Theor Appl Genet (2009) 118:433–441
DOI 10.1007/s00122-008-0909-x

ORIGINAL PAPER

Identification of a novel chimeric gene, *orf725*, and its use in development of a molecular marker for distinguishing among three cytoplasm types in onion (*Allium cepa* L.)

Sunggil Kim · Eul-Tai Lee · Dong Youn Cho · Taeho Han · Haejeen Bang · Bhimanagouda S. Patil · Yul Kyun Ahn · Moo-Kyoung Yoon

Received: 19 August 2008 / Accepted: 2 October 2008 / Published online: 21 October 2008
© Springer-Verlag 2008

Abstract A novel chimeric gene with a 5′ end containing the nearly complete sequence of the *coxI* gene and a 3′ end showing homology with chive *orfA501* was isolated by genome walking from two cytoplasm types: CMS-S and CMS-T, both of which induce male-sterility in onion (*Allium cepa* L.). In addition, the normal active and variant inactive *coxI* genes were also isolated from onions containing the normal and CMS-S cytoplasms, respectively. The chimeric gene, designated as *orf725*, was nearly undetectable in normal cytoplasm, and the copy number of the normal *coxI* gene was significantly reduced in CMS-S cytoplasm. RT-PCR results showed that *orf725* was not transcribed in normal cytoplasm. Meanwhile, the normal *coxI* gene, which is essential for normal mitochondrial function, was not expressed in CMS-S cytoplasm. However, both *orf725* and *coxI* were transcribed in CMS-T cytoplasm. The expression of *orf725*, a putative male-sterility-inducing gene, was not affected by the presence of nuclear restorer-of-fertility gene(s) in male-fertility segregating populations originating from the cross between a male-sterile plant containing either CMS-T or CMS-S and a male-fertile plant whose genotypes of nuclear restorer gene(s) might be heterozygous. The specific stoichiometry of *orf725* and *coxI* in the mtDNA of the three cytoplasm types was consistent among diverse germplasm. Therefore, a molecular marker based on the relative copy numbers of *orf725* and *coxI* was designed for distinguishing among the three cytoplasm types by one simple PCR. The reliability and applicability of the molecular marker was shown by testing diverse onion germplasm.

Communicated by R. Hagemann.

S. Kim (✉) · T. Han
Department of Plant Biotechnology,
Biotechnology Research Institute,
Chonnam National University,
Gwangju 500-757, South Korea
e-mail: dronion@jnu.ac.kr

E.-T. Lee
Mokpo Experiment Station, National Institute of Crop Science,
RDA, Muan 534-833, South Korea

D. Y. Cho
ONBREETECH Corp., Haenam 536-803, South Korea

H. Bang · B. S. Patil
Department of Horticultural Sciences,
Vegetable and Fruit Improvement Center,
Texas A&M University, College Station, TX 77845, USA

Y. K. Ahn · M.-K. Yoon
National Horticultural Research Institute,
RDA, Suwon 441-440, South Korea

Introduction

Cytoplasmic male-sterility (CMS) is an inability to produce viable pollen grains due to defects in mitochondrial function and thus, is inherited maternally. Up to now, CMS has been observed in more than 140 plant species, and the male-sterility inducing factors are exclusively present in mitochondrial genomes (Hanson 1991). Plant mitochondrial genomes have several peculiar features distinct from the small (15–18 kb) and circular animal mitochondrial genomes (Fauron et al. 1990; Budar et al. 2003; Hanson and Bentolila 2004; Knoop 2004).

First, the size of plant mitochondrial genomes is relatively large and variable. It was reported that *Brassica hirta* contained the smallest mitochondrial genome at 208 kb in size (Palmer and Herbon 1987), while muskmelon retained

2,400 kb of mtDNA (Ward et al. 1981). Second, the exact configuration of mitochondrial genomes is still elusive, although the complete sequences of master circles of mtDNAs have been obtained in *Arabidopsis* (Unseld et al. 1997), sugar beet (Kubo et al. 2000), rice (Notsu et al. 2002), rapeseed (Handa 2003), and tobacco (Sugiyama et al. 2005). Few plant mtDNAs are thought to have a single circular form. Rather, they appear to be multigenomic linear and circular molecules or head-to-tail concatemer structures (Oldenburg and Bendich 1998; Oldenburg and Bendich 2001). Third, dynamic mtDNA rearrangement produces a variety of gene organizations and even creates many novel chimeric genes. Nucleotide substitution mutations rarely occur in mitochondrial genomes (Palmer and Herbon 1988; Albert et al. 1998), but active mtDNA rearrangements drive the rapid evolution of mitochondrial genomes.

Repeat sequence-mediated homologous recombination may be responsible for mtDNA rearrangement and multipartite structures (Palmer 1988; Albert et al. 1998). Short repeats distributed throughout mtDNAs mediate frequent recombination and as a result, produce a variety of subgenomic mtDNA molecules (Palmer 1988; Small et al. 1989; Albert et al. 1998; Kim et al. 2007). The stoichiometry of subgenomic mtDNA molecules is diverse within a single plant species and is maintained throughout generations (Sakai and Imamura 1993; Bellaoui et al. 1998; Janska et al. 1998; Kim et al. 2007). However, the relative copy number of subgenomic molecules changes due to genomic shifting (Arrieta-Montiel et al. 2001), which is known to be triggered by tissue culture (Fauron et al. 1990) or nuclear genes (Mackenzie and Chase 1990; Janska et al. 1998; Abdelnoor et al. 2003). Genomic shifting produces a novel mitotype (mitochondrial genome variant), in which some low-copy-number molecules increase to express new phenotypes and the copy number of some genes on the master circle decreases, resulting in loss of function (Small et al. 1989). For instance, a low-copy-number chimeric gene inducing male-sterility may increase by genomic shifting to be expressed abundantly, leading to male-sterility phenotypes (Arrieta-Montiel et al. 2001).

