

Intraspecific variations of the cytoplasmic male sterility genes *orf108* and *orf117* in *Brassica maurorum* and *Moricandia arvensis*, and the specificity of the mRNA processing

Journal:	<i>Genome</i>
Manuscript ID	gen-2021-0011.R1
Manuscript Type:	Article
Date Submitted by the Author:	10-May-2021
Complete List of Authors:	Yamagishi, Hiroshi; Kyoto Sangyo University Hashimoto, Ayako; Kyoto Sangyo University Fukunaga, Asumi; Kyoto Sangyo Daigaku Bang, Sang Woo ; Utsunomiya University Terachi, Toru; Dept. Biotech., Fac. Eng., Kyoto Sangyo Univ.
Keyword:	Orf108 · Orf117 · Intraspecific variation · <i>Brassica maurorum</i> · <i>Moricandia arvensis</i>
Is the invited manuscript for consideration in a Special Issue? :	Not applicable (regular submission)

SCHOLARONE™
Manuscripts

Intraspecific variations of the cytoplasmic male sterility genes *orf108* and *orf117* in *Brassica maurorum* and *Moricandia arvensis*, and the specificity of the mRNA processing

Hiroshi Yamagishi^{1*} · Ayako Hashimoto² · Asumi Fukunaga¹ · Sang Woo Bang³ · Toru Terachi¹

1 Faculty of Life Sciences, Kyoto Sangyo University, Kamigamo, Kyoto 603-8555, Japan

2 Research Center of Botany, Kyoto Sangyo University, Kamigamo, Kyoto 603-8555, Japan

3 Faculty of Agriculture, Utsunomiya University, Utsunomiya 321-8505, Japan

* Corresponding author: hiyamagi@cc.kyoto-su.ac.jp

Abstract

The mitochondrial gene *orf108*, which is co-transcribed with *atp1* and causes cytoplasmic male sterility in *Brassica* crops, is widely distributed across wild species and genera of Brassicaceae. However, to date, intraspecific variations in the presence of *orf108* have not yet been studied, and the mechanisms for the wide distribution of the gene remain unclear. We analyzed the presence and sequence variations of *orf108* in two wild species, *Brassica maurorum* and *Moricandia arvensis*. After polymerase chain reaction amplification of the 5' region of *atp1* and the coding sequence of *orf108*, we determined the DNA sequences. *B. maurorum* and *M. arvensis* showed variations for the presence of *orf108* or *orf117* (*orf108^{V117}*) both between and within accessions, and were not fixed to the mitochondrial type having the male sterile genes. Sequencing of the amplicons clarified that *B. maurorum* has *orf108^{V117}* instead of *orf108*. Sequencing also indicated mitochondrial heteroplasmy in the two species; in particular, in *B. maurorum*, one plant possessed both the *orf108* and *orf108^{V117}* sequences. The results suggested that substoichiometric shifting of the mitochondrial genomes leads to the acquisition or loss of *orf108*. Furthermore, fertility restorer genes of the two species were involved in the processing of the mRNA of the male sterility genes at different sites.

Keywords: *Orf108*, *Orf117*, Intraspecific variation, *Brassica maurorum*, *Moricandia arvensis*

Draft

Introduction

Cytoplasmic male sterility (CMS) is a phenomenon in which plants fail to produce fertile pollen due to a mitochondrial gene. The phenotypic expression of CMS is suppressed by restorer of fertility (*Rf*) genes present in the nuclear genome. Thus, CMS is a useful tool for investigating the mechanisms of the interactions between nuclear and mitochondrial genes. CMS is also important practically for plant breeding. For F₁ breeding that employs hybrid vigor, the stable and efficient production of hybrid seeds is essential, and CMS enables such seed production.

Several types of CMS have been discovered in Brassicaceae plants, and they are used in the F₁ hybrid breeding of *Brassica* and *Raphanus* crops (Yamagishi and Bhat 2014). Among them, Ogura CMS, which was found in a Japanese radish (Ogura 1968) and is caused by the mitochondrial gene *orf138*, has been the most widely used. The mitochondrial gene *orf108*, on the other hand, induces CMS in alloplasmic *Brassica juncea* with the cytoplasm of wild species. *Orf108* is unique in that it is widely distributed across different species and genera of Brassicaceae (Kumar et al. 2012). Ashutosh et al. (2008) first found *orf108* in *Moricandia arvensis*. Thereafter, Kumar et al. (2012) surveyed the distribution of *orf108* in wild related species of *Brassica*, and detected *orf108* in *Brassica oxyrrhina*, *Diplotaxis assurgens*, *Diplotaxis berthautii*,

Diplotaxis catholica, *Diplotaxis cretacea*, *Diplotaxis erucoides*, *Diplotaxis siettiana*, *Diplotaxis tenuisiliqua*, *Enarthrocarpus lyratus*, and *Sinapis alba*. Among these species, *D. berthautii* provided an improved CMS line of *B. juncea* that showed high female fertility and better floral features (Bhat et al. 2008). However, there have been no studies to date on the intraspecific variations in the presence of *orf108* in these species.

In addition to the genus *Diplotaxis*, we previously examined the distribution of *orf108* in five species of *Sinapis*, including *S. alba* (Mukai et al. 2019). As hybrids are obtained by intergeneric crossing and embryo rescue among *S. alba*, *Sinapis arvensis*, *Sinapis turgida*, and *Brassica* crops, *Sinapis* species are useful genetic resources for the breeding of *Brassica* (Kaneko et al. 2009). We found that *S. arvensis* and *S. turgida* contain *orf108*, but five of the analyzed accessions of *S. alba* did not possess *orf108* (Mukai et al. 2019) even though *S. alba* has been reported to possess it (Kumar et al. 2012). Furthermore, in our preliminary experiment, one accession of *M. arvensis* did not possess *orf108*. These observations suggest that in addition to interspecific and intergeneric variations, there are intraspecific differences in the presence of *orf108* in these species. Thus, more detailed analyses within each species are needed to evaluate the usefulness of wild species as a source of CMS.

Orf108 in wild species is commonly found in the 5' region of the mitochondrial *atp1*

gene, and the two genes are co-transcribed (Naresh et al. 2016). Our sequencing of the complete mitochondrial genome of *Brassica maurorum* (GenBank: AP018931) revealed a novel gene, *orf117*, with sequences similar to *orf108* in the 5' region of *atp1*. In addition, it is known that the cytoplasm of *B. maurorum* induces CMS in alloplasmic radish (*Raphanus sativus*) (Bang et al. 2011).

To clarify the distribution of *orf108* in Brassicaceae species, and to infer the mechanisms underlying the acquisition or loss of *orf108* in each species, we examined the presence of *orf108* and the sequence variations in *orf108* in *M. arvensis* and *B. maurorum*.

Materials and methods

Plant materials

Seeds of *B. maurorum* (Ma) and *M. arvensis* (MOR-ARV) were obtained from Tohoku University *Brassica* Seed Bank (Sendai, Japan). After sowing the seeds, plants were grown in a glass house. The numbers of accessions and plants examined in this experiment for both species are shown in Table 1. DNA and RNA were isolated from leaves of an individual plant.

