




# Fine mapping of a male sterility gene *ms-3* in a novel cucumber (*Cucumis sativus* L.) mutant

Yike Han<sup>1,2</sup> · Fengyue Zhao<sup>3</sup> · Shang Gao<sup>3</sup> · Xianyun Wang<sup>3</sup> · Aimin Wei<sup>2</sup> · Zhengwu Chen<sup>2</sup> · Nan Liu<sup>2</sup> · Xueqiang Tong<sup>3</sup> · Xinmeng Fu<sup>3</sup> · Changlong Wen<sup>4</sup> · Zhenxian Zhang<sup>1</sup> · Ningning Wang<sup>3</sup> · Shengli Du<sup>2</sup> 

Received: 20 March 2017 / Accepted: 26 October 2017 / Published online: 13 November 2017  
© The Author(s) 2017. This article is an open access publication

## Abstract

**Key message** The cucumber male sterility gene *ms-3* was fine mapped in a 76 kb region harboring an *MMD1*-like gene *Csa3M006660* that may be responsible for the male sterile in cucumber.

**Abstract** A cucumber (*Cucumis sativus* L.) male sterile mutant (*ms-3*) in an advanced-generation inbred line was identified, and genetic analysis revealed that the male sterility trait was controlled by a recessive nuclear gene, *ms-3*, which was stably inherited. Histological studies suggested that the main cause of the male sterility was defective microsporogenesis, resulting in no tetrad or microspores being formed. Bulked segregant analysis (BSA) and genotyping of an F<sub>2</sub> population of 2553 individuals were employed used to fine map *ms-3*, which was delimited to a 76 Kb region. In this region, a single non-synonymous SNP was found in the *Csa3M006660* gene locus, which was predicted to result in an amino acid change. Quantitative RT-PCR analysis of *Csa3M006660* was consistent with the fact that it plays a role in the early development of cucumber pollen. The protein encoded by *Csa3M006660* is predicted to be homeodomain (PHD) finger protein, and the high degree of sequence conservation with homologs from a range of plant species further suggested the importance of the *ms-3* non-synonymous mutation. The data presented here provide support for *Csa3M006660* as the most likely candidate gene for *Ms-3*.

Communicated by Michael J. Havey.

Yike Han and Fengyue Zhao have contributed equally in this manuscript.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00122-017-3013-2>) contains supplementary material, which is available to authorized users.

✉ Shengli Du  
13512059996@163.com

<sup>1</sup> Beijing Key Laboratory of Growth and Developmental Regulation for Protected Vegetable Crops, College of Horticulture, China Agricultural University, Beijing 100193, China

<sup>2</sup> State Key Laboratory of Vegetable Germplasm Innovation, Tianjin Key Laboratory of Vegetable Breeding Enterprise, Tianjin Kernel Cucumber Research Institute, Tianjin 300192, China

<sup>3</sup> College of Life Sciences, Nankai University, Tianjin 300071, China

<sup>4</sup> Beijing Vegetable Research Center (BVRC), Beijing Academy of Agricultural and Forestry Sciences, Beijing Key Laboratory of Vegetable Germplasm Improvement, Beijing 100097, China

## Introduction

In the context of plant breeding, the phenomenon of male sterility represents a powerful means to improve hybrid seed production and to protect the commercial value of parental lines. It can greatly increase the effectiveness of F<sub>1</sub> hybrid seed production without manual pollination, and reduce the production cost dramatically. In general terms, male sterility can be divided into cytoplasmic male sterility (CMS) and genic male sterility (GMS). CMS, which is maternally inherited, is a result of mitochondrial gene function, and fertility can be restored by nuclear restorer-of-fertility (Rf) genes that can be introduced by crosses with restorer lines (Hanson and Bentolila 2004; Horn 2006). In contrast, GMS is caused by recessive or dominant nuclear genes (Ke et al. 2005). The majority of sterile crop plants have arisen from spontaneous mutants, although a few have been created by mutagen treatments (Kaul 1988; Budar and Pelletier 2001).

In breeding programs, CMS is preferred to GMS, because it allows easier identification of maintainers of male sterility. CMS has been reported in a large number of plant species and has been important in the development

of rice (*Oryza sativa*; Virmani 1994), maize (*Zea mays*), and wheat (*Triticum aestivum*) (Budar and Berthomé 2007), cotton (*Gossypium hirsutum*; Subhash 2015), and Brassicaceae crops (Yamagish and Bhat 2014). However, there are some disadvantages to this approach, such as sensitivity to temperature, incomplete sterility, and a limited availability of restorer lines (Fu and Tu 2002). GMS, in contrast, allows stable and complete male sterility and the genes underlying male sterility are relatively easy to transfer. Production of hybrids using GMS has been reported in rice (Kim et al. 2007), maize (Wu et al. 2016), cotton (Subhash 2015), tomato (*Solanum lycopersicum*; Atanassova 2000), and garlic (*Allium sativum*; Mayer et al. 2013), among others. One major drawback of the GMS system is that the progeny of maternal parents segregates as 50% male fertile and 50% male sterile plants, and this requires the eradication of male fertile plants from the maternal line. This problem can be addressed through marker-assisted selection (MAS), which facilitates the identification of different genotypes in segregating generations through the detection of molecular markers (Tanksley et al. 1989). Different types of genetic markers, closely linked to a GMS locus, have been used to identify genotypes in breeding populations, such as sequence tagged site (STS) markers in Chinese cabbage (*Brassica rapa*; Ying et al. 2003), cleavage amplified polymorphic sequence (CAPS) marker in chili pepper (*Capsicum annuum*; Lee et al. 2010), sequence characterized amplified regions (SCAR) and CAPS markers in lettuce (*Lactuca sativa*; Hayashi et al. 2011), and single-nucleotide polymorphism (SNP) markers in cotton (Feng et al. 2015).

Cucumber (*Cucumis sativus* L.), which belongs to the Cucurbitaceae family, is one of the most economically important vegetable crops worldwide (Li et al. 2013). For the flowers of cucumber which was monoecious and unisexual, male sterility systems can also be widely used in the breeding of cucumber. To date, four main types of male sterility have been reported in cucumber: (1) an apetalous sterile mutant (*ap*); (2) a pleiotropic pollen-aborted mutant (*ms-1*); (3) an aborted male flower genotype (*ms-2*); and (4) a closed-flower genotype (Qi Zhang et al. 1994). However, these cases of male sterility have not been used in hybrid cucumber seed production; because their inheritance is determined by nuclear genes and they are associated with undesirable traits, such as missing corolla, malformed ovaries, and closed female flowers (Grimbley 1980; Hutchins 1936).

There exists gynoeious sex expression in cucumber both genetically and chemically induction which provides alternative ways for commercial hybrids production. However, most current Chinese cucumber parent lines are genetically monoecious, and it is difficult to achieve 100% female flowers by chemical induction, and additional male flowers'

removal is needed. Gynoeious parental lines cannot be protected, because it can be selfed by chemical male flower induction.

We previously identified a novel male sterile cucumber mutant in the advanced-generation inbred line, YL-5, and here, we report that the inheritance of this male sterile trait is conferred by a single recessive nuclear gene. We designated this male sterile gene *ms-3*. Anatomical and histological studies showed that the main reason for the male sterility was the failure of meiosis, leading to no pollen being formed. Based on fine mapping, bulked segregant analysis (BSA), and genotyping of a large F<sub>2</sub> population using a Kbioscience Allele-specific Polymorphism (KASP) assay, the candidate *ms-3* gene, responsible for the male sterility, was identified. Finally, we mapped *ms-3* to a 76 Kb genomic DNA region.

