CRISPR/Cas9-based discovery of maize transcription factors regulating male sterility and their functional conservation in plants

Yilin Jiang1,†, Xueli An1,2,†, Ziwen Li1,2,†, Tingwei Yan1,†, Taotao Zhu1, Ke Xie1,2, Shuangshuang Liu1,2, Quancan Hou1,2, Lina Zhao1,2, Suowei Wu1,2, Xinze Liu1, Shaowei Zhang1, Wei He1, Fan Li1, Jingping Li2 and Xiangyuan Wan1,2,*

1Zhongzhi International Institute of Agricultural Biosciences, Biology and Agriculture Research Center of USTB, University of Science and Technology Beijing (USTB), Beijing, China
2Beijing Engineering Laboratory of Main Crop Bio-Tech Breeding, Beijing International Science and Technology Cooperation Base of Bio-Tech Breeding, Beijing Solidwill Sci-Tech Co. Ltd, Beijing, China

Received 5 February 2021; revised 9 March 2021; accepted 17 March 2021.
*Correspondence (Tel 86 186 0056 1850; fax 86 10 82346928; email wanxianguan@ustb.edu.cn)
†These authors contributed equally to this work.

Keywords: CRISPR, Cas9, transcription factor, anther and pollen development, genic male sterility, maize.

Summary
Identifying genic male-sterility (GMS) genes and elucidating their roles are important to unveil plant male reproduction and promote their application in crop breeding. However, compared with Arabidopsis and rice, relatively fewer maize GMS genes have been discovered and little is known about their regulatory pathways underlying anther and pollen development. Here, by sequencing and analysing anther transcriptomes at 11 developmental stages in maize B73, Zheng58 and M6007 inbred lines, 1100 transcription factor (TF) genes were identified to be stably differentially expressed among different developmental stages. Among them, 14 maize TF genes (9 types belonging to five TF families) were selected and performed CRISPR/Cas9-mediated gene mutagenesis, and then, 12 genes in eight types, including ZmbHLH51, ZmbHLH122, ZmTGA9-1/-2/-3, ZmTGA10, ZmMYB84, ZmMYB33-1/-2, ZmPHD11 and ZmLBD10/27, were identified as maize new GMS genes by using DNA sequencing, phenotypic and cytological analyses. Notably, ZmTGA9-1/-2/-3 triple-gene mutants and ZmMYB33-1/-2 double-gene mutants displayed complete male sterility, but their double- or single-gene mutants showed male fertility. Similarly, ZmLBD10/27 double-gene mutant displayed partial male sterility with 32.18% of aborted pollen grains. In addition, ZmbHLH51 was transcriptionally activated by ZmbHLH122 and their proteins were physically interacted. Molecular markers co-segregating with these GMS mutations were developed to facilitate their application in maize breeding. Finally, all 14-type maize GMS TF genes identified here and reported previously were compared on functional conservation and diversification among maize, rice and Arabidopsis. These findings enrich GMS gene and mutant resources for deeply understanding the regulatory network underlying male fertility and for creating male-sterility lines in maize.

Introduction
Male gametophyte development of flowering plants is a complex biological process requiring cooperative interactions between sporophytic and gametophytic tissues (Ma, 2005; McCormick, 1993). Genic male sterility (GMS) caused by nuclear gene alone can be generated by any disturbance in the process and is a useful agronomic trait for crop breeding and hybrid seed production, especially for crops with high heterosis such as maize and rice. The discovery of GMS genes and their corresponding mutants can facilitate understanding the molecular mechanism of anther development and developing biotechnology-based male-sterility (BMS) systems for hybrid seed production in crops (Wan et al., 2019; Wan et al., 2021).

So far more than 100 GMS genes have been identified in plants. These genes can be classified into four types based on their functions, including (1) genes encoding transcription factors (TFs), (2) lipid metabolic genes, (3) polysaccharide metabolic genes and (4) genes involved in other processes (Wan et al., 2019). Among them, GMS TF genes are reported to play crucial roles in multiple biological processes related to anther and pollen development. Currently, at least 18 and 15 GMS TF genes have been identified in Arabidopsis and rice, respectively (Wan et al., 2019) (the gene information was listed in Table S1). Notably, five of these Arabidopsis GMS TFs form a core genetic pathway (ADYT1-AtTDF1-AtAMS-AtMYB80-AtMs1) to regulate tapetum and anther development (Lou et al., 2018; Zhu et al., 2008), and the corresponding genetic pathway (OsUDT1-OsTDF1-OsTDR-
OsMYB80-OsPTC1) is proposed to be conserved in rice (Cai et al., 2015). Compared with Arabidopsis and rice, relatively few GMS TF genes have been identified in maize. Moreover, whether this core regulatory pathway with the five key GMS TFs exists in maize is unknown.

In maize, only six GMS TF genes have been reported to control male sterility so far. For example, OUTER CELL LAYER 4 (OCL4) encodes a homeobox-leucine zipper (HD-ZIP) TF regulating maize anther wall division and differentiation (Vernoud et al., 2009). ZmMs32 and ZmMs23 belong to basic helix-loop-helix (bHLH) TF family and function at anther premeiotic stages (Moon et al., 2013; Nan et al., 2017). ZmMs9 as the ortholog of AtTDF1 and OsTDF1 regulates early microspore development (Albertsen et al., 2016). Plant homeodomain (PHD) finger family gene ZmMs7 as the ortholog of AtMs1 and OsPTC1 controls postmeiotic anther development (An et al., 2020; Zhang et al., 2018), and GAMETO PHYTE 1 (G1) encoding a LATERAL ORGAN BOUNDARIES DOMAIN (LBD) TF is required for both female and male developments (Evans, 2007). Nevertheless, little is known about their roles on the regulatory network of anther development.

RNA sequencing (RNA-seq) provides a more precise and quantitative measurement of gene expression than microarrays (Wan and Li, 2019; Wang et al., 2009). Several RNA-seq transcriptomes have been reported for whole anther (An et al., 2020; Nan et al., 2017; Zhu et al., 2020) or anther single cells (Nelms and Walbot, 2019) in maize. Based on maize anther RNA-seq data of Oh43 line at six developmental stages (S5 to S9) and maize orthologs of the reported GMS genes identified in Arabidopsis and rice but without functional characterizations in maize, we predicted 62 putative maize GMS genes (Wan et al., 2019). Additionally, based on maize anther RNA-seq data of Oh43 and W23, we predicted 125 novel lipid metabolic GMS gene candidates in maize (Wan et al., 2020). However, whether the predicted GMS candidates are real GMS genes controlling maize male sterility is unknown. With the development of CRISPR/Cas9 technology widely used in many crops, it has been proved that the reverse genetic strategy is effective to discover and confirm gene functions in plants (Nadakuditi and Enciso-Rodríguez, 2021). Specifically, the CRISPR/Cas9 technology has been demonstrated to be efficient for gene editing in maize (Char et al., 2017; Feng et al., 2018; Qi et al., 2020), soybean (Li et al., 2015), sorghum (Jiang et al., 2013) and tomato (Cermak et al., 2015). Notably, a rice GMS-enriched mutant library has been successfully established by CRISPR/Cas9-mediated knock-outs of anther-specific expressed genes identified by microarray (Ma et al., 2019).

There are more than 39,000 protein-coding genes in maize genome (AGPV4) (Jiao et al., 2017), 2216 (5.7%) of them encode TFs (data extracted from plant TFDB). Furthermore, more than 20,000 protein-coding genes have detectable expression levels by microarray analysis (Ma et al., 2008). Similarly, more than 20,000 protein-coding genes are expressed in RNA-seq data of maize anthers, and 1220 (6.0%) of them are TF genes (Li et al., 2019). However, the genetic variations largely vary among different maize inbred lines. So far, there has been no report by which multiple sets of maize anther transcriptomes from different genetic backgrounds are compared and analysed to reveal potential GMS genes in maize. Moreover, RNA-seq analysis on expression patterns of TF genes covering anther whole developmental period could provide an important cue for understanding the regulatory mechanism of TF genes underlying anther development.

Here, we first sequenced maize anther transcriptomes from three inbred lines (B73, Zheng58 and M6007) across 11 developmental stages (S5 to S12) and found an overlapping set of 1100 stage differentially expressed (stage-DE) TF genes that were mainly grouped into six expression clusters. Then, combining the information of GMS genes reported in Arabidopsis and rice, we selected nine types including 14 TF genes and performed CRISPR/Cas9-mediated mutagenesis. As a result, eight types of them were demonstrated as new GMS genes in maize. Meanwhile, expression regulation and protein interaction between bHLH51 and bHLH122 were analysed, and molecular markers used for detecting GMS mutations were developed. Finally, we compared the functional conservation and diversification of all the identified GMS TF genes among maize, rice, and Arabidopsis.

