Sequence Analysis and Expression of *orf224* Gene Associated with Two Types of Cytoplasmic Male Sterility in *Brassica napus* L.

Jianmin Liu[a,b], Maoteng Li[a,b,*], Hao Wang[c], Longjiang Yu[a,b], and Dianrong Li[c,*]

[a] Institute of Resource Biology and Biotechnology, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, 430074, China.
E-mail: limaoteng426@163.com
[b] Key Laboratory of Molecular Biophysics, Ministry of Education, Wuhan, 430074, China
[c] Hybrid Rapeseed Research Center of Shaanxi Province, Dali, 715105, China.
E-mail: lidr@peoplemail.com.cn

* Authors for correspondence and reprint requests

Z. Naturforsch. **65c**, 395–402 (2010); received September 20/December 31, 2009

Polima and Shaan 2A are the two most widely used forms of cytoplasmic male sterility (CMS) in the utilization of heterosis of rapeseed (*Brassica napus*) in China. A previous study indicated that the mitochondrial gene, *orf224*, was the only gene with a differential expression pattern among the normal, sterile and fertility-restored lines in rapeseed. DNA sequences of *orf224*, including coding sequences from Shaan 2A and Polima CMS, were then amplified and analyzed. DNA sequence alignment indicated both the coding sequences were 675 bp in length and had 99.9 and 99% homology in nucleotides and amino acids, respectively, and shared certain similarity to homologues from other *Brassica* spp. and *Arabidopsis thaliana*. The probable promoter regions of *orf224* were conserved between *B. napus* and *A. thaliana*, but the upstream regions of probable promoter regions were completely divergent from each other. Additionally, analysis of the primary and secondary structure of the proteins encoded by *orf224* from the two lines predicted that the proteins contain a α-helix, extended strand, and random coil. After cloning a *in vitro* experiment showed that these two proteins could be expressed in *Escherichia coli* BL21.

Key words: Shaan 2A CMS, Polima CMS, *orf224* Gene

Introduction

Cytoplasmic male sterility (CMS) is a maternally inherited character, which results from abnormal microsporogenesis or microgametogenesis leading to no pollen production or dysfunctional pollen. The hybrid seed yield of rapeseed (*Brassica napus*) was reported to be as much as 60% higher than in parental lines therefore it was frequently used in commercial F_1 hybrid production (Horlow *et al.*, 1992; Lin *et al.*, 2006; Pascal and Francoise, 2004; Singh and Brown, 1991). Four types of CMS of rapeseed exist in nature, *i.e.* nap CMS, Polima CMS, Shaan 2A CMS, and Ogu CMS. Nap CMS is also known as the Shiga-Thompson system, and was first reported by Shiga and Baba (1973) and Thompson (1972). However, the sensitivity of this CMS to temperature is a critical restriction for its application. Ogu CMS, first observed by Ogura in a Japanese radish groups, also has serious defects in hybrid production due to its deficiency in chlorophyll and the absence of a restoring gene (Gourret *et al.*, 1992). Polima CMS, which might have originated from a Polish cultivar through genetic mutation, was discovered by Fu in 1972 (Fu, 1990). Shaan 2A CMS was discovered by Li in 1976 (Li, 1980) in the progeny of hybridization of S74 × (Fengshou 4 + 7207), and this CMS was first applied utilizing of rapeseed heterosis, and resulted in the world's first hybrid rapeseed 'Qinyou 2'.

At present, Polima and Shaan 2A CMS are the only two CMS widely applied in heterosis utilization in China. Although Shaan 2A CMS and Polima CMS have different origin, they are sometimes assumed to be the same type of CMS, because they share the same restorer lines and maintainer lines (Yang and Fu, 1991), and have the same cytoplasm as revealed by restriction fragment length polymorphism (RFLP) and mtDNA fragment polymorphism analyses (Yang *et al.*, 1998).

Previous studies of the molecular basis of Polima CMS have demonstrated that the *orf224/atp6* coding region is the only mitochondrial gene region expressed differently in normal, sterile and

fertility-restored lines (Handa and Nakajima, 1992; Singh *et al*., 1996; Yuan *et al*., 2003).