## Materials and methods

### Plant material

In June 2012, 36 male sterile cucumber plants that were spontaneous mutation were identified in the Northern Chinese cucumber fertile lines, YL-5, at the experimental farm of Tianjin kernel Cucumber Research Institute in Tianjin, China. Morphological characteristics, including plant height, leaf shape, and fruit traits, were documented. YL-5 male sterile plants were tagged and hand-pollinated with pollen from fertile YL-5 wild-type plants to generate an F<sub>1</sub> generation. In the autumn of 2012, the F<sub>1</sub> plants were selfed to generate the F<sub>2</sub> generation. At the same time, a test crossing was performed, where the YL-5 male sterile plants were hand-pollinated with the heterozygous fertile F<sub>1</sub> plants, generating C<sub>1</sub> seeds. The F<sub>2</sub> seeds and C<sub>1</sub> seeds were grown in the spring of 2013. Male flower fertility in the F<sub>1</sub>, F<sub>2</sub>, and C<sub>1</sub> generations was determined. All plants were grown in the greenhouses in Tianjin, China and were under long day-light exposure. The day and night temperature of the greenhouse was controlled at 25–30 and 15–18 °C, respectively.

### Stereo microscopic imaging of stamens

The male flower buds were divided into six stages, according to bud size: stage I, 2–3 mm; stage II, 3–6 mm; stage III, 6–9 mm; stage IV, 9–12 mm; stage V, 12–15 mm; stage VI, > 15 mm. Buds from male sterile and fertile plants were collected around 9:00 am. Corollas were carefully removed with a pair of tweezers and the stamens were then observed and photographed using an Olympus SZX16 stereoscopic microscope (Tokyo, Japan).

## Pollen observation

Sterile and fertile male flowers were collected and placed on wet filter paper in culture dishes, and then kept in an ice box. Anthers were removed and placed on a glass slide, 2–3 drops of distilled water were added, and the anthers were squashed gently with tweezers. Pollen grains were observed and photographed using an Olympus IX70 microscope (Tokyo, Japan).

## Paraffin sectioning and microscopy

Six different stages of buds from male sterile and fertile plants were collected and immersed into fresh FAA (5 mL formalin, 5 mL acetic acid, and 90 mL 70% ethanol) solution, and then kept in an ice box. Paraffin sections were made as previously described (Wang et al. 2013), The slice thickness is 8  $\mu\text{m}$ , using a Leica RM2016 microtome (Leica, Germany).

## Transmission electron microscopy (TEM)

Male flower buds (1–3 mm long) were collected from male sterile and fertile plants and placed in glutaraldehyde. Ultrathin sections were made as previously described (Vignolini et al. 2012), the slice thickness is 70 nm, and TEM images collected using a Hitachi H-7650 transmission electron microscope (Hitachi, Japan).

## Bulked segregant analysis

A cross was made between the YL-5 male sterile mutant (female parent, P<sub>1</sub>) and a genetically distinct male fertile line (D37-1, male parent, P<sub>2</sub>) to create an F<sub>1</sub> generation. An F<sub>2</sub> segregating population was then generated by F<sub>1</sub> self-pollination. Four DNA pools were constructed: the P<sub>1</sub> pool from the 20 YL-5 mutant (*ms-3/ms-3*) plants; the P<sub>2</sub> pool from the 20 D37-1 (*Ms-3/Ms-3*) plants; the male fertile (MF) pool from the 50 male fertile plants of the F<sub>2</sub> generation (*Ms-3/Ms-3* and *Ms-3/ms-3*); and the male sterile (MS) pool from the 50 male sterile plants of the F<sub>2</sub> generation (*ms-3/ms-3*). These DNA pools were sequenced using the Illumina HiSeq 2500 platform with paired-end reads of 100 bp (Illumina, USA).

## Sequencing data analysis

The raw reads from the four DNA pools were filtered and aligned to the cucumber genome sequence (Chinese long; v2) using the Burrows–Wheeler alignment tool (BWA) (Huang et al. 2009; Langmead and Salzberg 2012). GATK software was used to detect single-nucleotide polymorphisms (SNPs) and InDels (McKenna et al. 2010). The

SNP-index and the  $\Delta(\text{SNP-index})$  (DePristo et al. 2011) values were calculated to identify candidate genomic regions associated with male sterility. The  $\Delta(\text{SNP-index})$  was determined based on the difference in the SNP-index between the MF and MS pools. An average of  $\Delta(\text{SNP-index})$  of SNPs located in the given genomic interval was calculated using sliding window approach with 1 Mb window size and 10 kb increment. The  $\Delta(\text{SNP-index})$  of MF and MS pools and their corresponding SNP-index within the specified window size were plotted in a graph to generate SNP-index plots. The  $\Delta(\text{SNP-index})$  value should be significantly different from 0 if a genomic region harboring the target gene (Takagi et al. 2013). We calculated statistical confidence intervals of  $\Delta(\text{SNP-index})$  for all the SNP positions with given read depths and 99% confidence intervals were obtained. By examining the  $\Delta(\text{SNP-index})$ , the plot peak regions above the confidence value were defined as candidate regions for association with male sterility.

## SNP primer design and KASP genotyping

We employed KASP genotyping platform at the Beijing Vegetable Research Center (Beijing, China) to identify the SNP genotype in the segregation population and to construct a genetic map. According to the principle that the base type of the P<sub>1</sub> pool was the same as the MS pool and the P<sub>2</sub> pool was the same as the MF pool, all the SNP loci obtained from the sequencing data were filtered for primers design. For each SNP, two forward DNA primers were designed that matched the SNP allele. FAM and VIC fluorophores were used to label the 5' of the primers to distinguish between the different SNP genotypes, and two forward primers were used with one reverse primer (Table S1). SNP genotyping was conducted following the manufacturer's protocol (LGC genomics, UK). The "touch-down" PCR program for KASP SNP genotyping was: 95 °C for 15 min, 10 cycles at 94 °C for 20 s; 61 °C (– 0.6 °C/cycle) for 1 min; followed by 26 cycles of 94 °C for 20 s and 55 °C for 60 s.

## Mapping population

KASP was performed to identify the SNP genotypes in 2553 individuals from the F<sub>2</sub> population of the above-mentioned YL-5 mutant and D37-1 plants to construct a genetic map. Genomic DNA was extracted from young leaves using the CTAB method (Murray and Thompson 1980). To construct the genetic map, the KASP genotyping data were analyzed using the JoinMap 4.0 software (Stam 1993) with default parameters and an LOD score of 10 (Van Ooijen 2006).

## Sequence analysis and prediction of the candidate gene

The sequence and gene functional annotations were obtained from the Cucurbit Genomics Database (<http://www.icugi.org/cgi-bin/ICuGI/index.cgi>). Conserved domains were identified in the NCBI (<https://www.ncbi.nlm.nih.gov/>) and SMART (<http://smart.embl-heidelberg.de/>) databases. Amino acid sequences were aligned using DNAMAN7.0 (Lynnon Biosoft, Quebec, Canada) and ClustalX2.0 (Larkin et al. 2007) software packages.

## Extraction of nucleic acids and qRT-PCR

Total RNA was extracted using a MiniBEST Universal RNA Extraction kit (TaKaRa, Japan) and the first-strand cDNA was prepared using a PrimeScript RT Reagent Kit (TaKaRa, Japan). Quantitative RT-PCR was conducted using an SYBR Premix Ex Taq II kit (TaKaRa, Japan). The candidate gene primers for qRT-PCR were designed using Beacon Designer 7.9 software (PREMIER, Palo Alto, CA, USA) (forward primer: 5'AATCTCCTCCTCCGTGTAG3'; reverse primer: 5'CTGCGTATTCTAACCGTCTC3') and the cucumber Ubiquitin extension protein (UBI-ep) gene was used as the reference gene (forward primer: 5'CACCAAGCCCAAGAAGATC3'; reverse primer: 5'TAAACC TAATCACCACCAGC3'). Three biological and technical replicates were used for qRT-PCR. Average relative expression levels for each sample were calculated.