**Results**

**Stage differentially expressed TF genes under three maize genetic backgrounds during anther developmental stages 5 to 12**

In maize genome, there are 2216 protein-coding genes predicted as TF genes, six (ZmOCL4, ZmMs32, ZmMs9, ZmMs23, ZmMs7 and ZmG1) of them were reported as GMS genes in maize (Wan et al., 2019). To identify more GMS TF genes regulating maize anther development, we sequenced anther transcriptomes at 11 developmental stages (S5 to S12) in maize B73, Zheng58 and M6007 lines (Figures S1a, S2a and S3a), respectively, and then identified genome-wide stage-DE TF genes by analysing the anther transcriptome data. As a result, we found 21 560, 23 011 and 22 157 stage-DE genes (stage DEGs) in B73, Zheng58 and M6007 anthers, respectively, among which 1408, 1429 and 1340 stage DEGs encoded TFs (Figure 1a) and displayed obvious expression patterns across the investigated stages in each genetic background (Figure 1b). Moreover, the expression patterns of stage-DE TF genes in each line were grouped into 60 types that were further integrated into seven expression clusters (early/ premeiotic, middle/meiotic, late/postmeiotic, early-middle, early-middle, middle-late and other clusters; Figures S1b, S2b and S3b).

These results indicated that maize anther genome contains a number of TF genes with stage-DE expression patterns.

To identify the shared stage-DE TF genes among different maize lines, we compared and found 1100 overlapping stage-DE TF genes in the three lines (Figure 1c; Table S2). The ratios of the shared genes are 78% (1100/1408), 77% (1100/1429) and 82% (1100/1340) of the total stage-DE TF genes identified in B73, Zheng58 and M6007 anthers, respectively, indicating that most of detected stage-DE TF genes have stable expression patterns among the three lines. Moreover, all the six reported maize GMS TF genes were included in the 1100 genes (Table S2).

We further analysed the gene family components and co-expression patterns of the 1100 stage-DE TF genes and found the similar enrichment patterns between the totally shared gene set of 1100 genes and each of the three set stage-DE TF gene in B73, Zheng58 and M6007 (Figure 1d), indicating the gene family components were relatively consistent among the three lines. Furthermore, the expression patterns of the shared 1068 stage-DE TF genes were grouped into six co-expression gene clusters in B73 line (Figure 1e), including early stage (152 TF genes), middle stages (365 TF genes), late stages (205 TF genes), early-middle stages (95 TF genes), early-late stages (118 TF genes) and middle-late stages (133 TF genes).
Figure 1 Identification and analysis of stage differentially expressed TF genes during maize anther development (stages 5 to 12). (a) The amounts of identified stage-DE TF genes during maize anther development in three maize inbred lines (B73, Zheng58 and M6007). (b) The expression patterns of 1408, 1429 and 1340 identified stage-DE TF genes during maize anther developmental stages 5 to 12 in anther transcriptomes of B73, Zheng58 and M6007, respectively. (c) The identification of 1100 stage-DE TF genes shared by B73, Zheng58 and M6007 during maize anther developmental stages 5 to 12. (d) Gene family enrichment analysis on stage-DE TF genes. 42 gene families with more than or equal to 10 members were displayed. (e) 1100 stage-DE TF genes shared by B73, Zheng58 and M6007 during maize anther development were mainly grouped into six expression clusters. Grey backgrounds indicate middle or meiotic stages (S6 to S9-10) during maize anther development.
To investigate their potential functions in controlling maize male sterility, we selected 14 representative genes (i.e. two bHLH, four bZIP, five MYB, and two LBD TF genes) (Figures 2-5; S4-S8) homologous to GMS genes reported in Arabidopsis and/or rice but not in maize and performed knock-out mutagenesis by using the CRISPR/Cas9 strategy.

**CRISPR/Cas9-based generation of bHLH family gene mutants and protein interaction between ZmbHLH51 and ZmbHLH122**

In flowering plants, bHLH TFs belong to a large TF family with 213, 178 and 170 genes in maize, rice and Arabidopsis, respectively (Carretero-Paulet et al., 2010; Lin et al., 2014). Based on the sequence similarity, ZmbHLH51 (Zm00001d053895) is the ortholog of OsTDR and AtAMS, and ZmbHLH122 (Zm00001d017724) is the ortholog of OsEAT1 (Figure S9). To investigate their functions, the CRISPR/Cas9 system was used to generate bhlh51 and bhlh122 mutants. Two targets were designed in the first exon of ZmbHLH51 and ZmbHLH122, respectively, to construct the corresponding CRISPR/Cas9 vectors (Figure S6a, b), which were introduced into maize hybrid Hi II by Agrobacterium-mediated transformation. By using DNA sequencing, we identified the mutagenesis of ZmbHLH51 and ZmbHLH122 in primary T0 transgenic plants. Three homozygous ZmbHLH51 loss-of-function mutants with frasemesh mutations due to 43-bp deletion, 44-bp deletion and 24-bp Indel, respectively (Figure 2a1, a2, Table S3), were selected from independent T0 lines for further investigation. Similarly, three homozygous mutants of ZmbHLH122 with 2-bp Indel, 1-bp deletion and 23-bp Indel, respectively, were selected (Figure 2b1, b2; Table S3). All anthers of the above six homozygous T0 mutants produced no visible pollen compared with those of wild-type (WT) anther. To obtain stable and heritable mutants, we pollinated these T0 mutants with pollen grains from Zheng58 to harvest F1 seeds that can be divided into Cas9-positive (transgenic lines) and Cas9-negative (non-transgenic) segregation. To eliminate the interference of sgRNA and Cas9, we genotyped the Cas9-negative plants and selected F1 plants with Zheng58 genotype and male-sterility genotype to perform self-pollination to further generate F2 seeds. Using molecular markers co-segregating with mutations of ZmbHLH51 and ZmbHLH122 (Figure S8) to genotype the derived F2 plants, we identified three homozygous transgene-free mutants for each of the two TF genes, which are used for the following phenotypic and cytological observations.

Compared with WT anthers, the three bhlh51 mutants (ZmbHLH51-Cas9-1, ZmbHLH51-Cas9-2 and ZmbHLH51-Cas9-3) and three bhlh122 mutants (ZmbHLH122-Cas9-1, ZmbHLH122-Cas9-2 and ZmbHLH122-Cas9-3) displayed complete male sterility without exerted anthers and pollen grains (Figure 2a3, b3). In addition, we performed scanning electron microscopy (SEM) analysis. Compared with WT anthers, all the bhlh51 and bhlh122 mutants had smaller and wilterd anthers without visible pollen grains and displayed the smooth inner and outer surfaces of anther wall without knitting cuticle and Ubisch bodies (Figure 2a4, b4), demonstrating that both ZmbHLH51 and ZmbHLH122 are GMS genes required for male fertility in maize.

To investigate gene regulation and protein interaction between ZmbHLH51 and ZmbHLH122 with overlapping expression patterns (Figure 2c1), we performed transient dual-luciferase (LUC), Co-IP and BiFC assays. As a result, ZmbHLH51 promoter was significantly activated by ZmbHLH122, while ZmbHLH122 promoter was not activated by ZmbHLH51 (Figure 2c2), suggesting that ZmbHLH51 is transcriptionally activated by ZmbHLH122. Additionally, ZmbHLH122 physically linked ZmbHLH51 based on Co-IP assay (Figure 2c3), which was further confirmed by BiFC assay (Figure 2c4). Collectively, ZmbHLH122 can directly activate the expression of ZmbHLH51 and interact with ZmbHLH51 to form a heterodimer which may function in maize male fertility.

**CRISPR/Cas9-based generation of bZIP family gene mutants and their male-sterility phenotype evaluation**

AtTGA9/10 (Murmu et al., 2010) and OsTGA10 (Chen et al., 2018) are involved in archespore cell specification and male fertility in Arabidopsis and rice, respectively. There are three orthologs of AtTGA9 in maize genome, that is ZmTGA9-1 (Zm00001d052543), ZmTGA9-2 (Zm00001d042777) and ZmTGA9-3 (Zm00001d012294), and one maize ortholog (ZmTGA10, Zm00001d020938) of AtTGA10 and OsTGA10 (Figure S9). However, functions of the four maize TGA orthologs remain unclear.