In onion, two different CMS systems: CMS-S (Jones and Emsweller 1936) and CMS-T (Berninger 1965) have been discovered and used in F_1 hybrid production. Male-sterility caused by CMS-S cytoplasm is recovered to become fully male-fertile by a single nuclear locus (Jones and Clarke 1943). The CMS-S system has been widely used in F_1 hybrid cultivar development, because male-sterility is stable in diverse environmental conditions (Havey 2000). In addition, breeding schemes are less complicated and the frequency of maintainer lines would be higher because of the simple inheritance of a nuclear restorer gene (Jones and Clarke 1943). Meanwhile, the CMS-T system has been used primarily in Europe (Havey 2000). Unlike CMS-S, the inheritance of the nuclear restoration of fertility is complex. Three independent loci are involved in fertility restoration (Schweisguth 1973).

Identification of cytoplasm types, an initial key step in F_1 hybrid breeding, takes 4–8 years by progeny testing since onion is a biennial crop. However, molecular markers capable of distinguishing mitotypes at the DNA level enable breeders to save time and effort. Many molecular markers to distinguish normal and CMS-S mitotypes have been developed based on the variation of relative copy numbers of mtDNA molecules (Sato 1998) and polymorphic sequences of chloroplast genomes (Havey 1995). However, there has been only one report of a molecular marker for CMS-T mitotype identification (Engelke et al. 2003). The marker was designed on the basis of chive mtDNA sequence containing *orfA501*, which was unique to one of the chive CMS systems (Engelke and Tatlioglu 2002).

The onion sequence of the *orfA501* homolog and its flanking sequence were isolated in this study, and a novel chimeric gene was discovered from the isolated sequences. The differential stoichiometry of the chimeric gene among the three mitotypes and its implication in male-sterility induction were analyzed. Finally, an economical and efficient molecular marker was developed based on the novel chimeric gene organization.

Materials and methods

Plant materials

The cytoplasm types of 176 breeding lines and cultivars usually grown in Korea and Japan were initially analyzed using two previously reported molecular markers (Havey 1995; Engelke et al. 2003). Eleven accessions of each cytoplasm type were selected for gene isolation and molecular marker development. Each of four male-sterile and male-fertile individuals from the crosses between a single plant of male-sterile plants containing CMS-S or CMS-T cytoplasms and a single plant of male-fertile lines whose genotypes for the restorer gene(s) might be heterozygous were used for RT-PCR analysis.

DNA extraction and PCR amplification

Total genomic DNA was extracted from leaves, bulb scales, or peduncles, depending on availability of fresh tissues of breeding lines and cultivars using a CTAB method (Murray and Thompson 1980). PCR was performed in a 10 μL reaction mixture containing 0.05 μg template, 1 μL 10× PCR buffer, 0.2 μL forward primer (10 μM), 0.2 μL reverse primer (10 μM), 0.2 μL dNTPs (10 mM each), and

Table 1 Primer sequences used in RT-PCR analysis and molecular marker tests

Primer names	Primer sequences (5′ to 3′)	Experiments
RT-F	TGGGTGGGTAAAATCTTTGGTCGGACA	RT-PCR
RT-R1	CACTTTGTCCCACCCGAACTCCAACTC	RT-PCR
RT-R2	CCGTTCCGAAGGCGAATAAAAATTTGG	RT-PCR
Tubulin-F	GGAAGCATGTGCCCCGTGCTATATTTG	RT-PCR
Tubulin-R	ACAATCTGGATCGTGCGCTTCGTCTTT	RT-PCR
MK-F	CATAGGCGGGCTCACAGGAATA	Marker
MK-R1	AATCCTAGTGTCCGGGGTTTCT	Marker
MK-R2	CAGCGAACTTTCATTCTTTCGC	Marker

The positions of primers, except for tubulin, are depicted in Fig. 1a

0.1 μL polymerase mix (Advantage 2 Polymerase Mix, Clontech, Palo Alto, CA, USA). PCR amplification for sequencing was carried out with an initial denaturation step at 94°C for 5 min, followed by 40 cycles of 94°C for 30 s, 65°C for 30 s, and 72°C for 90 s, and a final 10 min extension at 72°C. As to the PCR amplification of the molecular marker developed in this study, the 60°C annealing temperature was used. The primer sequences used in this study for RT-PCR and molecular marker tests are presented in Table 1.

Genome walking for gene isolation and sequencing of PCR products

For genome walking, total genomic DNA of accessions containing CMS-T and normal cytoplasms was extracted from the leaf tissues of three-leaf-stage seedlings using a commercial DNA extraction kit (DNeasy Plant Mini Kit, QIAGEN, Valencia, CA, USA) according to the manufacturer's manual. Genome walking libraries were constructed using the Universal GenomeWalker Kit (Clontech) according to the manufacturer's protocol.

For DNA sequencing, the PCR products were purified using the QIAquick PCR Purification kit (QIAGEN). The purified PCR products were either sequenced directly or after cloning into a TOPO TA cloning vector for sequencing (Invitrogen, Carlsbad, CA, USA). Sequencing reactions were carried out using Big Dye (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocol, and the sequences were obtained using an ABI 3700 Genetic Analyzer (Applied Biosystems).

