

Fine Mapping and Candidate Gene Analysis For A Novel Male-Sterile Mutant Ms40 in Maize

Liu Xiaowei

Sichuan Agricultural University

Yue Yujing

Sichuan Agricultural University

Gu Zicheng

Sichuan Agricultural University

Huang Qing

Sichuan Agricultural University

Pan Zijin

Sichuan Agricultural University

Zhao Zhuofan

Sichuan Agricultural University

Zheng Mingmin

Sichuan Agricultural University

Zhang Zhiming

Shandong Agricultural University

Li Chuan

Sichuan Agricultural University

Moju Cao (✉ caomj@sicau.edu.cn)

Sichuan Agricultural University <https://orcid.org/0000-0002-5883-5882>

Research Article

Keywords: Maize, Male sterility, Tapetum, Map-based cloning, bHLH transcription factor

Posted Date: April 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-303363/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Plant Cell Reports on July 28th, 2021. See the published version at <https://doi.org/10.1007/s00299-021-02762-w>.

Abstract

Maize male sterile mutant 40 (*ms40*) was obtained from the progeny of ethyl methanesulfonate (EMS) treated inbred line RP125. Genetic analysis showed that it was controlled by a single recessive nuclear gene. Cytological observation of anthers revealed that abnormal cuticles and disappearing of Ubisch bodies presented in *ms40*. Moreover, its tapetum exhibited delayed degradation and blocked the formation of abnormal microspore. Using map-based cloning, *ms40* locus was located in a 282-kb interval on chromosome 4, five annotated genes were predicted within this region. PCR-based sequencing detected a single nonsynonymous SNP (G>A) which changed glycine (G) to arginine (A) in the seventh exon of *Zm00001d053895*, while no difference was found for the other four genes between *ms40* and RP125. *Zm00001d053895* encodes the bHLH transcription factor *bHLH51* which protein was located at nuclear. Phylogenetic analysis presented that *bHLH51* had the highest homology with *Sb04g001650*, a tapetum degeneration retardation (TDR) bHLH transcription factor in *Sorghum bicolor*. Co-expression analysis exposed a total of 1192 genes coexpressed with *Zm00001d053895* in maize, 647 out of 1192 were anther-specific genes. In summary, these findings are conducive to the marker-assisted selection of *ms40* in hybrid breeding and laid a foundation for further studies on the mechanisms of male fertility.

Key Messages

A novel genic male sterile mutant *ms40* was obtained by EMS treated RP125. The key candidate gene *bHLH51* located on chromosome 4 was identified by map based cloning. This study furtherly enriched the male sterile gene resource for both production application and theoretical study of abortion mechanism.

Introduction

Maize is one of the most important crops widely cultivated in the world and is one of the earliest crops to utilize heterosis. In the production process of maize hybrids seeds, artificial emasculation is the most common method but this method is time-consuming and laborious. Meanwhile, the purity of hybrids seeds is difficult to be guaranteed. However, using the male sterile lines in seed production will address these problems greatly. Maize male sterility is divided into cytoplasmic male sterility (CMS) and genic male sterility (GMS). There are some obvious problems in the application of CMS, such as the instability of sterility and difficulty in finding strong and stable restorer lines. For GMS, it is difficult to find completely maintainer lines, making it hard to apply directly in the hybrids seeds production of maize. The seed production technology (SPT) bring about hope for applying GMS in hybrid seeds production, therefore more and more GMS genes research have been reported and more attention from breeders were attracted. However, there are few male sterile mutants with independent intellectual property rights. So, it is particularly important to create male sterile mutants with independent intellectual property rights in China.

To date, approximately 19 genes of GMS mutants have been successfully cloned. At present, the reported maize male sterility genes encoded different protein types, including secretory protein, lipid transporter, redox protein, enzyme, and transcription factor. *MSCA1*(*MULTIPLE ARCHESPORIAL CELLS1*) encodes a plant specific glutathione reductase gene, the mutation *mzca1* had the deletion of GSH binding site, which maybe impact the initiation of archesporial cells (Marc et al. 2009), its homologous genes, *OsTDL1A* and *AtTPD1*, had been reported to be related to the development of anther (Wang et al. 2012). *MS6021*, *MS33*, *MS30* and *IPE1* (*IRREGULAR POLLEN EXINE1*) all belong to functional protein and participate in lipid or fatty acid metabolism (An et al. 2019; Chen et al. 2017; Tian et al. 2017; Xie et al. 2018). Both *MS26* and *APV1* (*ABNORMAL POLLEN VACUOLATION1*) encode cytochrome P450 monooxygenase (Djukanovic et al. 2013; Somaratne et al. 2017), *MS45* encodes hydroxyproline-rich glycoprotein family protein (Cigan et al. 2001), all of them are required for the formation of pollen exine and anther cuticles in maize. Some transcription factor had been reported to be associated with genic male sterility in maize. *OCL4* (*OUTER CELL LAYER*) encodes a HD-ZIP transcription factor, which plays a major role in the trichome differentiation and division of the anther cell wall in maize (Vernoud et al. 2009). *MS9* encodes a R2 / R3 plant specific MYB transcription factor (MC et al.). *IG1*(*INDETERMINATE GAMETOPHYTE1*) encodes a LOB domain protein which can regulate the proliferative phase of female gametophyte development (Evans 2007). *MS7* encodes a PHD-finger transcription factor, which was used to hybrid seed production by multicontrol sterility system (Zhang et al. 2017). *MS23* and *MS32* encode bHLH transcription factors responsible for tapetal development and PCD (Moon et al. 2013; Nan et al. 2017).

The bHLH transcription factors (TFs) in flowering plants consist of large families with 213 encoding genes annotated in maize (Lin et al. 2014), which is the most member transcription factor family in maize. *MS23* encodes the *bHLH16* transcription factor (Nan et al. 2017), which plays an important role in the differentiation of the endothecium and tapetum cells of anther and plays a direct or indirect role in the biogenesis of 24-phasiRNAs. *MS32* encodes the *bHLH66* transcription factor (Moon et al. 2013), and specifically expressed in anther of premeiotic stage. Moreover, *MS32* can interact with the protein encoded by *MAC1* to regulate the pericytosis of L2 layer cells and the differentiation of anthers sporogonia, thus affecting pollen development. Although these bHLH transcription factors had been cloned, the regulating mechanisms for pollen abortion have not been elucidated clearly. The discovery of more other bHLH transcription factors controlling maize male sterile mutants maybe helpful for clearing these bHLH transcription factors regulating relationship between each other.

It is a long process to transfer a specific male sterile gene from one genetic background into another elite inbred line, so the best strategy is to create male sterile mutant based on an elite inbred line background with a single base change, which can effectively accelerate the application of GMS gene. Maize inbred line RP125 cultivated by Sichuan Agriculture University is widely planted in southwest China, its characteristics of high combining ability, high yield, high resistance to northern leaf blight and southern leaf blight, moderate resistance to sheath blight and other major diseases in southwest corn production area, as well as the efficient utilization of phosphorus, making it to be one of the most popular parents in Southwest China in the 21st century.

In this study, we found a no-pollen male sterile mutant, *ms40*, which derived from maize inbred line RP125 by EMS mutagenesis treatment. Cytological observation showed that the tapetum of *ms40* anther exhibited abnormal expansion, and defective in Ubisch bodies and pollen exine aslo were observed. The sterile gene of *ms40* was located within a 282-kb interval on the chromosome 4 by map-based cloning, and *Zm00001d053895* was found to be the key candidate gene. This study provides a new genetic resource not only for the application of GMS in the hybrid seed production but also for the regulating mechanisn interpretation of maize anther development.

Materials And Methods

Plant materials

In the spring of 2015, the maize inbred line RP125 bred by Sichuan Agriculture University was planted in experimental field of Sichuan Agriculture University, Sichuan. Next, the pollen samples were treated by ethyl methanesulfonate (EMS) and then self-pollinated to produce M₁ seeds, then planted in Yunnan experimental field, in the autumn of 2015. The M₁ plants self-pollination were conducted and M₂ seeds were obtained. The M₂ seeds were planted at experimental field in spring of 2016, Sichuan. Among these, a male sterile mutant was found, termed *ms40*, pollinating with RP125 pollen to get (*ms40*×RP125)F₁ seeds. Two inbred lines B73, Mo17 were also used in this study.

Phenotype identification and genetic analysis

The (*ms40*×RP125)F₂ seeds were gained by the (*ms40*×RP125)F₁ self-pollinated, and then the field fertility identification and pollen grains stained with 1% (m/v) I₂-KI solution were adopted to evaluate the fertility of anthers. If the sterility phenotype of *ms40* can be inherited stably, then the inbred line B73 and Mo17 were used as test lines for genetic analysis, Chi-square test was used for the phenotype segregation analysis.

A Canon M3 digital camera and Olympus SZX16 stereomicroscope were used to take photographs of the plants and anthers. Pollen grains were stained with 1% I₂-KI (m/v) solution and photographed by a Lecia DM2000 microscope.

Semi-thin section and scanning electron microscopy of anthers

For scanning electron microscopy (SEM), the mature anthers of fertile and male sterile plants from the (*ms40*×RP125)BC₁F₁ population were dissected and fixed with glutamic dialdehyde. Then, the samples were dried about 18 hours with the freeze dryer (Freezong 2.5, Labconco) and examined under the scanning electron microscope (Inspect F50, FEI).

For cytological observation of anthers, anthers of fertile and male sterile plants from the sister cross population (*ms40/ms40*×*Ms40/ms40*) at different developmental stages were fixed in formaldehyde-acetic acid-ethanol (FAA) overnight and dehydrated by gradient concentrations of ethanol. Then the

anthers were infiltrated with a mixed gradient solution of ethanol and 7100 hardener II solution (Technovit 7100, Germany) and embedded in spur resin. Slicing was performed by a Leica DM2255 slicer and then stained with 0.1% (m/v) Toluidine Blue solution. A Lycra DM2000 microscope was used to observe and photograph sections.