To investigate their functions, we generated maize tga9 and tga10 mutants by using the CRISPR/Cas9 system. Since ZmTGA9 has three paralogs in maize genome, we constructed two vectors, each containing the Cas9 and two sgRNA expression cassettes (Figure S6c1, c2), which were co-introduced into maize hybrid Hi II to simultaneously edit the three paralogs. Three homozygous triple-gene mutants from independent T0 transgenic lines were thus obtained, including ZmTGA9-1/2/3-Cas9-1 (with 3-bp Indel in ZmTGA9-1, 74-bp insertion and transversion of T to G in ZmTGA9-2 and 1-bp insertion in ZmTGA9-3), ZmTGA9-1/2/3-Cas9-2 (with 2-bp Indel in ZmTGA9-1, 2-bp Indel in ZmTGA9-2 and 1-bp insertion in ZmTGA9-3) and ZmTGA9-1/2/3-Cas9-3 (with 215-bp deletion in ZmTGA9-1, 18-bp deletion in ZmTGA9-2 and 3-bp deletion in ZmTGA9-3) (Figure 3a1, a2; Table S3). All the three triple-gene T0 mutants displayed complete male sterility (Figure 3a3). In F2 generation, neither single-gene mutants nor double-gene mutants showed male sterility (Figure 34). Notably, the three triple-gene mutants were normal in vegetative organ growth and female fertility but displayed complete male sterility without exerted anthers and pollen grains compared with WT anthers (Figure 3a3), and exhibited the smooth inner surface without Ubisch bodies and fewer knitting cuticle on the outer surface of anther wall (Figure 3a4). Therefore, we can conclude that the redundant functions exist among the three ZmTGA9 paralogs in controlling maize male fertility. It may be the reason that ZmTGA9 has not been reported as a GMS gene so far since the simultaneous natural mutations of its three paralogs are greatly low probability.

To investigate function of ZmTGA10, two vectors were constructed to target the fourth and ninth exons of ZmTGA10, respectively (Figure S6d1, d2). We produced and selected three homozygous tga10 mutants including ZmTGA10-Cas9-1 (with 167-bp deletion in the ninth exon), ZmTGA10-Cas9-2 (with 162-bp deletion in the ninth exon) and ZmTGA10-Cas9-3 (with two 1-bp insertions in the fourth exon) (Figure 3b1, b2; Table S3). Compared with WT anthers, all the three tga10 mutants had exerted anthers but failed to dehisce (Figure 3b3). There was no obvious difference in anther and pollen grain sizes between WT and each of tga10 mutants after manually striping off mutant anthers (Figure 3b3). Nevertheless, SEM observation showed that some fragmentary materials were pasted on the pollen grain surface of tga10 mutants, and the sizes of Ubisch bodies on the inner surface of tga10 anther wall were slightly smaller than those of WT anther (Figure 3b4). These findings proved that ZmTGA10
Figure 2. The CRISPR/Cas9 mutagenesis and the derived mutant characterization of ZmbHLH51 and ZmbHLH122 as well as their protein interaction determination. (a) The CRISPR/Cas9 mutagenesis and mutant characterization of ZmbHLH51. (a1) Gene structure and mutation analysis of ZmbHLH51 in WT and three knock-out lines (ZmbHLH51-Cas9-1, ZmbHLH51-Cas9-2 and ZmbHLH51-Cas9-3) generated by the CRISPR/Cas9 technology. (a2) Identification of the mutation sites and fragments in three knock-out lines based on DNA sequencing. (a3) Phenotypic analysis of tassels, anthers and pollen grains stained with 1% I$_2$KI solution in WT and the three knock-out lines. (a4) SEM analysis of anther and pollen in WT and the three knock-out mutants at maize anther developmental stage 13. (b) The CRISPR/Cas9 mutagenesis and mutant characterization of ZmbHLH122. (b1) Gene structure and mutation analysis of ZmbHLH122 in WT and three knock-out lines (ZmbHLH122-Cas9-1, ZmbHLH122-Cas9-2 and ZmbHLH122-Cas9-3) generated by the CRISPR/Cas9 technology. (b2) Identification of the mutation sites and fragments in three knock-out lines based on DNA sequencing. (b3) Phenotypic analysis of tassels, anthers and pollen grains stained with 1% I$_2$KI solution in WT and the three knock-out lines. (b4) SEM analysis of anther and pollen in WT and the three knock-out mutants at maize anther developmental stage 13. (c) The relationship determination between ZmbHLH51 and ZmbHLH122. (c1) RT-qPCR expression analysis of ZmbHLH51 and ZmbHLH122 during maize anther 11 sequential development stages. (c2) Transient dual-luciferase assay of ZmbHLH51 and ZmbHLH122. (c3 and c4) Determination of protein interaction between ZmbHLH51 and ZmbHLH122 based on Co-IP (c3) and BiFC assays (c4) using a transient expression system in maize protoplast.
Figure 3  The CRISPR/Cas9 mutagenesis and the derived mutant characterization of ZmTGA9-1/-2/-3 and ZmTGA10. (a) The CRISPR/Cas9 mutagenesis and mutant characterization of ZmTGA9-1/-2/-3. (a1) Gene structure and mutation analysis of ZmTGA9-1/-2/-3 in WT and three triple-mutant lines (ZmTGA9-1/2/3-Cas9-1, ZmTGA9-1/2/3-Cas9-2 and ZmTGA9-1/2/3-Cas9-3) generated by the CRISPR/Cas9 technology. (a2) Identification of the mutation sites and fragments in the three triple-mutant lines based on DNA sequencing. (a3) Phenotypic analysis of tassels, anthers and pollen grains stained with 1% I₂-KI solution in WT and the three triple-mutant lines. (a4) SEM analysis of anther and pollen in WT and one tga9-1/2/3 triple mutant at maize anther developmental stage13. Notably, both the single and double mutants of ZmTGA9-1/-2/-3 display normal male fertility phenotype in maize (Figure S4). (b) The CRISPR/Cas9 mutagenesis and mutant characterization of ZmTGA10. (b1) Gene structure and mutation analysis of ZmTGA10 in WT and three knock-out lines (ZmTGA10-Cas9-1, ZmTGA10-Cas9-2 and ZmTGA10-Cas9-3) generated by the CRISPR/Cas9 technology. (b2) Identification of the mutation sites and fragments in the three knock-out lines based on DNA sequencing. (b3) Phenotypic analysis of tassels (with enlarged views of dehisced WT anthers and undehisced mutant anthers at pollen mature stage), anthers and pollen grains stained with 1% I₂-KI solution in WT and the three knock-out lines. (b4) SEM analysis of anther and pollen in WT and one tga10 mutant at maize anther developmental stage 13. Notably, the knock-out mutants of ZmTGA10 showed nearly normal pollen development, but the mutant anthers failed to dehisce.
functions in maize male fertility mainly by controlling anther dehiscence.

CRISPR/Cas9-based generation of MYB family gene mutants and their male-sterility phenotype evaluation

ZmMYB84 (Zm00001d025664) is the ortholog of OsMYB80 and AtMYB80 (Pan et al., 2020; Phan et al., 2012; Zhang et al., 2007), and ZmMYB33-1 (Zm00001d012544) and ZmMYB33-2 (Zm00001d043131) are the orthologs of OsGAMYB and AtMYB33/65 (Millar and Gubler, 2005; Tsuji et al., 2006) (Figure S9). Here, we produced myb84 and myb33 mutants via the CRISPR/Cas9 system to unveil their functions in maize male fertility.

