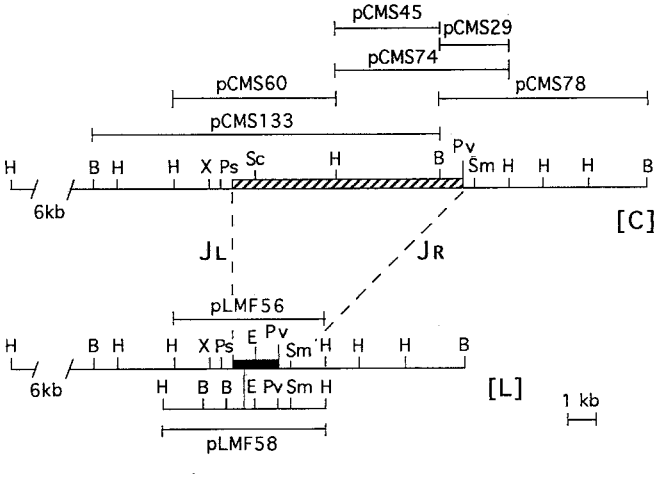


Fig. 2 Restriction maps of the CMS-associated region in CMS9B290 mtDNA and in fertile LPSB21 mtDNA. The non-homologous regions mapped in LPSB21 mtDNA [*L*] and in CMS9B290 [*C*] are represented by *solid* and *hatched rectangles* respectively. The *narrow lines* flanking the *boxed regions* represent homologous sequences. the location of the mtDNA fragments cloned in pUC19 and used as probes in hybridisation analyses are indicated by *solid bars* above the restriction maps. The *broken vertical and diagonal lines* represent the leftward and rightward junctions (J_L and J_R, respectively) delimiting the homologous and the non-homologous regions. All *Bam*HI and *Hin*dIII sites were mapped; in addition selected sites cleaved by other enzymes are also shown. *B* = *Bam*HI, *H* = *Hin*dIII, *E* = *Eco*RI, *P* = *Pst*I, *Pv* = *Pvu*II, *Sc* = *Sca*I, *Sm* = *Sma*I



ogy with the CMS9B290 mtDNA region described above) and to further extend the restriction maps of both mtDNA types, the 7.8-kb and 13.3-kb *Bam*HI fragments from CMS9B290 mtDNA were cloned as pCMS78 and pCMS133, respectively (Fig. 2), and used to probe filter-bound single and double digests of both types of mtDNA (data not shown). The restriction maps derived by this approach show that mtDNAs share homologous regions of 11 kb and 6 kb (the left and right arms of each map, respectively) which flank a non homologous region of 1.8 kb in fertile LPSB21 mtDNA and 9.6 kb in CMS9B290 mtDNA (Fig. 2). The 1.8-kb non-homologous region is located centrally within the 5.6-kb *Hin*dIII fragment (cloned as pLMF56) that is found exclusively in restriction digests of LPSB21 mtDNA. By contrast, the 9.6-kb non-homologous region in CMS9B290 mtDNA is divided between two adjacent *Hin*dIII fragments: these are the 6.0-kb and 7.4-kb fragments detected when a Southern blot containing *Hin*dIII-digested CMS9B290 mtDNA is probed with pLMF56 (Fig. 1).

The 5.8-kb *Hin*dIII restriction fragment that was detected in LPSB21 mtDNA by pLMF56 (Fig. 1) was sub-cloned to give pLMF58. When the restriction maps of pLMF56 and pLMF58 were compared, the rightward halves of both these cloned *Hin*dIII fragments (as diagrammed in Fig. 2) were found to be identical. However, the leftward halves had very different restriction maps suggesting that they contained non-homologous sequences. This was verified by a Southern-blot analysis using probes derived from the rightward and the leftward ends of pLMF56: the rightward probe detected both *Hin*dIII fragments (5.6 and 5.8 kb) in fertile *L. perenne* mtDNA but the leftward probe only detected the particular *Hin*dIII fragment from which the probe itself was derived (data not shown). This suggests that one or other of these *Hin*dIII fragments may have been generated by recombination between molecules containing similar sequences in the mitochondrial genome of LPSB21.