RT-PCR and rapid amplification of cDNA ends (RACE)

Total RNA was extracted from unopened flowers of male-sterile and male-fertile segregating individuals using an RNA extraction kit (RNeasy Plant Mini Kit, QIAGEN) following the manufacturer's instructions. cDNA was synthesized from total RNA using a commercial cDNA synthesis kit (SuperScript™ III first-strand synthesis system for RT-PCR, Invitrogen). RT-PCR amplification was performed with an initial denaturation step at 94°C for 3 min followed by 30 cycles of 94°C for 30 s, 65°C for 30 s, and 72°C for 2 min, and a final 10-min extension at 72°C. Primer sequences used in RT-PCR are presented in Table 1. The onion tubulin sequence that was used as a control was obtained from EST sequences (TC125) from the DFCI *Allium cepa* Gene Index (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/Blast/index.cgi). RACE was carried out with a commercial RACE kit (SMART RACE cDNA Amplification Kit; Clontech) according to the manufacturer's instructions.

Results

Identification of a novel chimeric gene predominantly present in mtDNA of male-sterile onions

A previous report (Engelke and Tatlioglu 2002) showed that an mtDNA fragment containing a putative open reading frame (*orfA501*) was exclusively present in one of the CMS cytoplasms in chives. Later, Engelke et al. (2003) noted that a primer pair designed for the chive *orfA501* produced PCR products of similar size only from CMS-T and CMS-S cytoplasms of onions. However, they did not mention the sequence of the *orfA501*-homolog from onions or information on the flanking sequences.

To develop more reliable molecular markers and to investigate the effect of the *orfA501*-homolog on male-sterility induction in onions, the *orfA501*-homolog and its flanking sequences were obtained via genome walking from CMS onions. The isolated homologous region of the *orfA501* showed 92% nucleotide sequence identity to that of chives, and the 3′ flanking region contained partial sequences of *atp6* and *atp9* genes as in the chive *orfA501* fragment (Fig. 1a). However, a completely different 5′ flanking sequence was found in onions. Nearly the entire *coxI* (cytochrome *c* oxidase subunit I) gene was connected to the *orfA501*-homolog sequence in onions, as compared

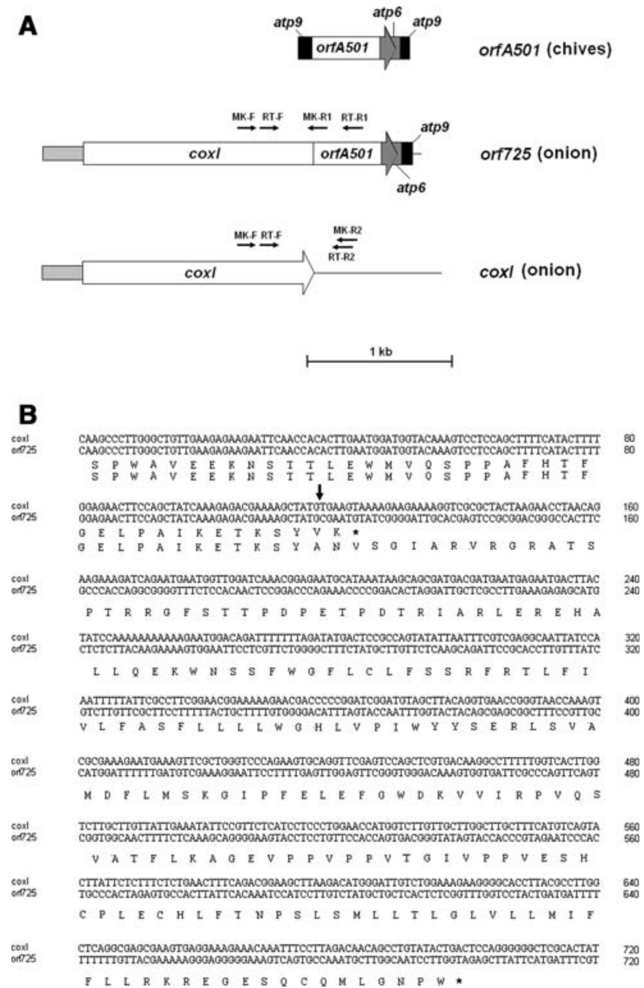

Fig. 1 The organization of a novel chimeric gene, *orf725*, and normal *coxI* isolated from onion mitochondrial genomes. **a** Comparison of gene structures of *orf725* and normal *coxI*. The previously reported chive chimeric gene, *orfA501*, is also depicted for comparison (Engelke and Tatlioglu 2002). *Arrow-shaped boxes* indicate the 5′-to-3′ direction. *Horizontal arrows* indicate primer-binding sites. **b** Alignment of nucleotide and deduced amino acid sequences of *orf725* and normal *coxI*. Partial sequences around the breakpoint are shown. *Asterisks* and the *vertical arrow* indicate the stop codons and a putative breakpoint of the mtDNA rearrangement, respectively

with the chive *orfA501*, which contained a partial sequence of the *atp9* gene at the 5′ end (Fig. 1a). Interestingly, the chimeric sequence consisting of *coxI* and the *orfA501*-homolog creates a continuous open reading frame of 2,175 bp. This novel open reading frame was designated *orf725*. The normal onion *coxI* gene was also isolated while performing genome walking from normal cytoplasms. The 281 bp 5′ putative promoter region and coding sequences of the *coxI* and *orf725* were exactly identical except for 8 bp at the 3′ end of the coding sequence which included a stop codon. The recombination may have happened 5 bp upstream of the stop codon, and as a result, the stop codon of the normal *coxI* disappeared, and the open reading frame was extended into the *orfA501*-homolog sequences (Fig. 1b).

