Identification of a candidate restorer gene for Nsa cytoplasmic male sterility

Jianyi HAO, Yunchang LI, Desheng MEI, Jia LIU,Yingde LI, Yusong XU, Qiong HU*

(Oil Crops Research Institute, Chinese Academy of Agricultural Sciences, National Center for Oil Crops Improvement, Key Laboratory for Biological Sciences of Oil Crops, Ministry of Agriculture, Wuhan 430062, China)

*Email: huqiong@oilcrops.cn

Abstract: CMS (cytoplasmic male sterility) system is of considerable value in facilitating efficient hybrid seed production. *Nsa* CMS is a novel type of rapeseed alloplasmic CMS developed by somatic cell fusion of *B. napus* and *Sinapis arvensis*. Previous study has revealed the distinctness of Nsa cytoplasm from the cytoplasm of other male sterility systems. A candidate gene was identified based on the homological sequence of known restorer genes in crops. Using Smart Race technology, the full length cDNA sequence of this gene was obtained. Sequence analysis showed that the gene belongs to *PPR* (pentatricopeptide repeat) gene family. This gene encodes 618 amino acids, containing 15 PPR motifs, with a mitochondrion transit signal located at the N terminal. The gene was highly homologous with *Rfo* restorer gene of *Ogu* CMS system. A gene specific marker developed based on the DNA sequence of this gene was closely linked to the fertility trait in *Nsa* CMS fertility segregation population. Genomic DNA fragment could only be amplified in the fertility restored plants, *Nsa* CMS restorers and *S. arvensis* parental accession by the gene specific primers. RT-PCR analysis showed that the candidate gene expressed only in flower buds, not in flower petals or leaves. Based on the sequence information, the specificity of occurrence and expression, it is highly possible that this gene is a candidate gene for *Nsa* CMS fertility restoration.

Key Words: alloplasmic cytoplasmic male sterility, candidate restorer gene, pentatricopeptide repeat protein

Introduction

CMS (cytoplasmic male sterility) system is of considerable value in facilitating efficient hybrid seed production. *Nsa* CMS is a novel type of rapeseed alloplasmic CMS developed by somatic cell fusion of *B. napus* and *Sinapis arvensis*. Previous study has revealed the distinctness of Nsa cytoplasm from the cytoplasm of other male sterility systems. The combination of CMS and a nuclear gene for restoration of fertility (Rf) is essential for breeding hybrid varieties and for hybrid seed production.

So far, *Rf* genes of CMS have been cloned in several crops, such as in maize, petunia, radish, rice and sorghum. The results showed that proteins of all cloned restorer genes, expect *Rf-2* in maize, had PPR motifs. The PPR motif is presumed to play a role in binding to macromolecules such as RNA, and it has been predicted that many *Arabidopsis* PPR proteins targeted to organelles. Members of the PPR family are thought to be involved in controlling organelle gene expression by processing or editing transcripts of CMS-associated genes and it is through this mechanism that some PPR proteins participating in the suppression of CMS. The objective of this study is to clone candidate *Rf* genes for *Nsa* CMS according to conserved amino acid sequence of PPR proteins, and to provide proofs for further elucidating the mechanism of fertility restoration.

Materials and methods

Plant material: *Nsa* CMS, the maintainer of *Nsa* CMS (Zhongshuang 4), the restore lines of *Nsa* CMS (Hui1, Hui2 and Hui3), *S. arvensis* (Yeyou 18), sterile and fertile individual plants of a backcross population, which derived from Zhongshuang 4 backcrossed with fertile plants of *Nsa* CMS×Hui1 for at least four generations. All materials were from Oil Crops Research Institute, Chinese Academy of Agricultural Sciences (Wuhan, China).