Map-based cloning of *ms40* male sterile gene

The (*ms40*×B73) F_2 was applied to the mapping population of *ms40*, and 134 InDel markers uniformly covering 10 chromosomes of maize were developed based on the differences between genome sequences of RP125 and B73, the genomic DNA was extracted using the CTAB (hexadecyl trimethyl ammonium bromide) method (Luan et al. 2008) with minor modifications from the original method. The bulk-segregant analysis (BSA) method was performed, the fertile DNA pool and the male sterile DNA pool were constructed using twenty fertile and twenty male sterile plants selected randomly from (*ms40*×B73) F_2 population by mixing equally. The 134 InDel markers were selected to detect polymorphisms between the two DNA pools. Next, the polymorphic markers were used to examine genotype of 115 male sterile individuals from the (*ms40*×B73) F_2 population to judge whether the sterility phenotype and the polymorphic markers were linked. Based on the primary mapping, 4 new polymorphic InDel markers were developed, and the 1230 sterile plants from the larger (*ms40*×B73) F_2 population was used. All markers information is provided in Table S1.

Key candidate gene prediction and analysis

Candidate gene predictions and functional annotations were obtained from the Gramene database (<http://ensembl.gramene.org/>). The conserved domains of candidate genes were predicted by the NCBI Conserved Domain Search tool (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>), and the data of the expression patterns were derived from an RNA-seq expression database (<https://www.maizegdb.org/>). The sequences of the candidate genes were amplified from *ms40* and RP125, then PCR products were sequenced and analysed.

Based on the sequence difference of *Zm00001d053895* among *ms40* and RP125, the relationship between the fertility and SNP of *Zm00001d053895* in *ms40* was analysed, the SNP marker was developed according to the flanking sequence of the mutation site (SNP-F: 5'-TGTCATTGTACGTACGGCGG-3', SNP-R: 5'-CGTGGGATGTACGGCGATG-3'). Co-segregation analysis of phenotypes and genotypes in the individuals of (*ms40*×Mo17) F_2 population and (*ms40*×RP125)BC $_1$ F $_1$ population with the SNP marker was performed. To detect whether the SNP of *Zm00001d053895* exists only in *ms40*, the sequence fragment including the mutation site was amplified with the SNP marker in 30 maize inbred lines and sequenced.

Rapid amplification of cDNA ends (RACE) assay

Total RNA was extracted from the anthers of RP125 at different developmental stages by using TRIzol reagent (Invitrogen). The N711 Kit (Vazyme, Nanjing, China) was used for the RACE assay according to

the manufacturer's instructions. GSP1 and GSP2 primers and the gene-specific primers used for RACE were designed according to the reference cDNA sequence of *Zm00001d053895*. The GC content of GSP1 and GSP2 primers was required at 50% - 70%, and its T_m were approximately 60°C (GSP1: 5'-ACCTGCCTCCATCAATCCAGCTCG-3', GSP2: 5'-AATGAGGTGGCAGTGCAGGCGGA-3'). GSP1 and GSP2 primers were used for 5' RACE and 3' RACE, respectively. The PCR products were cloned into the pEASY-Blunt cloning vector and sequenced. The CDS of *Zm00001d053895* was forecasted via the ORFfinder tool (<https://www.ncbi.nlm.nih.gov/orffinder/>).

RNA extraction and qRT-PCR

Total RNA, including leaves, roots, stems, and anthers of various developmental stages, were extracted from the RP125 plants using TRIzol reagent (Invitrogen, USA). Each sample contained three biological replicates. Total RNA was reverse transcribed using the Reverse Transcription Kit (Vazyme, China), and qPCR was performed using SYBR Green PCR Master Mix (TaKaRa, Japan). Biorad CFX96 system was used to detect the relative expression of *Zm00001d053895* with the primers q-51-F/R (5'-CTCTGGGTCCCCCTGCAT-3', 5'-GGTGGTGGTGGGGTGGGA-3'). Three biological replicates and four technical replicates were performed for each procedure. *ZmACTIN* was used as the internal control to normalize the expression data (Chen et al. 2017) and its amplifying primers were 5'-TCACCCTGTGCTGCTGACCG-3' and 5'-GAACCGTGTGGCTCACACCA-3'. Relative expression levels were calculated according to the $2^{-\Delta\Delta C_t}$ method, and all results were expressed as the mean \pm standard error of the mean (SEM).

Subcellular localization

For analysis of the subcellular location of *Zm00001d053895*, we first used the TargetP-2.0 Server (<http://www.cbs.dtu.dk/services/TargetP/>) to predict the putative subcellular location. Then, the CDS of *Zm00001d053895* was amplified from RP125 (without the stop codon) and cloned into the *pCAMBIA2300-eGFP* vector. The construct, as well as the empty construct, were cotransformed with the nuclear marker (p35S::NLS-RFP) (Wu et al. 2016) into tobacco (*Nicotiana benthamiana*) mesophyll cells, respectively. The fluorescence signals were observed by a laser scanning confocal microscope (Zeiss 800). eGFP and RFP fluorescence signals were tested at 488 nm and 561 nm, respectively.

Transactivation activity analysis

The *Zm00001d053895* CDS was inserted into the pGBKT7 vector using In-Fusion cloning method (Vazyme ClonExpress II One Step Cloning Kit, Vazyme Biotech, China). Then the recombinant was transformed into AH109 yeast strains (Tiandz, China) via the lithium acetate-mediated approach. The growth performances of cotransformants of positive clones were examined on SD/-Trp medium and SD/-His-Trp medium containing 50 mg l⁻¹ χ -gal (Coolaber, China) for 2-4 days at 28°C. The free pGBKT7 vector and pGBKT7-GAL4 AD (Liu et al. 2017) were used as positive and negative controls, respectively.

Phylogenetic analysis

For determination of the evolutionary relationship between *Zm00001d05395* and its homologs in various species, homologue sequences were searched in the NCBI database (<https://www.ncbi.nlm.nih.gov/>) using the *Zm00001d053895* amino acid sequence as the query, and 14 homologs from *Oryza sativa*, *Solanum lycopersicum*, *Brachypodium*, *Foxtail millet*, *Hordeum vulgare*, *A. thaliana*, *Solanum tuberosum*, *Sorghum* and *Triticum aestivum* were retrieved. Multiple sequence alignment was performed using CLUSTALW with default settings within MEGA 6 (HIGGINS 1996). We adopted MEGA 6 to construct an unrooted phylogenetic tree via the neighbour-joining method, which was tested by 1000 bootstrap replicates, and the phylogenetic tree was prettified using EvolView (<https://www.evolgenius.info/evolview/>) (Zhang et al. 2012).

Co-expression analysis

Co-expression analysis was used to identify potential interacting proteins of *Zm00001d053895*. The expression data of approximately 40,000 maize genes from 8 tissues of B73 were downloaded from the q-teller database (<http://www.qteller.com>), and the gene expression data were fragments per kilobase of exon per million fragments mapped (FPKM). Pearson correlation coefficient values (P-values) of each gene with *Zm00001d053895* were calculated based on expression data. The co-expressed genes with P-values > 0.8 and P-values < 0.05 were selected. The FPKM values of co-expressed genes were homogenized by $\log_2(\text{FPKM} + 1)$, and then, the Z-scores were calculated (Sekhon et al. 2011). A Z-score value larger than 2 was determined as the tissue-specific gene. For characterization of the putative function of *Zm00001d053895* co-expressed genes, GO terms for each coexpressed gene were obtained at Gramene (<http://www.gramene.org/>), and GO enrichment analysis was performed using OmicShare tools (<http://www.omicshare.com/tools>).

Results

Male sterile mutant *ms40* is controlled by a recessive nuclear gene

The pollens of maize inbred line RP125 were treated by EMS, M₁ plants were self-pollinated to obtain M₂. Male sterile mutant *ms40* was found among the M₂ generation, then RP125 pollens were used to fertilize *ms40*, all the individuals of (*ms40*×RP125)F₁ population were fertile, thereupon, the (*ms40*×RP125)F₁ individuals were conducted self-pollinated on the one hand, and reciprocal crossed with RP125 on the other hand, the male sterile phenotype were separated again no matter within the F₂ populations or the backcross populations, which implied that *ms40* was a genetic male sterile (GMS) mutant. Hence, some male sterile plants were selected and pollinated with the pollens of B73, Mo17 separately, all the individuals of (*ms40*×B73)F₁, (*ms40*×Mo17)F₁ presented fertile plants. Then self-pollination were conducted for (*ms40*×B73)F₁ and (*ms40*×Mo17)F₁, the F₂ seeds were planted both in Yunnan and Sichuan for the tassels fertility identification. The male sterile phenotype was separated in the F₂ population no matter planted in Sichuan or Yunnan. Moreover, the segregation ratio of fertile plants vs male sterile plants within (*ms40*×B73)F₂ population and (*ms40*×Mo17)F₂ population fit to the ratio of 3:1

with student *t* test (Tab 1). These results proved that the sterile phenotype of *ms40* was controlled by a single recessive nuclear gene. No obvious differences were found between RP125 and *ms40* for their agronomic traits (Fig. 1A, B). After tasseling, RP125 tassels and anthers presented normal-appearing, whereas *ms40* tassels failed to expose and no pollen shed, and its anthers were smaller and thinner compared with those of the RP125 (Fig. 1C-F). By means of I₂-KI staining, the pollen grains of RP125 were dark blue-stained with round-shaped staining, while no pollen grains were found for *ms40* (Fig. 1G-H).

The anther development presents a certain imperfection for *ms40*

In order to reveal the characteristics of *ms40* anthers, we examined epidermis and inner surface of anthers by scanning electron microscope (SEM). The anthers of *ms40* were significantly smaller, shorter and much more withering (Fig. 2A) in comparison with that of RP125 (Fig. 2B). In addition, the anthers of RP125 showed a latticed-waxy-crystal anther epidermal surface (Fig. 2C), but *ms40* was smooth without any cuticle (Fig. 2D). Meanwhile, we found that Ubisch bodies covered the whole inner surface of RP125 anthers (Fig. 2E), while no Ubisch bodies were observed on the inner surface of *ms40* anthers (Fig. 2F). From the broken anthers, we could find pollen grains with round-shape filled in RP125 anthers (Fig. 2G, I), but no pollen grains were found within *ms40* anthers (Fig. 2H), which further verified that *ms40* was a no-pollen type male sterile mutant. These results showed us that a series developmental defects of anther and pollen coupled with *ms40*.