## Phylogenetic analysis

Phylogenetic tree was constructed using the MEGA 7 software with a bootstrap method and 1000 replications (Kumar et al. 2016).

## Results

### Genetic analysis of male sterile lines

The male sterile mutant was characterized as having a normal corolla in both male and female flowers, normal fertility in the females, but an absence of pollen from male flowers that otherwise appeared normal. In segregation groups, all the  $F_1$  plants were fertile, whereas their  $F_2$  plants exhibited a segregation of 948 male fertile and 321 male sterile plants, giving a ratio close to 3:1.  $C_1$  plants from the test crossing exhibited a segregation of 103 male fertile and 104 male sterile plants, effectively corresponding to 1:1 ratio. The results above suggested that the male

sterility trait was controlled by a pair of nuclear recessive genes.

### Observation of stamens and pollens

Male flower buds at different stages were carefully dissected and observed by stereo microscopy. In the flowering stage, stamens from the male sterile plants were smaller, shriveled, transparent, and pale green. No pollen grains were observed on the anther surface and on the glass slide (Fig. 1a-1, a-2). In contrast, anthers from the male fertile plants were larger, more plump, opaque, and pale yellow (Fig. 1b-1). Many pollen grains were observed on the surface of the anthers and on the glass slide (Fig. 1b-1, b-2).

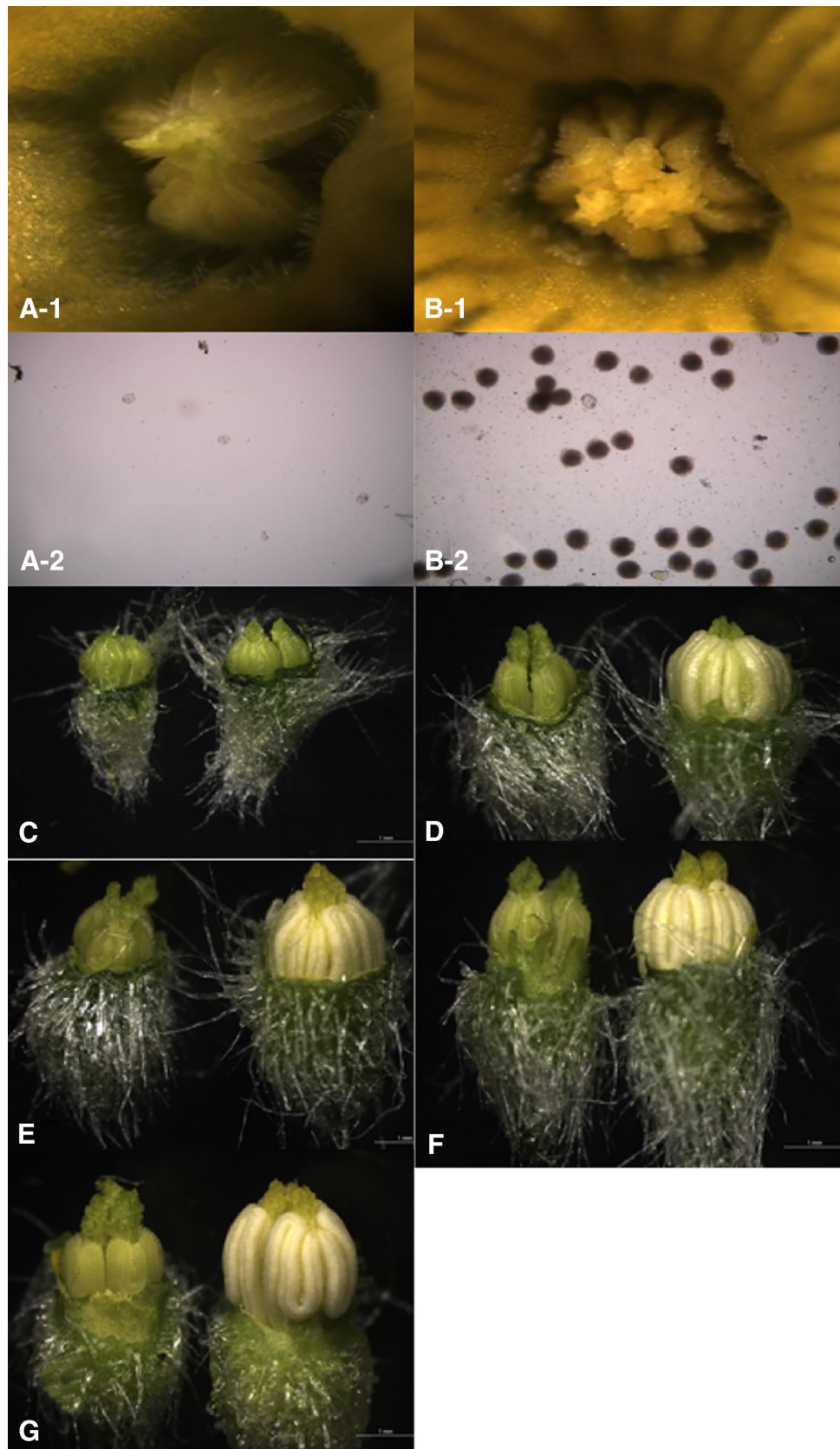
At stage I, no significant difference was found in stamens from male sterile and fertile plants; however, from stages II to V, significant differences were observed. Anthers from the male sterile plants were smaller, whereas anthers from the fertile plants were larger and plump (Fig. 1c–g).