Sequence analysis of the junctions between homologous and non-homologous mtDNA regions in fertile and CMS *L. perenne* mtDNA

To define more precisely the junctions between the homologous and non-homologous regions in the restriction maps of both mtDNA types, the nucleotide sequence of the following restriction fragments was determined: (1) a 3-kb *Xba*I-*Sma*I fragment containing both the leftward (J_L) and rightward (J_R) junctions delimiting the 1.8-kb non-homologous region in LPSB21 mtDNA and (2) a 1.85-kb *Xba*I-*Sca*I restriction fragment and a 1.5-kb *Bam*HI-*Sma*I restriction fragment containing sequences spanning J_L and J_R respectively in CMS9B290 mtDNA (Figs. 2 and 4 a).

An alignment of selected mtDNA nucleotide sequences spanning the various junction sites is presented in Fig. 3. For brevity, 885 bp from the non-homologous region in LPSB21 mtDNA and 8.5 kb from the corresponding region in CMS9B290 mtDNA, both of which begin at position 1777, have been omitted. (Neither of these segments detected any transcripts when used as probes on Northern blots of mtRNA). Extending rightwards from the common *Xba*I site (at nucleotide position 1) towards J_L both sequences are identical for the first 957 bases. This is followed by a 29-bp region containing six mismatches after which both sequences diverge completely. The sequence motif, 5′-TATA-TATAT-3′, found that position 978 in LPSB21 mtDNA defines the junction site J_L and is repeated imperfectly one and a half times on the non-homologous side of the junction in CMS9B290 mtDNA.

Extending leftwards from the *Sma*I site at position 2147 into the common homologous region, both sequences are identical for 263 nucleotides. This includes a pair of imperfect 9-bp inverted repeats (5′-AC-TTCAATC-3′), capable of forming a stem-loop structure, located immediately adjacent to J_R on the homologous side of the junction.

Identification of open reading frames (orfs)

A preliminary Northern-blot analysis of RNA from fertile and CMS mitochondria indicated that sequences in the common leftward ends of the 3-kb *Xba*I-*Sma*I fragment from LPSB21 mtDNA and the 1.8-kb *Xba*I-*Sca*I fragment from CMS9B290 gave rise to RNA transcript (data not shown). Consequently, a search for orfs was conducted by comparing the first 1776 nucleotides of the aligned sequences shown in Fig. 3 with mitochondrial gene sequences deposited in the Genbank and EMBL nucleotide-sequence databases. From this analysis, sequences encoding the *atp9* gene were identified within the common region, extending from nucleotide 651 to 878.

When compared with the *atp9* genes of maize (Dewey et al. 1985), rice (Kaleikau et al. 1990) and wheat (Schulte et al. 1989), the *L. perenne atp9* gene exhibited nucleotide-sequence homologies of 89, 90 and 92% respectively, within the coding region. A search for putative promoter elements identified a 245-bp region (nucleotides 104–349) showing 99% sequence identity with the promotor region of the co-transcribed *atpA-Atp9* genes of wheat (EMBL Accesion No X54387) (Covello and Gray 1991). In wheat, two *atpA-atp9* transcription start sites have been mapped to this region (the A and G nucleotides at positions 256 and 257) by direct sequencing of in vitro-capped mitochondrial transcripts. A highly conserved putative regulatory motif, 5′-TCGAAATAG-CGTATAA-3′, located immediately unstream of the transcription start site, is also found upstream of the transcription start sites of the wheat *coxII* and *orf25* genes (Covello and Gray 1991).