DNA extraction and PCR procedure: Total DNA was extracted from young leaves using CTAB method and amplified by touchdown PCR procedure. PCR were performed in a total volume of 20μL

reaction mix, containing 10μL TaqMIX (purchased from GeneStar), 1μL DNA template, 2.5μL (10μM /L) each primer, 4μL ddH2O. Thermocycling of touchdown PCR was: 95 ℃ 2 min, (94 ℃ 30 s, 55 ℃ - 40 ℃ 45s, 72 ℃ 2 min), 14 cycles, annealing temperature decreased 1℃ per cycle, and then (95 ℃ 30 s, 40 ℃ 45 s, 72 ℃ 2 min), 25 cycles. PCR products were analyzed on 2% (g / v) agarose gel and EB-stained.

Sequence analysis of PCR products: PCR amplification products of expected size were gel purified, ligated to pGEM-TEasy vector (Promega), and transformed into DH5α competent cells (Japan). Positive clones were sequenced by Shanghai Invitrogen.

Rapid amplification of cDNA ends analysis: Total RNA was extracted following the Trizol method (Invitrogen). Rapid amplification of cDNA ends (RACE) reactions were performed using the GeneRacer kit (Invitrogen) according to the procedures suggested in the specific manual. To get full sequence of the gene, 5' and 3' Race Primers were designed on the base of sequenced coding fragments. These primers were: 3'RACE GSP: GGC TAA GCA GAT GAT GGA CCT GAT GGC TAG CAA GGG, 5'RACE GSP: ATG TCC ACG AGA GGG GTG GTT GCC AAT ACA. Specific primers were designed on the base of 5'UTR and 3'UTR to amplify the full gene. Primers were synthesized by Shanghai Sangon.

Results

Specific amplification by degenerate primers and the characteristics of specific fragment: The comparison of amino acid sequences of two Arabidopsis PPR genes and Ogu CMS Rfo gene revealed that there were highly conserved regions among these three genes. Three forward and three reverse degenerate primers were designed depending on conserved amino acid regions (Fig. 1). Nine pairs of primer combination were used to amplify specific fragments. One of the primer combination (Forward primer: TTY GTN AAG GAR GGN AAG CT, Reverse primer: DAT RAG NGT RTT RTA NGT AAC) amplified specific fragments about 300bp in fertile plant and Yeyou 18, but not from Nsa CMS, Zhongshuang 4 and sterile plant (Fig. 2). DNA sequences of the specific fragment amplified from fertile plant and Yeyou 18 were both 309 bp in length and had 98% homology. The highest homology sequence on NCBI with the fragment is P2, a clone of a Brassica campestris accession which possesses the restoration of the fertility of Pol CMS. The 103 amino acid sequence was obtained by EBI transeq and it contained two PPR motifs (Fig. 3). The translated protein sequence was also submitted to NCBI BLAST for searching homology. BLAST results revealed that the homologous were all from plant restorer genes, including restorer genes of Ogu CMS in turnip and B. napus, restorer gene of Pol CMS in B. napus, restorer genes of rice CMS and petunia CMS, except some PPR genes in Arabidopsis.

Characterization of the full length cDNA: Using Smart Race technology, the full length cDNA sequence of this gene was obtained. Sequence analysis showed that the gene belongs to *PPR* gene family. This gene encodes 618 amino acids and thus was named *PPR-618*, containing 15 PPR motifs, with a mitochondrion transit signal located at the N terminal.The gene was highly homologous with *Rfo* restorer gene of *Ogu* CMS system. To amplify full length, forward primers and reverse primers were designed on the base of 5'UTR and 3'UTR. Using one of the primer combinations (Forward: ACG CAC TTT GCT TCG TGT CTT GTG TTC T, Reverse: TCG AAT TGC AAA CGC ACT TTG CTT CGT G), the 2.1kb DNA fragment could only be amplified in the fertility restored plants, *Nsa* CMS restorers and *S. arvensis* parental accession. It was absent from *Nsa* CMS, Zhongshuang 4 (maintainer of *Nsa CMS*), sterile plant and other CMS of *B. napus* (Fig. 4). Because the 2.1kb fragment was closely linked to the fertility trait in *Nsa* CMS fertility segregation population, it could be utilized as a specific marker for *Nsa* CMS restorer gene.