To analyze the cytoplasmic difference in Shaan 2A and Polima CMS, the *orf224* gene was selected for DNA sequence comparison. Cloning, expression in *Escherichia coli* BL21, and sequence analysis of the *orf224* gene in Shaan 2A and Polima CMS were performed. The results indicated only one base difference between them. The difference upstream of these two genes and the phylogenetic relationship between *B. napus* and *A. thaliana* were analyzed. Secondary structure analysis revealed that the orf224 protein contained an α-helix, extended strand, and random coil. The results provide theoretical support for producing other sterile crops.

Materials and Methods

Materials

Shaan 2A CMS and Polima CMS lines were provided by Dr. Dianrong Li, Chinese National Rapeseed Genetic Improvement Center (Shaanxi, Dali, China).

Bacterial strain and vectors

Escherichia coli DH5α was used as the host for plasmid amplification, and *E. coli* BL21 (DE3) (Novagen, Darmstadt, Germany) was used for gene expression. pMD18-T simple vector and pET-32a (+) plasmid were purchased from Takara Biotechnology Co. Ltd. (Dalian, China) and Novagen, respectively.

DNA isolation and primer design

Total DNA was isolated from leaf tissue according to Li *et al.* (2007). The primers for amplifying *orf224* were designed as follows, according to Lin *et al.* (2006):

orf224L: 5′-CGC↓CATGGATGCCTCAACTGGATAAATTCAC-3′ (*Nco*I),

orf224R: 5′-CGGAGCT↓CTCAGCGAAAGAGATCAAGGA-3′ (*Sac*I).

PCR amplification

PCR amplification was carried out in 20 μL of reaction system including 10 ng of total DNA, 20 mM Tris [2-amino-2-(hydroxymethyl)-1,3-propanediol]-HCl (pH 8.4), 20 mM KCl, 1.5 mM $MgCl_2$, 0.4 mM of each dNTP, 10 pM of each primer, and 1.5 U of Pfu DNA polymerase (Tiangen Biotech Co. Ltd., Beijing, China). PCR was performed in an i-Cycler thermal cycler (BIO-RAD, Hercules, CA, USA). The PCR procedure for *orf224* amplification was: 95 °C for 5 min, followed by 30 cycles at 94 °C for 60 s, 55 °C for 60 s, and 72 °C for 60 s, and then extension at 72 °C for 10 min.

DNA fragments purification and construction of pMD18-T-orf224 vector

The PCR product of the expected size was separated on 1.2% agarose gel and purified with 3S spin agarose gel DNA purification kit (Shanghai Biocolor Bioscience & Technology Co. Ltd., Shanghai, China) following the manufacturer's instructions. Subsequently, the purified PCR product was cloned into pMD18-T simple vector. The recombined vector (pMD18-T-orf224) was transformed into *E. coli* DH5α. Positive recombinants were verified by PCR, and sequenced at Beijing Sunbiotech Co. Ltd.

Bioinformatic analysis

The nucleotide sequence and inferred amino acid sequence of orf224 of Shaan 2A and Polima CMS were aligned with CLUSTAL X software and edited with the Jalview Multiple Alignment Editor V1.8. Sites containing gaps were excluded. Cluster trees were constructed with MEGA version 3.1 (Kumar *et al*., 2004), and the neighbour-joining method was applied to *p*-distance. The secondary structure prediction was performed online (http://bioinf.cs.ucl.ac.uk/psipred/psiform.html/).

Construction of pET32a (+)-orf224 plasmid

pMD18-T-orf224 and pET32a (+) were digested by *Nco*I and *Sac*I restriction enzymes according to the manufacturer's protocol (Fermentas, Vilnius, Lithuania). The prepared *Nco*I-orf224-*Sac*I insert was ligated into the treated pET32a (+) vector using T_4 DNA ligase at 16 °C overnight. The *E. coli* DH5α and *E. coli* BL21 (DE3) competent cells were prepared by the calcium chloride method and were used for transformation. The transformed bacteria were selected on media containing ampicillin. Meanwhile, the positive colonies were further analyzed by PCR, restriction enzyme digestion, and sequencing to ensure plasmid integrity.

Induction of orf224 expression with IPTG

Three 5-mL cultures of the recombinant *E. coli* BL21 were prepared in LB (Luria-Bertani) medium containing 100 μg/mL ampicillin. The cultures were incubated at 37 °C with shaking to an OD_{600} of 0.6–0.8. Then isopropyl-β-D-thiogalactopyranoside (IPTG) was added to the first two 5-mL cultures for a final concentration of 1 mM. The third culture was saved as control. All cultures were incubated while shaking at 37 °C. Samplings were done by removing 1 mL of cultures after 3 and 4 h, respectively. The cell pellets obrained after centrifugation at 9168 × *g* for 2 min were stored at –20 °C.