The most striking feature of the sequence alignment shown in Fig. 3 is the occurrence of a duplicated segment

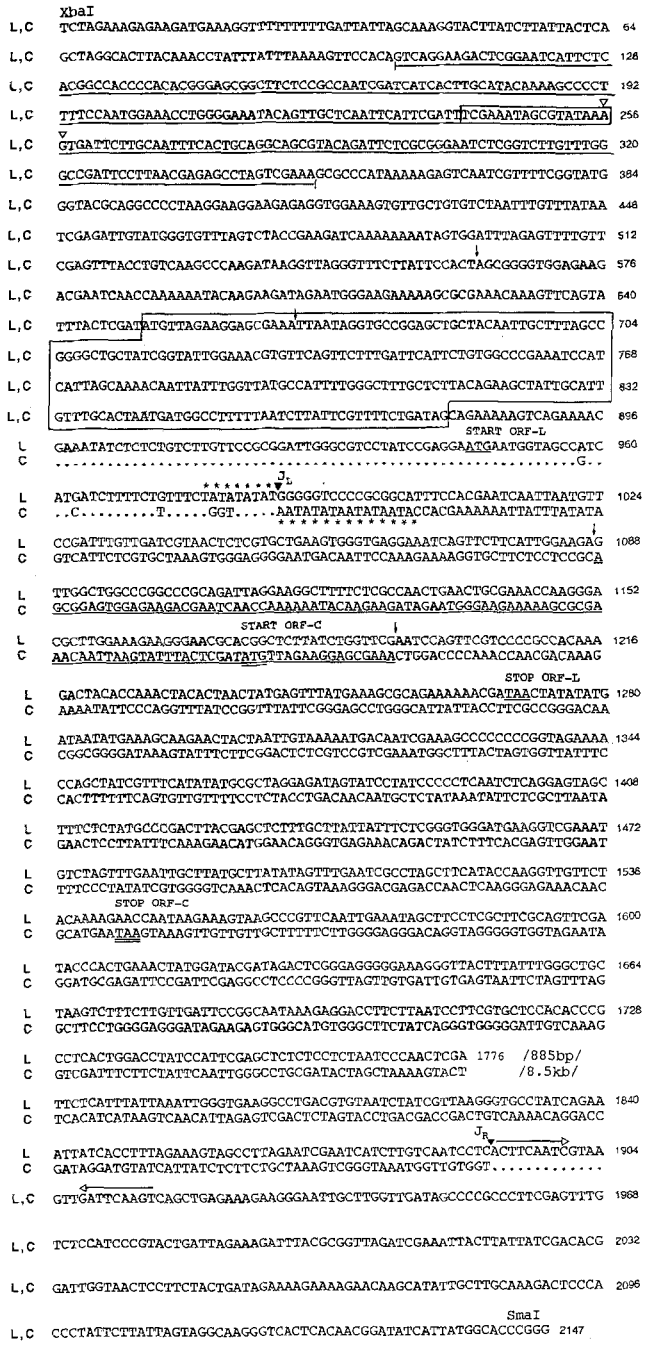


Fig. 3 Nucleotide-sequence analysis of the regions spanning J_L and J_R. An alignment of nucleotide sequence from the 3-kb *Xba*I-*Sma*I fragment of LPSB21 mtDNA (*L*) and the 1.8-kb *Xba*I-*Sca*I and 1.5-kb *Bam*HI-*Sma*I fragments of CMS9B290 mtDNA (*C*) is shown. Sequences in *C* that are identical to those in *L* are given either on the same line or are represented by *dots* below the *L* sequence. The *L* sequence is numbered from 1–2147. The 245-bp promotor region sharing high sequence homology with sequences upstream of the wheat *atpA-atp9* genes is *underlined* (nucleotides 104–349) and the putative regulatory motif flanking the transcription start site is *boxed*. The wheat *atpA-atp9*-homologous transcription start sites are indicated with *open arrow heads*. The *L. perenne atp9* gene is *boxed* (nucleotides 651–878). The 106-bp region containing *atp9* upstream and coding sequences (nucleotides 562–668) that is duplicated in *C* is delimited by *small vertical arrows* (and *underlined* in *C*). The rightward and leftward junctions between the homologous and non-homologous sequences (J_R and J_L, respectively) are indicated by *filled arrow heads*. The AT-rich repeat sequence motifs located on opposite sides of J_L in *L* and *C* are indicated by *. The start and stop codons of *orf-L* and *orf-C9* are *singly or doubly underlined*, respectively. The inverted repeat located on the homologous side of J_R is indicated by *horizontal arrows*. At position 1777 non-homologous, non-transcribed sequences totalling 885 bp in *L* and 8.5 kb in *C* have been omitted. Some restriction enzymes sites are indicated above their recognition sequences. The *L* and *C* nucleotide sequences will appear in the EMBL Nucleotide Sequence Database under the accession numbers Z50199 and Z50200, respectively