Mutant *ms40* exhibit the delayed degradation for anther tapetum

A variety of anther dysplasia were observed in various male sterile individuals. Understanding the cytological characteristics of pollen abortion is helpful for explain the mechanism of failure for a male sterile mutant. So, anthers of male sterile and fertile plant from (*ms40*×RP125)BC₁F₁ population at different stage were examined using semi-thin section. At sporogenous and pollen mother cell stages, no substantial difference was observed between the anthers of RP125 and *ms40* (Fig. 3A-D). At meiosis stage, the tapetum degradation began with paliform-shape in RP125, while the *ms40* tapetum remained almost intact, suggesting the degradation of *ms40* was delayed (Fig. 3E, F). Subsequently, the content of tapetum cell in RP125 began to concentrate, deepened staining, while the *ms40* tapetum was swelled seriously, and irregular microspores were observed (Fig. 3G, H). At large vacuolated stage, large vacuoles had formed in the centre of microspores, and tapetum cells were further concentrated and degraded in RP125, however the microspores of *ms40* began to shrink and were unable to form large vacuoles, tapetum cells were clearly visible with no signs of disintegration (Fig. 3I, J). At binucleate stage, the vacuolated microspores underwent asymmetric mitotic division and displayed falcate-shape, accompanied with tapetum completely disintegration in RP125, while the microspores of *ms40* gradually degraded, and the vacuolation of the tapetum was more obvious (Fig. 3K, L). At mature pollen grain stage, vast pollen grains filled with starch were observed in the anther locule of RP125, in contrast, the microspores of *ms40* almost completely degraded, leaving only remnants in their locules, and vacuolized

tapetum could be observed (Fig. 3M, N). In conclusion, male sterile mutant *ms40* exhibited significantly delayed disintegration of tapetum and no-pollen abortive type.

Male sterile gene of *ms40* was mapped on chromosome 4 within a 282-kb region

In this study, (*ms40*×B73) F₂ was taken as a mapping population, 134 InDel markers covering the whole maize genome were developed based on the whole genome re-sequencing data of RP125 and B73. The 134 InDel markers were used for polymorphism scanning between *ms40* and B73, then 73 polymorphic markers were obtained. Afterwards, the 73 polymorphic markers were applied to scan polymorphism between male sterile DNA pool and fertile DNA pool, then InDel markers umc1940 and umc1649 were selected for their polymorphism, interestingly, both umc1940 and umc1649 were located at bin 4.10 on maize chromosome 4. So, we developed some novel InDel markers between *ms40* and B73, among them, 6 polymorphic markers were obtained, then 115 male sterile individuals from (*ms40*×B73)F₂ population were genotyped with the 6 InDel markers, 1 and 7 recombinants were screened in the 115 male sterile individuals with X98 and X72, respectively, and no recombinant was detected with X76 among the 115 male sterile individuals. Therefore, *ms40* locus was located between X98 and X72 on chromosome 4 (Fig. 4A).

In order to shorten the mapping region, 1230 male sterile individuals derived from (*ms40*×B73)F₂ population were used for genotyping with X98 and X72, and a total of 11 and 6 recombinants were screened from the 1230 male sterile individuals, respectively. Then, we furtherly developed 4 polymorphic InDel markers within the mapping interval between X98 and X72. Among these InDel markers, only 1 recombinant were detected in the 1230 male sterile individuals for marker X214 and X242, respectively. No recombinant was detected in the 1230 male sterile individuals for marker X168, therefore the male sterile gene of *ms40* was mapped to chromosome 4 between X214 and X242, the physical distance between X214 and X242 was 282 kb (Fig. 4B). According to the MaizeGdb database, a total of 5 open reading frames (ORFs) were identified within the region of fine mapping (Tab 2). All markers information used in this study are provided at Table S1.

Zm00001d053895 is the key candidate gene of *ms40*

Then 5 ORFs were cloned from RP125 and *ms40*, next sequencing and sequence alignment were conducted. We found that only *Zm00001d053895* harboured an SNP (G to A) at position 2851 for *ms40* and RP125, no differences were found for the sequence of other 4 genes. The SNP locus was located at the seventh exon of *Zm00001d053895*, which led to a change of amino acid from Gly (GGG) to Arg (AGG) (Fig. 5A). *Zm00001d053895* was predicted to be an anther-specific gene that encodes a bHLH transcription factor (https://www.maizegdb.org/gene_center/gene/Zm00001d053895), the mutants (*tdr1* and *ams*) of its homologous genes *OsTDR* and *AtAms* manifested male sterile phenotype and displayed developmental defect in the anther tapetum (Ferguson et al. 2017; Li et al. 2006).

In order to reveal whether or not the SNP (G/A) at position 2851 of *Zm00001d053895* is related to the male fertility, we developed an SNP marker (SNP-F/SNP-R) based on the flanking sequence of the

mutation site in *ms40*. Then 1589 individuals from (*ms40*×*Mo17*) F_2 population including 1218 fertile plants and 371 male sterile plants, and 197 individuals from (*ms40*×*RP125*) BC_1F_1 population including 101 fertile plants and 96 male sterile plants were used for co-segregation analysis with the SNP marker (SNP-F: 5'-TGTCATTGTACGTACGGCGG-3', SNP-R: 5'-CGTGGGATGTACGGCGATG-3'). As a result, 1218 fertile plants of (*ms40*×*Mo17*) F_2 population carried either homozygous G/G or heterozygous A/G allele, 101 fertile plants of (*ms40*×*RP125*) BC_1F_1 population carried heterozygous A/G allele at position 2851 of *Zm00001d053895*, whereas homozygous A/A allele was detected in all the 467 male sterile plants, 371 sterile individuals from (*ms40*×*B73*) F_2 population and 96 sterile individuals from (*ms40*×*RP125*) BC_1F_1 population. All these results showed that the haplotype (A/A) at position 2851 was co-segregated with the male sterile phenotype of *ms40*. Moreover, 30 different inbred lines were examined using the SNP marker, and only *ms40* had the homozygous A/A allele at position 2851 of *Zm00001d053895*, the 30 inbred lines were all homozygous G/G for the corresponding locus of *Zm00001d053895* (Fig. 5B) (Table S2), suggesting that the 2851st nucleotide (G) should be a conserved nucleotide and the position 2851 of *Zm00001d053895* may be a key functional site.

Zm00001d053895 encodes a bHLH transcription factor and has the transcriptional activating ability

To illustrate the evolutionary relationship of *Zm00001d053895*, we performed phylogenetic analysis based on 14 orthologous genes from 10 plant species that shared high sequence similarity with *Zm00001d053895*. Through multiple sequence alignment, a classic HLH domain was found in all the 14 homologous genes (Fig. 6A), suggesting that *Zm00001d053895* is a typical bHLH transcription factor and the orthologs of *Zm00001d053895* might have conserved function among various plant species. Moreover, the mutation site of *Zm00001d053895* in *ms40* was located within the HLH conserved domain of the bHLH transcription factor. Phylogenetic analysis showed that these genes were divided into three clades, which indicates that their molecular functions had a degree of evolutionary conservation. *Zm00001d053895* shared the highest homology with *Sb04g001650* (81.2%) of *Sorghum* (Fig. 6B), a putative TDR bHLH transcription factor, which has connected with the development of anther tapetum, thus it can be seen that *Zm00001d053895* plays an important part in regulating tapetum development.

The transactivation activity assay was performed to investigate the transcriptional activating ability of *Zm00001d053895*, the transformant pGBKT7-*Zm00001d053895* were constructed, and free pGBKT7 and pGBKT7-GAL4 AD as negative and positive control respectively. All of them could grow well on the SD/-Trp medium. However, on the SD/-His-Trp medium containing 50 mg/l χ -gal, the free pGBKT7 transformant could not live, but the pGBKT7-GAL4 AD, pGBKT7-*Zm00001d053895* grew normally and turned the indicator blue (Fig. 7), which indicated that *Zm00001d053895* had the transcriptional activating ability.

Four transcripts of *Zm00001d053895* were identified in maize anther by RACE

In our study, the rapid amplification of cDNA ends (RACE) assay was performed for determine the structure of *Zm00001d053895* transcripts, four transcripts of *Zm00001d053895* were amplified with total

RNA of RP125 anthers from different development stages. The CDS was predicted according to the sequences of transcripts by the ORFfinder tool (<https://www.ncbi.nlm.nih.gov/orffinder/>), then four transcripts were identified and encoded 560aa (*Zm00001d053895-T001*), 597aa (*Zm00001d053895-T002*), 603aa (*Zm00001d053895-T003*) and 628aa (*Zm00001d053895-T004*), respectively (Fig. 8A). Among the four transcripts, *Zm00001d053895-T001*, *Zm00001d053895-T003* and *Zm00001d053895-T004* contained 7 exons and 6 introns, only *Zm00001d053895-T002* contains 8 exons and 7 introns. Through comparing with the shortest transcript *Zm00001d053895-T001*, we found that both *Zm00001d053895-T003* and *Zm00001d053895-T004* transcripts resulted from alternative 5' splice site (A5SS), while transcript *Zm00001d053895-T002* resulted from exon skipping (ES). By sequence alignment of the four transcripts, we also found that the stop codon position of the four transcripts were same, but the position of start codon was different, which account for the diversity of these CDS owing to the difference of transcription starts sites (TSS) of four transcripts.

To decipher the expression specificity of *Zm00001d053895*, qRT-PCR was performed for roots, stems, leaves and anthers from different developmental stages of RP125 with the specific primers q-51-F/R. *Zm00001d053895* was preferentially expressed at pollen mother cell and tetrad stage of anthers, and the expression level was low at uninucleate and binucleate stages. These results were consistent with the phenotypic differences observed at the meiotic stages between the RP125 and *ms40*. In contrast to the expression levels of anthers, the expression levels in roots, stems and leaves were very low (Fig. 8B), which suggested that *Zm00001d053895* was an anther-specific gene and played an important role in the development of anthers.