In addition, two lines each of CMS radish and CMS cabbage were grown. The CMS radishes ('(Ma) Kosena' and '(Ma) Riso') containing the cytoplasm of *B. maurorum* were bred by intergeneric hybridization between *B. maurorum* and radish cultivars, and successive repeated backcrosses to the radish cultivars. Similarly, the CMS cabbages ('(MOR-ARV) Hatsukoi' and '(MOR-ARV) Koiiji') were established by intergeneric hybridization between *M. arvensis* as a female parent and cabbage cultivars, and successive backcrosses to the cabbage cultivars. These CMS radishes and CMS cabbages were used for the isolation of RNA from leaves to assess the processing of the co-transcribed mRNA of *orf108-atp1*.

DNA and RNA isolation

Total DNA and RNA were isolated from young leaves using a DNeasy Plant Mini Kit (QIAGEN, CA, USA) and RNeasy Plant Mini Kit (QIAGEN), respectively. For determining the processing site of *orf117* mRNA, total RNA was isolated from the buds of *B. maurorum*.

Polymerase chain reaction (PCR) of the 5' region of *atp1*

Three primer sets were used for PCR. The first set included a forward primer (MF), which was designed for the region upstream of *orf108* and *atp1* based on the mitochondrial DNA sequence of *M. arvensis* (Naresh et al. 2016), and a reverse primer (R1), which was designed for the coding region of *atp1*. The second set included a forward primer (F2), which was designed for the sequence containing the start codon of *orf108* determined for *Brassica oxyrrhina* (Mukai et al. 2019), and R1; this primer set was used to examine whether *orf108* was present in the 5' region of *atp1*. The third set included a reverse primer (R2), which was designed for the 3' end of the coding sequence of *orf108*, and F2; this set was used to confirm whether each plant possessed *orf108*. The names, DNA sequences, and the positions of the primers are shown in Fig. 1.

Sequence determination for the *orf108-atp1* region

PCR with the MF/R1 primer set resulted in the amplification of DNA fragments for all of the plants studied. The sequences of the fragments were determined by cloning or direct sequencing. Cloning was performed according to the methods of Charoensakdi et al. (2007). For *B. maurorum*, the sequences of all 22 plants were examined. For *M. arvensis*, the numbers of plants for each accession used for sequencing were decided

based on the intra-accession differences in the results of PCR with the three primer sets.

Expression analysis of *orf108* and *orf117*

The mRNA processing of *orf108* and *orf117* was investigated using total RNA isolated from the young leaves of individual plants. In addition to alloplasmic lines of radish and cabbage, the RNA of *B. oxyrrhina* was used for comparisons, because *B. oxyrrhina* was reported to have *orf108* and the mRNA was processed at a novel site (Naresh et al. 2016). Using the RNA, northern hybridization was carried out with *atp1* as a probe. The hybridization procedure was based on the method of Naresh et al. (2016).

Detection of the processing site in the *orf117* mRNA of *B. maurorum*

Circular reverse transcription (RT)-PCR was conducted according to the methods described by Perrin et al. (2004) and Giancola et al. (2007). We used two *B. maurorum* plants (Ma-5-1, Ma-5-4) that had *orf117*, and isolated RNA from their buds. The isolated RNA was incubated with T4 RNA ligase (New England Biolabs, MA, USA), and cDNA was synthesized. The region containing the junction of the 5' and 3' extremities was amplified by RT-PCR. For circular RT-PCR, the following two primer pairs were used:

primer pair 1, forward (5'→3') = GGAATTGGCACAATATCGCG, reverse (5'→3') = CTAAACCCCCCGCCCGTTAAA; and primer pair 2, forward (5'→3') = CTGTGATCGAATGCCACTAGAC, reverse (5'→3') = CCAAGGTCATATCATCTCCAAG. The circular RT-PCR products amplified by either of the two primer pairs were cloned using the PGem®-T Easy vector system (Promega, Madison, WI, USA), and the sequences were determined.

Results

Presence of *orf108* in *B. maurorum* and *M. arvensis*

PCR with the MF/R1 primer pair amplified a DNA fragment of approximately 1 kbp in all individual plants of *B. maurorum* and *M. arvensis* (Table 2). Intraspecific differences in the possession of *orf108* were observed in *B. maurorum* (Table 2). Twelve plants of four accessions (Ma-1, -2, -5, -6) lacked *orf108*, whereas 10 plants of two accessions (Ma-3, -5) showed amplification by PCR with the F2/R1 and F2/R2 primer pairs (Table 2, Table 3), resulting in fragment sizes of approximately 400 bp and 320 bp respectively, as was expected. Among the five accessions, one (Ma-5) demonstrated intra-accession variation in the PCR amplification pattern with the F2/R1 primer set (Fig. 2). Similarly, although 17 out of 22 plants of *M. arvensis* showed the presence of *orf108* adjacent to *atp1*, five plants of three accessions (MOR-ARV-6, -13, -15) did not possess *orf108*

(Tables 2 and 4). Two accessions (MOR-ARV-13, -15) showed variable PCR patterns among individual plants within each accession. Fig. 3 shows the variations in PCR amplification among four plants within the MOR-ARV-15 accession.

Sequence variations in the 5' region of *atp1* in *B. maurorum*

As described above, *B. maurorum* showed differences both between and within accessions in the presence of *orf108*. Thus, we sequenced the amplicons obtained by PCR with the MF/R1 primer set for all 22 plants. All eight plants of Ma-3 and two plants of Ma-5 (Ma-5-1, Ma-5-4) produced the PCR amplicon with the F2/R2 primer pair (Table 3, Fig. 2). Sequencing demonstrated that all ten plants had identical sequences (3 in Fig. 4). They contained a G-to-A nucleotide substitution at the start codon of *orf108*. Thus, the coding sequence was elongated by 27 bp, resulting in a novel *orf117* (3 in Fig. 4), which we refer to as *orf108^{V117}*. Regardless of the substitution, PCR with F2/R1 and F2/R2 amplified the DNA fragment corresponding to *orf108* (Table 3, Fig. 2). Including the substitution, two plants of Ma-5 and eight plants of Ma-3 contained 27 nucleotide substitutions in the coding sequence of *orf108* when compared to *M. arvensis* (Table 3, Fig. 4). Among the 27 single nucleotide polymorphisms (SNPs), 21 were non-synonymous substitutions that induced the substitution of 15 amino acids. In

addition to the SNPs, the ten plants with *orf108^{V117}* all had a deletion of 3 bp and an insertion of 3 bp relative to *M. arvensis* (Table 3, Fig. 4). In contrast, 10 plants of three accessions (Ma-1, -2, -6) and two plants of one accession (Ma-5) lacked *orf108^{V117}* (Table 3). These 12 plants had the same DNA sequences, which were the same as those of *S. alba* and *Sinapis flexuosa* (Group 2) reported by Mukai et al. (2019), but different from that of *M. arvensis* reported by Naresh et al. (2016).