of the *atp9* gene in CMS9B290 mtDNA at nucleotide positions 1088–1193. The duplication is located 101 bp downstream from J_L in the non-homologous region and comprises a 106-bp interval corresponding to nucleotides 562–668 from the leftward common homologous region (underlined in Fig. 3). This effectively duplicates the putative promotor region of the *atp9* gene between-1 and -85 (relative to the ATG) and fuses the first six codons of *atp9* in-frame to a further 117 codons of unknown sequence. The resulting CMS-specific orf, designated *orf-C9*, extends from position 1175 to 1544 and potentially encodes an ATP9 fusion protein of 123 amino-acid residues with a predicted molecular weight of 14 kDa.

In fertile LPSB21 mtDNA an orf of 327 bp (nucleotides 944–1271), termed *orf-L*, which potentially encodes a 12.3-kDa polypeptide, is located 70 bp downstream from the termination codon of the *atp9* gene. The first 30 bp of *orf-L* are located in both mtDNAs within the common homologous region (immediately adjacent to J_L) and are thus also present in CMS9B290 mtDNA. However, in CMS9B290 mtDNA, *orf-L* sequences terminate (TAA) after 15 codons. A DNA sequence-homology search failed to reveal any significant regions of homology either between *orf-C9* (excluding the *atp9* region) or *orf-L*, and sequences held in the databases.

Transcriptional consequences of the CMS *L. perenne* mtDNA rearrangement

To ascertain whether the genomic rearrangements observed in CMS9B290 mtDNA resulted in transcriptional differences between fertile and CMS lines, Northern hybridisation were carried out with purified mtRNA from fertile and sterile plants. Probes spanning the rightward common region and 5.5 kb of the non-homologous region in CMS9B290 mtDNA (pCMS78 and pCMS45 respectively; Fig. 2) failed to detect any transcripts in mtRNA from fertile or sterile lines (data not shown). However, two probes, pCMS60 (Fig. 2), which includes sequences from the leftward common region and 3.8 kb from the non-homologous region in CMS9B290, and probe 2 (Fig. 4a), specific for *atp9* transcripts, each identified a 1.4-kb transcript in CMS9B290 mtRNA and a 1.0-kb transcript in LPSB21 mtRNA. Mitochondrial RNA from *F. pratensis* was included in the blot probed with pCMS60 and a transcript of approximately 1.0 kb was also detected in this species (Fig. 4b). Probe 3, which was specific for *orf-L*

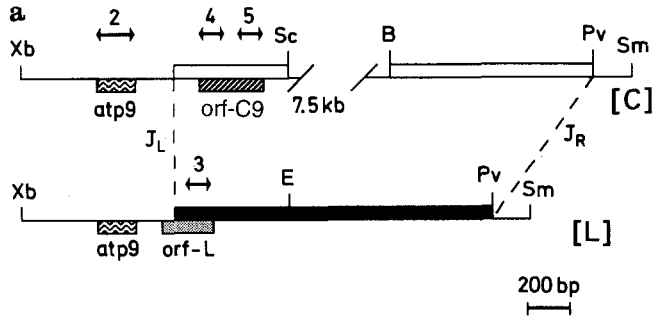
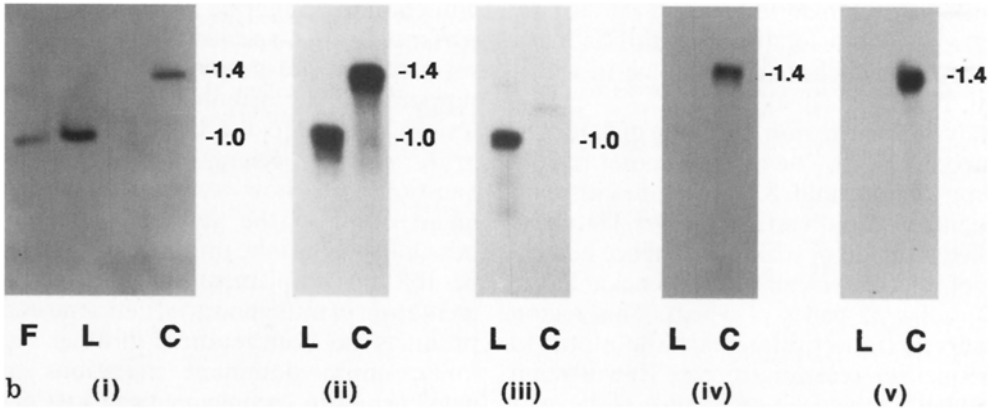