Zm00001d053895 was localized in the nucleus and coexpressed with some anther-specific genes

The TargetP-2.0 server were used to predict the putative subcellular location of *Zm00001d053895* firstly, then we performed transient expression assays in tobacco leaves. The vector p35S::*Zm00001d053895*-eGFP was constructed with *Zm00001d053895* coding sequence and fused to the N-terminus of eGFP driven by 35S, the free vector p35S-eGFP was used as a positive control. Vector p35S::NLS-RFP carried the nuclear location signal (NLS) was also constructed in this study. As expected, the nuclear localization signal (NLS) of vector p35S::NLS-RFP was distributed in the nucleus of tobacco mesophyll cells, the eGFP signal of vector p35S-eGFP was distributed throughout the entire cell, and the fluorescence signal merged in the nucleus when vector p35S::NLS-RFP and vector p35S-eGFP were cotransformed (Fig. 9A-D). When p35S::*Zm00001d053895*-eGFP and p35S::NLS-RFP were cotransformed, the fluorescence signals merged in nucleus (Fig. 9E-H), all the results suggested that *Zm00001d053895* was localized in the nucleus, which consistent with the location result predicted by TargetP-2.0 server tool.

We identified 1192 genes coexpressed with *Zm00001d053895* in the whole genome using the expression data of maize gene downloading from the q-teller database. Among them, the male sterile gene *Zm00001d02680* (*ms7*) (Zhang et al. 2017), and five putative GMS genes (Wan et al. 2019), *Zm00001d031312*, *Zm00001d033335*, *Zm00001d013732*, *Zm00001d013991*, and *Zm00001d035841*,

shared expression PCC values of 0.91, 0.98, 0.95, 0.97, 0.88 and 0.92 with *Zm00001d053895*, respectively. Such a high correlation suggested that *Zm00001d053895* may be related to male fertility.

Next, the GO terms of 1192 co-expressed genes were analysed. For the biological process category, metabolic process, cellular process and single-organism process were highly enriched in the GO classes. Of the cellular components, 107 GO terms were enriched, mostly in the membrane and membrane parts. For the molecular function category, binding and catalytic activity were the most abundant subcategories (Fig. 10A), and these terms had been reported that their functions were related to alterations in male fertility (Cunmin et al. 2015; Mei et al. 2016; Zhu et al. 2015). After homogenization of the expression data of co-expressed genes, we found that 647 genes were specifically expressed in anthers and pollen (Fig. 10B). These results indicated that *Zm00001d053895* highly possible was related to the development of anthers.

Discussion

Male sterile mutant *ms40*, which derived from the progeny of EMS-treated inbred line PR125, showed stable male sterility for the multi-year tests no matter planted in Sichuan or Yunnan locations. Mutant *ms40* exhibited no anther exerted and belonged to no pollen type. Genetic analysis showed that *ms40* is controlled by a recessive nuclear gene. Through map-based cloning, we had successfully located it on the long arm of chromosome 4 within a 282-kb region, which consists of five open reading frames. Based on cloning and sequencing analysis, an SNP from G to A change was found within the seventh exon of *Zm00001d053895*, which encodes a bHLH transcription factor (<https://www.maizegdb.org/gene-center/gene/Zm00001d053895>), its homologous genes *OsTDR1* and *AtAms* were reported to be related to the anther development (Li et al. 2006).

The intragenetic marker was developed according to the SNP locus and co-segregation analysis were conducted with two different fertility segregation populations, the results showed that all the male sterile individuals were A/A genotype, which furtherly supported that *Zm00001d053895* was the male sterile gene of *ms40*. In addition, 30 maize inbred lines were used as materials for the sequence conservative analysis of *Zm00001d053895*, as a result, all the inbred lines present the G/G genotype for gene *Zm00001d053895* through PCR amplification and sequencing analysis, only *ms40* individuals were A/A genotype at the same locus, these results indicating the G-A base change was responsible for the generation of *ms40* mutant. Therefore, *Zm00001d053895* is considered to be the key candidate gene of mutant *ms40*.

Understanding the failure characteristic of anther is necessary for exploring the abortion mechanism of a male sterile mutant. Tapetum located in the innermost layer of the anther is connected with the development of microspores, normal development of tapetum is vital to the formation of pollen. Cytological observation showed that the anther tapetum of *ms40* exhibited obvious vacuolization and delayed degradation. The abortive features of tapetum development in *ms40* are consistent with the previous reported function of bHLH transcription factors, such as *AtAMS*, *OsTDR1*, *OsUDT1*, *OsEAT1*,

ZmMS23, *ZmMS32* (Ferguson et al. 2017; Jung et al. 2005; Ko et al. 2014; Moon et al. 2013; Nan et al. 2017; Niu et al. 2013), all of these mutants manifested different degrees of tapetum abnormality. Thus, the cytological observation results of *ms40* also supported that *Zm00001d053895* was the key candidate gene of *ms40*. qRT-PCR analysis showed that *Zm00001d053895* expressed specifically in anther, and preferentially expressed at the pollen mother cell stage as well as tetrad stage of developing anthers. Consistently, the cytological differences of anther development between the RP125 and *ms40* occurred at the meiotic stages. Moreover, we also found that *ms40* had no cuticle on anther epidermal and no Ubisch bodies coating on its inner surface. Not all male sterile mutants were short of cuticles and Ubisch bodies in anther, such as *OsLSP1* (Luo et al. 2020) and *OsMS1* (Yang et al. 2019). Compared with RP125, the contents of tapetum cell in *ms40* anther were particularly thin, and then probably hindered the cuticles formation of anther epidermal as well as the Ubisch bodies formation adhering inner surface of anther.

Zm00001d053895 has the classical HLH domain of bHLH transcription factors, which is necessary for bHLH transcription factors forming homodimers or heterodimers and then regulating the expression of target genes (R et al. 2013), bHLH-bHLH complexes had been confirmed to be relevant to plant fertility. Our transcriptional activation assay showed that *Zm00001d053895* had the ability of transcriptional activation. However, the SNP site of G to A single nucleotide change was predicted to locate in the binding sites of *Zm00001d053895* with the NCBI-CDD tool. So different haplotype of the SNP within the seventh exon of *Zm00001d053895* may directly affect the ability of transcriptional activation, and then disturbed the pollen development through regulating the target genes expression.

It is well known that maize male sterile mutant plays an important part in hybrid seed production. Although various male sterile genes have been cloned, there is a certain distance from the discovery of a male sterile mutant to applying it in seed production for a particular cross combination. Therefore, the high yield and high combining ability excellent inbred line RP125 was selected as the basic materials for EMS-treatment. Fortunately, the male sterile mutant *ms40* was obtained from their progeny, thus we can utilize *ms40* for the hybrid seed production of some cross combinations which using PR125 as the female parent. Based on the different years and different locations experiments, no obvious difference was found between PR125 and *ms40* for their agronomic traits, other studies showed that PR125 can tolerate low phosphorus and resist to a variety of diseases. Moreover, exploring the abortion mechanism of *ms40* not only can contribute to the maize anther developmental biology research but also can substantially promote the application of using male sterile lines for hybrid seed production.

Declarations

Author Contributions: M.C. and Z.Z. conceived and designed the experiments. X.L., Y.Y., Z.G. and Q.H. performed the experiments. Z.Z., M.Z. and C.L. analyzed the data. X.L. wrote the manuscript. Z.P. directed the study and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Acknowledgements This research was financially supported by the National Key Research and Development Program of China (No. 2016YFD0102104) and Key Project of Education Development of Sichuan Province (No. 035Z2246).

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. References

2. An X, Dong Z, Tian Y et al (2019) ZmMs30 encoding a novel GDSL lipase is essential for male fertility and valuable for hybrid breeding in maize. *Mol Plant* 12:343-359. <https://doi.org/10.1016/j.molp.2019.01.011>
3. Chen X, Zhang H, Sun H et al (2017) IRREGULAR POLLEN EXINE1 is a novel factor in anther cuticle and pollen exine formation. *Plant Physiol* 173:307-325. <https://doi.org/10.1104/pp.16.00629>
4. Cigan AM, Unger E, Xu RJ et al (2001) Phenotypic complementation of ms45 maize requires tapetal expression of Sexual plant reproduction 14:135-142. <https://doi.org/10.1007/s004970100099>
5. Cunmin Q, Fuyou F, Miao L et al (2015) Comparative transcriptome analysis of recessive male sterility (RGMS) in sterile and fertile brassica napus lines. *PloS one* <https://doi.org/10.1371/journal.pone.0144118>
6. Djukanovic V, Smith J, Lowe K et al (2013) Male-sterile maize plants produced by targeted mutagenesis of the cytochrome P450-like gene (MS26) using a re-designed I-Crel homing endonuclease. *Plant J* 76:888-899. <https://doi.org/10.1111/tpj.12335>
7. Evans MMS (2007) The indeterminate gametophyte1 gene of maize encodes a LOB domain protein required for embryo sac and leaf development. *Plant Cell* 19:46-62. <https://doi.org/10.1105/tpc.106.047506>
8. Ferguson AC, Pearce S, Band LR et al (2017) Biphasic regulation of the transcription factor ABORTED MICROSPORES (AMS) is essential for tapetum and pollen development in Arabidopsis. *New Phytol* 213:778-790. <https://doi.org/10.1111/nph.14200>
9. HIGGINS DG (1996) Using CLUSTAL for multiple sequence alignments. *Methods in Enzymology* 266:384-402. [https://doi.org/doi:10.1016/S0076-6879\(96\)66024-8](https://doi.org/doi:10.1016/S0076-6879(96)66024-8)
10. Jung KH, Han MJ, Lee YS et al (2005) Rice undeveloped tapetum1 is a major regulator of early tapetum development. *Plant Cell* 17:2705-2722. <https://doi.org/10.1105/tpc.105.034090>
11. Ko SS, Li MJ, Ku SB et al (2014) The bHLH142 transcription factor coordinates with TDR1 to modulate the expression of EAT1 and regulate pollen development in rice. *Plant Cell* 26:2486-2504. <https://doi.org/10.1105/tpc.114.126292>
12. Li N, Zhang DS, Liu HS et al (2006) The rice tapetum degeneration retardation gene is required for tapetum degradation and anther development. *Plant Cell* 18:2999-3014. <https://doi.org/10.1105/tpc.106.044107>