Protein analysis by polyacrylamide gel electrophoresis (PAGE)

The cell pellets were resuspended in 100 μL of sterile distilled water. Then 100 μL of 2 × SDS loading buffer were added and mixed thoroughly. Samples were heated for 8 min at 100 °C and cooled immediately at 0 °C to denature the proteins and then analyzed by SDS-PAGE. The resolving and stacking gels used were 15% (w/v) and 5% (w/v), respectively (Laemmli, 1970). The gel was allowed to run for 1 h at 60 V and 2 h at 100 V, followed by staining with Coomassie Brilliant Blue R250. The low molecular weight standard was 14.4–94.0 kDa.

Results

Amplification of the orf224 gene from Shaan 2A and Polima CMS

The A_{260}/A_{280} and A_{260}/A_{230} of total DNA were assessed with a Beckman DU7500 instrument (Brea, CA, USA), and the ratios of both A_{260}/A_{280} and A_{260}/A_{230} were in the range 1.8–1.9, which indicated that the total DNA had high quality and could be used for PCR amplification.

The *orf224* gene of Shaan 2A and Polima CMS was amplified by the designed primers. The agarose gel analysis showed that both Shaan 2A and Polima CMS had a specific band of 500–750 bp length (Fig. 1a). The four purified PCR products were cloned into pMD18-T simple vector and then verified by *Nco*I and *Sac*I double digestion (Fig. 1b). The recombined vector was transformed into *E. coli* DH5α, and the positive recombinant clones were sequenced.

Sequence analysis and secondary structure of orf224 in Shaan 2A and Polima CMS

According to gene sequencing results, the length of *orf224* in both CMS lines was 675 bp (GenBank accession numbers for *orf224* of Shaan 2A and Polima CMS were EU254234 and EU254235, respectively); it encoded a polypeptide including 224 amino acids. The homology of the two genes in nucleotides and amino acids was 99.9 and 99%,

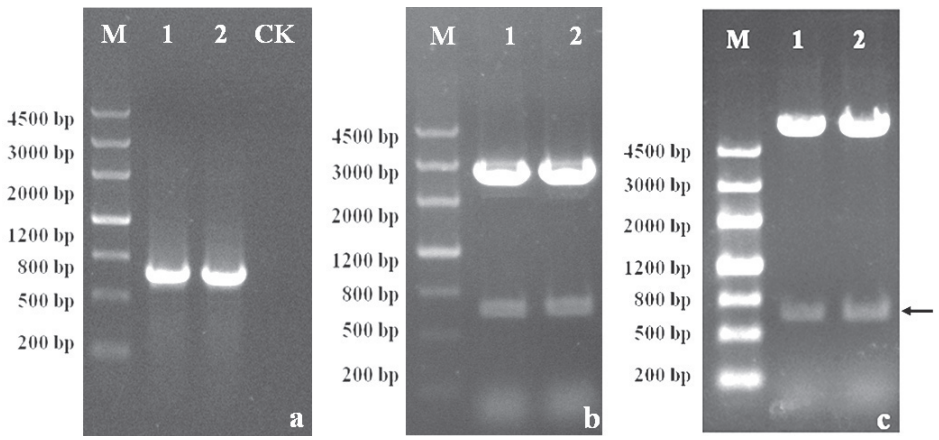

Fig. 1. PCR amplification and identification by restriction enzyme digestion. (a) PCR amplification of *orf224* gene; (b) pMD18-T-orf224 digested by *Nco*I and *Sac*I; (c) pET32a (+)-orf224 digested by *Nco*I and *Sac*I. Lane M, DNA marker; lane 1, Polima; lane 2, Shaan 2A; CK, negative control.