Cloning and sequencing demonstrated not only intraspecific variation, but also mitochondrial heteroplasmy in individual plants. All of the plants produced clones with one type of sequence, except for Ma-5-4; among the four obtained clones from Ma-5-4, one possessed *orf108^{V117}*, while the other three possessed *orf108* (4 in Fig. 4) and lacked the start codon of *orf108^{V117}*, similar to *M. arvensis* (Fig. 4). Furthermore, the clones had sequences that were identical to those of *M. arvensis* (1 in Fig. 4) in the coding and upstream regions of *orf108*, except for one nucleotide substitution that was 11 bp upstream of the start codon (Fig. 4). The plants contained two sequence types that were considered to be due to mitochondrial heteroplasmy: the major type was *orf108* resembling that of *M. arvensis*, and the other was *orf108^{V117}*, which is unique to *B. maurorum*.

Sequence variations in the 5' region of *atp1* in *M. arvensis*

PCR with the MF/R1 primer pair amplified DNA in all 22 plants of ten accessions of *M. arvensis* (Table 2). However, PCR using F2/R1 and F2/R2 did not amplify DNA in five plants belonging to three accessions, demonstrating that the plants did not possess *orf108* (Table 2). For these five plants (Table 4), we sequenced the DNA fragments amplified by PCR using MF/R1. Even though the MF primer was designed based on the sequence of *M. arvensis* (Naresh et al. 2016), they all had the same sequences as the *B. maurorum* plants lacking *orf108*^{V117}, which have sequences identical to those of *S. alba* belonging to Group 2 of Mukai et al. (2019).

PCR using F2/R1 and F2/R2 amplified DNA fragments in all of the other 17 plants of nine accessions (Table 2). The sequences of at least one plant among the nine accessions were determined. The sequences of the 12 sequenced plants exhibited two patterns: one pattern had only the sequence containing *orf108* (+ (Homo) in Table 4), and the other pattern had two kinds of sequences, namely with and without *orf108*, in a single plant (+ (Hetero) in Table 4). Six plants of six accessions exhibited the first pattern, and six plants of three accessions exhibited the second pattern (Table 4). It was suggested that the plants exhibiting the first pattern exclusively possessed the mitochondrial genome containing *orf108*, whereas the plants exhibiting the second pattern possessed the two

types of mitochondrial genomes in a substoichiometric ratio.

Among the six plants exhibiting the first pattern (+ (Homo) in Table 4), four had an *orf108* sequence identical to that of *M. arvensis* that was obtained by Naresh et al. (2016) (3-2, 4-5, 5-1, 10-1 in Table 4), and two had different sequences (9-1 and 15-2 in Table 4). In contrast, among the six plants exhibiting heteroplasmic mitochondrial genomes (+ (Hetero) in Table 4), only two (8-4, 12-1) possessed the same *orf108* as *M. arvensis* (Naresh et al. 2016), and the other four (8-1, 8-2, 8-3, 13-1) possessed a deletion and SNPs that induced a stop codon in *orf108* (Table 4, Fig. 4).

Processing of *orf108* and *orf108^{V117}* mRNA in *B. maurorum* and *M. arvensis*

For the northern hybridization of the CMS genes of *B. maurorum*, we used two CMS lines of radish as controls. Total RNA was isolated from the leaves. The CMS lines (Lanes 7, 8 in Fig. 5) produced longer *atp1* transcripts than *B. maurorum* plants (Lanes 1 to 4), suggesting that *orf108^{V117}* was co-transcribed with *atp1* in the male sterile plants (Fig. 5). The *B. maurorum* plants possessing *orf108^{V117}* demonstrated normal pollen fertility due to their own *Rf* gene(s). Both plants with *orf108^{V117}* (Ma-5-1 (Lane 1) and Ma-5-4 (Lane 4) in Fig. 5) possessed transcripts with a length similar to that in *B. oxyrrhina* (Lane 5 in Fig. 5), and they were shorter than the transcript of *M. arvensis*

(Lane 6 in Fig. 5). This suggested that the bicistronic transcript of *orf108^{V117}-atp1* in *B. maurorum* was cleaved at a different site from that in *M. arvensis*. In addition, *atp1* mRNA of the *B. maurorum* plants without *orf108^{V117}* (Lanes 2, 3 in Fig. 5) was unexpectedly processed to the same size (Fig. 5). The reason for this coincidence is unclear, but Naresh et al. (2016) also observed that the transcripts from cytoplasms with and without *orf108* had similar sizes when their blot was hybridized with an *atp1* probe.

In contrast, the mRNA of *M. arvensis* that possesses *orf108* but produces normal pollen (Lane 6 in Fig. 5 and Lane 1 in Fig. 6) was longer than the mRNA of *B. maurorum* and *B. oxyrrhina*. This difference suggested that there are variations in the mRNA processing sites due to the products of the *Rf* genes between *M. arvensis* and the other two species. As shown in Fig. 6, the transcript of the plant that had *orf108* heteroplasmically (MOR-ARV-8-2; Lane 2 in Table 4) showed processing to a similar size as that of *B. oxyrrhina* (Lane 3) when probed with *atp1*. Furthermore, the *atp1* transcripts produced from the leaves of CMS cabbages with the cytoplasm of *M. arvensis* (Lanes 4, 5) had the same length as that of *M. arvensis* (Lane 1 in Fig. 6). These results demonstrated that *orf108* was not expressed in the leaves of these CMS cabbages, corroborating the results of Gaikwad et al. (2006).

Processing sites in the *orf108^{V117}* mRNA of *B. maurorum*

To determine the processing sites, we conducted circular RT-PCR for the two plants (Ma-5-1 and Ma-5-4) possessing *orf108^{V117}*. Seven and eight clones were obtained from Ma-5-1 and Ma-5-4, respectively. The 5' end of the *orf108^{V117}* mRNA was placed mainly in two regions. The first region was at the 168th to 175th position from the start of *orf108^{V117}*, which is about 50 bp upstream of the site where the CMS gene is processed in *B. oxyrrhina* (Table 5). The other region was at the 322nd to 324th position from the start, which is about 100 bp downstream of the 5' end observed in *B. oxyrrhina* (Naresh et al. 2016) (Table 5). Both regions were unique to *B. maurorum* as they differed from those of *M. arvensis* and *B. oxyrrhina* (Naresh et al. 2016).

Discussion

Distribution of *orf108*

Although Kumar et al. (2012) observed a wide distribution of *orf108* across species and genera, they did not examine intraspecific variations in each species. The present study is the first to report sequence differentiation at the 5' region of mitochondrial *atp1*, which is co-transcribed with *orf108* in species containing CMS genes. Plants lacking

orf108 were observed in *B. maurorum* and *M. arvensis*, which have been used as breeding materials of *Brassica* and *Raphanus* crops, inducing CMS (Bang et al. 2011, Prakash et al. 1998). However, the results of this study suggested that the cytoplasm of these species is not fixed to the type causing CMS. It is therefore necessary to examine whether each plant of the species possesses the CMS gene before using it as breeding material.