### Microspore development

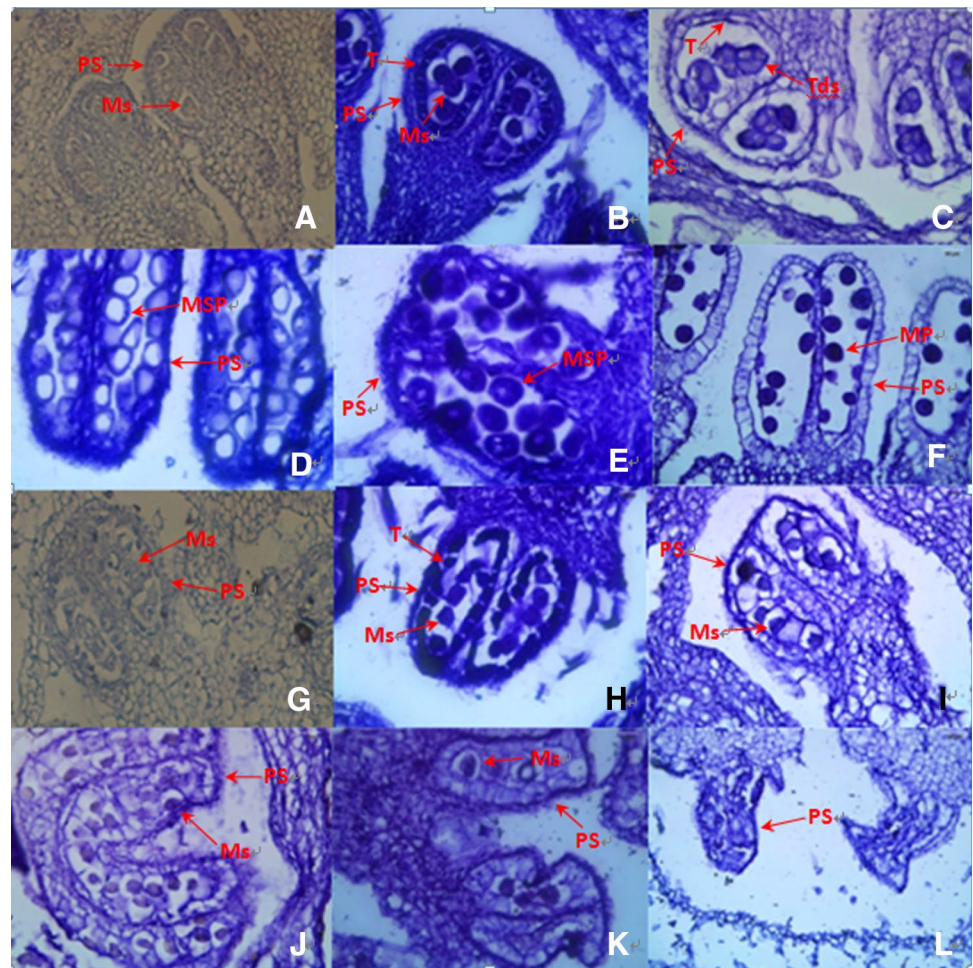
Paraffin embedded sections of anthers were observed using differential interference contrast microscopy (Fig. 2). In stage I, normal pollen sacs were observed in the fertile flowers (Fig. 2a), whereas in the male sterile flowers, pollen sacs were irregular (Fig. 2g). At stage II, the microsporocytes, the tapetal cells, and pollen sac cells were normal and well structured (Fig. 2b). In contrast, in the male sterile flowers, obvious morphological abnormalities, including irregularly shaped pollen sacs, microsporocytes and tapetal cells were observed (Fig. 2h). In stage III (Fig. 2c), normal tetrads were observed in the male fertile anthers, whereas no tetrads were observed in the male sterile anthers. The microsporocytes and pollen sacs were morphologically abnormal (Fig. 2i). From stage IV to stage VI, the tetrads developed microspores and matured pollens in fertile anthers (Fig. 2d–f). In contrast, no tetrads or microspores and pollens were observed in the male sterile plants, the tapetal cells and microsporocytes disintegrated and disappeared, and the pollen sacs were malformed and, eventually completely atrophic (Fig. 2j–l). These cytological observations revealed the obvious differences between fertile and sterile plants, and that the microsporocytes failed to produce microspores in the male sterile plants. The results further indicated that the onset of cucumber male sterility occurred in stage I.

Male flowers of 1, 2, and 3 mm in length were sectioned and observed by TEM (Figure S1). In bud length of 1 mm, normal sporogenic cells with visible nucleus were observed in both fertile and sterile flowers, and that the degree of cytoplasmic staining was similar. (Figure S1A, S1D). In bud length of 2 mm, the sporogenic cells developed to microsporocytes, and that the area around the nucleus was distinct,

**Fig. 1** Flowers of male sterile and fertile cucumber plants. **a-1** Sterile stamen and anther,  $\times 80$ ; **a-2** male sterile pollen,  $\times 200$ ; **b-1** fertile stamen and anther,  $\times 80$ ; **b-2** fertile pollen,  $\times 200$ ; **c** comparison of sterile stamen (left) and fertile stamen (right) in stage I,  $\times 20$ ; **d** comparison of sterile stamen (left) and fertile stamen (right) in stage II,  $\times 20$ ; **e** comparison of sterile stamen (left) and fertile stamen (right) in stage III,  $\times 20$ ; **f** comparison of sterile stamen (left) and fertile stamen (right) in stage IV,  $\times 20$ ; **g** comparison of sterile stamen (left) and fertile stamen (right) in stage V,  $\times 20$



**Fig. 2** Microstructural analysis of microspore development in cucumber male sterile and fertile lines. **a–f** Microspore development in male fertile plants; **g–l** microspore development in male sterile plants. *PS* pollen sac, *T* tapetum, *Ms* microsporocyte, *Tds* tetrads, *Msp* microspore, *MP* mature pollen;  $\times 500$



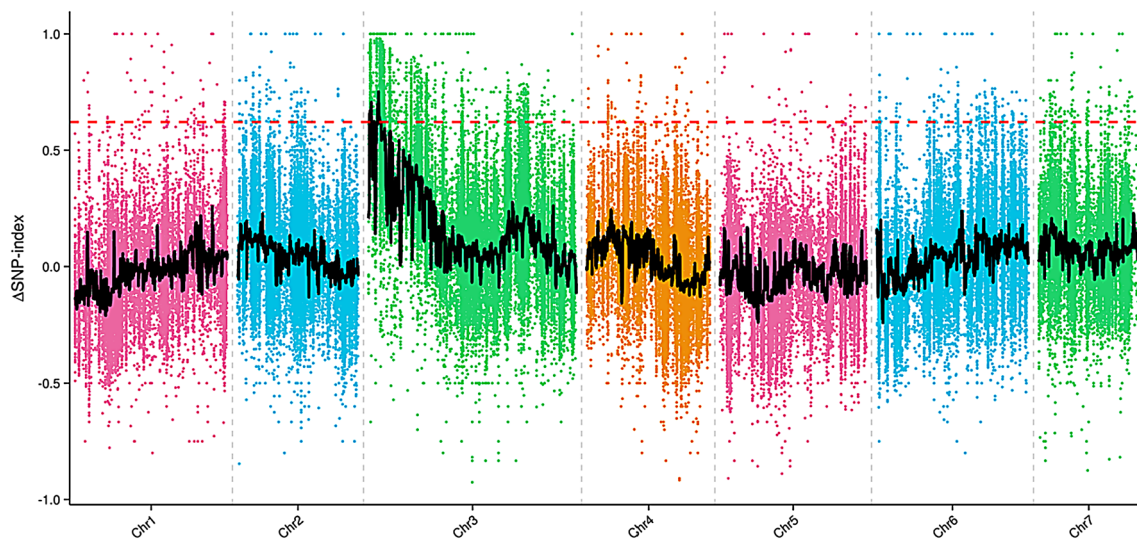
with larger nucleoli and smaller vacuoles in fertile flowers. In contrast, the area around the nucleus was indistinct, with smaller nucleoli and larger vacuoles in microsporocytes of sterile plants (Figure S1B, S1E). In bud length of 3 mm, the fertile microsporocytes developed well and were larger. In contrast, the sterile microsporocytes were seriously vacuolated and smaller (Figure S1C, S1F). The results indicated that the sporogenic cells failed to develop normal microsporocytes in the male sterile plants.

### BSA and genetic mapping

After resequencing, a total of 38.5 Gb were generated, with an average genome sequencing depth of  $19\times$  and a coverage of 98.75%. The sequence reads of the four DNA pools were aligned to the cucumber reference genome (cucumber Chinese long genome v2), and many SNPs loci and InDels were found. A total of 42,167 SNPs were found between male fertile (MF) and male sterile (MS) pool. The  $\Delta(\text{SNP-index})$  was calculated based on the SNP-index of the male fertile (MF) and male sterile (MS) pool.  $\Delta(\text{SNP-index})$  graphs were generated and the red line in

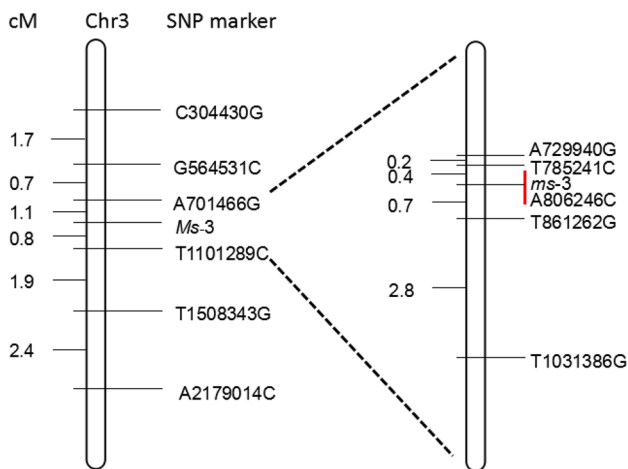
Fig. 3 shows the confidence value (99%). The  $\Delta(\text{SNP-index})$  value should be significantly different from 0 if a genomic region harboring the target gene. At 99% significance level, only two regions on the end of chromosome 3 were significantly different from 0, which spanned a total of 813 Kb (region 1 166,710–564,531, size 397 Kb; region 2: 1,954,776–2,371,279, size 416 Kb). These results indicated that *ms-3* was located at the two candidate regions.