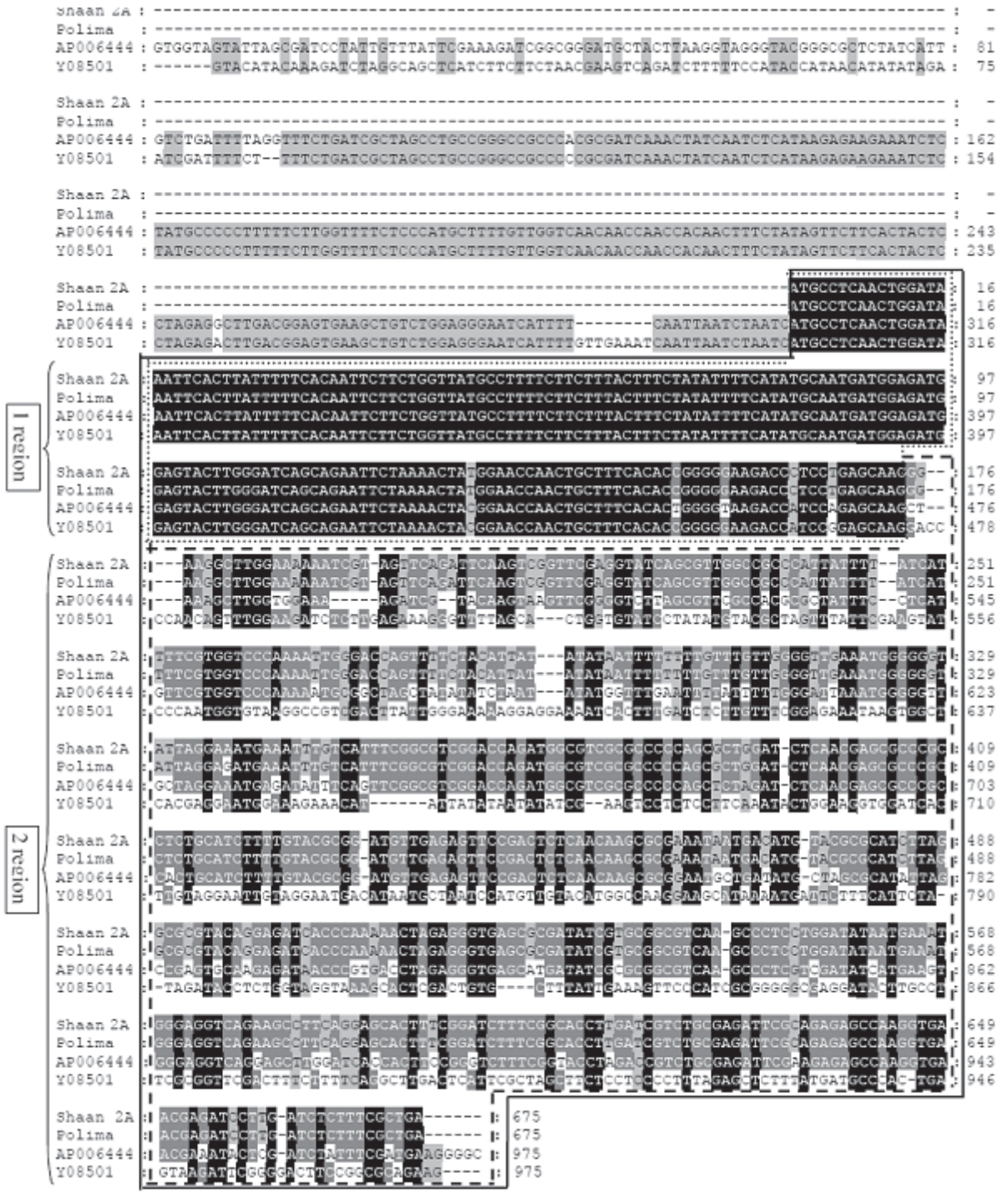

Fig. 2. Sequence alignment of the *orf224* gene in Shaan 2A and Polima CMS lines and a homologous sequence (including a upstream 300-bp fragment) in the mitochondrial genome of *B. napus* (GenBank number: AP006444) and *A. thaliana* (GenBank number: Y08501). The solid line frame indicates the orf224 gene regions.

```
Shaan 2A: MPQLDKFTYFSQFFWLCLFFFTFYIFICNDGDGVLGISRILKLWNQLLSHRGKTLLSKGR :  60
Polima  : MPQLDKFTYFSQFFWLCLFFFTFYIFICNDGDGVLGISRILKLWNQLLSHRGKTLLSKGR :  60

Shaan 2A: LGKNRSSDSSRFEVSALAAHYFIIFVVPKLGPVFYIIYNFFCLLGLKWGVLGNEICHFGV : 120
Polima  : LGKNRSSDSSRFEVSALAAHYFIIFVVPKLGPVFYIIYNFFCLLGLKWGVLGDEICHFGV : 120

Shaan 2A: GPDGVAPPALDLNERPPLHLLYADVESSDSQQARNNDMYAHLRRVQEITQKLEGERDIVR : 180
Polima  : GPDGVAPPALDLNERPPLHLLYADVESSDSQQARNNDMYAHLRRVQEITQKLEGERDIVR : 180

Shaan 2A: RQALLDIMKWEVRSLQEHFRIFRHLDRLRDSQRAKVNEILDLFR : 224
Polima  : RQALLDIMKWEVRSLQEHFRIFRHLDRLRDSQRAKVNEILDLFR : 224
```