As previously described, the mechanisms for the wide distribution of *orf108* are unclear. Concerning this point, it is of interest that several *M. arvensis* plants heteroplasmically possessed both sequences that contain and lack *orf108* (Table 4), although some plants possessing only the sequence with or without *orf108* were also observed. We also demonstrated that mitochondria with either *orf108* or *orf108*^{V117} were present in a heteroplasmic condition in one *B. maurorum* plant. To our knowledge, this is the first observation of the heteroplasmic coexistence of variants of CMS genes in one plant. Mitochondrial heteroplasmy is defined as the coexistence of divergent mitochondrial genotypes in a cell, and is currently considered to be common in plants (Woloszynska and Trojanowski 2009). The inclusion of different mitochondrial structures within one species suggests the presence of nuclear factors that control substoichiometric shifting and determine the predominant mitochondrial DNA

configuration in each plant. Examples of nuclear genes that control the shifting of mitochondrial genomes include *Fr* in common bean (Mackenzie and Chase 1990), *Psm* in cucumber (Havey et al. 2004), and *CHM* (Abdelnoor et al. 2003), *OSB1* (Zaegel et al. 2006), and *MSH1* and *RECA3* (Shedge et al. 2007) in *Arabidopsis thaliana*. To clarify the genetic mechanisms for the wide distribution of *orf108* in wild Brassicaceae species, it is necessary to identify such nuclear genes that influence substoichiometric shifting.

B. maurorum and *M. arvensis* may be good materials for such studies, because some of the plants exclusively possess one type of mitochondria, while some other plants heteroplasmically possess two types. Observation of the alterations or the persistence of mitochondrial structures over several generations in these two species would provide interesting information on the distribution and variations of *orf108* in wild species and genera.

Sequence variations of *orf108*

We found that *B. maurorum* contains intraspecific variations in the mitochondrial genomes. More than half of the plants lacked *orf108* or *orf108*^{V117}, and ten plants of two accessions possessed *orf108*^{V117} (Tables 2 and 3). The *orf108*^{V117} lacks the start codon of *orf108* that is present in *M. arvensis* and *B. oxyrrhina* (Fig. 4). Although *B. oxyrrhina*

was reported to possess *orf108* (Naresh et al. 2016), it has a start codon 27 bp upstream from the start of *orf108* (Fig. 4). If the upstream codon is the actual start codon, the CMS gene of *B. oxyrrhina* would be *orf108^{V117}*. In contrast, the sequence of the clone obtained from the *B. maurorum* plant with *orf108* (Ma-5-4) does not have this ATG codon (Fig. 4). On the other hand, the *B. maurorum* plants that lost the start codon of *orf108* had ATG codons at three amino acids upstream (*orf108^{V111}*) and nine amino acids upstream (*orf108^{V117}*) (Fig. 4). This type of sequence of *B. maurorum* with *orf108^{V117}* contained 27 nucleotide substitutions relative to *M. arvensis* that led to 15 amino acid substitutions. Despite the substitutions, the cytoplasm of *B. maurorum* induced male sterility in the nuclear background of the radish (Bang et al. 2011).

Among the 27 nucleotide substitutions, two were at positions where the product of the *Rf* gene of *B. oxyrrhina* processes the *orf108* mRNA (GG versus AC (Table 6, Fig. 4)) (Naresh et al. 2016). In addition to *M. arvensis*, all species having *orf108* possessed the nucleotides AC here (Mukai et al. 2019), and this sequence (GG) is unique to *B. maurorum*. On the other hand, all plants of *M. arvensis* and *B. maurorum* with *orf108* or *orf108^{V117}* had the same nucleotide (T) at the site (the 16th nucleotide) where the *Rf* gene product of *M. arvensis* processes the mRNA of *orf108* (Naresh et al. 2016) (Table 6, Fig. 4). However, northern hybridization analysis demonstrated that the *Rf* gene

product of *B. maurorum* does not cleave *orf108^{V117}* mRNA at this point (Fig. 5). The precise determination of the processing sites in *orf108^{V117}* mRNA of *B. maurorum* demonstrated that the mRNA was cleaved at two sites that differed from those of *M. arvensis* and *B. oxyrrhina* (Table 6).

Naresh et al. (2016) found that a single nucleotide substitution within the *orf108* coding region (T to G in Table 6) rendered the *Rf* gene of *M. arvensis* ineffective in *B. oxyrrhina*-based CMS. In the present study, it was observed that the *orf108^{V117}* of *B. maurorum*, which had nucleotide substitutions at the cleavage site by the *Rf* gene of *B. oxyrrhina*, was processed at sites that differed from that of *B. oxyrrhina* (Table 6). These data suggest that nucleotide substitutions at the cleavage site of CMS genes inhibit the suppression induced by the specific *Rf* gene. We recently found a similar relationship between the CMS gene (*orf138*) and the *Rf* genes in Ogura CMS of radish (Yamagishi et al. 2021). ORF687 encoded by *Rfo* binds to *orf138* mRNA, but a single nucleotide substitution of *orf138* in the binding site diminished the effect of the *Rfo* gene by drastically decreasing the binding affinity of ORF687, and the male sterility phenotype was expressed despite the presence of *Rfo*. The observations from the present study and in Ogura CMS suggest that mutations in the coding region of the CMS gene, where the specific *Rf* gene product binds, lower the binding affinity of the *Rf* gene

product and thereby maintain male sterility. However, although ORF687 in Ogura CMS has been identified as a member of the pentatricopeptide repeat protein family and characterized in detail, the *Rf* genes of *M. arvensis*, *B. oxyrrhina*, and *B. maurorum* have not yet been cloned. It is of interest to identify the *Rf* genes to understand the evolutionary relationship between CMS genes and *Rf* genes in these species, as in the case of CMS of wheat (Melonek et al. 2021). Furthermore, also from a practical viewpoint, it is of interest to clarify the association between CMS genes and the functions of *Rf* genes in more detail for each wild species as they may be useful as materials of F₁ breeding for *Brassica* and *Raphanus* crops.

Conflict of interest: The authors declare that they have no conflict of interest.

Author contributions: H. Y. and T. T. planned the experiments. H. Y., A. H., A. F., and S. W. B. carried out the experiments. H. Y. wrote the manuscript.

Funding: This work was supported by the fund of the Research Center of Botany, Kyoto Sangyo University (Grant No. F2001), and by a JSPS Grant-in-Aid for Scientific Research (19K05979 to H.Y.).

