An evolutionarily conserved mitochondrial orf108 is associated with cytoplasmic male sterility in different alloplasmic lines of *Brassica juncea* and induces male sterility in transgenic *Arabidopsis thaliana*

Pankaj Kumar*, Naresh Vasupalli, R. Srinivasan and Shripad R. Bhat†

National Research Centre on Plant Biotechnology, Indian Agricultural Research Institute, New Delhi 110012, India

* Present address: School of Biology, Sir Harold Mitchell Building, University of St Andrews, St Andrews, Fife, KY16 9TH, UK
† To whom correspondence should be addressed. E-mail: srbhat@nrcpb.org

Received 2 October 2011; Revised 22 December 2011; Accepted 22 December 2011

Abstract

Nuclear–mitochondrial gene interactions governing cytoplasmic male sterility (CMS) in angiosperms have been found to be unique to each system. Fertility restoration of three diverse alloplasmic CMS lines of *Brassica juncea* by a line carrying the fertility-restorer gene introgressed from *Moricandia arvensis* prompted this investigation to examine the molecular basis of CMS in these lines. Since previous studies had found altered *atpA* transcription associated with CMS in these lines, the *atpA* genes and transcripts of CMS, fertility-restored, and euplasmic lines were cloned and compared. *atpA* coding and downstream sequences were conserved among CMS and euplasmic lines but major differences were found in the 5’ flanking sequences of *atpA*. A unique open reading frame (ORF), orf108, co-transcribed with *atpA*, was found in male sterile flowers of CMS lines carrying mitochondrial genomes of *Diplotaxis berthaultii*, *D. catholica*, or *D. erucoides*. In presence of the restorer gene, the bicistronic orf108–*atpA* transcript was cleaved within orf108 to yield a monocistronic *atpA* transcript. Transgenic expression of orf108 with anther-specific *Atprx18* promoter in *Arabidopsis thaliana* gave 50% pollen sterility, indicating that Orf108 is lethal at the gametophytic stage. Further, lack of transmission of orf108 to the progeny showed for the first time that mitochondrial ORFs could also cause female sterility. orf108 was found to be widely distributed among wild relatives of *Brassica*, indicating its ancient origin. This is the first report that shows that CMS lines of different origin and morphology could share common molecular basis. The gametic lethality of Orf108 offers a novel opportunity for transgene containment.

Key words: *atpA*, *Brassica juncea*, circularized RNA reverse-transcription PCR, cytoplasmic male sterility, orf108, transgene containment.

Introduction

Cytoplasmic male sterility (CMS) trait is widespread in angiosperms and is the basis of hybrid seed production in many crops. Investigations of CMS in diverse plant species have shown that the male sterility-inducing factors are present on the mitochondrial genome whereas the male fertility-restoring genes are located in the nuclear genome (Chase, 2007; Pelletier and Budar, 2007). A CMS system is thus an example of mitochondrial–nuclear genome interaction and comprises three lines, designated A, B, and R. The A line carries a sterility-inducing mitochondrial genome and is male sterile whereas the R line carries the fertility-restoring nuclear gene and is male fertile. The B line is iso-nuclear to the A line but is fertile because it carries a normal (compatible) mitochondrial genome. Organelle genomes are present in multiple copies, display low intraspecific variation, and are unparentally inherited from the maternal parent in angiosperms. Hence, CMS is occasionally found in crop species but is commonly encountered in alloplasmic
lines derived from interspecific hybridization. Each CMS system studied so far has been found to be unique for the male sterility-causing mitochondrial gene defect, which is overcome by the action of specific nuclear restorer gene(s) (Schnable and Wise, 1998). Accordingly, different CMS sources are traditionally distinguished based on genetics of fertility restoration. Molecular studies with a number of CMS systems have shown that male sterility is associated with altered organization and/or expression of one or more mitochondrial genes (Schnable and Wise, 1998). Differences between male sterile and normal ( euplasmic) lines have been recorded for mitochondrial restriction fragment length polymorphisms (RFLPs), transcript size, transcript editing, occurrence of novel transcripts (open reading frames, or ORFs), and/or proteins (Schnable and Wise, 1998; Chase, 2007; Fujiy and Toriyama, 2008). Fertility restoration is usually accompanied with lowering or disappearance of the altered transcript or protein. In agreement with these observations, majority of the fertility-restorer genes cloned so far have been found to belong to pentatricopeptide repeat (PPR) protein family involved in organellar transcript processing (Chase, 2007).