The pCas9-ZmMYB84 was designed and constructed with two targets in the first exon of ZmMYB84 and ZmMYB33-1/-2 generated by the CRISPR/Cas9 technology. (a) Gene structure and mutation analysis of ZmMYB84 in WT and three knock-out lines (ZmMYB84-Cas9-1, ZmMYB84-Cas9-2, and ZmMYB84-Cas9-3) generated by the CRISPR/Cas9 technology. (b) Identification of the mutation sites and fragments in the single knock-out lines of ZmMYB33-1/-2 based on DNA sequencing. (b1) Gene structure and mutation analysis of ZmMYB33-1/-2 in WT and the single knock-out lines (ZmMYB33-1-Cas9 and ZmMYB33-2-Cas9) generated by the CRISPR/Cas9 technology. (b2) Identification of the mutation sites and fragments in the single knock-out lines of ZmMYB33-1/-2 based on DNA sequencing. (b3) Phenotypic analysis of tassels, anthers and pollen grains stained with 1% I2-KI solution in WT and the single and double-mutant lines of ZmMYB33-1/-2. (b4) SEM analysis of anther and pollen in WT and the double-mutant myb33-1/2 at maize anther developmental stage13.
Figure 5 The CRISPR/Cas9 mutagenesis and the derived mutant characterization of ZmPHD11, ZmLBD27 and ZmLBD10/27. (a) The CRISPR/Cas9 mutagenesis and mutant characterization of ZmPHD11. (a1) Gene structure and mutation analysis of ZmPHD11 in WT and three knock-out lines (ZmPHD11-Cas9-1, ZmPHD11-Cas9-2 and ZmPHD11-Cas9-3) generated by the CRISPR/Cas9 technology. (a2) Identification of the mutation sites and fragments in the three knock-out lines based on DNA sequencing. (a3) Phenotypic analysis of tassels, anthers and pollen grains stained with 1% I$_2$KI solution in WT and the three knock-out lines. (a4) SEM analysis of anther and pollen in WT and one phd11 mutant at maize anther developmental stage 13. (b) The CRISPR/Cas9 mutagenesis and mutant characterization of ZmLBD27. (b1) Gene structure and mutation analysis of ZmLBD27 in WT and three knock-out lines (ZmLBD27-Cas9-1, ZmLBD27-Cas9-2 and ZmLBD27-Cas9-3) generated by the CRISPR/Cas9 technology. (b2) Identification of the mutation sites and fragments in the three knock-out lines based on DNA sequencing. (b3) Phenotypic analysis of tassels, anthers and pollen grains stained with 1% I$_2$KI solution in WT and the three knock-out lines. (b4) SEM analysis of anther and pollen in WT and one lbd27 mutant at maize anther developmental stage 13. (c) The CRISPR/Cas9 mutagenesis and mutant characterization of ZmLBD10/27. (c1) Gene structures and mutation analyses of ZmLBD10 and ZmLBD27 in the single knock-out lines (ZmLBD10-Cas9 and ZmLBD27-Cas9), respectively. (c2) Phenotypic analysis of tassels, anthers and pollen grains stained with 1% I$_2$KI solution in the double-mutant lbd10/lbd27 at maize anther developmental stage 13. (c3) The proportions of normal and aborted pollen grains measured by staining with 1% I$_2$KI solution in WT and the single/double mutants of ZmLBD10 and ZmLBD27 at stage 13 (n = 4726 to 4800).
and 54-bp deletions, respectively (Figure 4a1, a2; Table S3). The three myb84 mutants in F2 generation displayed complete male sterility with normal vegetative organ growth and female fertility. Compared with WT anthers, myb84 anthers were smaller without pollen grains (Figure 4a3). SEM assay revealed that the outer surface of myb84 anther wall had a denser cuticle layer, while the inner surface of myb84 anther wall was smooth due to lack of Ubisch bodies (Figure 4a4). Therefore, ZmMYB84 is essential for maize pollen development.

Since ZmMYB33 has two paralogs in maize genome, we constructed two vectors with each one simultaneously target-editing the two paralogs (Figure 5e6) and produced one homozygous myb33 double mutant (ZmMYB33-1/2-Cas9) with 1-bp deletion in ZmMYB33-1 and 1-bp insertion in ZmMYB33-2 (Figure 4b1, b2; Table S3). In F2 generation, the segregated single-gene loss-of-function mutants of either ZmMYB33-1 or ZmMYB33-2 have exerted anthers and normal pollen grains (Figure 4b3), while the double-gene mutant exhibited complete male sterility with smaller anthers and without visible pollen grains compared with WT anther (Figure 4b3). SEM observation further revealed that the outer and inner surfaces of anther wall of myb33-1/2 mutant were smooth and glossy (Figure 4b4). Therefore, ZmMYB33-1 and ZmMYB33-2 act redundantly, and the two MYB33 paralogs as a whole are critical for maize male sterility.

CRISPR/Cas9-based production of PHD and LBD family gene mutants and their male-sterility phenotype evaluation

ZmPHD11 (Zm00001d013416) is the maize ortholog of the reported GMS genes OsTIP3 and AtMMADM1 encoding PHD finger TFs (Yang et al., 2003; Yang et al., 2019), while ZmLBD27 (Zm00001d013732) and ZmLBD10 (Zm00001d033335) are the maize orthologs of GMS gene AtLBD10 in Arabidopsis (Kim et al., 2015) (Figure S9). To identify their roles in maize male fertility, we generated phd11, lbd10, lbd27 and lbd10/27 mutants in maize.

We constructed a pcas9-ZmPHD11 vector with two targets in the first exon of ZmPHD11 (Figure S6g) and then employed three homozygous phd11 mutants with 165-bp, 173-bp and 32-bp deletions, respectively (Figure 5a1, a2; Table S3). The vegetative organ growth and female development were normal in the three mutants; however, they had smaller and shrunk anthers with very few and aborted pollen grains compared with WT anthers (Figure 5a3). SEM assay revealed defects of phd11 mutant anthers that lack the knitting cuticle and Ubisch bodies (Figure 5a4). Therefore, ZmPHD11 is required for maize male fertility via controlling the formation of anther cuticle and Ubisch bodies.

Two sgRNAs targeting the exon of ZmLBD27 were designed (Figure S6h), by which we obtained three ZmLBD27 loss-of-function homozygous mutants (with 112-bp deletion, 51-bp deletion and 2-bp insertion, respectively) edited by the CRISPR/Cas9 system (Figure 5b1, b2; Table S3). The three mutant lines showed slight male sterility with 9.53% of aborted pollen grains (Figure 5b3, c3). SEM analysis revealed that partial shrivelled pollen grains existed in mutant anthers, while the outer and inner surfaces of anther walls were largely consistent between mutant and WT anthers (Figure 5b4). These results indicated that ZmLBD27 may not be completely necessary for maize anther cuticle and tapetum development; nevertheless, it still has an impact on the reduced ratio of normal pollen grains. Additionally, three ZmLBD10 loss-of-function homozygous mutants ZmLBD10-Cas9-1, ZmLBD10-Cas9-2 and ZmLBD10-Cas9-3 (with 248-bp and 1-bp deletions, 9-bp indel and 1-bp insertion, respectively) (Figures S5a, S7a) were found to be nearly normal in another development and pollen formation (Figure S5b).

Notably, although the ZmLBD10/27 double-gene mutant (with 1-bp insertions in ZmLBD10 and ZmLBD27, respectively) (Figure 5c1) edited by the pcas9-ZmLBD10/27 vector (Figure S7a) was largely normal in tassel and anther development (Figure 5c2), lbd10/27 had a relatively larger proportion of aborted pollen grains (32.18%) than those of lbd27 (9.53%) and lbd10 (2.02%) (Figure 5c3). These results indicated that the two paralogs act redundantly, and their combination has a significant effect on viable pollen formation in maize.

The functional conservation and diversification of 14-type maize GMS TF genes among maize, rice and Arabidopsis

To further investigate the roles of maize GMS TF genes in controlling anther development and pollen formation, we analysed expression patterns of the 12 GMS TF genes (belonging to eight types) identified here (Figures 2–5) and six-type GMS TF genes reported previously (Figure S9b). Their expression patterns can be classified into four categories, including (1) six genes highly expressed at early stage (before stage 6), (2) five at middle stages (S6 to S9), (3) one at late stages (since S10), and (4) six genes with two expression peaks during stages 5 to 12 (Figure 6a). Each of the 18 TF genes showed similar expression patterns among transcriptome data from B73, Zheng58 and M6007 lines, and their expression patterns were further verified by using RT-qPCR analysis (Figure 6b).

Maize has the similar anther developmental process with Arabidopsis and rice that can be divided into 14 stages, belonging to four developmental phases: (1) archesporial cell specification (stages 1 to 2), (2) anther somatic cell division (stages 3 to 5), (3) tapetum development and cell meiosis (stages 6 to 9) and (4) mature pollen formation and anther dehiscence (stages 10 to 14; Figure 6c; Wan et al., 2019). Based on the expression patterns and functional characteristics reported here and previously, the 18 maize GMS TF genes corresponding to 14 types were assigned to different anther developmental phases (Figure 6c). For example, ZmMs32, as the ortholog of AtDYT1 and OsUDT1, is located to the second phase of anther somatic cell division (Figure 6c), consistent with their functions in early anther development (Moon et al., 2013). ZmMs9, ZmbHLHS1, ZmMYBB4 and ZmMs7 located to the third phase may function as their orthologs of AtTDF1/OsTDF1, AtAMS/OsTDR, AtMYB80/OsMYB80 and AtMs1/OsPTC1 in Arabidopsis and rice, respectively, and play comparable or relatively conserved roles in tapetum development and cell meiosis in plants (Figure 6c).