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
- Ashutosh, Kumar P, Kumar VD, Sharma PC, Prakash S, Bhat SR (2008) A novel *orf108* co-transcribed with the *atpA* gene is associated with cytoplasmic male sterility in *Brassica juncea* carrying *Moricandia arvensis* cytoplasm. *Plant Cell Physiol* 49: 284-289
- Bang SW, Tsutsui K, Shim S, Kaneko Y (2011) Production and characterization of the novel CMS line of radish (*Raphanus sativus*) carrying *Brassica maurorum* cytoplasm. *Plant Breed* 130: 410-412
- Bhat SR, Kumar P, Prakash S (2008) An improved cytoplasmic male sterile (*Diplotaxis berthautii*) *Brassica juncea*: identification of restorer and molecular characterization. *Euphytica* 159: 145-152
- Charoensakdi R, Murakami S, Aoki K, Rimphanitchayakit V, Limpaseni T (2007) Cloning and expression of cyclodextrin glycosyltransferase gene from *Paenibacillus* sp. T16 isolated from hot spring soil in northern Thailand. *Biochem Mol Biol* 40: 333-340

Gaikwad K, Baldev A, Kirti PB, Mohapatra T, Bhat SR, Chopra VL, Prakash S (2006)

Organization and expression of the mitochondrial genome in CMS (*Moricandia*)

Brassica juncea: nuclear-mitochondrial incompatibility results in differential expression of the mitochondrial *atpa* gene. Plant Breed 125: 623-628

Giancola S, Rao Y, Chaillou, S, Hiard S, Martin-Canadell A, Pelletier G, Budar F (2007)

Cytoplasmic suppression of Ogura cytoplasmic male sterility in European natural populations of *Raphanus raphanistrum*. Theor Appl Genet 114: 1333-1343

Havey MJ, Park YH, Bartoszewski G (2004) The *Psm* locus controls paternal sorting of the cucumber mitochondrial genome. J Hered 95:492-497

Kaneko Y, Bang SW, Matsuzawa Y (2009) Distant hybridization. In Gupta SK (ed): Biology and Breeding of Crucifers. CRC Press (Boca Raton): 207-248

Kumar P, Vasupalli N, Srinivasan R, Bhat SR (2012) 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*. J Exp Bot 63: 2921-2932

Mackenzie SA, Chase CD (1990) Fertility restoration is associated with loss of a portion of the mitochondrial genome in cytoplasmic male-sterile common bean. Plant Cell 2:905-912

Melonek J et al. (2021) The genetic basis of cytoplasmic male sterility and fertility restoration in wheat. *Nature Com* 12: 1036

Mukai A, Jikuya M, Tsujimura M, Terachi T, Yamagishi H (2019) Complete mitochondrial genome sequence of *Brassica oxyrrhina* and comparative analysis of the region containing *orf108*, a male sterility gene for mustard (*Brassica juncea*), among *B. oxyrrhina*, *Diplotaxis erucoides* and *Sinapis* species. *Plant Breed* 138: 614-623

Naresh V, Singh SK, Watts A, Kumar P, Kumar V, Rao KRSS, Bhat SR (2016) Mutations in the mitochondrial *orf108* render *Moricandia arvensis* restorer ineffective in restoring male fertility to *Brassica oxyrrhina* based cytoplasmic male sterile line of *B. juncea*. *Mol Breed* 36: 67-75

Ogura H (1968) Studies on the new male-sterility in Japanese radish, with special reference to the utilization of this sterility towards the practical raising of hybrid seeds. *Mem Fac Agri Kagoshima Univ* 6: 39-78

Perrin R, Meyer EH, Zaepfel M, Kim Y-J, Mache R, Grienberger J-M, Gualberto M, Gagliardi D (2004) Two exoribonucleases act sequentially to process mature 3'-ends of *atp9* mRNAs in *Arabidopsis* mitochondria. *J Biol Chem* 279: 25440-25446

Prakash S, Kirti PB, Bhat SR, Gaikwad K, Kumar VD, Chopra VL (1998) A *Moricandia*

arvensis-based cytoplasmic male sterility and fertility restoration system in *Brassica juncea*. Theor Appl Genet 97: 488-492

Shedge V, Arrieta-Montiel M, Christensen AC, Mackenzie SA (2007) Plant mitochondrial recombination surveillance requires unusual *RecA* and *MutS* homologs. Plant Cell 19: 1251-1264

Woloszynska M, Trojanowski D (2009) Counting mtDNA molecules in *Phaseolus vulgaris* sublimons are constantly produced by recombination via short repeats and undergo rigorous selection during substoichiometric shifting. Plant Mol Biol 70: 511-522.

Yamagishi H, Bhat SR (2014) Cytoplasmic male sterility in Brassicaceae crops. Breed Sci 64: 38-47

Yamagishi H, Jikuya M, Okushiro K, Hashimoto A, Fukunaga A, Takenaka M, Terachi T (2021) A single nucleotide substitution in the coding region of Ogura male sterile gene, *orf138*, determines effectiveness of a fertility restorer gene, *Rfo*, in radish. Mol Genet and Genom 296:705-717

Zaegel V, Guermann B, Le Ret M, Andres C, Meyer D, Erhardt M, Canaday J, Gualberto JM, Imbault P (2006) The plant-specific ssDNA binding protein OSB1 is involved in the stoichiometric transmission of mitochondrial DNA in *Arabidopsis*.

Plant Cell 18: 3548-3563

Draft

Table 1 Species used in this experiment

Species (Code)	Number of accessions	Number of plants
<i>Brassica maurorum</i> (Ma)	5	22
<i>Moricandia arvensis</i> (MOR-ARV)	10	22

Draft

Table 2 DNA amplification by PCR with primer pairs designed to amplify the 5' region of mitochondrial *atp1*.

Species (Code)	PCR amplification ^a			Number of accessions	Number of plants
	MF/R1	F2/R1	F2/R2		
<i>Brassica maurorum</i> (Ma)	+	+	+	2	10
	+	-	-	4	12
<i>Moricandia arvensis</i> (MOR-ARV)	+	+	+	9	17
	+	-	-	3	5

a +; a DNA fragment was amplified by PCR. -; no DNA fragment was amplified by PCR.

Table 3 Presence of mitochondrial *orf108^{V117}* in *B. maurorum* and sequence variations in comparison to the *orf108* coding region of *M. arvensis* (Naresh et al. 2016).

Accession	Number of plants	Presence of <i>orf108^{V117a}</i>	Sequence variations in comparison to the <i>orf108</i> of <i>M. arvensis</i> (Naresh et al. 2016)
Ma-1	4	-	
Ma-2	2	-	
Ma-3	8	+	ins (3 bp), del (3 bp), SNP (27)
Ma-5	2	+	ins (3 bp), del(3 bp), SNP (27)
Ma-5	2	-	
Ma-6	4	-	

a + and -: *orf108^{V117}* is present (+) or absent (-).
ins and del: insertions and deletions, respectively, of nucleotides with the numbers shown in parentheses.
SNP: single nucleotide polymorphisms with the numbers shown in parentheses.

Table 4 Sequence variations of mitochondrial *orf108* in *M. arvensis* in comparison to the sequence determined by Naresh et al. (2016).