Several CMS lines of Brassica juncea have been developed through sexual or somatic hybridization with wild related species of Brassica. Bhat et al. (2005, 2006, 2008) reported previously that the R line’s restorer gene introgressed from Moricandia arvensis is capable of conferring male fertility to four CMS lines having mitochondrial genomes of the wild species: M. arvensis, Diplotaxis berthautii, D. catholica, or D. erucoides. These CMS lines, however, differ from one another for the mitochondrial RFLPs and, in some cases, for the male sterility phenotype. For example, anthers are converted into petal-like structures in the CMS line carrying D. catholica mitochondrial genome whereas in the other three CMS lines, anther development is nearly normal except for the male sterility phenotype. For example, anthers are converted into petal-like structures in the CMS line carrying D. catholica mitochondrial genome whereas in the other three CMS lines, anther development is nearly normal except for the sterility pollen. Analyses of mitochondrial transcripts among A, B, and R lines in these four CMS systems had revealed altered atpA transcription associated with male sterility. Occurrence of a bicistronic orf108–atpA transcript in flower buds of the M. arvensis CMS line and monocistronic atpA transcript in the fertility-restored line pointed to a direct link between atpA and CMS (Ashutosh et al., 2008).

The present study carried out a detailed molecular analyses of atpA and the transcripts in the three CMS lines carrying Diplotaxis sp. cytoplasm with the aim of uncovering the molecular mechanisms of CMS in B. juncea. It is shown that orf108 co-transcribed with atpA is associated with male sterility in all these systems. In presence of the fertility-restorer gene, the bicistronic atpA transcript was cleaved within orf108 to yield a monocistronic atpA. The direct link between orf108 and male sterility was demonstrated through transgenic experiments with Arabidopsis thaliana. Further, orf108 was found to be widely distributed among wild allies of Brassica, suggesting an ancient origin. Finally, plants expressing orf108 under the control of the Atrpx18 promoter failed to transmit the transgene to the progeny, which showed that the Atrpx18::orf108 cassette could be useful in devising a novel transgene containment system particularly relevant to clonally propagated plants.

Materials and methods

Plant material
Male sterile and fertility-restored lines of Brassica juncea (L.) Czern. carrying cytoplasm of Diplotaxis berthautii (Bhat et al., 2008), D. catholica (Pathania et al., 2003), or D. erucoides (Bhat et al., 2006) and euplasmic B. juncea cv. Pusa Bold were used in the present study. The wild allied species of Brassica used in this study are from germplasm collections of (late) Dr. Shyam Prakash and date back to more than 30 years. Plants were grown in the field or net house and were checked for their phenotype at flowering.

Nucleic acid isolation
Mitochondrial DNA was isolated from unopened flower buds and total cellular DNA was isolated from young leaves following the method described earlier (Kirti et al., 1995). Mitochondrial RNA isolation was done according to Pathania et al. (2003).

Sequence accession numbers
The sequences reported in this paper were submitted to GenBank at the National Center for Biotechnology Information (NCBI). Their accession numbers are FJ626737, FJ626738, EF483940, JN645148, JN793449, and JN661706–JN661710.

Isolation of atpA–HindIII genomic region
Isolation of atpA from CMS M. arvensis and B. juncea has been reported previously (Gaikwad et al., 2006). The atpA coding and flanking sequences of CMS B. juncea carrying D. catholica or D. berthautii cytoplasm were isolated following PCR amplification. Mitochondrial DNA (1 μg) was digested with the restriction endonuclease HindIII and a 200-ng aliquot was ligated into 70 ng of HindIII-digested Phagemid vector pBluescript SK+ (Fermentas, USA) in a total 40 μl ligation reaction. Ligated product (4 μl) was used as a template for PCR with the primer for the T3 promoter of the vector (5'-ATTAAACCCTCTACTAAAGGG-3') and primer P1 (5'-CTATTTGACCTTTGGAATAATGATTCG-3') to amplify the 5’ flanking sequence. Similarly, the atpA-encoding and 3’ flanking sequence was amplified using the primer combination T3 + P2 (5’-GGAATTGGAATTATCCAGAAGG-3’). PCR products were cloned into pGEM-T Easy Vector (Promega, USA) and sequenced using an automated DNA sequencer (ABI Prism 310, Applied Biosystems, USA). Sequence reads were assembled and analysed using DNASTAR version 4.05 (Lasergene, http://www.dnastar.com).