Fig. 3. Alignment of a deduced amino acid sequence of the *orf224* gene in Shaan 2A and Polima CMS lines. The arrow points to a different amino acid.

respectively (Figs. 2 and 3), and they had 83% homology with a related gene in the mitochondrial DNA of *B. napus* (AP006444) (Handa, 2003) and 57% homology with an associated gene in the mitochondrial genome of *A. thaliana* (Y08501) (Unseld *et al.*, 1997). Meanwhile, sequence alignment showed that *orf224* gene regions were conserved in *B. napus*, and the probable promoter regions were conserved between *B. napus* and *A. thaliana* (Fig. 4). However, the upstream regions of probable promoter regions were completely divergent from each other. The results were coincident with the experimental analysis of Handa (2003).

The orf224 protein included 224 amino acids, with a molecular mass of 26 kDa. A theoretical p*I* of 9.08 for Shaan 2A and a p*I* of 8.9 for Polima CMS were deduced in the ExPaSy web server (http://expasy.org/tools/protparam.html). There were 24 negatively charged residues (Asp + Glu) and 29 positively charged residues (Arg + Lys) of the orf224 protein of Shaan 2A CMS, but the respective numbers were 25 and 29 residues for Polima CMS. The instability index (II) was computed to be 53.17, which showed that the protein was unstable. Grand average of hydropathicity was −0.09. The secondary structure had seven

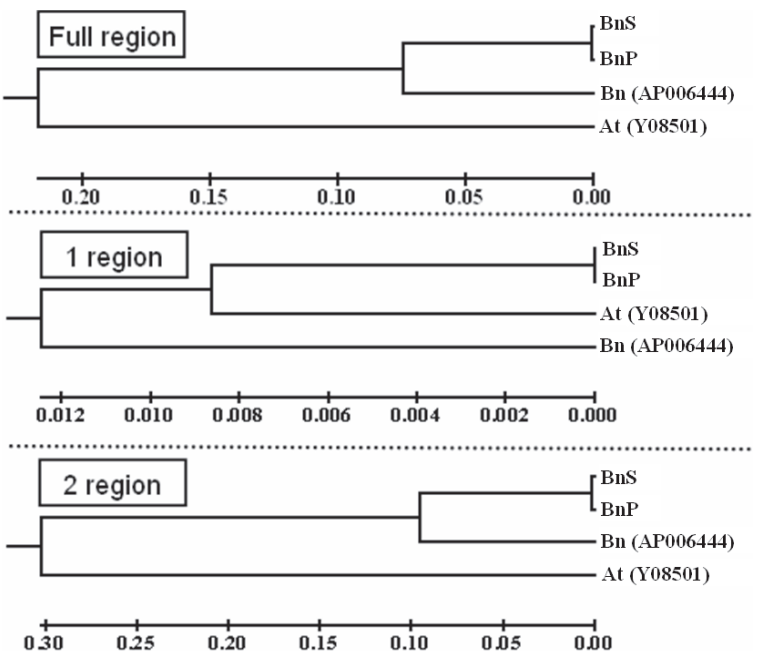

Fig. 4. Cluster tree based upon the *orf224* sequences from full region, 1 region and 2 region of BnS (*B. napus* Shaan 2A), BnP (*B. napus* Polima), Bn (AP006444) (*B. napus*) and At (Y08501) (*A. thaliana*) in MEGA format. The numbers below the line are branch lengths.