Fig. 4a, b Northern-blot analyses of mitochondrial transcripts of LPSB21 and CMS9B290. **a** Restriction maps of the regions for which nucleotide sequences were determined in CMS9B290 mtDNA [*C*] and in fertile LPSB21 mtDNA [*L*]. Homologous flanking regions are indicated by *narrow lines*; non-homologous regions are represented by *rectangles* above the lines. J_L and J_R represent the leftward and rightward junctions respectively, between homologous and non-homologous sequences. The locations of open reading frames for *atp9*, *orf-C9* and *orf-L* are indicated by *rectangular boxes* below the lines. The regions from which the Northern-blot probes 2, 3, 4, 5 were derived are indicated by the *bars* shown above the maps. All probes were gel-purified restriction fragments. Probe 2 (*atp9*-specific) was a 480-bp *Sau*3A1 fragment from pLMF56, probe 3 (*orf-L*-specific) was a 400-bp *Sau*3A1 fragment from pLMF56; probes 4 and 5 (*orf-C9*-specific) were 300-bp and 310-bp *Stu*I-*Spe*I fragments, respectively, from pCMS60. **b** Northern-blotted purified mtRNA was hybridised with (*i*) pCMS60 (see Fig. 2), (*ii*) probe 2, (*iii*) probe 3, (*iv*) probe 4 and (*v*) probe 5. *F. pratensis* mtRNA (*F*) was included in **a**



(but did not include sequences common to fertile and CMS mtDNAs), hybridised to a 1.0-kb transcript in LPSB21 mtRNA but failed to show significant hybridization to transcripts in mtRNA of the CMS line (the faint hybridisation signals in CMS9B290 mtRNA are due to incomplete removal of a previous probe) (Fig. 4 b). Probes 4 and 5 representing 5′ and 3′ portions of the unique (i.e. non-*atp9*) sequences in *orf-C9* respectively, both hybridised to a 1.4-kb transcript in CMS9B290 mtRNA but failed to hybridise to LPSB21 mtRNA (Fig. 4 b).

Each of the Northern blots shown in Fig. 4 b) was re-probed with the *atp 9*-specific probe 2. Alignment of autoradiographs from the initial and subsequent (*atp 9*-specific) hybridisations showed that all of the bands sized at 1.4 kb occupied a single position on the filter as did those sized at 1.0 kb, suggesting that *orf-C9* and *orf-L* sequences are co-transcribed with *atp 9* sequences. These data are consistent with transcription initiation occurring at the putative (wheat) transcription start site at nucleotide 256 (Fig. 3) and terminating downstream from *orf-L* in LPSB21 mitochondria, or downstream from *orf-C9* in CMS9B290 mitochondria.