Circularized RNA reverse-transcription PCR analysis
Mitochondrial transcripts often undergo processing at the 5’ end. While the primary transcripts contain 5’-triphosphate, processed transcripts have 5’-monophosphate. Thus, primary transcripts fail to circulate upon treatment with RNA ligase, which requires 5’-monophosphate and 3’ OH for ligation. However, 5’-triphosphate is processed to 5’-monophosphate by treatment with tobacco acid phosphatase (TAP). Thus circularized RNA reverse-transcription (RT) PCR without or with a prior TAP treatment was employed to identify the 5’ and 3’ ends of primary and processed atpA transcripts (Ashutosh et al., 2008). Briefly, mitochondrial RNA from flower buds was divided into two 10–15-μg aliquots; one aliquot was treated with 25 units of TAP (Epicentre Biotechnologies, USA) at 37 °C for 60 min. After TAP treatment, the RNA was extracted with phenol/chloroform, precipitated with ethanol, and dissolved in water. Subsequently, both lots of RNA (TAP-treated and untreated) were self-ligated with 40 units of T4 RNA ligase (New England Biolabs, USA) in a 20-μl reaction mixture. After ligation, samples were again purified through phenol/chloroform extraction and dissolved in water. First-strand cDNA synthesis was done with 5 μg mitochondrial RNA using primer P1.
Binary vector construction for stable plant transformation

Three gene cassettes, 35S::preatp9–orf108, Atpx18::preatp9–orf108, and Atpx18::orf108, were constructed and cloned into binary vector pBinAR (Höfgen and Willmitzer, 1992) for stable transformation of A. thaliana (Col-0) with orf108. For constitutive expression, orf108 was placed under the control of the cauliflower mosaic virus (CaMV) 35S promoter and targeted to mitochondria by translational fusion of the atp9 precursor of Neurospora crassa (locus tag NCU02250, nucleotide positions 106–325). The cDNA encoding the mitochondrial precursor of atp9 was amplified using the forward primer 5'-ATTGGATCCGTA-TACCAACCAGTACGAA-3' (containing the BamHI recognition sequence, in bold, to facilitate cloning) and the reverse primer 5'-AGATTTCACTG6CCGCCCTTCGGAA-3' (containing the BssHI recognition sequence, in bold, and the sequence complementary to the forward primer of orf108, underlined). The orf108 sequence was amplified from a 5.8-kb genomic clone of CMS M. arvensis using the forward primer 5'-AGAAGGCGGCCAAT- TACTATCAACCACATCTTC-3' (containing the BssHI recognition sequence, in bold, and the sequence complementary to the forward primer of orf108, underlined) and the reverse primer 5'-TTAATGCAGCTAAACCCCTCCCGCCGTAA-3' (containing the SalI recognition sequence, in bold). The amplicons were mixed, denatured at 95 °C for 5–10 min, and allowed to cool. This reaction mixture was used for second round of PCR to obtain a translational fusion of preatp9–orf108 using the forward primer of the atp9 precursor and the reverse primer for orf108. The 550-bp amplicon of preatp9–orf108 was ligated into the BamHI and SalI restriction sites of the binary vector pBinAR to obtain 35S::preatp9–orf108. For another-specific expression of orf108, the CaMV 35S promoter of 35S::preatp9–orf108 was replaced with the promoter of the A. thaliana peroxidase gene (Atpx18, AT2G24800) (Thakare et al., 2006; Kumar et al., 2011). Peroxidase promoter sequences were amplified from A. thaliana genomic DNA using the forward primer 5'-GAATTCGTTAGTGAATTTCCGAGTCCA-3' (containing the EcoRI recognition sequence, in bold) and the reverse primer 5'-GGTACCAAGAGAAGTGTATGTTG-3' (containing the KpnI recognition sequence, in bold). The Atpx18::preatp9–orf108 was further modified by replacing the preatp9–orf108 sequence with the orf108 sequence within the BamHI and SalI restriction sites to retain ORF108 in the cytoplasm. This construct was designated Atpx18::orf108.

Transformation of A. thaliana (Col-0)

A. thaliana (Col-0) plants were transformed by the floral dip method (Clough and Bent, 1998) with the three binary vector constructs mobilized into Agrobacterium tumefaciens (GV3101). Transgenic plants were selected by germinating T1 seeds on MS agar (Murashige and Skoog, 1962) containing kanamycin.

Pollen viability test by fluorescein diacetate staining

Pollen viability was assessed through the fluorescein diacetate (FDA) test in accordance to McCabe et al. (1997). A 0.1% (w/v) FDA stock solution was prepared in acetone. This was diluted to 0.002% with 0.5 M mannitol immediately prior to use. Anthers were placed on a slide in a few drops of FDA working solution and covered with a cover glass. Pollens were released by gently tapping the cover slip with a fine needle. After incubation in dark for 5 min, pollens were observed under fluorescence microscope (Carl Zeiss, Germany) fitted with an excitation filter (B-2A, 450–490 nm) and a barrier filter (BA 520). Pollen grains that were fluorescing green were counted in four microscopic fields and photographed. Also, pollen grains that were visible with normal light in the same field were counted to determine the percentage of viable pollen.