13. Lin J, Yu C, Chang Y et al (2014) Maize and millet transcription factors annotated using comparative genomic and transcriptomic data. 15:818. <https://doi.org/10.1186/1471-2164-15-818>
14. Liu Y, Li J, Wei G et al (2017) Cloning, molecular evolution and functional characterization of ZmbHLH16, the maize ortholog of OsTIP2 (OsHHLH142). *Biology Open* 6:1654-1663. <https://doi.org/10.1242/bio.026393>
15. Luan F, Delannay I, Staub JE (2008) Chinese melon (*Cucumis melo* L.) diversity analyses provide strategies for germplasm curation, genetic improvement, and evidentiary support of domestication patterns. *Euphytica* 164:445-461. <https://doi.org/10.1007/s10681-008-9699-0>
16. Luo T, Zou T, Yuan G et al (2020) Less and shrunken pollen 1 (LSP1) encodes a member of the ABC transporter family required for pollen wall development in rice (*Oryza sativa* L.). *The Crop Journal* 8:492-504. <https://doi.org/10.1016/j.cj.2019.09.001>
17. Marc A, Tim F, Mary T (2011) MSCA1 nucleotide sequences impacting plant male fertility and method of using same. US patent US007915478B2
18. MC A, T F, A L et al (2016) Cloning and use of the ms9 gene from maize. US patent US20160024520A1
19. Mei S, Liu T, Wang Z (2016) Comparative transcriptome profile of the cytoplasmic male sterile and fertile floral buds of radish (*Raphanus sativus* L.). *International Journal of Molecular Sciences* 17:42. <https://doi.org/10.3390/ijms17010042>
20. Moon J, Skibbe D, Timofejeva L et al (2013) Regulation of cell divisions and differentiation by MALE STERILITY32 is required for anther development in maize. *Plant J* 76:592-602. <https://doi.org/10.1111/tpj.12318>
21. Nan GL, Zhai J, Arikiti S et al (2017) MS23, a master basic helix-loop-helix factor, regulates the specification and development of the tapetum in maize. *Development* 144:163-172. <https://doi.org/10.1242/dev.140673>
22. Niu N, Liang W, Yang X et al (2013) EAT1 promotes tapetal cell death by regulating aspartic proteases during male reproductive development in rice. *Nat Commun* 4:1445. <https://doi.org/10.1038/ncomms2396>
23. R E, L T, R Z (2013) The Role of Id2 in the Regulation of Chromatin Structure and Gene Expression. *Chromatin Remodelling*. 5:91-116.
24. Sekhon RS, Lin H, Childs KL et al (2011) Genome-wide atlas of transcription during maize development. *Plant J* 66:553-563. <https://doi.org/10.1111/j.1365-313X.2011.04527.x>
25. Somaratne Y, Tian Y, Zhang H et al (2017) ABNORMAL POLLEN VACUOLATION1 (APV1) is required for male fertility by contributing to anther cuticle and pollen exine formation in maize. *Plant J* 90:96. <https://doi.org/10.1111/tpj.13476>
26. Tian Y, Xiao S, Liu J et al (2017) MALE STERILE6021 (MS6021) is required for the development of anther cuticle and pollen exine in maize. *Sci Rep* 16736. <https://doi.org/10.1038/s41598-017-16930-0>

27. Vernoud V, Laigle G, Rozier F et al (2009) The HD-ZIP IV transcription factor OCL4 is necessary for trichome patterning and anther development in maize. *Plant J* 59:883-894.
<https://doi.org/10.1111/j.1365-313X.2009.03916.x>
28. Wan X, Wu S, Li Z et al (2019) Maize genic male-sterility genes and their applications in hybrid breeding: Progress and Perspectives. *Mol Plant* 12:321-342.
<https://doi.org/10.1016/j.molp.2019.01.014>
29. Wang CJR, Nan GL, Kelliher T et al (2012) Maize multiple archesporial cells 1 (*mac1*), an ortholog of rice TDL1A, modulates cell proliferation and identity in early anther development. *Development* 139:2594-2603. <https://doi.org/10.1242/dev.077891>
30. Wu F, Liu Z, Jie X et al (2016) Molecular evolution and association of natural variation in *ZmARF31* with low phosphorus tolerance in maize. *Front Plant Sci* 7:1076.
<https://doi.org/10.3389/fpls.2016.01076>
31. Xie K, Wu S, Li Z et al (2018) Map-based cloning and characterization of zea mays male sterility33 (*ZmMs33*) gene, encoding a glycerol-3-phosphate acyltransferase. *Theor Appl Genet* 131:1363-1378.
<https://doi.org/10.1007/s00122-018-3083-9>
32. Yang Z, Liu L, Sun L et al (2019) *OsMS1* functions as a transcriptional activator to regulate programmed tapetum development and pollen exine formation in rice. *Plant Mol Biol* 99:175-191.
<https://doi.org/10.1007/s11103-018-0811-0>
33. Zhang D, Wu S, An X et al (2017) Construction of a multicontrol sterility system for a maize male-sterile line and hybrid seed production based on the *ZmMs7* gene encoding a PHD-finger transcription factor. *Plant Biotechnol J* 16:459-471. <https://doi.org/10.1111/pbi.12786>
34. Zhang H, Gao S, Lercher MJ et al (2012) *EvoView*, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Res* W1:569-572. <https://doi.org/10.1093/nar/gks576>
35. Zhu Q, Song Y, Zhang G et al (2015) De novo assembly and transcriptome analysis of wheat with male sterility induced by the chemical hybridizing agent SQ-1. *Plos One* 10:e123556.
<https://doi.org/10.1371/journal.pone.0123556>

Tables

Tab. 1 Genetic segregation of fertility in F₂ populations of *ms40*

Population	Total	Fertile plants	Sterile plants	Theoretical ratio (fertile: sterile)	c^2_c	Location and years
(<i>ms40</i> × <i>B73</i>)F ₂	194	155	39	3:1	2.227	Sichuan, 2017
(<i>ms40</i> × <i>B73</i>)F ₂	582	354	129	3:1	0.663	Sichuan, 2018
(<i>ms40</i> × <i>B73</i>)F ₂	1276	978	298	3:1	1.757	Yunnan, 2019
(<i>ms40</i> × <i>B73</i>)F ₂	1208	911	297	3:1	0.089	Sichuan, 2020
(<i>ms40</i> × <i>Mo17</i>)F ₂	1589	1218	371	3:1	2.226	Sichuan, 2020

Note: $c^2_{(0.05,1)} = 3.84$

Tab. 2 The genes information of the fine mapping interval

No.	Gene ID	Location	Description
1	<i>Zm00001d053890</i>	Chr4:243050411-243053629	SBP-Transcription factor 20
2	<i>Zm00001d053891</i>	Chr4:243143280-243151346	No annotation
3	<i>Zm00001d053894</i>	Chr4:243184586-243227697	No annotation
4	<i>Zm00001d053895</i>	Chr4:243230105-243233634	bHLH-Transcription factor 51
5	<i>Zm00001d053896</i>	Chr4:243263176-243265641	Alpha-trehalose-phosphate synthase [UDP-forming] 5

Figures

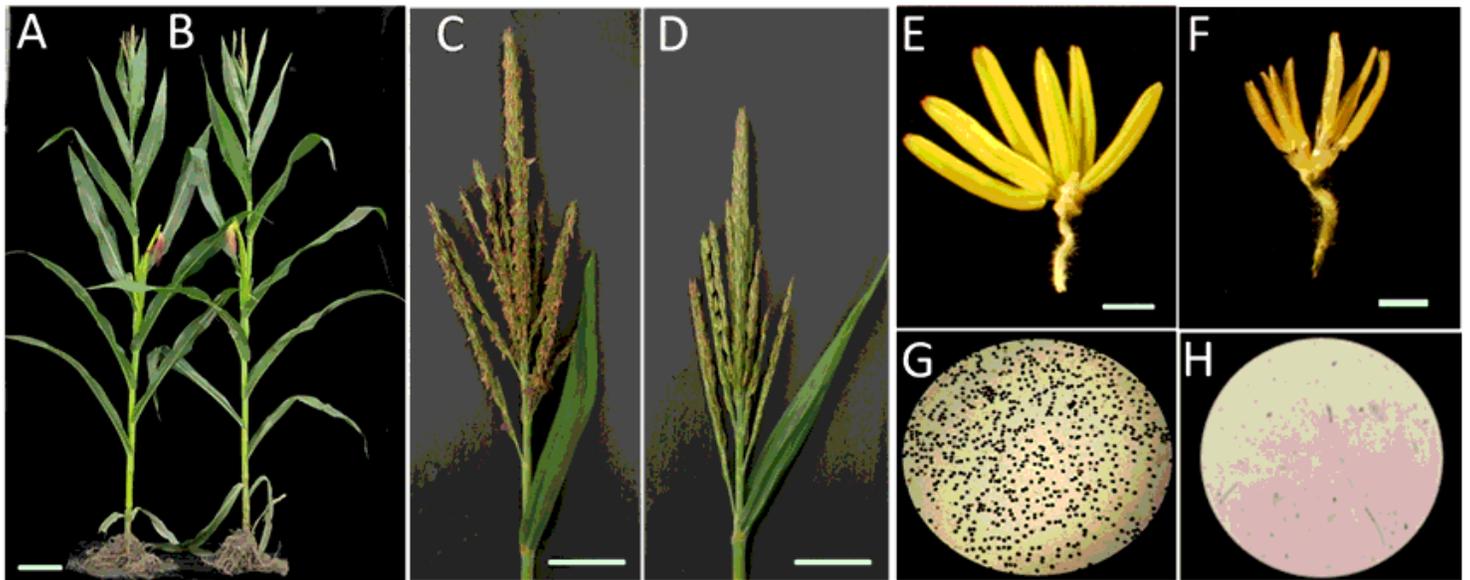


Figure 1

Phenotypes comparison between RP125 and ms40. (A-B) The plants of RP125 (A) and ms40 (B) at flowering stage. (C-D) The tassels of RP125 (C) and ms40 (D). (E-F) The anthers of RP125 (E) and ms40 (F). (G-H) The pollen grains of RP125 (G) and ms40 (H) with I2-KI staining. Scale bars: 20 cm in (A, B); 5 cm in (C, D); 1 mm in (E, F).