Given that the core regulatory pathways controlling tapetum and pollen development in Arabidopsis (AtDYT1-AtTDF1-AtAMS-AtMYB80-AtMs1) and rice (OsUDT1-OsTDF1-OsTDR-OsMYB80-OsPTC1) have been established and proposed, respectively (Cai et al., 2015; Lou et al., 2018), together with our results in this study, we proposed that a similar core regulatory pathway may also exist in maize anther, that is ZmMs32-ZmMs9-ZmbHLHS1-ZmMYBB4-ZmMs7 (Figure 6c, genes in yellow boxes). Besides the five maize GMS TFs in the core regulatory pathway, three-type GMS TF genes identified here (ZmTGAI10, ZmPHD11 and ZmMYB33-1/2) are the maize orthologs of AtTGAI10/OsTGA10, AtMMADM1/OsTIP3 and AtMYB33/65/OsGAMYB in Arabidopsis and rice, respectively (Figure 6c), and may play the similar and conserved roles in controlling male sterility in plants.

© 2021 The Authors. Plant Biotechnology Journal published by Society for Experimental Biology and The Association of Applied Biologists and John Wiley & Sons Ltd., 1–16
Figure 6 Expression patterns of 18 stage differentially expressed GMS TF genes and their functional conservation and diversification for male sterility in maize, rice and Arabidopsis. (a) Expression patterns of 18 GMS TF genes in anther transcriptomes of B73, Zheng58 and M6007, respectively. The 18 investigated genes were clustered into four groups including six genes with expression peaks at premeiotic stages (black), five genes with expression peaks at meiotic stages (blue), six genes each with two expression peaks (red) and one gene with an expression peak at postmeiotic stages (green). (b) Expression patterns of the 18 GMS TF genes were confirmed by the RT-qPCR assay. The colours and order of the 18 TF genes were the same as those in a. (c) 18 GMS TF genes identified in maize, rice and/or Arabidopsis were localized at specific stages of maize anther developmental process based on their expression peaks. Maize anther development process includes 14 stages and can be divided into four phases as described previously (Wan et al., 2019). The five sets of key GMS TF genes identified in maize, rice and Arabidopsis, including ZmMs32/OsUDT1/AtDYT1, ZmMs9/OsTDF1/AtTDF1, ZmMbh/LHS1/OsTDR/AtAMS, ZmMYB84/OsMYB80/AtMYB80 and ZmMs7/OsPTC1/AtMs1, were highlighted by yellow boxes.
On the other hand, maize ZmOCL4 and ZmIG1 are characterized as GMS genes (Evans, 2007; Vernoud et al., 2009), while their orthologs in rice, Arabidopsis and other plants have been not reported to control male sterility. Similarly, the maize orthologs of some rice and Arabidopsis GMS TF genes have not been reported as GMS genes so far (Figure S9b). Additionally, a single GMS TF gene in rice or Arabidopsis may correspond to more than one ortholog in maize (Figure S9b), such as AtTGA9 to ZmTGA9-1/2/3, AtLBD10 to ZmLBD10/ZmLBD27, and OsGAMYB to ZmMYB33-1/2. In these cases, we found only maize multi-gene mutants displayed complete male sterility, but not their single- or double-gene mutants (Figures 3a, 4b, 5c, 6). These results suggested that besides the functional conservation as described above, the molecular mechanisms regulating male reproduction may be diversified at some aspects among maize, rice and Arabidopsis.

Discussion

The integration of RNA-seq analysis and CRISPR/Cas9 mutagenesis is an effective strategy to discover new male-sterility genes in maize

So far, more than 100 GMS genes have been cloned in plants, and most of them are investigated in model plants Arabidopsis and rice, while less than 20 GMS genes have been reported in maize (Wan et al., 2019). Notably, most of the reported GMS genes are identified via the forward genetic approaches, such as map-based cloning, T-DNA or transposon tagging and MutMap methods (Wan and Wu, 2020), which are dependent on the existing GMS mutants. As the GMS genes often show anther-specific or anther-preferential expression patterns, it is reasonable to predict GMS genes by using anther transcriptomic analysis (i.e., RNA-seq or microarray assay) in plants (Dhaka et al., 2020; Ma et al., 2019; Wan et al., 2019; Wan et al., 2020). Furthermore, the reverse genetic strategies (e.g., CRISPR/Cas9 and RNAi) have been used to identify GMS genes in different plants, especially for those species with very limited GMS mutants or difficulty to isolate genes by map-based cloning, such as wheat TaMs26 and TaMs45 (Singh et al., 2017; Singh et al., 2018), barley HvMs1 (Fernández Gómez and Wilson, 2014) and tomato SMYB33 (Zhang et al., 2020). However, the large-scale discovery of GMS genes by using the combined strategy of RNA-seq analysis and CRISPR/Cas9 mutagenesis has less been reported in plants. Here, based on the transcriptomic analysis on three sets of another RNA-seq data under three genetic backgrounds, we predicted the shared 1100 stage-DE TF genes with the similar anther-specific or anther-preferential expression patterns, which largely excluded the effects of maize different genetic background. Interestingly, these stage-DE TF genes include the six cloned maize GMS TF genes and nearly all the orthologs of GMS TF genes reported previously in Arabidopsis, rice and other plants (Figure 1; Tables S1, S2), indicating the reliability of these predicted TF genes possibly involving in maize male fertility. To confirm functions of the predicted TF genes, the CRISPR/Cas9 technology was used to produce knock-out mutants for nine-type genes orthologous to the cloned GMS TF genes in other plants (Table S3). The results showed that null mutations of eight-type TF genes, including ZmbHLH51, ZmbHLH1122, ZmTGA9-1/2/3, ZmTGA10, ZmMYB84, ZmMYB33-1/2, ZmPHD11 and ZmLBD10/27, resulted in male sterility with different kinds of anther and pollen defects (Figures 2–5).

Collectively, it is high efficient to discover new GMS TF genes in maize by using the integrated strategy of anther RNA-seq analysis and CRISPR/Cas9 system, which has at least three advantages: (1) the effects of different genetic background on identification of specifically functional genes can be excluded by using multiple sets of RNA-seq data under different inbred lines; (2) the regulators controlling the large part of anther development process can be found based on the RNA-seq data covering 11 developmental stages; and (3) the GMS TF genes with functional redundancy can be efficiently identified through the reverse genetic strategy, as the multiple mutation sites can be simultaneously generated in different paralogous genes, of which single or double mutants may display normal phenotype (e.g. ZmTGA9-1/2-3 and ZmMYB33-1/2) (Figures 3 and 4). However, by using the natural GMS mutants or artificially mutated GMS mutants, it is relatively difficult to identify multiple GMS genes with functional redundancy by the forward genetic strategies, since their single-gene mutations display normal phenotype but their multi-gene mutations hardly take place naturally. Therefore, this study will enrich GMS gene and mutant resources for deeply understanding the regulatory network underlying male fertility and for creating GMS lines in maize.

Functional conservation and diversification of GMS TF genes required for anther and pollen development in different plants

Lots of orthologous GMS TF genes play relatively conserved roles in anther and pollen development between monocot and dicot species (Cai et al., 2015; Gomez et al., 2015). For example, ZmMs32 and its orthologs AtDYT1 and OsUDT1 play similar roles in controlling anther periclinal division and male fertility in maize, Arabidopsis and rice (Jung et al., 2005; Moon et al., 2013; Nan et al., 2017; Zhang et al., 2006). ZmMs9 and its orthologs AtTDF1 and OsTDF1 are essential for regulating early tapetum development and male fertility (Albertsen et al., 2016; Cai et al., 2015; Zhu et al., 2008). ZmMs7 and its orthologs AtM1s and OsPTC1 function as transcription activators in tapetal development and pollen exine formation (An et al., 2020; Ito et al., 2007; Li et al., 2011; Zhang et al., 2018). Based on the CRISPR/Cas9 strategy, we found that ZmbHLHS1 like its orthologs of AtAMs (Xu et al., 2010) and OsTDR (Li et al., 2006), ZmMYB84 like its orthologs of AtMYB80 (Higginson et al., 2003; Zhang et al., 2007) and OsMYB80 (Phan et al., 2012), ZmbHLH122 like its ortholog of OsEAT1 (Niu et al., 2013) and ZmPHD11 like its orthologs of AtMMMD1 (Yang et al., 2003) and OsTIP3 (Yang et al., 2019) play relatively conserved roles in controlling anther development and male fertility, since their knock-out mutants showed complete male sterility (Figures 2, 4–6). The conserved functions may be due to their critical and fundamental roles in the genetic networks controlling male reproduction among different plant species.