Amplification of orf108 from wild Brassica spp

Presence of orf108 in wild relatives of Brassica was checked by amplification using genomic DNA as template and the primers P4 (5’-CTTCACCTACTCTAGAGGCT-3’) and P1.

Results

Comparison of atpA and its flanking sequences

This study isolated atpA and its flanking sequences from D. berthaultii, D. catholica, and D. erucoides. PCR was used to clone atpA and its flanking sequences from CMS D. berthaultii and D. catholica. The atpA coding sequence was amplified using primers based on the atpA sequence of euplasmic B. juncea. A 1.5-kb amplicon was obtained in all the three CMS lines. The 5’ flanking, atpA coding, and 3’ flanking sequences were amplified utilizing sequence information of the atpA coding region and the phagemid pBlueScript SK+. The HindIII-digested mitochondrial DNA was ligated into pBlueScript SK+ and used as template for amplification. In CMS D. catholica and D. berthaultii, the primer combinations T3+P1 and T3+P2 amplified 1.35-kb and 4.5-kb DNA fragments, respectively (Fig. 1A, B). Following sequencing and assembly, a 5.7-kb sequence containing atpA of CMS D. catholica (FJ626738) and D. berthaultii (JN645148) were obtained. The atpA coding sequence of CMS D. erucoides was
amplified utilizing atpA sequence information of B. juncea and 5′ flanking sequences were amplified using primer P1+P4 (data not shown here). Thus, a 2.0-kb sequence containing atpA coding and 5′ flanking sequences of CMS D. erucoides (JN793449) was assembled. The atpA coding sequences of all the three CMS systems are identical and matched with the atpA coding sequences of CMS M. arvensis (EF483940) and euplasmic B. juncea (FJ626737). Similarly, approximately 1659-bp sequences downstream of atpA were also identical between euplasmic B. juncea and the CMS lines having the mitochondrial genomes of M. arvensis, D. berthautii, or D. catholica (Fig. 2). Except for a few indels and a single-nucleotide polymorphism, sequence identity was found to extend ~580 bp upstream to the atpA of the CMS lines (Fig. 2). The consensus nonanucleotide mitochondrial promoter sequence AAAATATCATTAAGAGAAG found upstream of the atpA start codon in M. arvensis (714 bp) was not detected in the three CMS lines carrying the mitochondrial genome of Diploptaxis spp. The atpA upstream sequences of the CMS lines derived from Diploptaxis spp. differed completely from the euplasmic B. juncea line. Thus, the three CMS lines carrying mitochondria of Diploptaxis spp. were found to show a high degree of identity in the entire atpA region. On the other hand, except for a short stretch of a 650-bp upstream sequence, M. arvensis CMS line was found to be similar to euplasmic B. juncea in the atpA region. A novel orf108 detected upstream to the atpA coding sequence in M. arvensis CMS line (Ashutosh et al., 2008) was also found in CMS D. berthautii and D. erucoides. Although present in D. catholica, orf108 was interrupted by a stop codon and hence is designated as pseudo-orf108 (see Fig. 2).

Fig. 2. Comparison of mitochondrial genome sequences encompassing the mitochondrial atpA gene of euplasmic Brassica juncea and CMS Moricandia arvensis, Diploptaxis catholica, D. berthautii, and D. erucoides. Regions with identical or nearly identical sequences are represented by identical boxes. Numbers denote nucleotides (1 represents the 5′ nucleotide of the HindIII recognition sequence in all cases except D. erucoides). Closed triangles indicate the locations of primers used in this study.

atpA transcripts in CMS and fertility-restored plants

Previous studies based on northern blot analysis had revealed size differences in atpA transcripts among the three CMS lines carrying the Diploptaxis mitochondrial genome and their fertility-restored lines. Therefore, the atpA transcripts were cloned to identify the exact differences. Complete atpA mRNA sequences were amplified through RT-PCR after identification of transcription initiation and termination sites. The circularized RNA RT-PCR method (Kuhn and Binder, 2002) was used to identify the 5′ and 3′ termini of atpA.