Accession	Plant No.	Presence of <i>orf108</i> ^a	No. of clones ^b	Sequence variations in comparison to the <i>orf108</i> of <i>M. arvensis</i> ^c	No. in Fig. 4	Effect of mutation
MOR- ARV-3	3-2	+ (Homo)	2	-	5	
MOR- ARV-4	4-5	+ (Homo)	3	-	5	
MOR- ARV-5	5-1	+ (Homo)	3	-	5	
MOR- ARV-6	6-1	- (Homo)	1, D			
MOR- ARV-8	8-1	+ (Hetero)	D	del (2 bp), SNP(11)	6	stop
	8-2	+ (Hetero)	1, D	del(1 bp), SNP(4)	7	stop
	8-3	+ (Hetero)	6	del (2 bp), SNP(11)	8	stop
	8-4	+ (Hetero)	2	-	5	
MOR- ARV-9	9-1	+ (Homo)	1, D	SNP(1)	9	K→N
MOR- ARV-10	10-1	+ (Homo)	2	-	5	
MOR- ARV-12	12-1	+ (Hetero)	4	-	5	
MOR- ARV-13	13-1	+ (Hetero)	2	del (2 bp), SNP(12)	10	stop
	13-2	- (Homo)	1, D			
MOR- ARV-15	15-1	- (Homo)	1, D			
	15-2	+ (Homo)	1, D	del (2 bp), SNP(12)	11	stop
	15-3	- (Homo)	2			
	15-4	- (Homo)	2			

a + (Homo) and - (Homo): DNA sequences containing *orf108* (+ (Homo)) and lacking *orf108* (- (Homo)) were exclusively observed. + (Hetero): Both the sequences with and without *orf108* were observed.

b D: The sequence was determined by direct sequencing.

c -: identical to the *orf108* of *M. arvensis* (Naresh et al. 2016). del: deletions of nucleotides with the numbers shown in parentheses. SNP: single nucleotide polymorphisms with the numbers shown in parentheses.

Table 5 Position of the 5' end of the *orf108^{V117}* mRNA produced by the processing in *B. maurorum*.

	Position from the start of <i>orf108^{V117}</i>					
	168 to 175	187	219 ^a	316	322 to 324	333
Number of clones	5	1	0	1	6	2

^a The 5' end of mRNA of the CMS gene of *B. oxyrrhina* counted from the start of *orf108^{V117}* according to Naresh et al. (2016)

Draft

Table 6 Processing site of the *orf108^{V117}* mRNA of *B. maurorum* in comparison to *M. arvensis* and *B. oxyrrhina*^a

Species	Nucleotides and cleavage position in <i>orf108^{V117}</i>			
	43rd ^c	168th to 175th	218th/219th	322nd to 324th
<i>M. arvensis</i>	▼ T	CGCCTTTT	AC	ATT
<i>B. oxyrrhina</i>	G	CACCTTTT	▼ AC	ATT
<i>B. maurorum</i>	T	(2) ▼ (2) ▼ (1) ▼ CAACTTTT	GG	(1) ▼ (1) ▼ (4) ▼ ATT

^a Although *M. arvensis* possesses *orf108* instead of *orf108^{V117}*, the nucleotide positions were shown from the start of *orf108^{V117}* for the comparisons with the two other species.

^b Black arrowheads indicate the processing site in *M. arvensis* and *B. oxyrrhina* (Naresh et al. 2016), while white arrowheads show the processing sites in *B. maurorum*, and numbers in parentheses indicate the clone numbers.

^c The mRNA of *M. arvensis* is actually cleaved at the 16th position of *orf108*.

Figure legends

Fig. 1 Positions and sequences of the primers designed to amplify the 5' region of *atp1*.

Fig. 2 Intra-accession variations in DNA amplification by PCR with the F2/R1 primer pair in *Brassica maurorum* (Ma-5). M: Molecular size marker (ϕ x174/*Hae*III).

Fig. 3 Intra-accession variations in DNA amplification by PCR with the F2/R1 primer pair in *Moricandia arvensis* (MOR-ARV-15). M: Molecular size marker (ϕ x174/*Hae*III).

Fig. 4 DNA sequence of the mitochondrial *orf117* of *B. maurorum* in comparison to the *orf108* of *M. arvensis* and *B. oxyrrhina*.

1: *M. arvensis* (Narsh et al. 2016), 2: *B. oxyrrhina* (Narsh et al. 2016), 3: *B. maurorum* (AP018931 and this article), 4: Substoichiometric clones obtained from a plant of *B. maurorum* (Ma-5-4). 5: MOR-ARV-3-2 etc., 6: MOR-ARV-8-1, 7: MOR-ARV-8-2, 8: MOR-ARV-8-3, 9: MOR-ARV-9-1, 10: MOR-ARV-13-1, 11: MOR-ARV-15-2.

Upward arrowhead indicates the site where the *Rf* gene of *M. arvensis* functions for the processing of *orf108* mRNA.

Downward arrowhead indicates the site where the *Rf* gene of *B. oxyrrhina* functions for the processing of *orf108* mRNA.

Underlines of Nos. 6, 7, 8, 10, and 11 show the stop codon.

Fig. 5 Northern hybridizations of the total RNA of *B. maurorum*, *B. oxyrrhina*, *M. arvensis*, and CMS radishes with the cytoplasm of *B. maurorum* hybridized with the mitochondrial gene probe of *atp1*.

1: Ma-5-1, 2: Ma-5-2, 3: Ma-5-3, 4: Ma-5-4, 5: *B. oxyrrhina*, 6: MOR-ARV-3-2, 7: CMS radish, '(Ma) Kosena', 8: CMS radish, '(Ma) Riso'

orf108/orf108^{V117}: + = present, - = absent

Pollen fertility: F = fertile, S = sterile

Fig. 6 Northern hybridizations of the total RNA of *M. arvensis*, *B. oxyrrhina*, and CMS cabbages with the cytoplasm of *M. arvensis* hybridized with the mitochondrial gene probe of *atp1*.

1: MOR-ARV-3-2, 2: MOR-ARV-8-2, 3: *B. oxyrrhina*, 4: CMS cabbage, '(MOR-ARV) Hatsukoi', 5: CMS cabbage, '(MOR-ARV) Koji'

orf108: + = present, - = absent

Pollen fertility: F = fertile, S = sterile

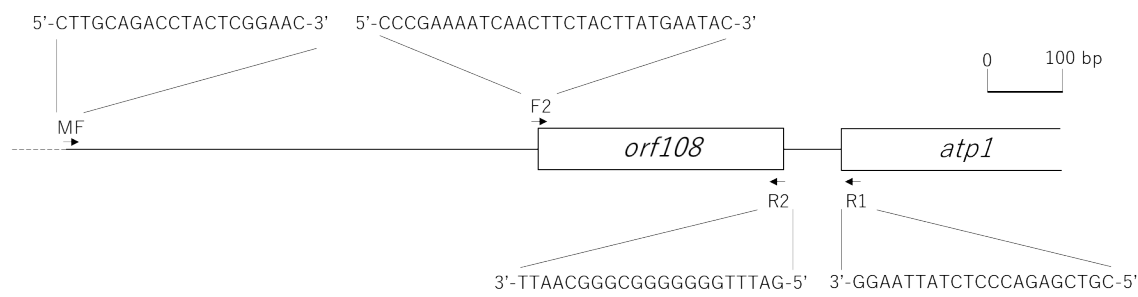


Fig. 1 Positions and sequences of the primers designed to amplify the 5' region of *atp1*.