Discussion

The male-sterile *L. perenne* line CMS9B290 was originally derived from an intergeneric sexual hybridisation between two fertile species, *L. perenne* and *F. pratensis*, followed by backcrossing to *L. perenne* over nine generations (Connolly and Wright-Turner 1984). Hence it was expected that differences between the mtDNA restriction-fragment profiles of the CMS9B290 and fertile *L. perenne* lines might segregate with the CMS phenotype. The most obvious single difference between the mitochondrial genomes of sterile and fertile lines was the absence in CMS9B290 of a unique 5.6-kb *Hin*dIII fragment that was observed in the mtDNA of fertile *L. perenne* (Kiang et al. 1993). Southern-blot analysis of CMS9B290 mtDNA, using the cloned 5.6-kb *Hin*dIII fragment (pLMF56) from fertile *L. perenne* as a probe, revealed additional *Hin*dIII fragments of 6.0 kb and 7.4 kb that were diagnostic for the CMS9B290 mitochondrial genome, but were not observed on ethidium bromide-stained agarose gels due to the masking effect of co-migrating fragments common to both mtDNA types. Restriction mapping of the region encompassing the 5.6-kb *Hin*dIII fragment in the fertile line and the corresponding region in CMS9B290 mtDNA showed that they differed in both their sequence content and transcription pattern.

The restriction-mapping data suggests a model for the origin of the CMS-associated rearrangement in which recombination across repeated sequences at J_L and J_R resulted in the deletion of 1.8 kb in CMS9B290 relative to LPS21 mtDNA and its replacement with 9 kb of non-homologous sequence. Both the junctions defining the site of the deletion/insertion event are located

within the 5.6-kb *Hin*dIII fragment in fertile LPSB21 mtDNA. Evidence for the presence of a recombinationally active region within this fragment is suggested by the restriction maps of pLMF56 and pLMF58. Both these cloned *Hin*dIII fragments contain identical restriction sites in their rightward halves but differ completely in their leftward halves suggesting that one may have been derived from the other by recombination. Furthermore, the junctions between homologous and non-homologous regions in the CMS and fertile *L. perenne* mitochondrial genomes contain sequence elements of the type known to be recombinogenic: short directly repeated A-T rich sequences at J_L and an inverted repeat sequence at J_R. Recombination within sequences as short as 7 bp has been invoked to account for the origin of the CMS-associated chimaeric gene *urf-rmc* in rice (Kadowaki et al. 1990) and for tissue culture-induced rearrangement of the mitochondrial genome in wheat (Hartmann et al. 1994).

Sequence analysis of the region upstream of J_L identified the gene encoding *atp9*. The *atp9* promoter region, the entire coding region, and 83 bp of downstream sequences are identical in both mtDNA types. Database searches identified a region of striking sequence homology upstream of the *L. perenne apt9* gene and the co-transcribed *atpA-atp9* genes of wheat. This region includes a conserved transcription start site motif. In CMS9B290, sequence rearrangements downstream from J_L have resulted in the deletion of most of the *orf-L* sequences found downstream from the *atp9* gene in fertile *L perenne* mtDNA and simultaneously created a chimaeric gene, *orf-C9*, containing the first six codons of *atp9* fused to downstream sequences encoding an additional 117 amino acids. In *Petunia*, a similar chimaeric gene, the S-*pcf* gene, assembled from promotor and coding regions of *atp9*, the coding regions of the *coxII* gene and additional unidentified sequences, has been implicated in the CMS phenotype (Young et al. 1986; Young and Hanson 1987).

Assuming that *atp9* transcripts in *L. perenne* originate at the conserved wheat *atpA-atp9* transcription start-site motif, then the Northern-blot analyses suggests that *orf-C9* is co-transcribed with the only normal copy of the *atp9* gene in CMS9B290. Co-transcription of CMS-associated orfs with upstream or downstream coding sequences is not unusual. The S-*pcf* gene in *petunia* is co-transcribed with the downstream genes *nad3* and *rps12* (Nivison and Hanson 1989) and in sunflower the CMS-associated *orf-522* is co-transcribed with an upstream *atpA* gene (Kohler et al. 1991; Laver et al. 1991). In general, the presence of CMS-associated sequences does not appear to affect the correct expression of the co-transcribed normal gene. However, in rice carrying the *cms-bo* cytoplasm, co-transcription of CMS-associated sequences with a duplicated but normal copy of the *atp6* gene (B-*atp6*) results in abnormal processing and editing of the *atp6* portion of the transcript (Iwabuchi et al. 1993). If translated, these incorrectly edited *atp6* transcripts would be expected to give rise to abnormal ATP6 polypeptides. Restoration to fertility by the rice nuclear Rf-1 gene is accompanied by restoration of a normal *atp6* editing pattern which presumably prevents translation of abnormal ATP6 polypeptides (Iwabuchi et al. 1993). In the case of CMS *L. perenne*, it remains to be seen whether the co-transcribed *orf-C9* sequences interfere with the correct post-transcriptional processing or the translation of the unstream *atp9* region of the transcript.