Different stoichiometry of *orf725* and *coxI* genes among three onion mitotypes

As expected, *orf725* was only amplified from CMS-T and CMS-S mitotypes (Fig. 2), but faint bands of the expected size were detected in some accessions when the PCR cycles were increased to more than 40 cycles (data not shown). This result indicates that the *orf725* might exist in normal cytoplasm at the substoichiometric level. The identities of the faint bands from normal mitotypes were confirmed to be *orf725* by sequencing.

Meanwhile, the *coxI* gene was likely to exist on the master circles in normal and CMS-T mitotypes, but the copy number was significantly reduced in the CMS-S mitotype (Fig. 2). When the PCR products from CMS-S mitotypes were sequenced, surprisingly, the sequence contained a variant of the *coxI* gene. There was 98% nucleotide sequence identity between the normal and variant *coxI* genes throughout a 2,767 bp region including the 5′ and 3′ flanking and coding sequences. The sequence analysis of the variant *coxI* gene showed that it contained a premature stop codon and 20 amino-acid changing point mutations (Fig. 3). A cladogram and sequence comparison with the *coxI* gene of *Asparagus officinalis* showed that all point mutations accumulating in the variant *coxI* gene began to be produced within onions or *Allium* species (Fig. 4a). The unusually high rate of point mutations in the *coxI* gene, an essential mitochondrial gene, may be due to loss of function immediately after duplication. The many amino-acid changes in the conserved region of the *coxI* genes among distantly related plant species support this hypothesis (Fig. 4b). Therefore, this variant of *coxI* gene must be an inactive pseudogene.

At first, this *coxI* variant identified from only CMS-S mitotypes was thought to be a counterpart of the normal *coxI* in the normal and CMS-T mitotypes. However, not all CMS-S mitotypes exclusively contained the variant *coxI* gene. Some accessions containing CMS-S cytoplasm were shown to have normal *coxI* (data not shown). Even the

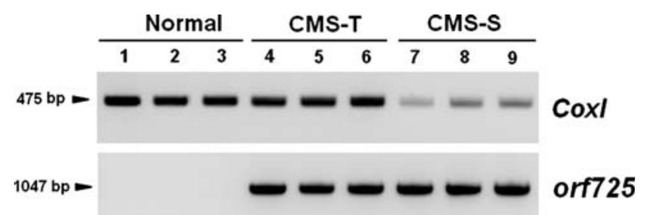

Fig. 2 PCR products of *coxI* and *orf725* from three onion mitotypes. 1–3, 4–6, 7–9 representative cultivars or breeding lines for each mitotype

Fig. 3 Alignment of the deduced amino acid sequences of normal and variant *coxI* genes isolated from onions. Polymorphic sequences are enclosed in *empty rectangular boxes*. *Triangles* above the sequences indicate the sequences changed by RNA editing. *Asterisk and vertical arrow* represent the position of a pre-mature stop codon. '**:**' and '**.**' represent conserved and semi-conserved substitution, respectively

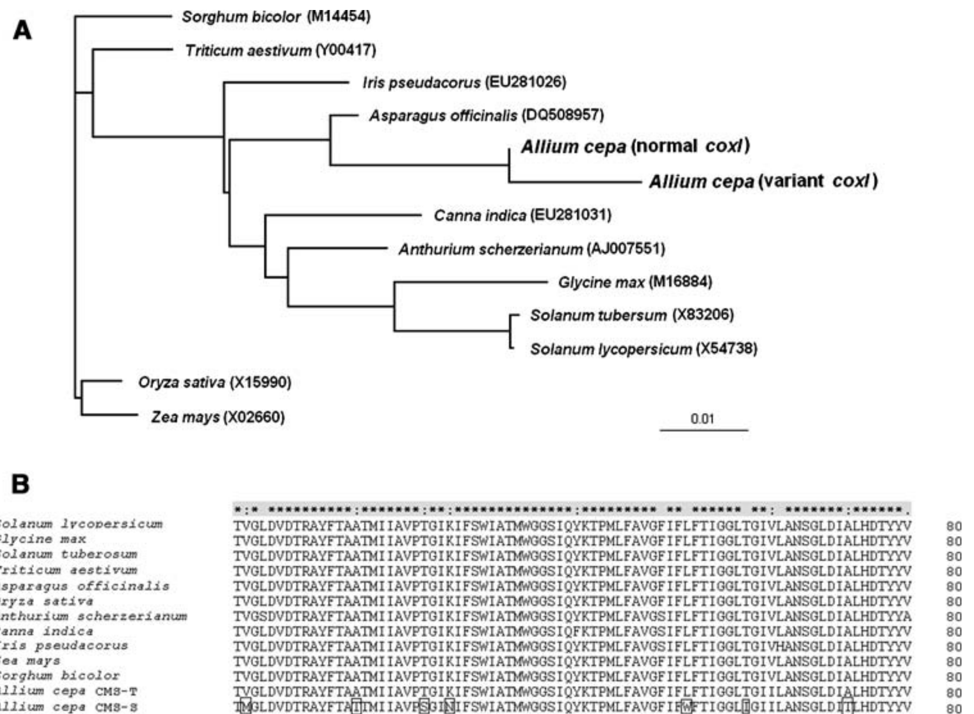

Fig. 4 Phylogenetic relationship between normal and variant *coxI* genes. **a** Cladogram showing genetic distances among *coxI* genes of diverse plant species. The tree was produced using ClustalX and TreeView softwares. The GenBank accession numbers of *coxI* genes of each species are shown in the *parenthesis*. **b** Alignment of conserved amino acid sequences of *coxI* genes in diverse plant species with normal and variant *coxI* genes isolated from onions. The polymorphic sequences specific to the variant *coxI* gene are enclosed in *empty rectangular boxes*. '*' all identical amino acids, '**:**'conserved substitution, '**.**'semi-conserved substitution

progeny originating from a single female plant containing CMS-S cytoplasm included two individuals with both forms of the *coxI* gene, although most of the progeny contained the variant *coxI* gene (Table 2).