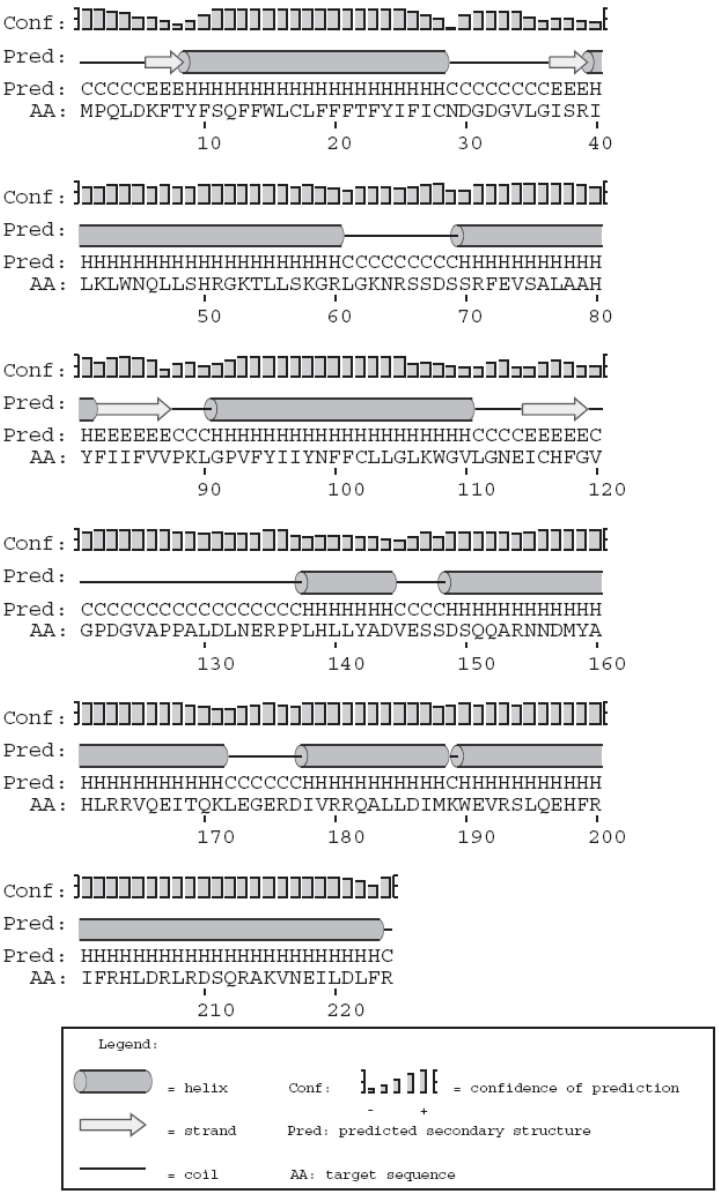

Fig. 5. The secondary structure of the orf224 protein.

α-helices, four extended strands, and eight random coils. However, no β-turn was observed (Fig. 5).

Construction of recombinant plasmid and protein expression analysis

The expression vector pET32a (+) was used to construct pET32a (+)-orf224. *Nco*I and *Sac*I cloning sites were chosen for insertion. The constructed plasmid was amplified in *E. coli* DH5α and expressed in *E. coli* BL21 (DE3). Ligation of the *orf224* gene insert into pET32a (+) and subsequent transformation of *E. coli* DH5α yielded ampillicin-resistant colonies. The selected colonies were purified to obtain a plasmid, and then double digests on the plasmid with *Nco*I and

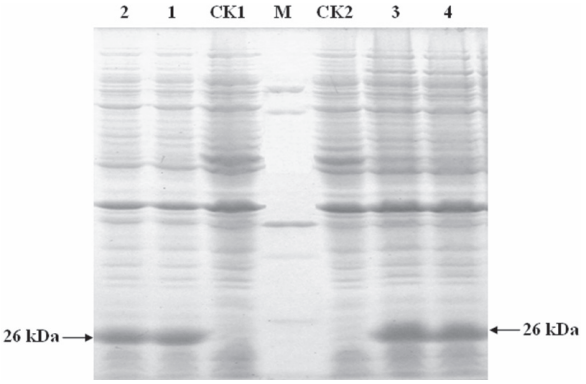

Fig. 6. SDS-PAGE analysis of *orf224* gene expression from the recombinants in *E. coli* BL21 (DE3) induced by 1 mM IPTG. Lanes 1 and 2, expression induction of pET32a (+)-orf224 of Shaan 2A at 37 °C by addition of 1 mM IPTG for 3 h and 4 h; lanes 3 and 4, expression induction of pET32a (+)-orf224 of Polima at 37 °C by addition of 1 mM IPTG for 3 h and 4 h; lanes CK1 and CK2, negative control of pET32a (+); lane M, protein molecular weight standard marker (from up to down are 94, 66.2, 45, 33, 26 kDa, respectively). Expression products are indicated by arrows.