Draft

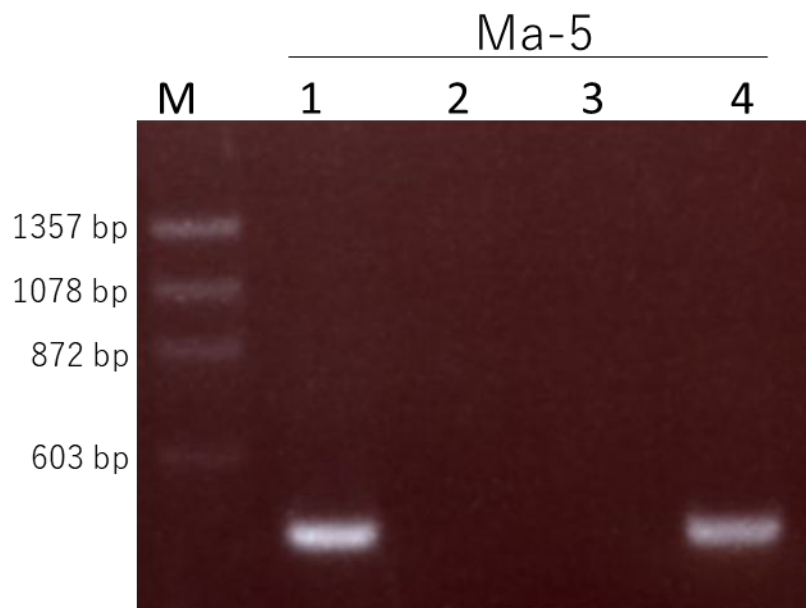


Fig. 2 Intra-accession variations in DNA amplification by PCR with the F2/R1 primer pair in *Brassica maurorum* (Ma-5).

M: Molecular size marker (ϕ x174/*Hae* III).

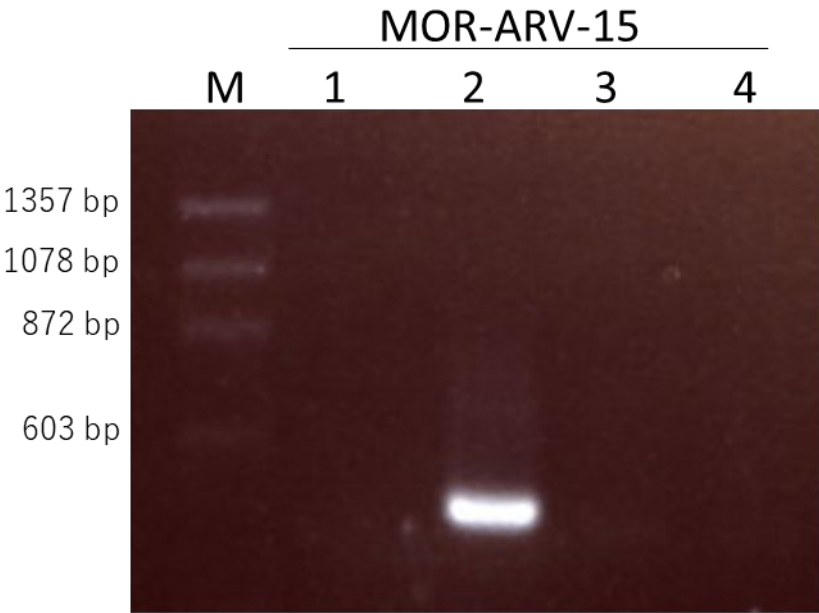


Fig. 3 Intra-accession variations in DNA amplification by PCR with the F2/R1 primer pair in *Moricandia arvensis* (MOR-ARV-15).

M: Molecular size marker (ϕ x174/*Hae* III).