The effect of restorer-of-fertility gene(s) on the expression of *orf725* and *coxI*

The full-length cDNA of normal *coxI* was obtained by RACE PCR from normal mitotypes. It consisted of a 52 bp-long 5′UTR, 1,599 bp-long coding sequence, and 123 bp-long 3′UTR. In addition, we found four C-to-U RNA editing positions in the coding regions. The four editing sites caused amino acid changes as indicated in Fig. 3. Meanwhile, the full-length cDNA of the *orf725* gene was composed of a 52 bp-long 5′UTR, 2,175 bp-long coding region, and 151 bp-long 3′ UTR and was obtained from CMS-T mitotypes. The RNA editing positions of the *orf725* gene were exactly the same as those for *coxI*.

To investigate the effect of restorer-of-fertility gene(s) on the expression of *orf725* and *coxI*, RT-PCR was carried out using total RNAs extracted from floral buds of male-fertility segregating populations containing CMS-T or CMS-S cytoplasms. In the case of the CMS-T mitotype, both *coxI* and *orf725* were highly expressed in all individuals. The identities of the *coxI* genes were confirmed to be all normal *coxI* genes, not the inactive variant. There was no significant differential expression between male-sterile and male-fertile individuals, suggesting that the restorer-of-fertility gene(s) may have no effect on the transcription of

Table 2 Detection of normal and variant *coxI* from the progeny originating from the cross of a male-sterile plant containing CMS-S cytoplasm and a male-fertile breeding line containing normal cytoplasm and heterozygous restorer-of-fertility gene(s)

	Progenies from the cross (HNR × MOM8-1)[a]									
	Male-sterile					Male-fertile				
	1	2	3	4	5	6	7	8	9	10
orf725	+[b]	+	+	+	+	+	+	+	+	+
Normal *coxI*		+	−	−	−	−	+	−	−	−
Variant *coxI*	+	+	+	+	+	+	+	+	+	+

[a] HNR: male-sterile line containing the CMS-S cytoplasm, MOM8-1: male-fertile line containing the normal cytoplasm

[b] '+' indicates the presence of the PCR products and '−' indicates the absence of the PCR products. The existence of the PCR products was determined according to patterns of sequencing chromatograms

both genes (Fig. 5a). On the other hand, no RT-PCR product of the *coxI* gene was observed for the CMS-S mitotype, but *orf725* showed a high level of expression. Like the CMS-T mitotype, no differential expression of *orf725* between male-sterile and male-fertile individuals was observed in the CMS-S mitotype (Fig. 5b). In contrast, no expression of *orf725* was observed, and *coxI* was highly expressed in normal mitotypes (Fig. 5c).

Development of a molecular marker for distinction of the three mitotypes

Until now, no molecular markers which can distinguish the three onion mitotypes by one simple PCR have been developed. To develop a simple PCR-based molecular marker for distinguishing the three mitotypes, the aforementioned *orf725* and *coxI* gene organization and the difference of their relative copy numbers in the three mitotypes were utilized.

One common forward primer (MK-F) binding to the *coxI* coding sequence was designed, and reverse primers binding to regions unique to *orf725* (MK-R1) or *coxI* (MK-R2) were designed (Fig. 1a). To avoid amplification of the inactive *coxI* variant in CMS-S mitotypes, two and one mismatches for the variant *coxI* sequence were contained in the 3′ end of the common forward primer and the *coxI* reverse primer, respectively. One single band for normal *coxI* appeared in normal mitotypes, because the *orf725* gene exists at an undetectable level. Two bands amplified from both normal *coxI* and *orf725* were detected in CMS-T mitotypes since the copy numbers of both *coxI* and *orf725* were likely to be similar in CMS-T mitotypes. One band of *orf725* was amplified in CMS-S mitotypes due to the relatively low copy number of *coxI* and the frequent presence of inactive *coxI* variant (Fig. 6a).

The developed molecular markers were tested in a variety of breeding lines and cultivars to prove to have reliable application in diverse germplasm. Mitotypes of each tested onion sample had been verified using two previously reported molecular markers (Havey 1995; Engelke et al. 2003). The results showed that the marker can clearly distinguish between the three mitotypes in all tested plant materials, proving its applicability and reliability (Fig. 6b).