6 SNP markers were selected from the region 166,710–2,371,279 and tested for linkage to *ms-3* in 938 individuals of the  $F_2$  population, based on the sequence data. Two SNP markers, A701466G (with four recombinants) and T1101289C (with eight recombinants), were found to be the closest flanking markers for *ms-3*. Next, 5 SNP markers between A701466G and T1101289C were selected for fine mapping of the *ms-3* in 2553  $F_2$  population and the closest flanking markers for *ms-3* were found to be T785241C (with 2 recombinants) and T861262G (with 4 recombinants) (Fig. 4). T785241C and T861262G were 1.1 cM apart, corresponding to physical distance of 76 Kb, based on the cucumber genome sequence (Chinese long; v2), and these were located in the middle of the two



**Fig. 3**  $\Delta$ (SNP-index) graph of MF (male fertile) and MS (male sterile) pool from the bulked segregant analysis (BSA). The X-axis represents the chromosome position. The Y-axis represents the  $\Delta$ (SNP-index) value. The spots represent  $\Delta$ (SNP-index) values calculated by the formula: [ $\Delta$ (SNP-index) = SNP-index\_Largest-SNP-

index\_Smallest]. The black line corresponds to the fitted value of the  $\Delta$ (SNP-index) and the red line represents the confidence value (99%). Two candidate regions (region 1 166,710–564,531, size 397 Kb; and region 2 1,954,776–2,371,279, size 416 Kb) above the confidence value were identified on cucumber chromosome 3



**Fig. 4** Fine mapping of the cucumber *ms-3* gene. Single-nucleotide polymorphism (SNP) markers were selected from the region (166,710–2,371,279), and genotyping in an F<sub>2</sub> population of 2553 individuals delimited *ms-3* to a 76 Kb region with flanking marker T785241C and T861262G, respectively, with two and four recombinants. One non-synonymous mutation was detected in *Csa3M006660* between WT and *ms-3*

candidate regions. SNP marker information is provided in Supplementary Table S1.

Using the cucumber genome database as a reference, 11 genes were predicted to be located between T785241C and T861262G. The positions and putative functions of these predicted genes are shown in Supplementary Table S2. Three SNPs and one InDel were existed between T785241C

and T861262G between male fertile (MF) and male sterile (MS) pool, however, only non-synonymous SNP was detected in *Csa3M006660* between WT (YL-5 wild type) and *ms-3* (YL-5 mutant type). Taken together, our results suggested that *Csa3M006660* was the candidate *Ms-3* gene.

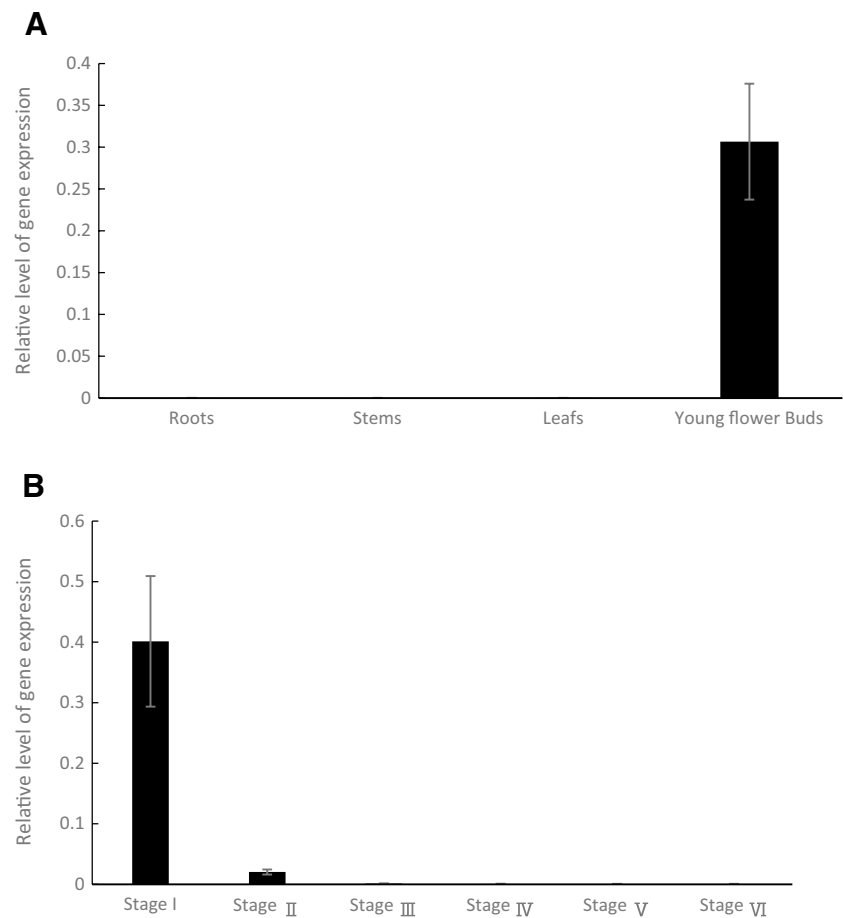
### Candidate gene screening and expressing analysis

We cloned *Csa3M006660* from both WT and *ms-3*. The gene annotation suggested the presence of 3 exons and 2 introns (Supplementary Figure S2) and the predicted coding sequence (CDS) of the cloned *Csa3M006660* cDNA was 1995 bp, with a predicted corresponding protein length of 664 amino acids. Alignment of the *Csa3M006660* gene sequence between the wild type and the mutant revealed a single non-synonymous T → G mutation in the third exon, resulting in the conversion of a tyrosine (Tyr) to an aspartic acid (Asp) at residue 420. According to the Cucurbit Genomics Database, *Csa3M006660* was predicted to encode a homeodomain (PHD) finger protein that is typically characterized as Cys4HisCys3.

To investigate the conservation of the male sterile mutant locus, 28 genetically distinct fertile cucumber inbred lines were selected and their *Csa3M006660* sequences were determined. All fertile lines shared the WT sequence in this locus (Supplementary Figure S3).

The expression of *Csa3M006660* was investigated in WT roots, stems, leaves, and young flower buds and we observed that it was only expressed in young flower buds (Fig. 5a). We then investigated expression levels in the different stages

**Fig. 5** Expression analysis of the *Csa3M006660* gene in cucumber organs. *Csa3M006660* expression in **a** roots, stems, leaves, and young flower buds of wild type (WT); and **b** different stages of flower bud development: stage I (shorter than 3 mm), stage II (3–6 mm), stage III (6–9 mm), stage IV (9–12 mm), stage V (12–15 mm), and stage VI (longer than 15 mm). Expression levels were determined using q-PCR, and values were normalized using UBI-ep as the reference gene. Error bars indicate standard deviation from three biological replicates



of flower bud development and showed that *Csa3M006660* expression in stage I (bud length 2–3 mm) was significantly higher than in the other stages (Fig. 5b).