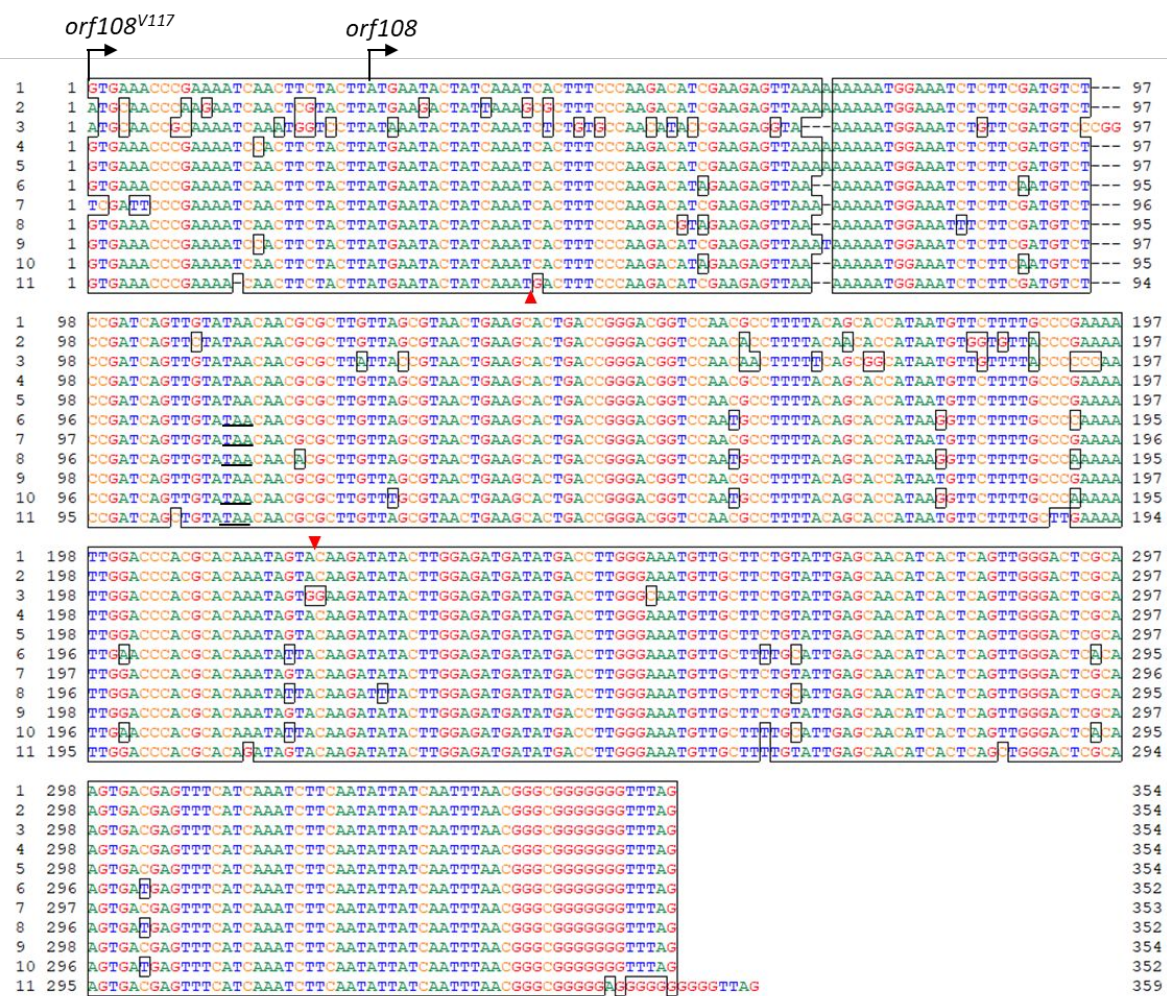


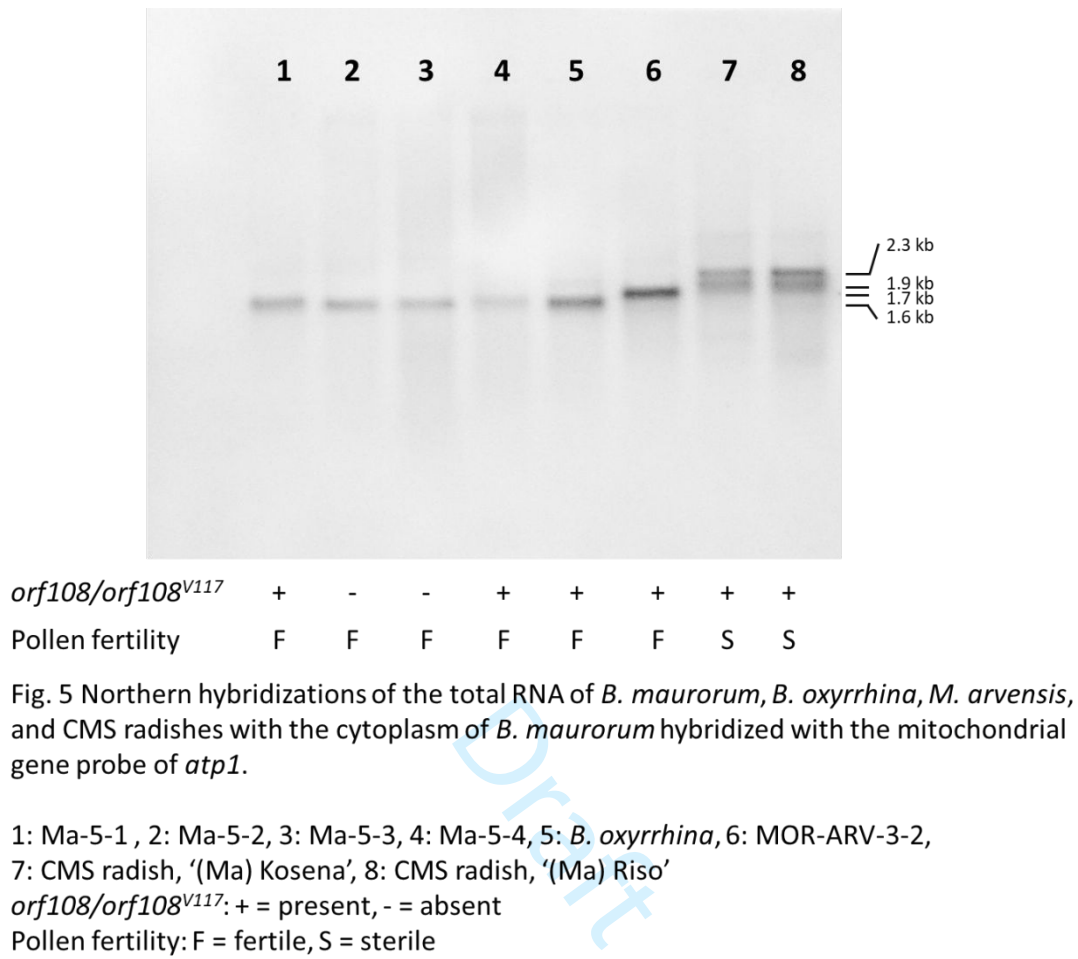
Fig. 4 DNA sequence of the mitochondrial *orf117* of *B. maurorum* in comparison to the *orf108* of *M. arvensis* and *B. oxyrrhina*.

1: *M. arvensis* (Narsh et al. 2016), 2: *B. oxyrrhina* (Narsh et al. 2016), 3: *B. maurorum* (AP018931 and this article), 4: Substoichiometric clones obtained from a plant of *B. maurorum* (Ma-5-4). 5: MOR-ARV-3-2 etc., 6: MOR-ARV-8-1, 7: MOR-ARV-8-2, 8: MOR-ARV-8-3, 9: MOR-ARV-9-1, 10: MOR-ARV-13-1, 11: MOR-ARV-15-2.

Upward arrowhead indicates the site where the *Rf* gene of *M. arvensis* functions for the processing of *orf108* mRNA.

Downward arrowhead indicates the site where the *Rf* gene of *B. oxyrrhina* functions for the processing of *orf108* mRNA.

Underlines of Nos. 6, 7, 8, 10, and 11 show the stop codon.



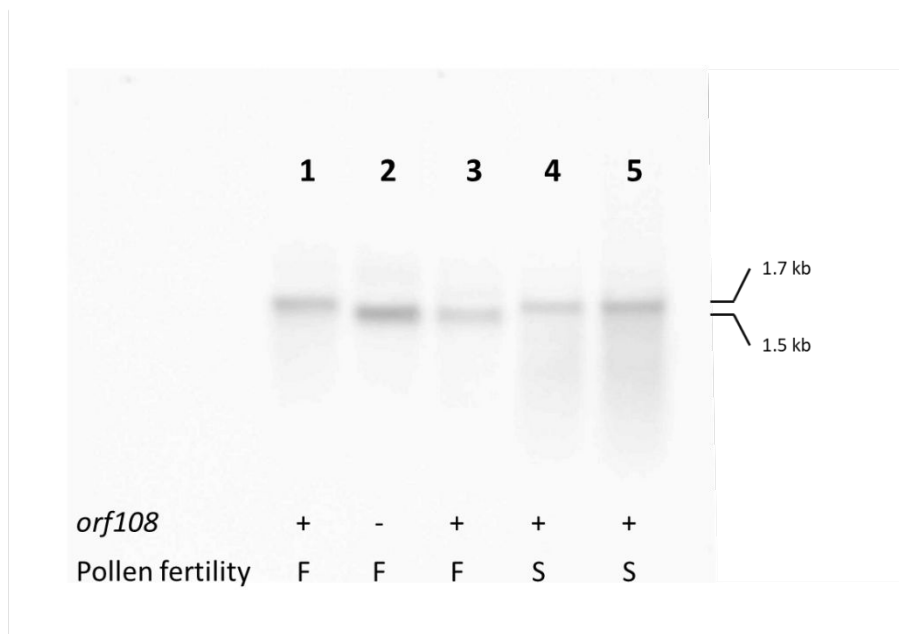


Fig. 6 Northern hybridizations of the total RNA of *M. arvensis*, *B. oxyrrhina*, and CMS cabbages with the cytoplasm of *M. arvensis* hybridized with the mitochondrial gene probe of *atp1*.

1: MOR-ARV-3-2, 2: MOR-ARV-8-2, 3: *B. oxyrrhina*, 4: CMS cabbage, '(MOR-ARV) Hatsukoi',
 5: CMS cabbage, '(MOR-ARV) Koiiji'
orf108: + = present, - = absent
 Pollen fertility: F = fertile, S = sterile