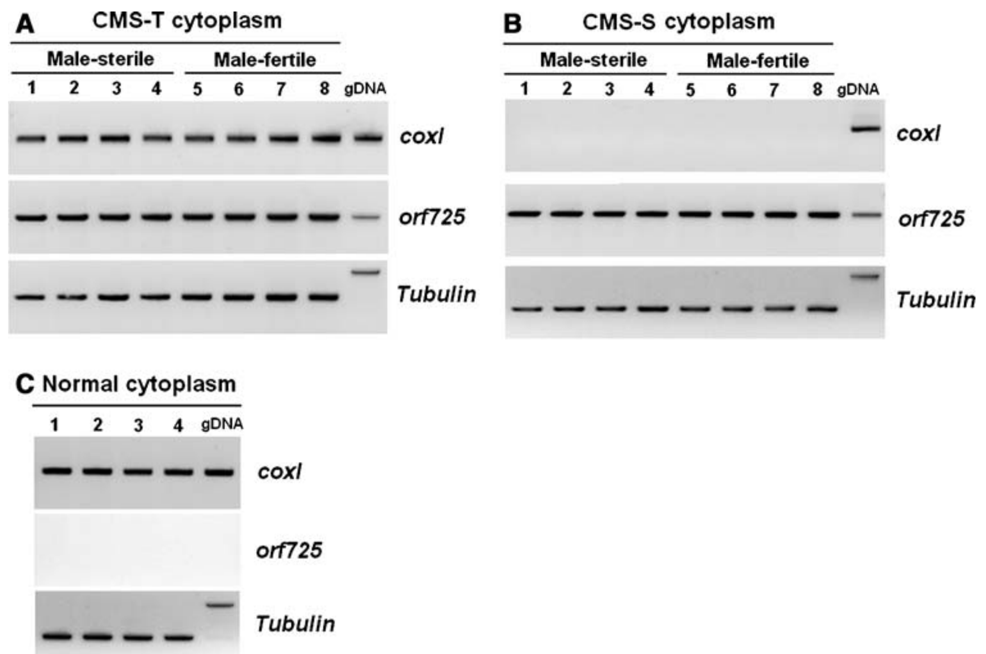

Fig. 5 RT-PCR amplification of *coxI* and *orf725* from three onion mitotypes. **a** CMS-T mitotype. *1–8* segregating progeny from the cross between a male-sterile plant containing CMS-T cytoplasm and a male-fertile plant. **b** CMS-S mitotype. *1–8* segregating progeny from the cross product of a CMS-S-containing male-sterile plant and a male-fertile plant. **c** Normal mitotype. *1–4* four representative accessions containing normal cytoplasms

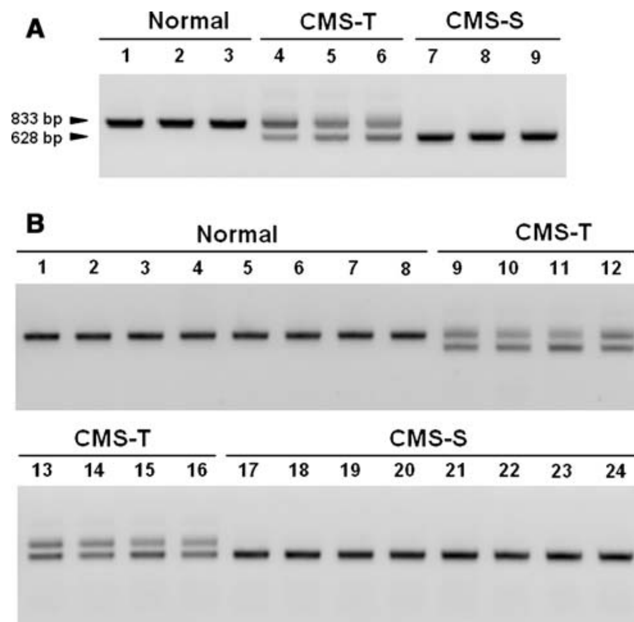

Fig. 6 Molecular markers developed in this study for distinguishing between the three onion mitotypes. **a** PCR patterns of three representative accessions of each mitotype. **b** PCR pattern of the molecular marker tested on diverse onion germplasm

Discussion

Identification of a novel chimeric gene, *orf725*, from male-sterile cytoplasms in onion

In this study, a novel chimeric gene, *orf725*, was identified from onions containing male-sterility inducing cytoplasms. This chimeric gene was probably created by short repeat sequence-mediated homologous recombination, although no known repeat sequence was identified around the breakpoint of the recombination due to insufficient mtDNA sequence information in onion and other *Allium* species. Mitochondrial genome rearrangement via short repeat sequences has been well documented in other plant species (Small et al. 1989; Albert et al. 1998; Knoop 2004; Kim et al. 2007). Nevertheless, the junction connecting the *coxI* gene and *orfA501* homolog sequence might be a hotspot for mtDNA recombination in *Allium* species since the positions of the junction of two chimeric genes: onion *orf725* and chive *orfA501* were identical (Fig. 1a). A similar hot spot for mtDNA recombination was reported in the *rpl5-rps14* region in *Solanum* species (Scotti et al. 2004).

The recombination that created *orf725* may have happened very recently in *Allium cepa* or a common ancestor of very closely related *Allium* species. The identical sequence of the *coxI* gene-harboring 5′ end of *orf725* with the sequence of the corresponding normal *coxI* gene supports this very recent creation of the chimeric gene. On the contrary, there were 8% polymorphic sequences between chive *orfA501* and its onion homolog. It has been known that onion and chives are close relatives among more than 700 *Allium* species (Havey 1991; Havey 1992; van Raamsdonk et al. 2003).

Isolation of the inactive *coxI* variant from CMS-S mitotypes led us to speculate that normal and CMS-S mitotypes diverged a long time ago or the CMS-S mitotype may have been introduced from other *Allium* species via interspecific hybridization as Havey (1993) proposed. Like *orfA501* regions of onion and chives, there were 8% polymorphic sequences between the normal and variant *coxI* genes. However, detection of the normal *coxI* gene from some accessions containing CMS-S cytoplasm and identification of individuals possessing both normal and variant *coxI* genes enabled us to exclude the hypothesis of ancient divergence. We identified a premature stop codon and many amino acid-changing point mutations within the highly conserved blocks in *coxI* genes (Fig. 4b). These occurrences of random point mutations indicate that the variant *coxI* may have been inactivated right after it had been duplicated from the active *coxI* gene and has accommodated all arising natural mutations. The cladogram and sequence alignment of both normal and variant *coxI* in onion and the *coxI* gene of *Asparagus officinalis* showed that all polymorphic nucleotide sequences between normal and variant *coxI* genes of onions were conserved between normal onion and *Asparagus officinalis coxI* genes, meaning that all mutations in variant *coxI* arose after the divergence of *Asparagus officinalis* and onions (Fig. 4a). However, the 8% nucleotide substitution rate is still surprising considering that the mitochondrial genome has a relatively low frequency of point mutations compared with nuclear and chloroplast genomes (Palmer and Herbon 1988; Albert et al. 1998).