*Sac*I determined the presence of the *orf224* gene (Fig. 1c).

Plasmid pET32a (+)-orf224 was transformed into *E. coli* BL21 (DE3) competent cells, and one of the positive colonies was cultured in LB medium, containing 100 μg/mL ampicillin, and grown at 37 °C. When the OD_{600} value reached 0.6–0.8, IPTG was added to a final concentration of 1 mM, and incubation was continued for 3 and 4 h. The cells were centrifuged at 9168 × *g* for 2 min. The expression protein was analyzed by SDS-PAGE. Compared with the uninduced cells, the IPTG-induced cells had an additional protein band of 26 kDa on SDS-PAGE gel (Fig. 6), which corresponded to the calculated molecular weight of the protein. The amount of additional protein induced after 3 and 4 h was equal.

Discussion

CMS occurs widely in higher plants and is widely used in plant heterosis. Much research has revealed that mitochondria have the closest relationship with CMS (L'Homme and Brown, 1993; L'Homme *et al*., 1997; Schnable and Wise, 1998; Song and Hedgcoth, 1994). Erickson *et al*. (1986) found a remarkable difference between mtDNA of Polima CMS and of cam CMS. Further research revealed that some regions of the mitochondrial genome were absolutely necessary to CMS (Schnable and Wise, 1998). It was shown that the *orf224/atp6* region was correlated with the Polima CMS (Wang *et al*., 1995) and the *orf222* was related to nap CMS (L'Homme *et al*., 1997); the homology of nucleotides and proteins of these two regions was 85 and 79%, respectively (L'Homme *et al*., 1997). *orf138* was only found in Ogu cytoplasm (Bellaoui *et al*., 1997). Wan *et al*. (2007) compared hau CMS with pol CMS, nap CMS, ogu CMS, and tour CMS in the fields of genetics, cytology, and molecular biology. They found that the restorer and maintainer relationship of hau CMS differed from other CMS lines, and they could distinguish between other five CMS lines using RFLP. Shaan 2A CMS and Polima CMS have the same restorer lines and maintainer lines. Many researchers thought that these two CMS belong to the same line, according to research on the general shape, cytology and restriction maps (Liu *et al*., 1998). Wang *et al*. (2002) analyzed the sequence of the *orf224* gene of Shaan 2A CMS and Polima CMS, and found only one nucleotide difference. However, Lin *et al*. (2006) found differences in five bases between Shaan 2A CMS and Polima CMS.

The present research on the sequence of the *orf224* gene found only one base difference, which led to one amino acid change. The results confirmed that Shaan 2A and Polima CMS were not of the same CMS. The secondary structure of the orf224 protein was predicted, and it was observed that the orf224 protein was expressed immediately after induction with IPTG in *E. coli* BL21 (ED3). Although this is a preliminary conclusion, it is very important for studying the relationship between the *orf224* gene and CMS.

Acknowledgements

This work was supported by the National High Technology Research and Development Program of China (2009AA101105), the New Century Talents Support Program of the Ministry of Education of China (M2009059), and the National Science Foundation of Shaanxi Province (2009JM3001-4).

Bellaoui M., Pelletier G., and Budar F. (1997), The steady-state level of mRNA from the Ogura cytoplasmic male sterility locus in *Brassica* cybrids is determined post-transcriptionally by its 3' region. Embo J. **16**, 5057–5068.

Erickson L., Grant I., and Beversdorf W. (1986), Cytoplasmic male sterility in rapeseed (*Brassica napus* L.). Theor. Appl. Genet. **72**, 151–157.

Fu T. D. (1990), General statement on the studies and utilization of rapeseed heterosis in China. Crop Res. **4**, 1–4.

Gourret J. P., Delourme R., and Renard M. (1992), Expression of ogu cytoplasmic male sterility in cybrids of *Brassica napus*. Theor. Appl. Genet. **83**, 549–556.

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 genome of rapeseed and *Arabidopsis thaliana*. Nucleic Acids Res. **31**, 5907–5916.

Handa H. and Nakajima K. (1992), Different organization and altered transcription of the mitochondrial atp6 gene in the male sterile cytoplasm of rapeseed (*Brassica napus* L.). Curr. Genet. **21**, 153–159.

Horlow C., Defrance M. C., Pollien J. M., Goujaud J., Delon R., and Pelletier G. (1992), Transfer of cytoplasmic male sterility by spontaneous androgenesis in tobacco (*Nicotiana tabacum* L.). Euphytica **66**, 45–53.