### Phylogenetic analysis

To better understand the relationship between *Csa3M006660* protein and its close homologs, we searched public databases NCBI using BLAST with the *Csa3M006660* amino acid sequence and used an alignment of the closest homologs from 17 additional species and 8 commercial crops for a phylogenetic analysis (Fig. 6). The resulting neighbor-joining tree showed that *Csa3M006660* from cucumber grouped together with homologs from *Morus notabilis*. We found that the site where the non-synonymous mutation occurred was highly conserved and that all the proteins had a typical PHD domain, which is a Cys4HisCys3 protein conserved motif (Supplementary Figure S4), including some commercial crops, such as *Hordeum vulgare*, *Triticum aestivum*, and so on.

The results of the alignment revealed that the *Csa3M006660* protein shared 48.96% sequence identity with *Arabidopsis thaliana* MMD1, and both shared the same PHD domain and nucleotide in the site where the

non-synonymous mutation was located (Supplementary Figure S5). These data presented indicate that *Csa3M006660* is the most likely candidate gene for *Ms-3* and that a non-synonymous SNP is responsible for the male sterile phenotype.

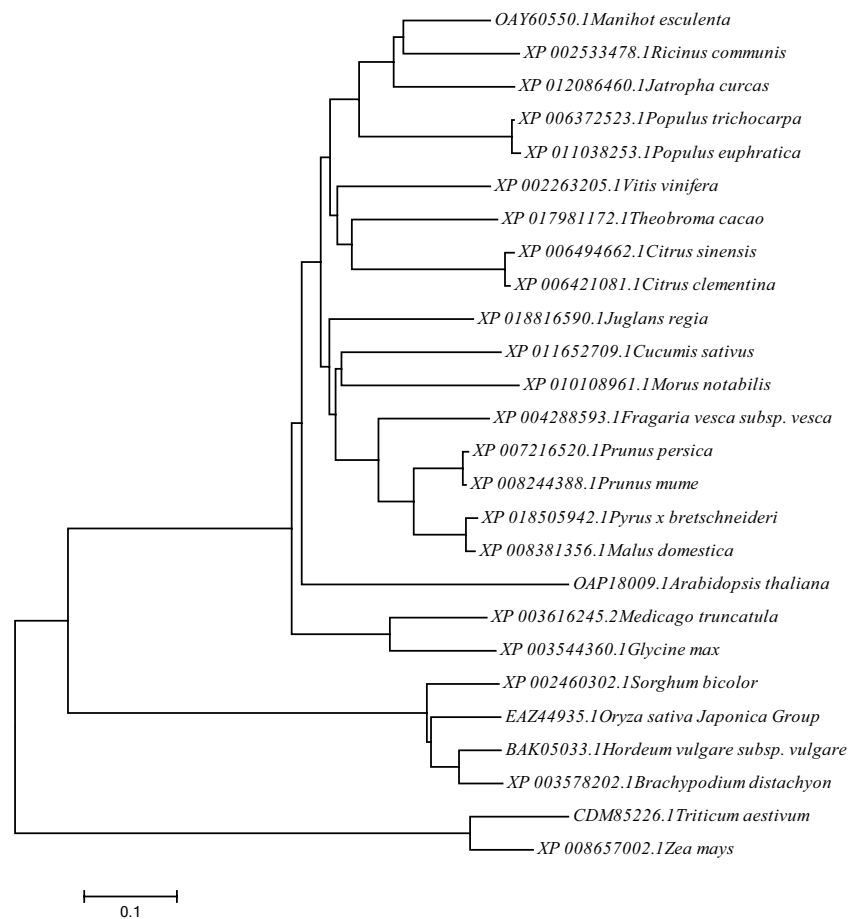
### Discussion

#### Abnormal meiosis is the cause of male sterility in *ms-3*

According to pollen abortion variants and the main features of male sterility, its underlying causes can be classified into different types, including: abnormal meiosis (Nonomura et al. 2004; Xiao 2016); abnormal callose metabolism (Ostergaard et al. 2002); early or late degeneration of tapetum cells (Li et al. 2006); critical chemical changes in pollen wall development (Jung et al. 2006); failure in anther dehiscence (Steiner-Lange et al. 2003), and other types (Kaneko et al. 2004). In this study, the male sterile sporogenous cells appeared to be normal and could generate middle layer cells, tapetal cells, and microsporocytes. However, the microsporocytes were highly vacuolated, with fewer cytoplasmic ribosomes and plastids, and could not produce tetrads and



**Fig. 6** Phylogenetic analysis of *Ms-3* and its homologs in 25 other plant species. Evolutionary relationships were inferred using the neighbor-joining method. The ID numbers refer to the gene IDs in the NCBI database and the names of the species are given. The bootstrap test values (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic relationships



microspores. Thus, the cucumber male sterile *ms-3* mutant belongs to the type of male sterility caused by abnormal meiosis. Two types of pollen abortion male sterility have been reported: (1) a pleiotropic pollen-aborted mutant (*ms-1*), whose most distinguishing external characteristic was the failure of anthesis in the staminate flowers. The pollen sterility varied from 30 to 90%, and was found to be controlled by a pair of recessive genes (Shifriss 1950); and (2) male sterile-2 (*ms-2*), whose staminate flower buds usually aborted when they were approximately three quarters of the way through development, and where pollen abortion did not occur until after the mitotic division of the pollen grain nucleus (Barnes 1960). The *ms-2* mutation was found to be in a single recessive gene (Whelan 1972). In contrast, the cucumber male sterile mutant described in this current study is completely pollen sterile, and the onset of sterility occurs early in the pollen mother cell, which is different from the above-mentioned *ms-1* and *ms-2* mutants. We designated the male sterile gene as *ms-3*.

In *A. thaliana* and rice, anther development has been well studied and many of the genes involved in meiosis have been identified. For example, rice *PAIR1*, *PAIR2*, and *PAIR3*, and *A. thaliana* *SYN1/DIF1* are important for chromosome condensation and pairing (Nonomura et al. 2004, 2006; Yuan

et al. 2009; Bai et al. 1999). *A. thaliana* *SWITCH1/DYAD* and rice *REC8* are essential for sister chromatid cohesion and bivalent formation (Mercier et al. 2003; Shao et al. 2011), and *A. thaliana* *ASK1*, *ATK1*, and *MPS1* play important roles in spindle formation and chromosome segregation (Yang et al. 1999; Chen et al. 2002; Jiang et al. 2009). In addition, *A. thaliana* *MMD1/DUET* was shown to be involved in regulating meiotic cell cycle progression (Yang et al. 2003). Mutations in these genes also cause defective meicyte development and male sterility.

### ***Csa3M006660* is the most likely candidate gene underlying the *ms-3* mutation**

Through BSA and genotyping, we identified the *ms-3* gene in a 76 Kb candidate region at the end of chromosome 3, which contained 11 predicted genes. In *Csa3M006660*, a single non-synonymous mutation was detected in the third exon. Further sequence alignment and expression analysis suggested that the non-synonymous mutation site of *Csa3M006660* is highly conserved and that the protein may play an important role in the early development of cucumber pollen. Thus, we propose that *Csa3M006660* is the candidate gene underlying the male sterility phenotype.

Amino acid alignments suggested that the Csa3M006660 protein shares 48.96% sequence identity with *A. thaliana* MMD1, and that the proteins have the same PHD domain and amino acid in the site of the non-synonymous *ms-3* mutation.