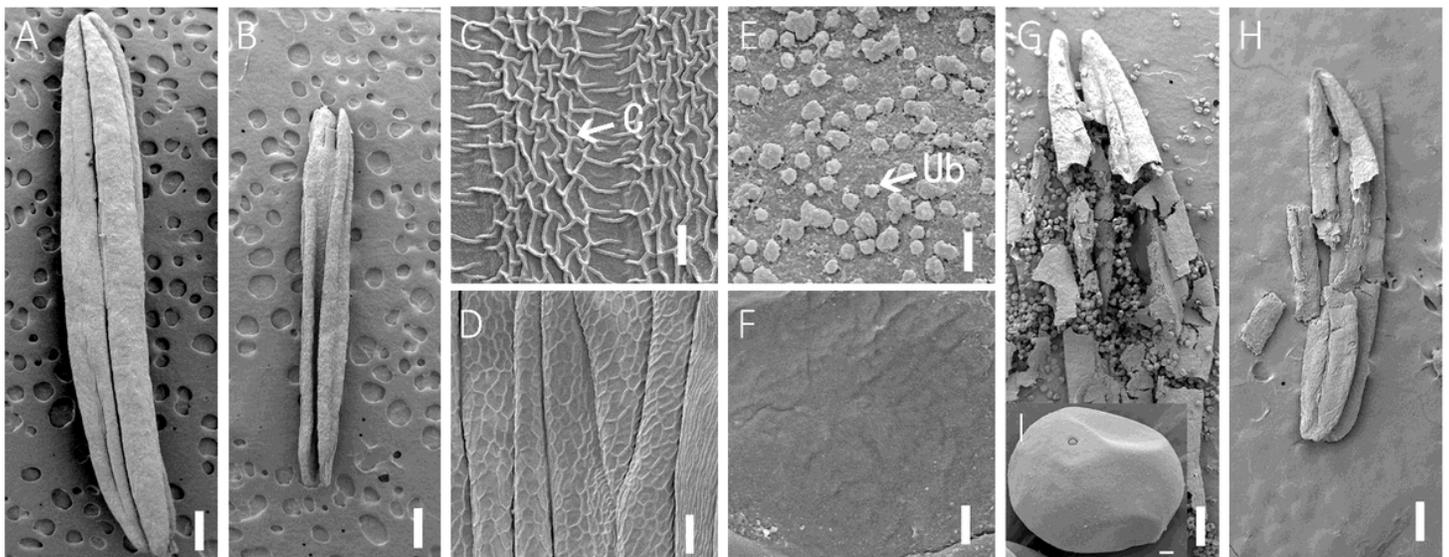


Figure 2

SEM analysis of RP125 and ms40 anthers at mature pollen grain stage. (A, B) The anthers of RP125 (A) and ms40 (B). (C, D) The anthers epidermis of RP125 (C) and ms40 (D). (E, F) The anthers inner surfaces of RP125 (E) and ms40 (F). (G, H) The broken anthers of RP125 (G) and ms40 (H). (I) The mature pollen grains of RP125. C, cuticle; Ub, Ubisch body. Scale bars: 400 μ m in (A, B); 6 μ m in (C, I).

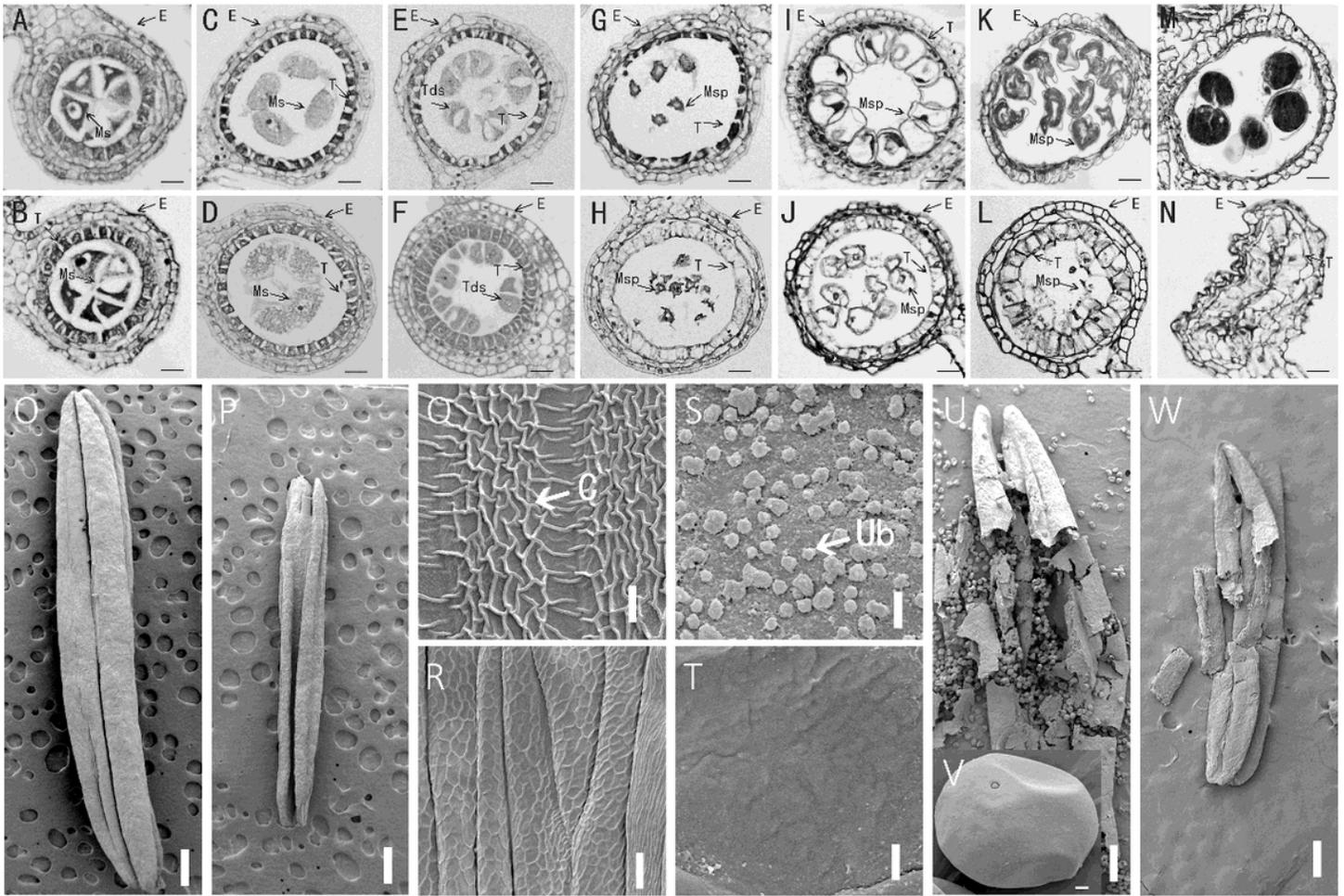


Figure 3

Transverse section of RP125 and ms40 anthers and TUNEL assay of DNA fragmentation in RP125 and ms40 tapetum. (A, B) RP125 (A) and ms40 (B) anthers at the sporogenous stage. (C, D) RP125 (C) and ms40 (D) anthers at the pollen mother cell stage. (E, F) RP125 (E) and ms40 (F) anthers at the tetrad stage. (G, H) RP125 (G) and ms40 (H) anthers at the microspore release stage. (I, J) RP125 (I) and ms40 (J) anthers at the large vacuole stage. (K, L) RP125 (K) and ms40 (L) anthers at the binucleate stage. (M, N) RP125 (M) and ms40 (N) anthers at the mature pollen stage.

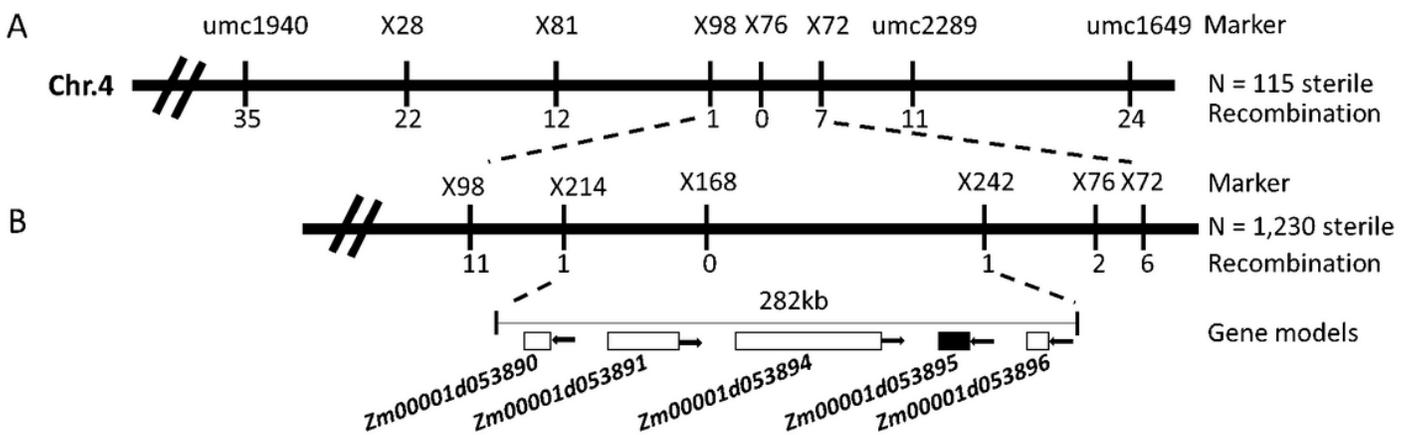


Figure 4

Map-based cloning of *ms40*. (A) Primary mapping of *ms40* using the (*ms40*×*B73*)F₂ population including 115 male sterile individuals. (B) Fine mapping of *ms40* using the (*ms40*×*B73*)F₂ population including 1230 male sterile individuals. *ms40* was mapped to an interval of approximately 282 kb flanked by the X214 and X242 markers.