Nevertheless, this study also demonstrated distinctive features of mutant phenotypes and functional redundancies of some orthologous TF genes among different plant species. For example, the three paralogous genes ZmTGA9-1, ZmTGA9-2 and ZmTGA9-3 in maize are the orthologs of AtTGA9, which is redundantly required for anther development with AtTGA10 in Arabidopsis (Murmu et al., 2010). Here, both single and double mutants of three ZmTGA9 paralogs displayed male fertility, whereas triple-gene mutants of ZmTGA9-1/2/3 exhibited male sterility (Figures 3 and S4). Conversely, Arabidopsis tga9 mutation did not result in any anther defect (Murmu et al., 2010). Notably, ZmTGA10 mutants displayed male sterility with indehiscient
anthers (Figure 3), which is different from its orthologous attga10 and ostga10 mutants (Chen et al., 2018; Murmu et al., 2010). In zmtga10 mutants, anthers were filled with mature pollen grains but failed to dehisce at stage 13 (Figure 3), while the ostga10 endothecium cell wall thickening was impaired with fewer pollen grains in anther and only 10% of seed setting rate (Chen et al., 2018). Arabidopsis tga10 mutant exhibited male fertility, and only tga9tga10 double mutant showed defects in anther development and dehiscence (Murmu et al., 2010). Therefore, ZmTGA10 and its orthologs OsTGA10 and AtTGA10 may function distinctively in anther and pollen development among different plant species. Similarly, two paralogs of ZmMYB33 are orthologous to OsGAMYB (Aya et al., 2009) and AtMYB83/AtMYB65 (Millar and Gubler, 2005). Single mutants of ZmMYB33-1 or ZmMYB33-2 exhibited male fertility, whereas the double mutant displayed male sterility (Figure 4), indicating that ZmMYB33-1 and ZmMYB33-2 redundantly regulate anther and pollen development in maize, similar to the roles of AtMYB33 and AtMYB65 in Arabidopsis (Millar and Gubler, 2005), but different from that of OsGAMYB in rice (Aya et al., 2009). In addition, ZmLBD27 and ZmLBD10 are the orthologs of AtLBD10 which acts redundantly with AtLBD27 in controlling male sterility (Kim et al., 2015). Loss of ZmLBD27 function led to partial male sterility, while ZmLBD10 mutants displayed male fertility (Figure S5). The diversified functions between orthologous GMS TF genes may be associated with the differences in plant species in the phylogenetic positions (monocot or dicot plants), the natural habitats (high or low temperature) and even the pollination types (self- or cross-pollination), which need to be investigated in future.

Taken together, both the conserved and diversified functions exist in the GMS TF orthologous genes among different plant species, which will promote our understanding on molecular regulation mechanism underlying plant anther development and male fertility.

The relatively conserved core regulation pathway mediated by GMS TF genes for anther and pollen development in different plants

In Arabidopsis, the five key TFs that are essential for anther tapetum development form the core genetic regulation pathway (AtDYT1-AtTDF1-AtAMS-AtMYB80-AtMs1) (Gu et al., 2014; Li et al., 2017; Zhu et al., 2008; Zhu et al., 2011). In this pathway, AtDYT1 directly regulates AtTDF1 and indirectly regulates other three TFs. AtDYT1, AtTDF1 and AtAMS are sequentially activated to regulate early tapetum development, while AtMYB80 and AtMs1 are subsequently activated for the late tapetum development and pollen wall formation (Gu et al., 2014). The similar regulation pathway (OsUDT1-OsTDF1-OsTDR-OsMYB80-OsPTC1) has been proposed in rice (Cai et al., 2015). However, the rice genetic pathway is slightly different from that in Arabidopsis. For example, OsTIP2 acts upstream of OsTDR and OsEAT1 and directly regulates their gene expression, and the three crucial bHLH TFs may form a complex to fulfill the function of AtAMS for rice tapetal development (Fu et al., 2014; Ko et al., 2014).

Compared with Arabidopsis and rice, fewer GMS TF genes have been investigated in maize. To discover more GMS TF genes, we firstly found that ZmbHLH51, ZmbHLH122 and ZmMYB84, orthologous to AtAMS/OsTDR, OsEAT1 and AtMYB80/OsMYB80, respectively, are essential for anther development and male fertility in maize (Figures 2 and 4). Additionally, ZmbHLH122 directly binds to promoter of ZmbHLH51 and activates its expression, and ZmbHLH51 and ZmbHLH122 can form a heterodimer (Figure 2). Based on these results obtained here and previous reports, we proposed the similar and relatively conserved genetic pathway (ZmMs32-ZmMs9-ZmbHLH51/122-ZmMYB84-ZmMs7) in maize (Figure 6) to control tapetum development and pollen wall formation. Nevertheless, the regulatory pathway still needs to be proved by using more experimental data.

Methods

Plant materials, growth conditions and phenotype observation

Inbred lines B73 and M6007 corresponding to the WT lines of maize GMS mutants ms1-ab1 and ms7-6007, respectively, were obtained from the Maize Genetics Cooperation Stock Center. All materials were planted in the experimental stations of the University of Sciences and Technology Beijing. T2 transgenic plants were grown in glasshouse with a photoperiod of 16 h/8 h (day/night) at 26 °C/22 °C. The images of tassels and anthers were taken with a Canon EOS 700D digital camera and a SZX2-ILLB stereomicroscope, respectively. Pollen grains stained with 1% I2-KI solution were imaged using a BX-53 microscope.

RNA-seq analysis

Twenty fresh anthers similar in length were collected for each sample, three of them were fixed in FAA solution for confirming the developmental stage by transverse section assay, and the rest were frozen immediately in liquid nitrogen and stored at −80°C until processing. Total RNAs of 32 samples from B73 (S5 to S12, 11 stages), 30 samples from Zheng58 (S5 to S12, 11 stages) and 31 samples from M6007 (S5 to S12, 11 stages) were isolated using TRizol reagent (Invitrogen), respectively. Two to four biological replications were used for anther samples at each developmental stage in each genetic background (Figures S1-S3). RNA-seq libraries were constructed using NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, USA). mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. 150 bp paired-end reads were generated by the Illumina Hiseq 4000 platform. Clean reads were mapped to the B73 reference genome (version 4) using TopHat2.0 with default parameters (Trapnell et al., 2009). Gene expression levels were estimated by RSEM and edgeR (Liao et al., 2014; Robinson et al., 2010). Genes with more than a twofold change at the expression level and false discovery rate (FDR) < 0.05 were identified as DEGs. Maize TF gene family information was obtained from the Plant Transcription Factor Database (PlantTFDB, http://planttfdb.cbi.pku.edu.cn/). Gene co-expression analysis was performed by the K-means method in R (https://www.r-project.org/).

Plasmid construction

For mutagenesis of putative GMS TF genes, CRISPR-P 2.0 (http://crispr.hzau.edu.cn/CRISPR2/) was used to choose specific sgRNAs that targeted genes. Off-target analysis of sgRNA sequences was carried out on a website (http://www.rgenome.net/cas-offinder/). The pBUE411 vector (Xing et al., 2014) was used to construct CRISPR/Cas9 plasmid, and two 19-bp fragments in the exons of target gene were introduced into the vector. For assembly of two gRNAs, PCR fragments of each gene were amplified based on pCPC-MTT2 vector (Xing et al., 2014) with gene-specific primer pair (Table S5), and then, purified PCR fragments were digested with BsaI and cloned into the pBUE411 vector. The CRISPR/Cas9
plasmids were confirmed by DNA sequencing and then used for maize transformation.

Maize transformation

Maize hybrid genotype Hi II was used as the transformation receptor. Agrobacterium-mediated transformation of the immature embryos of Hill was performed at the Plant Transformation Platform of USTB following previous protocols (Frame et al., 2002).

DNA extraction and genotyping

Genomic DNA of the maize seedlings was extracted using cetyltrimethyl ammonium bromide (CTAB) method (Murray and Thompson, 1980). The examination of transgenic lines was performed by PCR with specific primer pair of Bar gene (Table S5). Specific primers spanning the target sites (Table S5) were designed to determine the mutations in the target gene of T_0 and F_1 generation plants. The PCR amplicons flanking the target sites were performed with high-fidelity DNA polymerase, and purified PCR products were cloned into the pEASY-T1 vector (TransGen Biotech, China) and followed by Sanger sequencing. For the determination of transgene-free plants in F_1 generation, Bar receptor. Maize hybrid genotype Hi II was used as the transformation construct. The plasmids were confirmed by DNA sequencing and then used for maize transformation.