The simultaneous absence of *orf-L* transcripts and the presence of chimaeric *orf-C9* transcripts in mitochondria of the CMS line suggest a number of scenarios which might account for the CMS phenotype in *L. perenne*. Should the putative ORF-L polypeptide be indispensable for fertility, its absence in CMS mitochondria might be sufficient to cause sterility. The presence of *orf-L*-specific transcripts in fertile *F. pratensis*, a related grass species, adds some circumstantial support to this possibility. However, nuclear genes that restore fertility to the CMS line have been identified in *L. perenne* (Kiang et al. 1993). This raises the question as to how restoration of an *of-L* deficiency might occur in the absence of an *orf-L* gene. One possibility is that the products of a nuclear Rf gene may be able to complement an *orf-L* deficiency. Complementation of mitochondrial deficiencies by nuclear gene products has been reported in other systems. In yeast, for example, dominant mutations in the nuclear *nam2* gene can complement the loss of the *cob* intron 4 maturase by activating a latent maturase activity encoded by the fourth intron of *coxI* (Dujardin et al. 1983).

Evidence from other CMS systems, however, strongly implicates the products of chimaeric orfs as causative agents in CMS. Assuming that *orf-C9* transcripts are translated, the predicted polypeptide, which has an ATP9 amino-terminus, might be able to compete with the authentic ATP9 polypeptide for incorporation into the mitochondrial F_0-ATPase complex. In this capacity, the ORF-C9 product might function as a dominant inhibitor of mitochondrial function during critical stages of microsporogenesis, as is thought to be the case for T-URF13 in CMS-T maize (Levings 1990). Typically, where the translation products of chimaeric orfs are implicated in the CMS phenotype, nuclear Rf genes act to decrease the abundance of the chimaeric polypeptide. With respect to the putative *L. perenne* ORF-C9 protein, this remains to be investigated.

Analysis of the effects of nuclear restorer genes and reversion to fertility on mtDNA structure and on *orf-L* and *orf-C9* transcripts and their translation products are in progress and should facilitate a more complete understanding of the mechanism responsible for the CMS phenotype in *L. perenne*.

Acknowledgements We thank V. Connolly for providing the *L. perenne* genotypes used in this study. We are grateful to D.M. Lonsdale for technical advice on mtDNA isolation, K. Wolfe for

assistance with DNA sequence analysis and C.J. Leaver for helpful discussions.

References

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

Covello PS and Gray MW (1991) Sequence analysis of wheat mitochondrial transcripts capped in vitro: definitive identification of transcription initiation sites. Curr Genet 20:245–251

Dewey RE, Schuster AM, Levings CS III, Timothy DH (1985) Nucleotide sequence of F_0-ATPase proteolipid (subunit 9) gene of maize mitochondria. Proc Natl Acad Sci USA 82:1015–1019

Dujardin G, Labouesse M, Netter, P, Slonimski PP (1983) Genetic and biochemical studies of the nuclear supressor NAM2: extraneous activation of a latent pleiotropic maturase. In: Schweyen RJ, Wolf K, Kaudewitz F (eds) Mitochondrial 1983– Nucleo-Mitochondrial Interactions. Walter de Gruyter, Berlin, pp 233–250

Grelon M, Budar F, Bonhomme S, Pelletier, G (1994) Ogura cytoplasmic male sterility (CMS)-associated orf138 is translated into a mitochondrial membrane polypeptide in male-sterile *Brassica* cybrids. Mol Gen Genet 243:540–547

Hanson MR (1991) Plant mitochondrial mutations and male sterility. Annu Rev Genet 25:461–486