RT-PCR analysis on different tissue: RT-PCR analysis showed that the candidate gene expressed only in flower buds of fertile plants, not in flower petals or leaves. This result further indictes that *PPR-618* is related to fertility restoration .

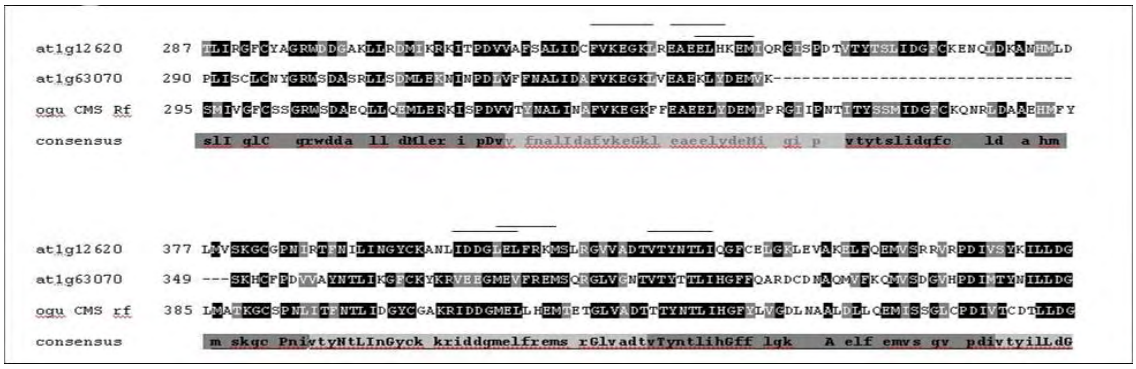

Fig.1 Alignment of PPR proteins and the conserved regions

Note：The figure was a part of the whole alignment, sequences with different colors indicate conserved PPR motif. Lines show the positions for primers.

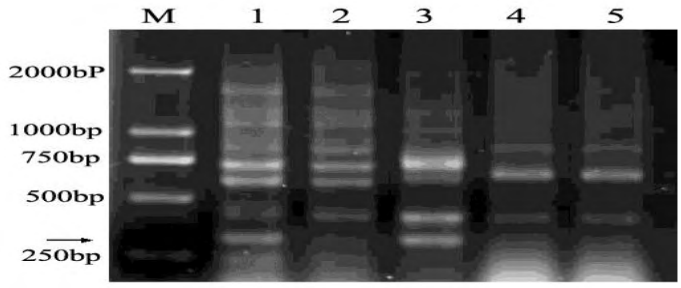

Fig. 2 Specific amplification of targeted fragment by degenerate primers

Note: 1 Fertile plant; 2 Sterile plant; 3 Yeyou 18 (*S. arvensis*); 4 Zhongshuang 4 (*B. napus*, maintainer); 5 Sterile line. Arrow indicates specific fragment.

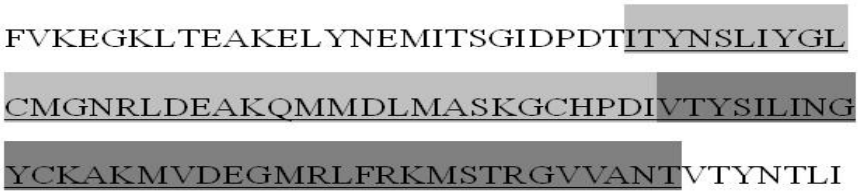

Fig. 3 PPR motifs predicted by SMART based on the specific fragments

Note: Shadowed regions are PPR motifs.