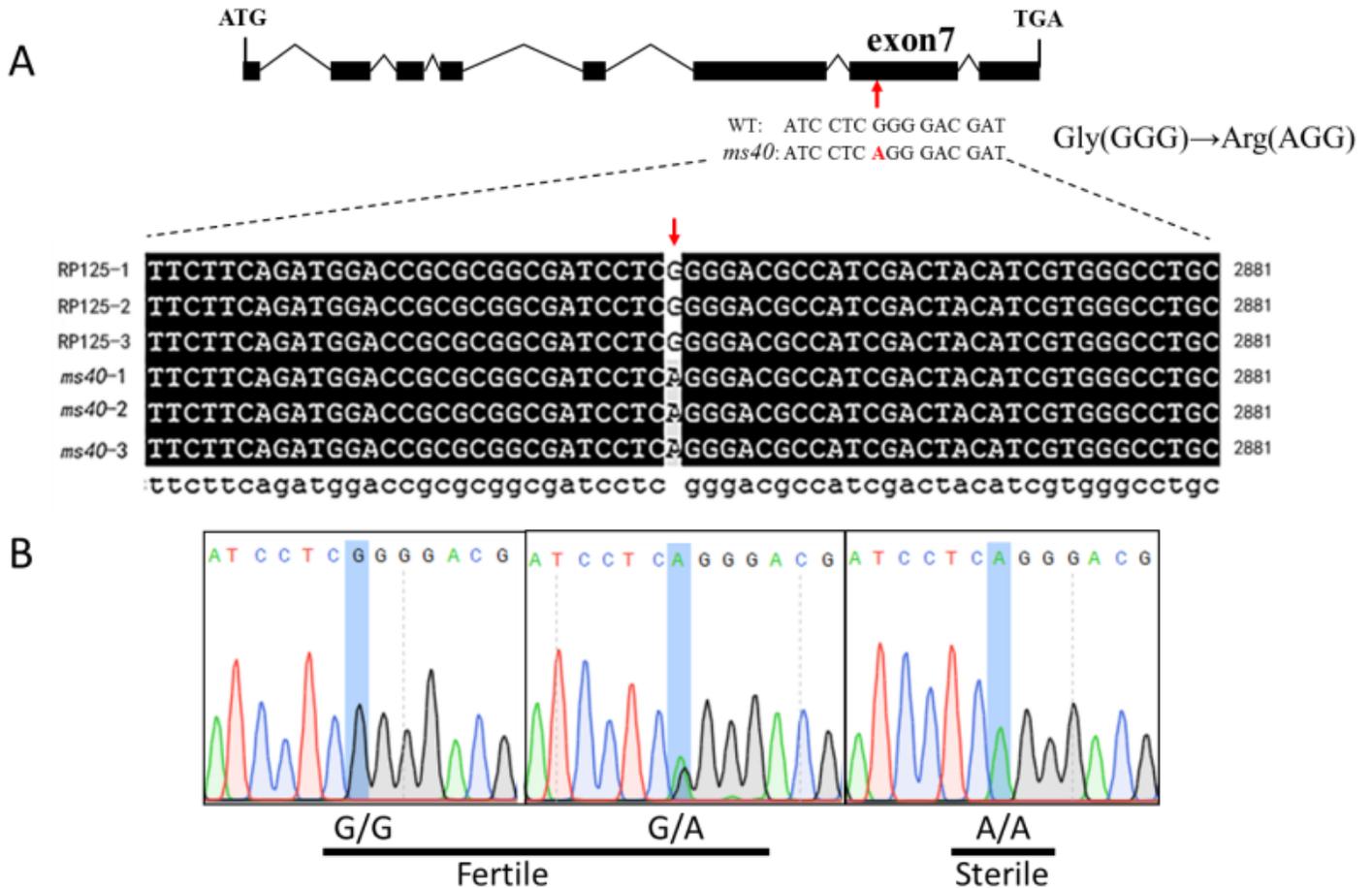


Figure 5

Mutation site analysis of key candidate genes. (A) The gene structure and sequence alignment of Zm00001d053895 between fertile plants (RP125-1/-2/-3) and sterile plants (*ms40*-1/-2/-3). Black boxes and black bent lines represent exons and introns, respectively. The red arrow indicates the mutation site and mutation site was located in exon 7 of Zm00001d053895 which result in a change in amino acid from Gly (GGG) to Arg (AGG). (B) Sequencing diagrams of the mutation site between the fertile and sterile plants, the G/G and G/A represent fertile plants, A/A represents sterile plants, the blue box represents sequencing diagrams of the mutation site.

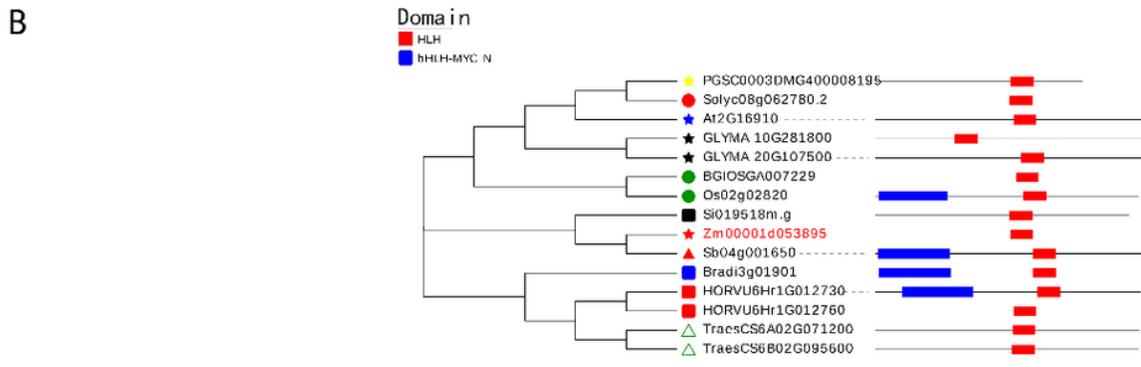
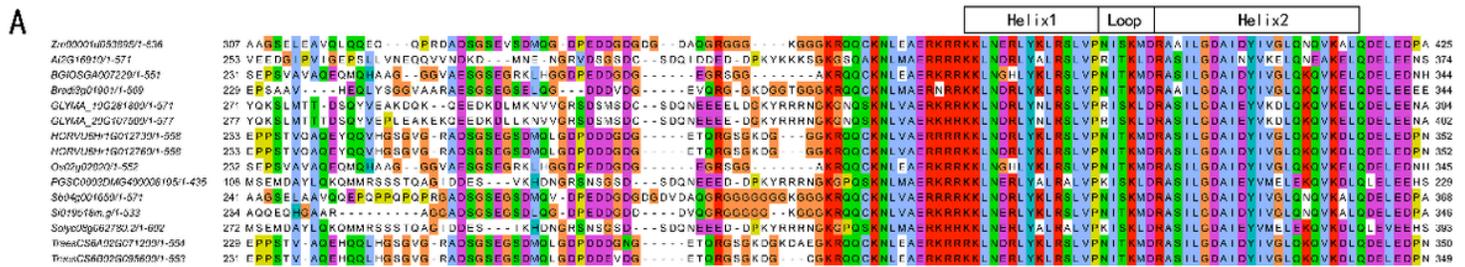


Figure 6

Protein sequence and phylogenetic analysis. (A) Multiple sequence alignment of the amino acids from Zm00001d053895 and related proteins of other species (only part of the comparison results are shown). The black frame indicates the HLH domain region. The red arrow indicates mutation site of ms40. (B) Phylogenetic tree of Zm00001d053895 and its homologs in other plant species. The percentage identities represent the sequence similarity between the corresponding proteins and Zm00001d053895. The yellow star represents *Solanum tuberosum*, the red circle represents *Solanum lycopersicum*, the blue star represents *Arabidopsis thaliana*, the black star represents *Glycine max*, the green circle represents *Oryza sativa*, the black rectangle represents *Foxtail millet*, the red star represents *Zea mays*, the red triangle represents *Sorghum*, the blue rectangle represents *Brachypodium*, the red rectangle represents *Hordeum vulgare*, and the green triangle represents *Triticum aestivum*.

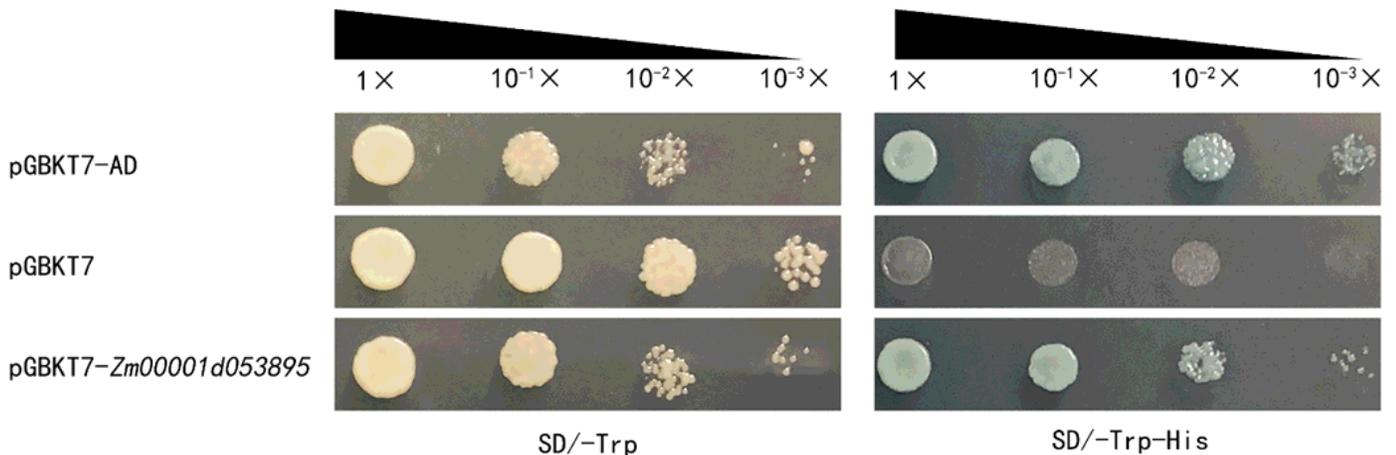


Figure 7

transcriptional activation assay of Zm00001d053895 in the AH109 yeast strain. Free pGBKT7 and pGBKT7-GAL4 AD as negative and positive control respectively. Yeast transformants were spotted on to control medium (SD/-Trp) and selective medium (SD/-Trp/-His).