In conclusion, all three onion mitotypes probably contain *orf725*, normal, and variant *coxI* genes in populations of mtDNA molecules, but the relative copy numbers of the three genes vary depending on mitotypes. Two major genomic shifting events, which are known to be a mechanism of changing the stoichiometry of mtDNAs (Small et al. 1989; Arrieta-Montiel et al. 2001), may produce three different mitotypes. The first genomic shifting in which the copy number of *orf725* increased and the copy number of normal *coxI* decreased resulted in creation of the CMS-S mitotype. The second genomic shifting, which may have happen more recently, increased *orf725* without decreasing normal *coxI* gene, resulting in creation of CMS-T mitotypes. The recent creation of the CMS-T mitotype is inferred from the fact that no polymorphism on cpDNA has been identified between normal and CMS-T cytoplasm to date (Havey 1993).

The role of the chimeric gene, *orf725*, in male-sterility induction

Several mitochondrial genes responsible for male-sterility have been cloned in some plant species and most of them have been revealed as chimeric genes created through mtDNA rearrangement (Hanson and Bentolila 2004). The chimeric gene *orf725* identified in this study would be a candidate male-sterility inducing gene in onions. Although the transcription of *orf725* was not affected by the presence of dominant alleles of the restorer-of-fertility gene(s) as shown in Fig. 5, the restorer gene may affect the translation of *orf725* or mediate special protein–protein interactions at the post-translational level. Indeed, accumulation of male-sterility inducing mitochondrial gene products in some plants was affected by nuclear restorer genes at the post-transcriptional level (Nivison and Hanson 1989; Kadowaki et al. 1990; Koizuka et al. 2000). Therefore, further studies on the effect of restorer gene(s) on the accumulation of *orf725* gene products at the translational and post-translational level are required to clarify the role of *orf725* in male-sterility induction in onions.

Another intriguing aspect of *orf725* is its role in assembling cytochrome *c* oxidase (COX) in CMS-S cytoplasm, in which transcription of the normal *coxI* gene was not detected at all (Fig. 5b). The cytochrome *c* oxidase complex is the terminal electron acceptor in the mitochondrial respiratory chain. In mammals, the COX complex consists of 13 subunits, some of which are encoded by the mitochondrial genome and others which are encoded by the nuclear genome (Taanman 1997). Since mitochondrial-encoded subunits comprise the catalytic core of the enzyme, the role of *coxI* in encoding cytochrome *c* oxidase subunit I is essential in the normal function of COX complex. Although *orf725* contained almost the entire coding sequence of *coxI*, the extended C-terminal regions containing the *orfA501*-homolog may disturb optimal assembling of the complex. In humans, several defects of nuclear-encoded factors involved in the assembly of COX complex have been reported to cause many clinical symptoms (Pecina et al. 2004). The symptoms are known to be prominent in high energy demanding tissues such as muscle, brain, and heart (Wallace 1992). Similarly, anthers in plants where functional pollen grains are produced also require a high level of energy. Therefore, it is possible that dysfunction of the COX complex caused by *orf725* fails to produce sufficient energy above the threshold level required to make functional pollen grains. Accordingly, specific nuclear-encoded assembly factors which stabilize the deformed COX complex in male-sterile cytoplasm would be candidates of nuclear restorer-of-fertility genes in onions.

The specific stoichiometry of *orf725* and *coxI* was used in developing economical and reliable molecular markers for distinguishing the three onion mitotypes in this study, but the role of *orf725* in male-sterility induction and COX assembly would be an interesting topic for future studies.

Acknowledgments This study was supported by the "Specific Joint Agricultural Research-promoting Projects (Project No. 200803A01081219)", RDA, Republic of Korea, and partially supported by the Biotechnology Research Institute, Chonnam National University, Republic of Korea. We thank Seung-Hee Kim and Seon Chong for their dedicated technical help.

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 100:5968–5973

Albert B, Godelle B, Gouyon PH (1998) Evolution of the plant mitochondrial genome: dynamics of duplication and deletion of sequences. J Mol Evol 46:155–158

Arrieta-Montiel M, Lyznik A, Woloszynska M, Janska H, Tohme J, Mackenzie S (2001) Tracing evolutionary and developmental implications of mitochondrial stoichiometric shifting in the common bean. Genetics 158:851–864

Bellaoui M, Martin-Canadell A, Pelletier G, Budar F (1998) Low-copy-number molecules are produced by recombination, actively maintained and can be amplified in the mitochondrial genome of Brassicaceae: relationship to reversion of the male sterile phenotype in some cybrids. Mol Gen Genet 257:177–185

Berninger E (1965) Contribution à l'étude de la sterilité mâle de l'oignon (*Allium cepa* L.). Ann Amélior Plant 15:183–199

Budar F, Touzet P, De Paepe R (2003) The nucleo-mitochondrial conflict in cytoplasmic male sterilities revised. Genetica 117:3–16