Kumar S., Tamura K., and Nei M. (2004), MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform. **5**, 150–163.

Laemmli U. K. (1970), Cleavage of structural proteins during assembly of head of bacteriophage T_4. Nature **227**, 680–685.

L'Homme Y. and Brown G. G. (1993), Organizational differences between cytoplasmic male sterile and male fertile *Brassica* mitochondrial genomes are confined to a single transposed locus. Nucleic Acids Res. **21**, 1903–1909.

L'Homme Y., Stahl R. J., Li X.-Q., Hameed A., and Brown G. G. (1997), *Brassica nap* cytoplasmic male sterility is associated with expression of a mtDNA region containing a chimeric gene similar to the *pol* CMS-associated *orf224* gene. Curr. Genet. **31**, 325–335.

Li D. R. (1980), Preliminary report on three-line breeding of *Brassica napus*. J. Shanxi Agric. Sci. **1**, 26–29.

Li M. T., Liu J. M., Wang Y. T., Yu L. J., and Meng J. L. (2007), Production of partial new-typed *Brassica napus* by introgression of genomic components from *B. rapa* and *B. carinata*. J. Genet. Genomics **34**, 460–468.

Lin B. G., Huang H., Zhang L., and Zhang M. L. (2006), Sequence comparison on gene orf224 of Polima CMS and Shaan2A CMS in *Brassica napus* L. Sci. Agric. Sin. **39**, 1282–1286.

Liu G. S., Zhao H., Liu J., and Cao M. Q. (1998), Application of molecular marker in *Brassica* SSP. Chin. Bull. Bot. **15**, 67–72.

Pascal T. and Francoise B. (2004), Unveiling the molecular arms race between two conflicting genomes in cytoplasmic male sterility? Trends Plant Sci. **9**, 568–570.

Schnable P. S. and Wise R. P. (1998), The molecular basis of cytoplasmic male sterility and fertility restoration. Trends Plant Sci. **3**, 175–180.

Shiga T. and Baba S. (1973), Cytoplasmic male sterility in oilseed rape, *B. napus* L. and its utilization to breeding. Jpn. J. Breed. **23**, 187–197.

Singh M. and Brown G. G. (1991), Suppression of cytoplasmic male sterility by nuclear genes alters expression of a novel mitochondrial gene region. Plant Cell **3**, 1349–1362.

Singh M., Hamel M., Menassa R., Li X. Q., Jean Y. M., Landry B. S., and Brown G. G. (1996), Nuclear genes associated with a single *Brassica* CMS restorer locus influence transcripts of three different mitochondrial gene regions. Genetics **143**, 505–516.

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

Thompson K. F. (1972), Cytoplasmic male sterility in oilseed rape. Heredity **29**, 253–257.

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

Wan Z. J., Fu T. D., and Tu J. X. (2007), Genetic classification of a newly identified cytoplasmic male sterility *hau* CMS system in *Brassica napus* L. In: The 12th International Rapeseed Congress "Sustainable development in cruciferous oilseed crops production", Wuhan, March 26–30, pp. 18–24.

Wang H. M., Ketela T., Keller W. A., Gleddie S. C., and Brown G. G. (1995), Genetic correlation of the orf224/atp6 gene region with Polima CMS in *Brassica* somatic hybrids. Plant Mol. Biol. **27**, 801–807.

Wang Y. F., Ma S. M., Wang M., Zheng X. Q., Gu M., and Hu S. W. (2002), Sequence analysis of the gene correlated with cytoplasmic male sterility (CMS) in rapeseed (*Brassica napus*) Polima and Shaan 2A. Chin. Sci. Bull. **47**, 122–126.

Yang G. S. and Fu T. D. (1991), A preliminary study on the restoring-maintaining relationship in rapeseed (*Brassica napus* and *Brassica campestris*). Acta Agron. Sin. **17**, 151–156.

Yang G. S., Fu T. D., and Brown G. G. (1998), The genetic classification of cytoplasmic male sterility systems in *Brassica napus* L. Sci. Agric. Sin. **31**, 27–31.

Yuan M., Yang G. S., Fu T. D., and Li Y. (2003), Transcriptional control of *orf224/atp6* by pol CMS restorer *Rfp* gene in *Brassica napus* L. Acta Genet. Sin. **30**, 469–473.