CMS D. catholica gave a single 1050-bp amplicon in the TAP(+) circularized RNA RT-PCR reaction with the primer combination P3+P1 (see Fig. 2 for primer locations) whereas no amplicon was detected in TAP(–) treatment (Fig. 3A). This indicated that the atpA transcript in the male sterile line of D. catholica does not undergo any processing and the amplicon obtained in the TAP(+) reaction represents the primary transcript. In the male fertile line, the circularized RNA RT-PCR reaction gave similar size products in TAP(+) and TAP(–) treatments (Fig. 3A). These amplicons were cloned and sequenced. In the male sterile line, the G residue located 670 bp upstream to the atpA initiation codon was identified as the 5′ terminus. In the male fertile line, amplicons of both TAP(+) and TAP(–) reactions had identical sequences. The atpA transcript in the male fertile line was found to start at the T residue located 386 bp upstream to the atpA start codon. Identical 5′ ends in TAP(+) and TAP(–) treatments indicated that the atpA transcript is processed in the fertility-restored lines. Full-length atpA transcripts of both the male sterile (JN661709) and the fertility-restored (JN661710) lines were found to encode only atpA (Fig. 4).
In CMS *D. berthaultii*, two amplicons were obtained in circularized RNA RT-PCR with TAP treatment using the primer combination P3+P1. A 550-bp amplicon was found in both TAP(+) and TAP(–) treatments whereas a longer 1050-bp amplicon was present only in the TAP(+) treatment (Fig. 3B). In contrast, the amplification pattern was nearly identical in TAP(+) and TAP(–) treatments of the male fertile lines (Fig. 3B). The circularized RNA RT-PCR nearly identical in TAP(+) and TAP(–) treatments of the transcript extremities from male sterile and fertility-restored lines of *D. berthaultii* (B), LRT, linear RT-PCR; CRT(+), circularized RNA RT-PCR with tobacco acid phosphatase treatment; CRT(–), circularized RNA RT-PCR without tobacco acid phosphatase treatment; M, 100-bp DNA ladder (Fermentas).

The presence of *orf108* sequence upstream to the *atpA* coding region in the male sterile lines and its loss upon transcript processing in the R lines suggested a role for *orf108* in causing male sterility. To test whether *orf108* is directly responsible for male sterility, this study generated transgenic *A. thaliana* plants expressing *orf108* either constitutively or specifically in the anthers. The CaMV 35S promoter was used for constitutive expression and the *A. thaliana* peroxidase gene (*Atprx18*, *At2G24800*) promoter (Thakare et al., 2006; Kumar et al., 2011) for another specific expression of *orf108*. As mitochondrial transformation is not feasible in plants, to test the function of the mitochondrial gene *orf108* through nuclear transformation, a translational fusion of the *Neurospora crassa* *atp9* mitochondrial targeting presequence with the *orf108* coding sequence was made to ensure the delivery of the ORF108 protein to the mitochondria. *orf108* lacking the mitochondrial targeting presequence served as a control. Thus three plant binary vector cassettes, namely 35S::preatp9–*orf108*, Atprx18::preatp9–*orf108* and Atprx18::orf108 were constructed (Fig. 6A) and used to transform *A. thaliana* through *Agrobacterium*-mediated floral-dip transformation (Clough and Bent, 1998). Transgenic plants were identified by selection on kanamycin-containing medium.

Plants with 35S::preatp9–*orf108* cassette were indistinguishable from the wild type except for partial pollen sterility, which ranged from 10–30% in independent transgenic lines as evidenced by the FDA test. If *orf108* expression in anthers impairs pollen development through its effect on sporophytic tissues (premeiotic stage tissues or the tapetum), then complete
or high degree of pollen sterility is expected in these plants. Partial pollen sterility is often observed in transgenic lines of *A. thaliana* (Xing and Zachgo 2007; Ge et al., 2008). Since the CaMV 35S promoter is not very effective in driving gene expression in the anthers (Ge et al., 2008), pollen sterility observed in plants carrying 35S::preatp9–orf108 could not be conclusively ascribed to *orf108*. Further, when pollen from T1 plants were used to cross wild-type *A. thaliana* plants, kanamycin-positive progenies were recovered, which confirmed that the 35S::orf108 gene cassette was transmitted via