Hartmann C, Recipon H, Jubier M-F, Valon C, Delcher-Besin E, Henry Y, De Buyser J, Lejeune B, Rode A (1994) Mitochondrial DNA variability detected in a single wheat regenerant involves a rare recombination event across a short repeat. Curr Genet 25:456–464

Iwabuchi M, Kyozuka J, Shimamoto K (1993) Processing followed by complete editing of an altered mitochondrial atp6 RNA restores fertility of Cytoplasmic male-sterile rice. EMBO J 12:1437–1446

Kadowaki K-I, Suzuki T, Kazama S (1990) A chimeric gene containing the 5′ portion of *atp6* is associated with the cytoplasmic male sterility of rice. Mol Gen Genet 224:10–16

Kaleikau EK, Andre CP, Walbot V (1990) Sequence of the Fo-ATPase proteolipid (*atp9*) gene from rice mitochondria. Nucleic Acids Res 18:370

Kiang A-S, 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 A-S, 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

Kohler RH, Horn R, Lossl A, Zetsche K (1991) Cytoplasmic male sterility in sunflower is correlated with the co-transcription of a new open reading frame with the *atpA* gene. Mol Gen Genet 227:369–376

Krishnasamy S, Makaroff CA (1994) Organ-specific reduction in the abundance of a mitochondrial protein accompanies fertility restoration in cytoplasmic male-sterile radish. Plant Mol Biol 26:935–946

Laver HK, Reynolds SJ, Moneger F, Leaver CJ (1991) Mitochondrial genome organization and expression associated with cytoplasmic male sterility in sunflower (*Helianthus annus*). Plant J 1:185–193

Levings CSIII (1990) The Texas cytoplasm of maize: cytoplasmic male sterility and disease susceptibility. Science 250:942–947

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

Moneger F, Smart CJ, Leaver CJ (1994) Nuclear restoration of cytoplasmic male sterility in sunflower is associated with tissue-specific regulation of a novel mitochondrial gene. EMBO J 13:8–17

Murphy G, Kavanagh TA (1988) Speeding-up the sequencing of double-stranded DNA. Nucleic Acids Res 16:5198

Nivison HT, Hanson MR (1989) Identification of a mitochondrial protein associated with cytoplasmic male sterility in petunia. Plant Cell 1:1121–1130

Pruitt KD, Hanson MR (1991) Transcription of the *Petunia* mitochondrial cms-associated *pcf* locus in male-sterile and fertility restored lines. Mol Gen Genet 227:348–355

Rottmann WH, Brears T, Hodge TP, Lonsdale DM (1987) A mitochondrial gene is lost via homologous recombination during reversion of CMS T maize to fertility. EMBO J 6:1541–1546

Rouwendal GJA, Creemers-Molenaar, J, Krens FA (1992) Molecular aspects of cytoplasmic male sterility in perennial ryegrass (*Lolium perenne* L.): mtDNA and RNA differences between plants with male sterile and fertile cytoplasm and restriction mapping of their *atp6* and *coxI* homologous regions. Theor Appl Genet 83: 330–336

Schulte E, Staubach S, Laser B, Kuch U (1989) Wheat mitochondrial DNA: organization and sequence of *atpA* and *atp9* genes. Nucleic Acids Res 17:7531

Song J, Hedgcoth C (1994) A chimeric gene (*orf256*) is expressed as protein only in cytoplasmic male-sterile lines of wheat. Plant Mol Biol 26:535–539

Wise RP, Pring DR, Gengenbach BG (1987) Mutation to male fertility and toxin insensitivity in Texas (T)-cytoplasm maize is associated with a frameshift in a mitochondrial open reading frame. Proc Natl Acad Sci USA 84:2858–2862

Wit F (1974) Cytoplasmic male sterility in rye grasses (*Lolium* spp.) detected after intergeneric hybridisation. Euphytica 23:31–38

Young EG, Hanson MR (1987) A fused mitochondrial gene associated with cytoplasmic male sterility is developmentally regulated. Cell 50:41–49

Young EG, Hanson MR, Dierks PM (1986) Sequence and transcription analysis of the *Petunia* mitochondrial gene for the ATP synthase proteolipid subunit. Nucleic Acids Res 14:7995–8006