In *A. thaliana*, MMD1 (also known as DUTE) encodes a PHD finger protein required for normal male meiosis (Reddy et al. 2003), and an *mmd1* mutation was shown to cause abnormal cell death of male meiocytes (Yang et al. 2003). MMD1 plays a role in meiotic I-independent cell division by regulating the expression of two genes related to meiosis, *TDM* and *JAS*, which have critical functions in cell cycle transition and spindle organization (Andreuzza et al. 2015). The PHD finger can interact with different modified histone peptides and is required for specific binding to H3K3me2 (Patel and Wang 2013). A recent study showed that MMD1 is required for progressive compaction of prophase I chromosome to metaphase I bivalents, by ensuring the progression of male meiotic chromosome condensation, a process in which the PHD domain is essential (Wang et al. 2016). MMD1 regulates the expression of the condensin gene, *CAP-D3*, by directly binding to its promoter region in male meiocytes (Wang et al. 2016). Our analyses suggested that Csa3M006660 and MMD1 may have similar functions.

Taken together, the data presented above provide support for *Csa3M006660* as the most likely candidate gene for *Ms-3*. However, more evidence is needed to definitively demonstrate this.

### Application of the novel male sterile mutant in cucumber breeding

The traditional breeding system using GMS is very laborious and time-consuming as visible characteristics linked to male sterility have not been found so far and the phenotype is recognized only after flowering. To breed GMS cucumber maternal parent lines, consecutive backcrosses are performed to introduce the recessive *ms-3* gene. MAS has been utilized to select the GMS individuals. Based on our result, the SNP marker A806246C can be used to detect the presence of *ms-3* gene by KASP genotyping, which makes it easier and more efficient for the selection of *ms-3* individuals in the breeding population.

In GMS system, the progeny, which from the cross of homozygous recessive male sterile and heterozygous male fertile plant, will segregate 50% male fertile and 50% male sterile. This problem precludes the application of GMS system in hybrid seed production. An effective approach to overcome the problem has been reported in maize using the Seed Production Technology (SPT) (Wu et al. 2016). An SPT construct was employed in this technology, which contains a complementary wild-type male fertility gene, a pollination disruption gene, and a seed screenable marker

gene. A heterozygous SPT maintainer line could be generated by transformed the SPT construct into a homozygous recessive male sterile line. The SPT maintainer could only produce pollen grains without SPT transgenes as expression of the pollination disruption gene blocks the pollen germination. Hence, it is capable of propagating non-transgenic and homozygous nuclear male sterile lines for use as female parents in hybrid production. In cucumber, we can also construct an SPT maintainer through transgenic method as in maize. Crossing of SPT maintainer with the male sterile plants will produce homozygous male sterile lines for hybrid production. We infer that the SPT process may be the most effective, practical, and safe method to utilize the cucumber *ms-3* gene for hybrid seed production.

**Author contribution statement** YH, FZ, SG, NW, and SD designed the experiments. YH, XT, XF, and NL conducted genetic analysis and developed the F<sub>2</sub> populations. FZ, SG, and XW analyzed the sequencing data. FZ performed genetic and physical mapping, and conducted gene expression analysis and phylogenetic analysis. SG, ZC, and AW conducted observation of stamens, pollens, and microspore development. CW performed SNP marker development and KASP genotyping. YH, FZ, SG, ZZ, and SD wrote the paper. NW and SD supervised the experiments. All authors read and approved the final manuscript.

**Acknowledgements** This work was financed by The National Key Research and Development Program of China (2016YFD0101705), the Tianjin Natural Science Foundation Key Program of China (14JCZDJC33900), and Innovation Team of Tianjin Vegetables Research System (ITTVRS)201701. We also thank PlantScribe (<http://www.plantscribe.com>) for editing this manuscript.

### Compliance with ethical standards

**Conflict of interest** The authors declare no conflicts of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

### References

- Andreuzza S, Nishal B, Singh A, Siddiqi I (2015) The chromatin protein DUET/MMD1 controls expression of the meiotic gene TDM1 during male meiosis in Arabidopsis. *PLoS Genet* 11:e1005396
- Atanassova B (2000) Functional male sterility in tomato (*Lycopersicon esculentum* Mill.) and its application in hybrid seed production. *Acta Physiol Plant* 22(3):221–225