BiFC assay

For protein–protein interaction studies, BiFC assay was performed as previously described (Walter et al., 2004). Full-length cDNA of ZmbHLH51 was fused with the N-terminal of YFP (nYFP: 1-155 aa) and cloned into the vector pUC19-35S-FLAG-RBS (Li et al., 2005) in front of FLAG coding sequence by homologous recombination to generate 35S-ZmbHLH51-nYFP.Flag-plasmid. Full-length cDNA of ZmbHLH122 was fused with the C-terminal half of YFP (cYFP: 156-239 aa) and 3 × Myc coding sequence. The resulted ZmbHLH122-cYFP-Myc fragment was cloned into the vector pUC19-35S-FLAG-RBS by replacing FLAG coding sequence to generate 35S-ZmbHLH122-cYFP-Myc plasmid. The BiFC constructs were cotransformed into maize protoplasts. The protoplasts were incubated at 28 °C for 12–16 hr in the dark. YFP fluorescence was detected using a confocal laser-scanning microscope (Leica TCS SP8) with 514 nm/525–565 nm excitation/emission wavelengths. The primers used are listed in Table S5.

Co-IP assay

For Co-IP assays using protoplast transient expression system, we created a 35S-nYFP-Flag vector by homologous recombination on the basis of 35S-ZmbHLH51-nYFP-Flag vector. The vectors 35S-nYFP-Flag, 35S-ZmbHLH51-nYFP-Flag and 35S-ZmbHLH122-cYFP-Myc in various recombinations were cotransformed into maize leaf protoplasts. The protoplasts were incubated at 28 °C for 16 hr in the dark. The extraction of total protein was performed as described previously (Ma et al., 2018). The total protein immunoprecipitated with anti-c-Myb affinity gel (E6654, Sigma-Aldrich) according to the manufacturer’s instructions. The eluted immunoprecipitates were immunoblotted with anti-c-Myb (9E10, sc-40, Santa Cruz, 1:1000) and anti-FLAG (FLA-1, M185-3, MBL, 1:10 000) antibodies. Secondary goat anti-mouse-IgG-HRP antibody was used at 1:10 000 dilutions in PBS. The signals were detected using a SuperSignal West Pico kit (34080, Thermo Scientific).

Acknowledgements

This research was funded by the National Key Research and Development Program of China (2017YFD0102001, 2018YFD0100806, 2017YFD0101201), the National Natural Science Foundation of China (31771875, 31971958, 31871702), the Fundamental Research Funds for the Central Universities of China (06500136, FRF-TP-18-049A1) and the Beijing Science & Technology Plan Program (Z201100006820114).

Conflicts of interest

The authors declare no conflicts of interest.

Author contributions

X.W. designed research; Z.L. and X.A. performed RNA-seq analysis; Y.J., X.A., T.Y., T.Z., X.L., K.X., S.L., W.H. and F.L.
performed plasmid construction, genetic transformation, CRISPR/Cas9 assay and molecular markers development; Y.J., X.A. and T.Z. performed phenotypic and cytological observations; Q.H. and L.Z., analysed gene expression regulation and protein interaction; Y.J., X.A., Z.L., S.W. and X.W. wrote and revised the paper; X.W., X.A. and J.L. administered project; X.W. supervised project.

References

Albertsen, M., Fox, T., Leonard, A., Li, B., Loveland, B. and Trimnell, M. (2016) Cloning and use of the ms9 gene from maize. US patent US20160024520A1.

An, X., Dong, Z., Tian, Y., Xie, K., Wu, S., Zhu, T., Zhang, D. et al. (2019) ZmMs30 encoding a novel GDSL lipoase is essential for male fertility and valuable for hybrid breeding in maize. Mol. Plant. 12, 343–359.

An, X., Ma, B., Duan, M., Dong, Z., Liu, R., Yuan, D., Hou, Q. et al. (2020) Molecular regulation of ZmMs7 required for maize male fertility and development of a dominant male-sterility system in multiple species. Proc. Natl. Acad. Sci. USA, 117, 23499–23509.

Aya, K., Ueguchi-Tanaka, M., Kondo, M., Hamada, K., Yano, K., Nishimura, M. and Matsuoka, M. (2009) Gibberellin modulates another development in rice via the transcriptional regulation of GAMYB. Plant Cell. 21, 1453–1472.

Cai, C.F., Zhu, J., Lou, Y., Guo, Z.L., Xiong, S.X., Wang, K. and Yang, Z.N. (2015) The functional analysis of OsTDF1 reveals a conserved genetic pathway for tapetal development between rice and Arabidopsis. Science Bull. 60, 1073–1082.

Carretero-Paulet, L., Galstyan, A., Roig-Villanova, I., Martinez-Garcia, J.F., Bellasio-Castro, J.R. and Robertson, D.L. (2010) Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in Arabidopsis, poplar, rice, moss, and algae. Plant Physiol. 153, 1398–1412.

Cermak, T., Baltes, N.J., Cegan, R., Zhang, Y. and Voytas, D.F. (2015) High-frequency, precise modification of the tomato genome. Genome Biol. 16, 232.

Char, S.N., Neelakandan, A.K., Nahapun, H., Frame, B., Main, M., Spalding, M.H., Becraft, P.W. et al. (2017) An Agrobacterium-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. Plant Biotechnol. J. 15, 257–268.

Chen, Z.S., Liu, X.F., Wang, D.H., Chen, R., Zhang, X.L., Xu, Z.H. and Bai, S.N. (2015) Flagellin induces innate immunity in nonhost interactions that is suppressed by Pseudomonas syringae effectors. Proc. Natl. Acad. Sci. USA, 112, 12990–12995.

Chi, L., An, X., Zhu, T., Yan, T., Wu, S., Tian, Y. and Li, J. et al. (2019) Discovering and constructing cellRNA-miRNA-target gene regulatory networks during Atther development in maize. Int. J. Mol. Sci. 20, 3480.

Li, Z., Liu, Z.B., Xing, A., Moon, B.P., Koehler, J.P., Huang, L., Ward, R.T. et al. (2015) Cas9-guide RNA directed genome editing in soybean. Plant Physiol. 169, 960–970.

Liao, Y., Smyth, G.K. and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930.

Lin, J.I., Yu, C.P., Chang, Y.M., Chen, S.C. and Li, W.H. (2014) Male and millet transcription factors annotated using comparative genomic and transcriptomic data. BMC Genom. 15, 818.

Lou, Y., Zhou, H.S., Han, Y., Zeng, Q.Y., Zhu, J. and Yang, Z.N. (2018) Positive regulation of ZmMs7 required for maize male fertility and tapetum development. Plant Cell, 1–16. published by Society for Experimental Biology and The Association of Applied Biologists and John Wiley & Sons Ltd., 1–16.

Ma, H. (2005) Molecular genetic analyses of microsporogenesis and microgametogenesis in flowering plants. Annu. Rev. Plant Biol. 56, 393–434.

Ma, J., Skibbe, D.S., Fernandes, J. and Walbot, V. (2008) Male reproductive development: gene expression profiling of maize anther and pollen ontogeny. Genome Biol. 9, R181.

Ma, K., Han, J., Hao, Y., Yang, Z., Chen, J., Liu, Y.G., Zhu, Q. et al. (2019) An effective strategy to establish a male sterility mutant mini-library by CRISPR/Cas9-mediated knockout of anther-specific genes in rice. J. Genet. Genom. 46, 273–275.

McCormick, S. (1993) Male gametophyte development. Plant Cell 5, 1265–1275.

Millar, A.A. and Gubler, F. (2005) The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-mediated genes that redundantly facilitate anther development. Plant Cell, 17, 705–721.

Moon, J., Skibbe, D., Timofejeva, L., Wang, C.J., Kellner, T., Kremling, K., Walbot, V. et al. (2013) Regulation of cell divisions and differentiation by...
MALE STERILITY32 is required for Anther development in maize. Plant J. 76, 592–602.

Murmu, J., Bush, M.J., Delong, C., Li, S., Xu, M., Khan, M., Malcolmson, C. et al. (2010) Arabidopsis basic leucine-zipper transcription factors TGA9 and TGA10 interact with floral glutaredoxins ROXY1 and ROXY2 and are redundantly required for Anther development. Plant Physiol. 154, 1492–1504.

Murray, M.G. and Thompson, W.F. (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8, 4321–4325.

Nadakuditi, S.S. and Enciso-Rodriguez, F. (2021) Advances in genome editing with CRISPR systems and transformation technologies for plant DNA manipulation. Front. Plant Sci. 11, 637119.