---

**Fig. 4.** Alignment of c-DNA sequences upstream to the *atpA* coding region of CMS *Diplotaxis catholica*, *D. berthaultii*, and *D. erucoides*. The arrow and bar indicate the start and stop codons, respectively, of *orf108*. A frameshift due to a two-nucleotide deletion in *D. catholica* (asterisk) abolishes *orf108*. The triangle shows the start of the *atpA* coding region. The complete coding sequence of *atpA* is not shown here.
pollen (Supplementary Fig. S1, available at JXB online). Floral dip transformation with Atprx18::orf108 or Atprx18::preatp9–orf108 constructs yielded very few transgenic plants. Only three kanamycin-positive plants were recovered with each of these constructs. All these six plants showed about 50% pollen sterility (Fig. 6B). As T1 plants are hemizygous for the transgene, this suggested that orf108 expression in pollen might cause male sterility. To verify this, T1 plants of Atprx18::preatp9–orf108 were selfed and also crossed as the pollen parent to wild-type A. thaliana (Col-0). In the F1, no kanamycin-positive plants were recovered, which clearly showed that pollen expressing orf108 were non-viable. Further, selfing of T1 plants resulted in very low seed set (Fig. 6C). When these seeds were selected for kanamycin resistance, no green plants were recovered in the T2 progenies which suggested that orf108 is not transmitted from the female side. However, when T2 seeds were plated on MS medium without kanamycin, green seedlings were obtained, indicating the viability of seeds (Supplementary Fig. S2). The promoter of Atprx18 has been shown to be active in anthers (Thakare et al., 2006; Kumar et al., 2011). But the current results suggest that it might also be active in the female side. These results clearly show that orf108 is capable of inducing male sterility and targeting the protein to the mitochondria is not necessary for inducing pollen sterility.

Survey of orf108 in wild allied species of Brassica

Since orf108 sequence was found in widely divergent species such as M. arvensis and three Diplotaxis species, its distribution was tested in 17 wild relatives of Brassica, namely Capsella bursa-pastoris, Brassica oleracea, Camelina sativa, Diplotaxis assurgentis, D. catholica, D. cretacea, D. erucoides, D. gomez-campoi, D. muralis, D. siettiana, D. siifolia, D. tenusilique, D. viminalis, Enarthocarpus lyratus, Erucastum canariense, Orychophragmus violaceus, and Sinapis alba. Primers P4 and P1 from the flanking region were used to amplify orf108 sequences from the wild species. Out of 17 species, a 550-bp amplicon encompassing orf108 was amplified in nine species namely B. oleracea.
D. assurgens, D. catholica, D. cretacea, D. erucoides, D. siettiana, D. tenuisilique, E. lytus, and S. alba (Fig. 7). Sequence analysis of the 550-bp amplicon revealed the presence of orf108 in four species, namely B. oxyrhina, D. erucoides, D. tenuisilique, and S. alba. In addition, an orf108-like sequence was also present in E. lytus. However, the ORF was disrupted by a frameshift mutation after 39 bp of the orf108 start codon and it also carried few dissimilar bases (Fig. 8). These results showed that the orf108 sequence is conserved and widely distributed among members of the Brassica caenop subspecies.

Discussion

Molecular analyses of CMS systems of various plant species have revealed that each system is unique with respect to the mitochondrial gene(s) or ORF(s) associated with male sterility and the mechanism of fertility restoration (Schnable and Wise, 1998). While more than one restorer gene, each operating in a different way, have been reported for a given CMS line (He et al., 1995; Wen et al., 2003; Wang et al., 2006), examples of a single restorer gene conferring male fertility to CMS lines of diverse origin are rare (Bhat et al., 2005, 2006, 2008; Jordan et al., 2011). Hence, fertility restoration of four alloplasmic CMS lines of B. juncea carrying mitochondrial genomes from M. arvensis, D. berthaultii, D. catholica, or D. erucoides by a B. juncea line carrying introgression from M. arvensis suggested two possibilities: (i) the restorer locus might contain multiple Rf genes each capable of conferring fertility to different CMS lines; and (ii) a common molecular mechanism might operate in these CMS lines. Since mitochondrial atpA transcripts were altered in all these four CMS lines, the second alternative was tested.

Comparison of the mitochondrial atpA gene region of the CMS lines with that of the euplasmic B. juncea line revealed a highly conserved organization of atpA gene region in all the CMS lines. However, a few differences on the 5’ region were found, which could be useful in identifying these cytoplasmics. Identical atpA coding and 3’ sequences in the CMS and euplasmic B. juncea lines suggested that atpA per se might not be the cause of CMS. A novel orf108, reported earlier in the M. arvensis CMS line, was the distinguishing feature of all the CMS lines from euplasmic B. juncea.