- Bai X, Peirson BN, Dong F, Xue C, Makaroff CA (1999) Isolation and characterization of SYN1, a RAD21-like gene essential for meiosis in *Arabidopsis*. *Plant Cell* 11:417–430
- Barnes WC (1960) A male sterile cucumber. *Am Soc Hortic Sci* 77:415–416
- Budar F, Berthomé R (2007) Cytoplasmic male sterilities and mitochondrial gene mutations in plants. *Plant Mitochondria: Annu Plant Rev* 31:278–307
- Budar F, Pelletier G (2001) Male sterility in plants: occurrence, determinism, significance and use. *Comptes Rendus de l'Académie des Sciences-Series III-Sciences de la Vie* 324(6):543–550
- Chen C, Marcus A, Li W, Hu Y, Calzada J-PV, Grossniklaus U, Cyr RJ, Ma H (2002) The *Arabidopsis* ATK1 gene is required for spindle morphogenesis in male meiosis. *Development* 129:2401–2409
- DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, Del Angel G, Rivas MA, Hanna M (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43:491–498
- Feng X, Keim D, Wanjugi H, Coulibaly I, Yan F, Schwarz J, Huesgen S, Cho S (2015) Development of molecular markers for genetic male sterility in *Gossypium hirsutum*. *Mol Breeding* 35:141
- Fu T, Tu J (2002) Present situation and prospects on the research and utilization of hybrid rapeseed. *Analects of Crop Breeding China Agricultural University Press, Beijing*, pp 235–250
- Grimblly P (1980) An apetalous male sterile mutant in cucumber. Report, Cucurbit Genetics Cooperative
- Hanson MR, Bentolila S (2004) Interactions of mitochondrial and nuclear genes that affect male gametophyte development. *Plant Cell* 16:S154–S169
- Hayashi M, Ujiie A, Serizawa H, Sassa H, Kakui H, Oda T, Koba T (2011) Development of SCAR and CAPS markers linked to a recessive male sterility gene in lettuce (*Lactuca sativa* L.). *Euphytica* 180:429–436
- Horn R (2006) Recombination: cytoplasmic male sterility and fertility restoration in higher plants. In: *Progress in botany*. Springer, Berlin, Heidelberg, pp 31–52
- Huang S, Li R, Zhang Z, Li L, Gu X, Fan W, Lucas WJ, Wang X, Xie B, Ni P (2009) The genome of the cucumber, *Cucumis sativus* L. *Nat Genet* 41:1275–1281
- Hutchins A (1936) The inheritance of a green flowered variation in *Cucumis sativus*. *Proc Am Soc Hort Sci* 33:513
- Jiang H, Wang FF, Wu YT, Zhou X, Huang XY, Zhu J, Gao JF, Dong RB, Cao KM, Yang ZN (2009) MULTIPOLAR SPINDLE 1 (MPS1), a novel coiled-coil protein of *Arabidopsis thaliana*, is required for meiotic spindle organization. *Plant J* 59:1001–1010
- Jung K-H, Han M-J, D-y Lee, Lee Y-S, Schreiber L, Franke R, Faust A, Yephremov A, Saedler H, Kim Y-W (2006) Wax-deficient anther1 is involved in cuticle and wax production in rice anther walls and is required for pollen development. *Plant Cell* 18:3015–3032
- Kaneko M, Inukai Y, Ueguchi-Tanaka M, Itoh H, Izawa T, Kobayashi Y, Hattori T, Miyao A, Hirochika H, Ashikari M (2004) Loss-of-function mutations of the rice GAMYB gene impair  $\alpha$ -amylase expression in aleurone and flower development. *Plant Cell* 16:33–44
- Kaul MLH (1988) Male sterility in higher plants. Springer, Berlin
- Ke L, Sun Y, Hong D, Liu P, Yang G (2005) Identification of AFLP markers linked to one recessive genic male sterility gene in oilseed rape, *Brassica napus*. *Plant Breed* 124:367–370
- Kim SS, Jung JY, Jeong SK, Lee DS, Chen LJ, Suh HS (2007) Use of herbicide-resistant genic male sterility in hybrid rice seed production. *Euphytica* 156:297–303
- Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33(7):1870–1874
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359
- Larkin MA, Blackshields G, Brown N, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R (2007) Clustal W and clustal X version 2.0. *Bioinformatics* 23:2947–2948
- Lee J, Yoon JBok, Han J-H, Lee WP, Kim SH, Park HG (2010) Three AFLP markers tightly linked to the genic male sterility ms 3 gene in chili pepper (*Capsicum annum* L.) and conversion to a CAPS marker. *Euphytica* 173:55–61
- Li N, Zhang D-S, Liu H-S, Yin C-S, X-x Li, W-q Liang, Yuan Z, Xu B, Chu H-W, Wang J (2006) The rice tapetum degeneration retardation gene is required for tapetum degradation and anther development. *Plant Cell* 18:2999–3014
- Li YH, Wen CL, Weng YQ (2013) Fine mapping of the pleiotropic locus B for black spine and orange mature fruit color in cucumber identifies a 50 kb region containing a R2R3-MYB transcription factor. *Theor Appl Genet* 126:2187–2196
- Mayer ES, Winiarczyk K, Błaszczyk L, Kosmala A, Rabinowitch HD, Kamenetsky R (2013) Male gametogenesis and sterility in garlic (*Allium sativum* L.): barriers on the way to fertilization and seed production. *Planta* 237:103–120
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303
- Mercier R, Armstrong SJ, Horlow C, Jackson NP, Makaroff CA, Vezon D, Pelletier G, Jones GH, Ying M, Dreyer F, Cai D, Jung C (2003) Molecular markers for genic male sterility in Chinese cabbage. *Euphytica* 132:227–234
- Murray M, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* 8:4321–4326
- Nonomura K-I, Nakano M, Fukuda T, Eiguchi M, Miyao A, Hirochika H, Kurata N (2004) The novel gene HOMOLOGOUS PAIRING ABERRATION IN RICE MEIOSIS1 of rice encodes a putative coiled-coil protein required for homologous chromosome pairing in meiosis. *Plant Cell* 16:1008–1020
- Nonomura K-I, Nakano M, Eiguchi M, Suzuki T, Kurata N (2006) PAIR2 is essential for homologous chromosome synapsis in rice meiosis I. *J Cell Sci* 119:217–225
- Ostergaard L, Petersen M, Mattsson O, Mundy J (2002) An *Arabidopsis* callose synthase. *Plant Mol Biol* 49:559–566
- Patel DJ, Wang Z (2013) Readout of epigenetic modifications. *Annu Rev Biochem* 82:81–118
- Reddy TV, Kaur J, Agashe B, Sundaresan V, Siddiqi I (2003) The DUET gene is necessary for chromosome organization and progression during male meiosis in *Arabidopsis* and encodes a PHD finger protein. *Development* 130:5975–5987
- Shao T, Tang D, Wang K, Wang M, Che L, Qin B, Yu H, Li M, Gu M, Cheng Z (2011) OsREC8 is essential for chromatid cohesion and metaphase I monopolar orientation in rice meiosis. *Plant Physiol* 156:1386–1396
- Shifriss O (1950) Spontaneous mutations in the american varieties of *Cucumis sativus* L. *Proc Am Soc Hortic Sci* 55:351–357
- Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: Join Map. *Plant J* 3:739–744
- Steiner-Lange S, Unte US, Eckstein L, Yang C, Wilson ZA, Schmelzer E, Dekker K, Saedler H (2003) Disruption of *Arabidopsis thaliana* MYB26 results in male sterility due to non-dehiscent anthers. *Plant J* 34:519–528
- Subhash M (2015) Constraints of hybrid seed production in upland and cultivated diploid cottons: will different male sterility systems rescue? A review. *J Cotton Res Dev* 29(2):181–211
- Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S (2013) QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J* 74:174–183

- Tanksley S, Young N, Paterson A, Bonierbale M (1989) RFLP mapping in plant breeding: new tools for an old science. *Nat Biotechnol* 7:257–264
- Van Ooijen J (2006) JoinMap 4: software for the calculation of genetic linkage maps in experimental populations. Wageningen, Kyazma BV
- Vignolini S, Rudall PJ, Rowland AV, Reed A, Moyroud E, Faden RB, Baumberg JJ, Glover BJ, Steiner U (2012) Pointillist structural color in Pollia fruit. *Proc Natl Acad Sci* 109(39):15712–15715
- Virmani SS (1994) Hybrid rice technology: new developments and future prospects: selected papers from the international rice research conference. *Int. Rice Res. Inst*
- Wang S, Yang F, Jiu L, Zhang W, Zhang W, Tian Z, Wang F (2013) Plant regeneration via somatic embryogenesis from leaf explants of *Muscari armeniacum*. *Biotechnol Biotechnol Equip* 27(6):4243–4247
- Wang J, Niu B, Huang J, Wang H, Yang X, Dong A, Makaroff C, Ma H, Wang Y (2016) The PHD finger protein MMD1/DUET ensures the progression of male meiotic chromosome condensation and directly regulates the expression of the condensin gene CAP-D3. *Plant Cell* 28:1894–1909
- Whelan EDP (1972) A cytogenetic study of a radiation-induced male sterile mutant of cucumber. *J Am Soc Hortic Sci* 97(4):506–509
- Wu Y, Fox TW, Trimnell MR, Wang L, Xu RJ, Cigan AM, Huffman GA, Garnaat CW, Hershey H, Albertsen MC (2016) Development of a novel recessive genetic male sterility system for hybrid seed production in maize and other cross-pollinating crops. *Plant Biotechnol J* 14:1046–1054
- Xiao H (2016) Defective callose walls and cell plates during abnormal meiosis cause male-sterility in the oat mutant zbs1. *J Integr Agric* 15(2):241–248
- Yamagishi H, Bhat SR (2014) Cytoplasmic male sterility in Brassicaceae crops. *Breed Sci* 64:38–47
- Yang M, Hu Y, Lodhi M, McCombie WR, Ma H (1999) The Arabidopsis SKP1-LIKE1 gene is essential for male meiosis and may control homologous chromosome separation. *Proc Natl Acad Sci* 96:11416–11421
- Yang X, Makaroff CA, Ma H (2003) The Arabidopsis MALE MEIOCYTE DEATH1 gene encodes a PHD-finger protein that is required for male meiosis. *Plant Cell* 15:1281–1295
- Ying M, Dreyer F, Cai D, Jung C (2003) Molecular markers for genic male sterility in Chinese cabbage. *Euphytica* 132:227–234
- Yuan W, Li X, Chang Y, Wen R, Chen G, Zhang Q, Wu C (2009) Mutation of the rice gene PAIR3 results in lack of bivalent formation in meiosis. *Plant J* 59:303–315
- Zhang Q, Gabert AC, Baggett JR (1994) Characterizing a cucumber pollen sterile mutant: inheritance, allelism, and response to chemical and environmental factors. *J Am Soc Hortic Sci* 119(4):804–807