Engelke T, Tatlioglu T (2002) A PCR-marker for the CMS_1 inducing cytoplasm in chives derived from recombination events affecting the mitochondrial gene *atp9*. Theor Appl Genet 104:698–702

Engelke T, Terefe D, Tatlioglu T (2003) A PCR-based marker system monitoring CMS-(S), CMS-(T) and (N)-cytoplasm in the onion (*Allium cepa* L.). Theor Appl Genet 107:162–167

Fauron CMR, Havlik M, Brettell RIS (1990) The mitochondrial genome organization of a maize fertile cmsT revertant line is generated through recombination between two sets of repeats. Genetics 124:423–428

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

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

Hanson MR, Bentolila S (2004) Interactions of mitochondrial and nuclear genes that affect male gametophyte development. Plant Cell 16:S154–S169

Havey MJ (1991) Phylogenetic relationships among cultivated *Allium* species from restriction enzyme analysis of the chloroplast genome. Theor Appl Genet 81:752–757

Havey MJ (1992) Restriction enzyme analysis of the nuclear 45 s ribosomal DNA of six cultivated Alliums (*Alliaceae*). Pl Syst Evol 181:45–55

Havey MJ (1993) A putative donor of S-cytoplasm and its distribution among open-pollinated populations of onion. Theor Appl Genet 86:128–134

Havey MJ (1995) Identification of cytoplasms using the polymerase chain reaction to aid in the extraction of maintainer lines from

open-pollinated populations of onion. Theor Appl Genet 90:263–268

Havey MJ (2000) Diversity among male-sterility-inducing and male-fertile cytoplasms of onion. Theor Appl Genet 101:778–782

Janska H, Sarria R, Woloszynska M, Arrieta-Montiel M, Mackenzie SA (1998) Stoichiometric shifts in the common bean mitochondrial genome leading to male sterility and spontaneous reversion to fertility. Plant Cell 10:1163–1180

Jones HA, Clarke A (1943) Inheritance of male sterility in the onion and the production of hybrid seed. Proc Am Soc Hort Sci 43:189–194

Jones HA, Emsweller SL (1936) A male-sterile onion. Proc Am Soc Hort Sci 34:582–585

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

Kim S, Lim H, Park S, Cho K, Sung S, Oh D, Kim K (2007) Identification of a novel mitochondrial genome type and development of molecular makers for cytoplasm classification in radish (*Raphanus sativus* L.). Theor Appl Genet 115:1137–1145

Knoop V (2004) The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. Curr Genet 46:123–139

Koizuka N, Imai R, Iwabuchi M, Sakai T, Imamura J (2000) Genetic analysis of fertility restoration and accumulation of ORF125 mitochondrial protein in the kosena radish (*Raphanus sativus* cv. Kosena) and a *Brassica napus* restorer line. Theor Appl Genet 100:949–955

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

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

Murray MG, Thompson WF (1980) Rapid isolation of high-molecular-weight plant DNA. Nucleic Acids Res 8:4321–4326

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

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

Oldenburg DJ, Bendich AJ (1998) The structure of mitochondrial DNA from the Liverwort, *Marchatia polymorpha*. J Mol Biol 276:745–758

Oldenburg DJ, Bendich AJ (2001) Mitochondrial DNA from the Liverwort *Marchantia polymorpha*: circularly permuted linear molecules, head-to-tail concatemers, and a 5′ protein. J Mol Biol 310:549–562

Palmer JD (1988) Intraspecific variation and multicircularity in *Brassica* mitochondrial DNAs. Genetics 118:341–351

Palmer JD, Herbon LA (1987) Unicircular structure of the *Brassica hirta* mitochondrial genome. Curr Genet 11:565–570

Palmer JD, Herbon LA (1988) Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J Mol Evol 28:87–97

Pecina P, Houstkova H, Hansikova H, Zeman J, Houstek J (2004) Genetic defects of cytochrome *c* oxidase assembly. Physiol Res 53:S213–S223

Sakai T, Imamura J (1993) Evidence for a mitochondrial sub-genome containing radish *atpA* in a *Brassica napus* cybrid. Plant Sci 90:95–103

Sato Y (1998) PCR amplification of CMS-specific mitochondrial nucleotide sequences to identify cytoplasmic genotypes of onion (*Allium cepa* L.). Theor Appl Genet 96:367–370

Schweisguth B (1973) Étude d'un nouveau type de stérilité male chez l'oignon, *Allium cepa* L. Ann Amélior Plant 23:221–233

Scotti N, Marechal-Drouard L, Cardi T (2004) The *rpl5-rps14* mitochondrial region: a hot spot for DNA rearrangements in *Solanum* spp. Somatic hybrids. Curr Genet 45:378–382

Small I, Suffolk R, Leaver CJ (1989) Evolution of plant mitochondrial genomes via substoichiometric intermediates. Cell 58:69–76

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

Taanman JW (1997) Human cytochrome *c* oxidase: structure, function, and deficiency. J Bioenerg Biomembr 29:151–163

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

van Raamsdonk LWD, Ensink W, van Heusden AW, Vrielink-van Ginkel M, Kik C (2003) Biodiversity assessment based on cpDNA and crossability analysis in selected species of *Allium* subgenus *Rhizirideum*. Theor Appl Genet 107:1048–1058

Wallace DC (1992) Disease of the mitochondrial DNA. Annu Rev Biochem 61:1175–1212

Ward BL, Anderson RS, Bendich AJ (1981) The mitochondrial genome is large and variable in a family of plants (Cucurbitaceae). Cell 25:793–803