Analysis of atpA transcripts of CMS and euplasmic B. juncea lines revealed that the transcription of bicistronic atpA starts with a G or A residue in D. catholica and D. berthaultii, respectively. Most the A. thaliana mitochondrial gene transcripts have been shown to start with either A or G residues (Kuhn et al., 2005). Moreover, a well-defined conserved nonanucleotide motif type 2 (5’-CGTATAA-3’) motif of mitochondrial promoters sequences (Kuhn et al., 2005) was identified just upstream to the transcription start site of atpA in CMS D. berthaultii and D. catholica. In dicots, two conserved motifs of mitochondrial promoters have been characterized. The conserved nonanucleotide motif type 1 (5’-CRTAAGAGA-3’) and type 2 contain a well-conserved 5’-CRTA-3’ tetranucleotides (Binder and Brennicke, 1993; Binder et al., 1994). Thus, atpA transcription in male sterile lines of D. berthaultii and D. catholica seems to use the well-conserved transcription initiation mechanisms of dicots. In contrast, atpA transcription initiation in euplasmic B. juncea does not fit with transcription initiation found in other species. However, atpA transcripts of all CMS lines and the euplasmic B. juncea terminate at the same site (at the T residue 160 bp after atpA stop codon). Interestingly, the 3’ end of atpA mRNA of A. thaliana was found to terminate at the same site (Forner et al., 2007). Therefore, while the 5’ ends of atpA transcripts varied in different lines, the 3’ ends were conserved. Complete mitochondrial transcriptome mapping in A. thaliana also revealed that the 3’ ends of transcripts of different genes are identical (Forner et al., 2008). Besides, the three CMS lines showed nearly identical bicistronic atpA transcripts in the CMS lines and a monocistronic atpA in the fertility-restored lines.

The facts that atpA is bicistronic in CMS lines and is processed in a similar fashion in fertile lines strongly point to a common molecular basis of CMS and fertility restoration. Since the atpA transcripts in fertility-restored lines are processed in an almost identical manner, the restorer gene appears to be common. Bicistronic transcripts combining normal mitochondrial respiratory pathway genes and novel ORFs have been found to be associated with male sterility in a number of CMS systems such as pol (orf224–atp6; Singh and Brown, 1991) and nap (orf222–nad5c–orf139; Brown, 1999) CMS of B. napus and Chinsura Boro CMS of rice (orf79–atp6; Wang et al., 2006). Since the orf108 sequence in D. catholica was interrupted by the stop codon TAA after 81 bp from the ATG codon, it was not clear whether orf108 is directly responsible for CMS or does so indirectly by constraining efficient translation of atpA. Since mitochondrial translation machinery does not always use standard genetic code, this study could not infer whether orf108 is functional or a pseudo-allele in D. catholica. Therefore, the transgenic approach was used to test whether orf108 could induce male sterility.

A. thaliana plants constitutively expressing orf108 targeted to the mitochondria showed partial male sterility in a few plants. If orf108 causes male sterility through its action
on sporophytic tissues of anthers, one would expect nearly all pollen to be non-viable. Since the CaMV 35S promoter is not very effective in driving gene expression in floral organs (Ge et al., 2008) these results were inconclusive. When orf108 was expressed with the anther-specific Atprx18 promoter, which is active during both sporophytic and gametophytic stages of anther development (Kumar et al., 2011), 50% of the pollen was non-viable. If ORF108 is lethal to the gametes, one would expect 50% pollen sterility in T1 transgenics which would be hemizygous for the transgene. Interestingly, male sterility was found in plants carrying Atprx18::preatp9–orf108 or Atprx18::orf108 constructs, which clearly showed that targeting of ORF108 to mitochondria is not essential for causing pollen sterility. The current results are comparable to those of He et al. (1996) who reported male sterility in tobacco plants expressing Phaseolus vulgaris CMS-associated orf239. Similarly, in rice, expression of orf79 led to male sterility (Wang et al., 2006). In contrast, targeting of the CMS-associated orf456 protein to mitochondria was essential for inducing male sterility in chilli (Capsicum) (Kim et al., 2007). Thus the results reported here show that orf108 could be directly responsible for gametophytic male sterility in B. juncea CMS lines carrying M. arvensis, D. berthautii, or D. erucoides mitochondrial genomes.

Transgenic plants carrying Atprx18::preatp9–orf108 or Atprx18::orf108 constructs failed to yield kanamycin-resistant progeny. If orf108 expression causes only pollen sterility, one would expect 50% of the selfed progeny to inherit the selectable marker gene, nptII, via the female side. Hence, the present results could be explained if the Atprx18 promoter is active in ovules and ORF108 is also lethal on the female side. Transgenic expression of CMS-associated ORFs has not previously been reported to induce female sterility. Thus this study shows for the first time that CMS-associated ORFs could also cause female sterility. Low female fertility was found in a D. catholica-based CMS of B. juncea (Pathania et al., 2003). Although Atprx18 promoter shows strong and specific expression in developing anthers (Thakare et al., 2006; Kumar et al., 2011), the above results indicate it is active in the ovules. However, no seed sterility was observed in lines where orf108 was driven by the 35S promoter, which suggested that ORF108 might be deleterious only at the gametophytic stage. Ovules are the targets in floral-dip transformation in A. thaliana (Ye et al., 1999; Bechtold et al., 2000; Desfeux et al., 2000) and