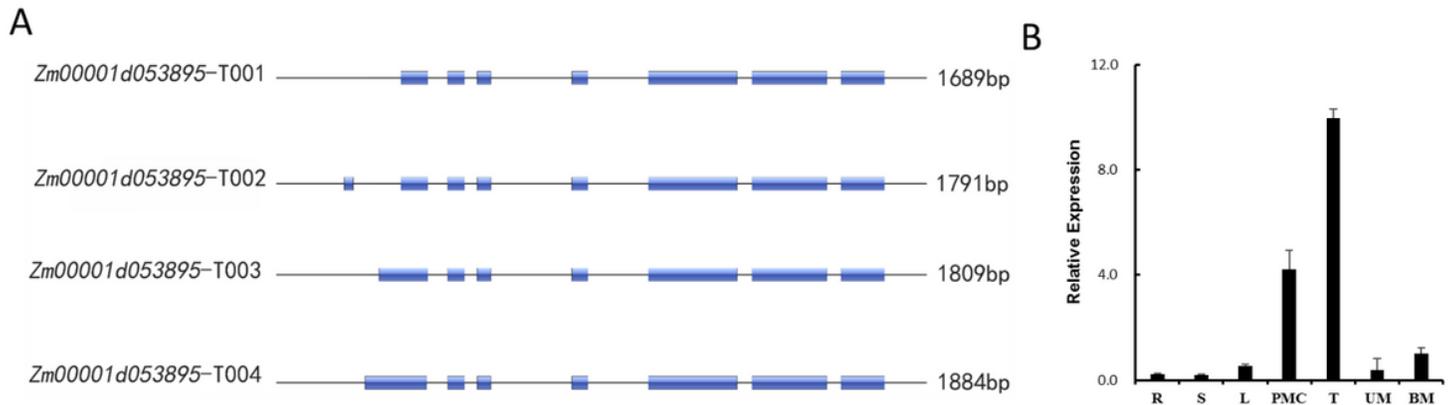


Figure 8

Transcripts analysis and expression pattern analysis of Zm00001d053895. (A) Schematic representation of Zm00001d053895 transcripts, the blue box represents exon, Zm00001d053895 have four transcripts, named Zm00001d053895-T001 (seven exon), Zm00001d053895-T002 (eight exon), Zm00001d053895-T003 (seven exon), Zm00001d053895-T004 (seven exon), respectively. (B) The data are given as the mean \pm SEMs of at least three biological replicates. R, root; S, stem; L, leaf; PMC, pollen mother cell; T, Tetrad; UM, uninucleate microspore; BM, bicellular microspore.

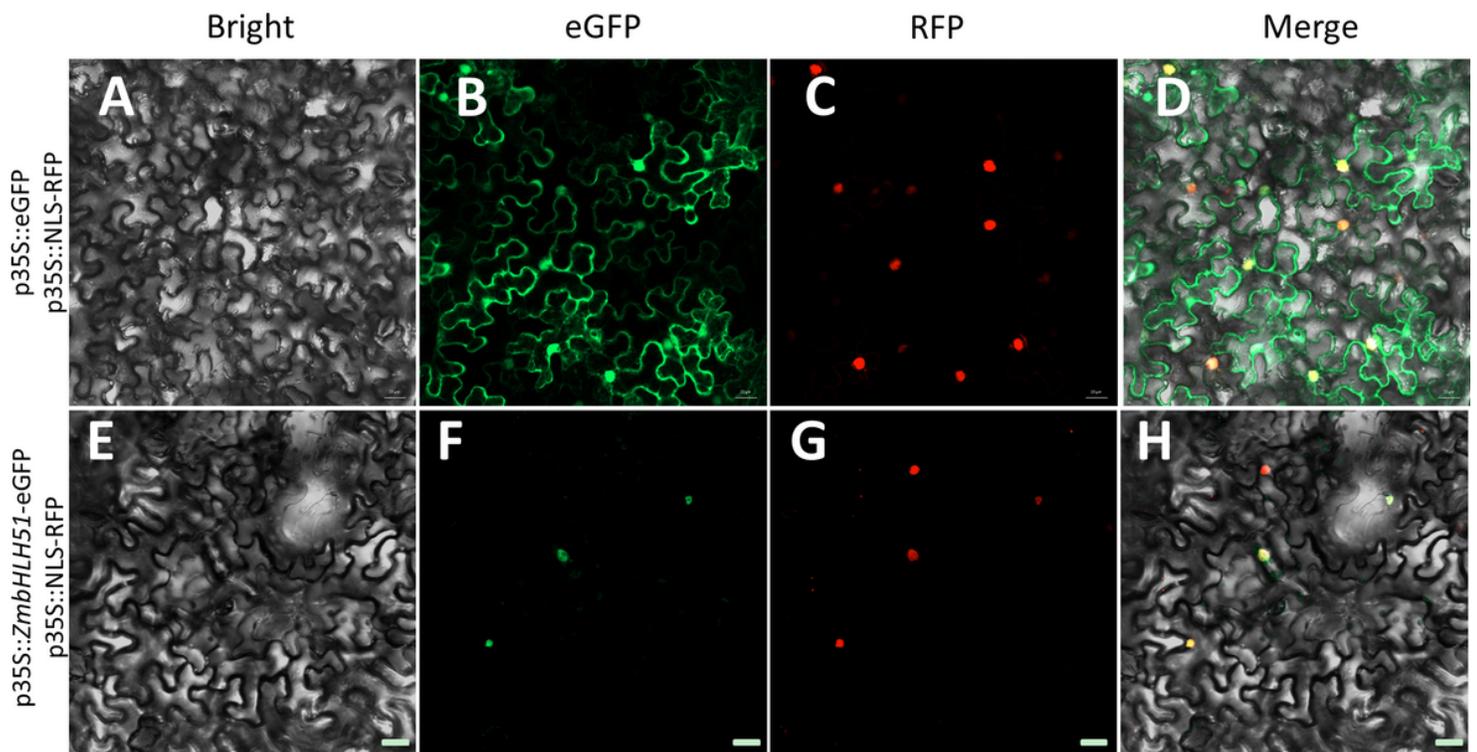


Figure 9

Subcellular localization of Zm00001d053895. (A-D) Tobacco leaves transformed with p35S::NLS-RFP and p35S::eGFP as controls. (E-H) Zm00001d053895-eGFP targeted the cell nucleus in tobacco leaves. The bright-field (A, E), eGFP fluorescence (green) (B, F), RFP fluorescence (red) (C, G), and merged images (D, H) are shown. Scale bars: 20 μ m.

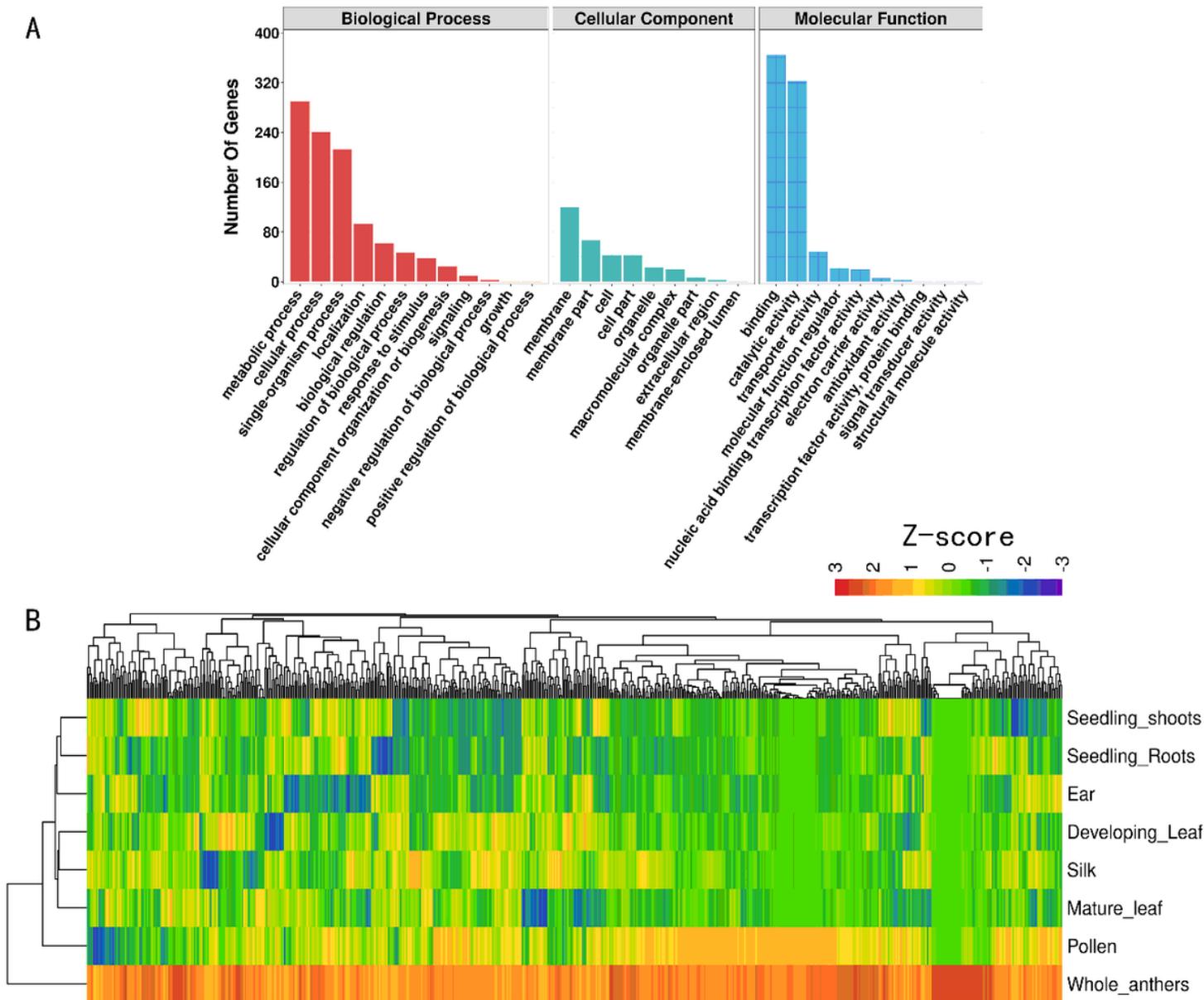


Figure 10

GO annotations and expression patterns of co-expressed genes. (A) GO analysis of 1192 genes coexpressed with Zm00001d053895. (B) Cluster heatmap of 647 genes specifically expressed in anthers. The bar indicates the relative gene expression level. These gene expression data were retrieved from the qTeller database and log₂-normalized (original data+1).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Tab.S1.docx](#)
- [Tab.S2.docx](#)