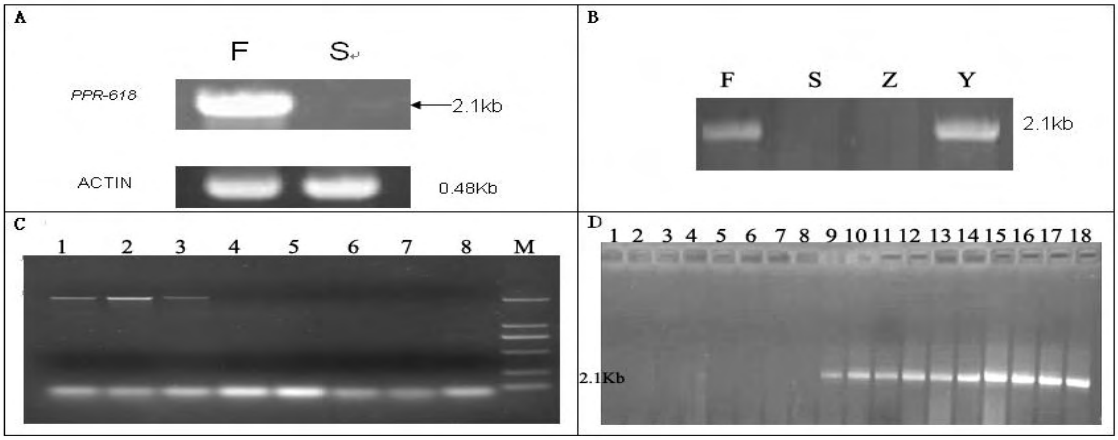

Fig. 4 Specific amplification of *PPR-618* in different plants.

A: RT PCR of PPR-618 in fertile and sterile plants. F, fertile plants; S, sterile plants; ACTIN: House keep gene as a control. B: PCR of PPR-618 of genomic DNA from fertile, sterile, *B. napus* and *S. arvensis* plants. F, fertile plants; S, sterile plants, Z: Zhongshuang 4, the maintainer of *Nsa* CMS; Y: Yeyou 18. C: PCR of PPR-618 of genomic DNA from different lines. 1-3, Three different restorer lines of *Nsa* CMS; 4, Restorer of *Pol* CMS; 5, Maintainer of *Pol* CMS; 6-7: Two restorer lines of *ogu* CMS; 8.Maintainer of *Nsa* CMS; M: DNA marker-DL2000. D: PCR of PPR-618 of genomic DNA from fertility segregation population. 1-8, Sterile individuals; 9-18, fertile individuals.

.

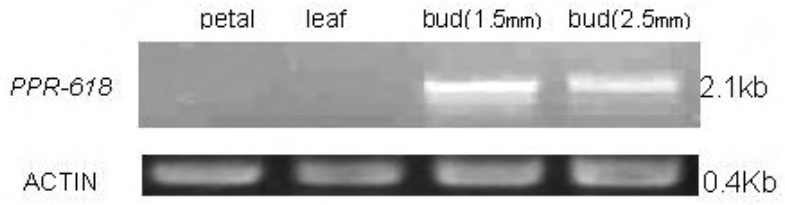

Fig. 5 Special expression analysis of *PPR-618* in different tissue

Discussion

Using degenerate primers and Smart Race technology, the full length sequence of *PPR-618* was obtained. Sequence analysis showed that it contained 15 PPR motifs, with a mitochondrion transit signal located at the N terminal. The gene was highly homologous with *Rfo* restorer gene of *Ogu* CMS system. RT-PCR analysis showed that the candidate gene expressed only in flower buds, not in flower petals or leaves. It is highly possible that *PPR-618* is a candidate gene for *Nsa* CMS fertility restoration. However, the molecular mechanism underlying the candidate gene, and the cytoplasmic–nuclear interaction in *Nsa* CMS system remains unclear. On this basis, further study for specificity expression model of the genes in tissues by in situ hybridization and gene function verification by transgenic complementation experiment is underway.

Acknowledgements

This work was supported by Chinese National 863Hi-tech R & D program (2011AA10A104) and National Sciences Foundation program (31000725).

References are omitted due to space limitation.