Nan, G.L., Zhai, J., Arkit, S., Morrow, D., Fernandes, J., Mai, L., Nguyen, N. 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.

Nelms, B. and Walbot, V. (2019) Defining the developmental program leading to meiosis in maize. Science 364, 52–56.

Niu, N., Liang, W., Yang, X., Jin, W., Wilson, Z.A., Hu, J. and Zhang, D. (2013) EAT1 promotes tapetal cell death by regulating asparatic proteases during male reproductive development in rice. Nat. Commun. 4, 1445.

Pan, X., Yan, W., Chang, Z., Xu, Y., Luo, M., Xu, C., Chen, Z. et al. (2020) OsMYB80 regulates anther development and pollen fertility by targeting multiple biological pathways. Plant Cell Physiol. 61, 988–1004.

Phan, H.A., Li, S.F. and Parish, R.W. (2012) MYB80, a regulator of tapetal and pollen development, is functionally conserved in crops. Plant Mol. Biol. 78, 171–183.

Qi, X., Zhang, C., Zhu, J., Liu, C., Huang, C., Li, X. and Xie, C. (2020) Genome editing enables next-generation hybrid seed production technology. Mol. Plant. 13, 1262–1269.

Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140.

Singh, M., Kumar, M., Thilges, K., Cho, M.-J. and Cigan, A.M. (2017) MS26/CYP704B2 is required for anther and pollen development in wheat (Triticum aestivum L.) and combining mutations in all three homeologs causes male sterility. PLoS One 12, e0177632.

Singh, M., Kumar, M., Albertsen, M.C., Young, J.K. and Cigan, A.M. (2018) Concurrent modifications in the three homeologs of M445 with CRISPR/Cas9 lead to rapid generation of male sterile dwarf wheat (Triticum aestivum L.), Plant Mol. Biol. 97, 371–383.

Trapnell, C., Pachter, L. and Salzberg, S.L. (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105–1111.

Tsai, H., Aya, K., Ueguchi-Tanaka, M., Shimada, Y., Nakazono, M., Watanabe, R., Nishizawa, N.K. et al. (2006) GAMYB controls different sets of genes and is differentially regulated by microRNA in anther cells and anthers. Plant J. 47, 427–444.

Vernoud, V., Laigle, G., Rozier, F., Meeley, R.B., Perez, P. and Rogowsky, P.M. (2009) The HD-ZIP IV transcription factor OCL4 is necessary for trichome patterning and anther development in maize. Plant J. 59, 883–894.

Walter, M., Chaban, C., Schütze, K., Batistic, O., Weckermann, K., Nake, C., Blazevic, D. et al. (2004) Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. Plant J. 40, 428–438.

Wan, X. and Li, Z. (2019) Plant comparative transcriptomics reveals functional mechanisms and gene regulatory networks involved in anther development and male sterility. In Transcriptome analysis (Blumenberg, M., ed), pp. 39–60. London, UK: IntechOpen.

Wan, X., Wu, S., Li, Z., Dong, Z., An, X., Ma, B., Tian, Y. et al. (2019) Maize genic male-sterility genes and their applications in hybrid breeding: progress and perspectives. Mol. Plant. 12, 321–342.

Wan, X. and Wu, S. (2020) Molecular cloning of genic male-sterility genes and their applications for plant heterosis via biotechnology-based male-sterility systems. In Synthetic Biology - New Interdisciplinary Science (Nagao, M.L., Boldura, O.-M., Baltà, C. and Enany, S., eds), pp. 75–102. London, UK: IntechOpen.

Wan, X., Wu, S., Li, Z., An, X. and Tian, Y. (2020) Lipid metabolism: critical roles in male fertility and other aspects of reproductive development in plants. Mol. Plant. 13, 955–983.

Wan, X., Wu, S. and Li, X. (2021) Breeding with dominant genic male-sterility genes to boost crop grain yield in post-heterosis utilization era. Mol. Plant. 14, https://doi.org/10.1016/j.molp.2021.02.004.

Xing, H.L., Dong, L., Wang, Z.P., Zhang, H.Y., Han, C.Y., Liu, B., Wang, X.C. et al. (2014) A CRISPR/Cas9 toolkit for multiple genome editing in plants. BMC Plant Biol. 14, 327.

Xu, J., Yang, C., Yuan, Z., Zhang, D., Gondwe, M.Y., Ding, Z., Liang, W. et al. (2010) The ABORTED MICROSPIORES regulatory network is required for postmeiotic male reproductive development in Arabidopsis thaliana. Plant Cell. 22, 91–107.

Yang, X., Makaroff, C.A. and Ma, H. (2003) The Arabidopsis MEI10CYTE DEATH1 gene encodes a PHD-finger protein that is required for male meiosis. Plant Cell. 15, 1281–1295.

Yang, Z., Sun, L., Zhang, Y., Zhang, Y., Yu, P., Liu, L., Abbas, A. et al. (2019) TDR INTERACTING PROTEIN 3 encoding a PHD-finger transcription factor regulates Ubisch bodies and pollen wall formation in rice. Plant J. 99, 844–861.

Zhang, D., Wu, S., An, X., Xie, K., Dong, Z., Zhou, Y., Xu, L. et al. (2018) Construction of a multicontrol sterility system for a maize male-sterile line and hybrid seed production based on the ZmM7 gene encoding a PHD-finger transcription factor. Plant Biotechnol. J. 16, 459–471.

Zhang, W., Sun, Y., Timofeeva, L., Chen, C., Grossniklaus, U. and Ma, H. (2006) Regulation of Arabidopsis tapetum development and function by DYSFUNCTIONAL TAPETUM1 (DYT1) encoding a putative bHLH transcription factor. Development 133, 3085–3095.

Zhang, Y., Yang, B., Yang, T., Zhang, J., Liu, B., Zhan, X. and Liang, Y. (2020) The GAMYB-like gene SIMYB3 mediates flowering and pollen development in tomato. Hortic Res. 7, 133.

Zhang, Z., Zhu, J., Gao, J.F., Wang, C., Li, H., Li, H., Zhang, H.Q. et al. (2007) Transcription factor AtMYB103 is required for anther development by regulating tapetum development, callose dissolution and exine formation in Arabidopsis. Plant J. 52, 528–538.

Zhu, J., Chen, H., Li, H., Gao, J.F., Jiang, H., Wang, C., Guan, Y.F. et al. (2008) Defective in Tapetal development and function 1 is essential for anther development and tapetal function for microspore maturation in Arabidopsis. Plant J. 55, 266–277.

Zhu, J., Lou, Y., Xu, X. and Yang, Z.N. (2011) A genetic pathway for tapetum development and function in Arabidopsis. J. Integr. Plant Biol. 53, 882–900.

Zhu, T., Li, Z., An, X., Long, Y., Xue, X., Xie, K., Ma, B. et al. (2020) Normal structure and function of endothecium chloroplasts maintained by ZmMs33-mediated lipid biosynthesis in tapetal cells are critical for anther development in maize. Mol. Plant. 13, 1624–1643.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 RNA-Seq data analysis of B73 anthers during developmental stages 5 to 12.

Figure S2 RNA-Seq data analysis of Zheng58 anthers during developmental stages 5 to 12.

Figure S3 RNA-Seq data analysis of M6007 anthers during developmental stages 5 to 12.

Figure S4 The CRISPR/Cas9 mutagenesis and the derived single and double mutant characterization of three ZmTGA9 homologous genes (ZmTGA9-1, –2 and –3).

Figure S5 The CRISPR/Cas9 mutagenesis and the derived mutant characterization of ZmLB10.

Figure S6 The physical maps and target-site information of CRISPR/Cas9 constructs for editing eight types of GMS TF genes controlling maize male sterility.
**Figure S7** The physical maps and target-site information of CRISPR/Cas9 constructs for editing two-type TF genes and mutant phenotype of ZmMYB60 and ZmMYB44 gene.

**Figure S8** The molecular markers co-segregating with male-sterility mutations.

**Figure S9** The previously cloned and newly identified GMS TF genes in maize and their conversed and diversified relationships among maize, rice and Arabidopsis.

**Table S1** Previously reported GMS TF genes in Arabidopsis and rice.

**Table S2** 1100 stage-DE TF genes shared by B73, Zheng58 and M6007 inbred lines.

**Table S3** CRISPR/Cas9-induced mutations of maize TF genes investigated in this study.

**Table S4** Functional molecular markers for detecting CRISPR/Cas9 editing mutations (excluding marker sequences listed in Figure S8).

**Table S5** Primers used in this study.