Fig. 8. Sequence alignment of orf108 of Diplotaxis berthautii and Enarthocarpus lyratus. The arrow and bar indicate the start and stop codons of orf108, respectively. The E. lyratus sequence shows some nucleotide mismatches (asterisks) after the 39th nucleotide.
transformed ovules would perish if ORF108 is lethal at the gametophytic stage. This explains the low frequency of transformation observed with the Atprx18::orf108 and Atprx18::preatp9–orf108 constructs. Since petaloid anther phenotype was not found in transgenic Arabidopsis plants expressing orf108 (either constitutively or in anthers), the precise molecular basis of male sterility in B. juncea CMS line with D. catholica cytoplasm needs further investigation. As anther development is energy demanding (Lee and Warnke, 1979), any disturbance of oxidative phosphorylation pathway is said to lead to male sterility. Thus, inefficient translation of atpA from the bicistronic transcript containing pseudo-orf108 sequences might lead to petaloid anther development in the D. catholica-based CMS system of B. juncea. Alternatively, additional, as-yet undiscovered, mitochondrial dysfunction might be the cause of petaloidy in CMS D. catholica.

Among the CMS systems studied so far, the nap and pol CMS systems share high degree of similarity for the CMS-associated ORFs orf222 and orf224 and their fertility-restored lines are found to be allelic (Singh et al., 1996). Likewise, Kosena and Ogura CMS cytoplasms of radish have nearly identical ORFs (orf138 and orf125, respectively) and their restorers are identical (Brown et al., 2003; Desloire et al., 2003; Koizuka et al., 2003). However, these are examples of intraspecific variants of cytoplasm. In the current case, the cytoplasms are evolutionarily wide apart, yet they were found to share common male-sterility-inducing orf108.

This PCR-based survey of orf108 sequences among wild allies of Brassica indicated its wide distribution across genera. However, orf108 was not detected in all species of a genus. This could be due to mutations in the primer-binding sites or the absence of orf108 sequences. The distribution of orf108 among diverse wild allies of Brassica suggests that it arose quite early and is under dynamic evolution. These results are comparable to those of the common bean pvs–orf239, which has been found in several Phaseolus species and accessions but with high variation in copy number (Arrieta-Montiel et al., 2001). The presence of orf108 in several species also suggests that restorer genes could be present in different species.

Concerns of transgene spread to wild relatives and traditional varieties have spurred innovations to check the movement of transgenes (Daniell, 2002; Hills et al., 2007). Since expression of orf108 under the control of Atprx18 promoter leads to pollen and ovule sterility, it provides a unique opportunity for transgene containment. Transforming plants with Atprx18::orf108 linked to the gene of interest would confer the desired trait, but the transgenes would not be transmitted to the progeny. Such plants would nevertheless produce seeds as half of the gametes will be wild type and viable. This is particularly relevant to clonally propagated plants and tree species.

In conclusion, the present study clearly established a direct link between orf108 and CMS in three CMS lines of B. juncea carrying the mitochondrial genome of M. arvensis, D. berthautii, or D. erucoides. Also, as far as is known for the first time, these results show that CMS systems of diverse origin could have a common molecular basis. This study underscores the need for incisive molecular studies to establish the diversity of CMS systems based on functional differences rather than differences in mitochondrial genome organization, genetics of fertility restoration, or even phenotype. Only CMS lines that are functionally diverse would be meaningful for insuring against vulnerability of hybrids based on a few CMS sources to diseases or pests.

Supplementary material

Supplementary data are available at JXB online.

Supplementary Fig. S1. Detection of nptII through PCR in F1 A. thaliana plants of the cross 3S::preatp9–orf108 × wild type.

Supplementary Fig. S2. Testing the transmission of the orf108 gene cassette to T2 progenies.

Acknowledgements

The authors are grateful to Dr KV Prabhu, Manager, National Phytotron Facility, IARI, New Delhi for extending the Phytotron facility for work with Arabidopsis. PK is grateful to Council of Scientific and Industrial Research, Government of India for the award of a senior research fellowship.

References


Molecular mechanism of cytoplasmic male sterility in Brassica juncea
Brassica CMS restorer locus influence transcripts of three different mitochondrial gene regions. Genetics 143, 505–516.


