Theor Appl Genet (2015) 128:2341–2350
DOI 10.1007/s00122-015-2589-7

ORIGINAL ARTICLE

Multiple loci not only *Rf3* involved in the restoration ability of pollen fertility, anther exsertion and pollen shedding to S type cytoplasmic male sterile in maize

Yang Feng[1] · Qi Zheng[1] · Hui Song[1] · Yi Wang[2] · Hui Wang[2] · Lijing Jiang[2] · Jianbing Yan[1] · Yonglian Zheng[1] · Bing Yue[1]

Received: 23 March 2015 / Accepted: 18 July 2015 / Published online: 29 July 2015

Abstract

Key message **Thirty loci for fertility restoration of pollen fertility, anther exsertion and pollen shedding to maize CMS-S were identified by GWAS.**

Abstract S type cytoplasmic male sterile (CMS-S) is the main type of CMS in maize; poor understanding of the genetic architecture of fertility restoration to CMS-S is one of the reasons to impede its utility in hybrid breeding. In this study, genome-wide identification of genetic loci for fertility restoration ability to CMS-S was firstly conducted with a set of testcrossing association mapping panel in three environments. A total of 19, 3 and 8 significant loci ($P < 1.8 \times 10^{-6}$, $\alpha = 1$) for pollen fertility, anther exsertion and pollen shedding were identified, respectively, and individual locus explained up to 28.26 % of phenotypic variation. Of them, only *Rf3*, the main restorer-fertility gene of CMS-S, was identified for the three traits simultaneously. In addition, 83 candidate genes within the 100 kb extension regions of these loci were predicted. These results revealed that besides *Rf3* multiple genetic loci and mechanisms are involved in the fertility restoration ability to CMS-S. Results in this study would provide important information for understanding the genetic architecture of fertility restoration to CMS-S in maize.

Communicated by T. Lubberstedt.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-015-2589-7) contains supplementary material, which is available to authorized users.

✉ Bing Yue
yuebing@mail.hzau.edu.cn

[1] National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China

[2] Industrial Crops Research Institution, Heilongjiang Academy of Land Reclamation of Sciences, Harbin, China

Introduction

Cytoplasmic male sterility (CMS) caused by incompatibility between cytoplasmic and nuclear gene products is a common phenomenon in plants. The fertility of CMS can be restored by nuclear fertility restorers (*Rf*). The understanding of *Rf* and CMS is of great commercial importance for crop hybrid breeding and the usage of CMS can avoid extra efforts for artificial emasculation. Moreover, it is an attractive scientific topic for studying the interactions between proteins encoded by cytoplasmic and nuclear genes. Since *Rf2* for T type CMS in maize was firstly cloned (Cui et al. 1996), 11 more *Rf* genes have been cloned in petunia, radish, sorghum, Kosena rapeseed and rice (Bentolila et al. 2002; Brown et al. 2003; Kazama and Toriyama 2003; Koizuka et al. 2003; Komori et al. 2004; Wang et al. 2006; Fujii and Toriyama 2009; Itabashi et al. 2011; Hu et al. 2012; Tang et al. 2014). Except for the *Rf2* in maize (encoding a mitochondrial aldehyde dehydrogenase) (Cui et al. 1996), the *Rf2* in rice [encoding a mitochondrial glycine-rich protein (GRP)] (Itabashi et al. 2011) and the *Rf1* in sugar beet (Matsuhira et al. 2012), the rest of *Rf* genes cloned all encode proteins containing pentatricopeptide repeat (PPR) motifs. PPR proteins are characterized by the existence of tandem repeats of a highly degenerate 35-amino-acid motif with function of reducing the accumulation of CMS-associated RNAs. They can be divided into P and PLS classes according to the different motif structures in their C terminals (Schmitz-Linneweber and Small 2008). Most of the *Rf*-related PPRs belong to the P type. Recently, Hu et al. (2012) reported that a PPR protein, RF5,

needed to cooperate with other factors to restore the fertility of rice Hong-Lian CMS lines.

For maize, three major types of CMS, T (*T*exas), C (*C*harrua) and S (U*S*DA) were identified. The S type CMS (CMS-S) is the main type with the most members. Fertility restoration to CMS-S was considered mainly governed by *Rf3*, which has been mapped to the long arm of chromosome 2 more than 30 years ago (Laughnan and Gabay 1978). Since then, however, no great progress on fine-mapping of *Rf3* is made for the complex sequences in the chromosomal region harboring *Rf3* (Kamps and Chase 1997; Tie et al. 2006; Zhang et al. 2006). Several years later, Xu et al. (2009) fine-mapped *Rf3*, and identified three candidate genes with PPR motifs, *PPR-814a*, *PPR-814b* and *PPR-814c*. Besides the main nuclear restorer, *Rf3*, other loci for fertility restoration to CMS-S, such as *RfIII* and *RfIV* (Laughnan and Gabay 1973), and *Rf9* (Gabay-Laughnan et al. 2009), have been reported. Tie et al. (2006) also identified two minor QTLs for fertility restoration to CMS-S on chromosomes 6 and 9, which explained 9.2 and 10.4 % of phenotypic variation, respectively. Polygenic pattern of fertility restoration was also observed for the C type CMS in maize (Kohls et al. 2011). To get a better understanding of the genetic architecture of fertility restoration to CMS-S in maize, it is important to unveil its genetic structure and to clone the *Rf* genes as far as possible, especially *Rf3*.

With regard to phenotyping to measure, pollen fertility as well as anther exsertion and pollen shedding/anther dehiscence are options to better dissect the overall restoration ability. As in maize, female and male flowers are separated from each other; the degree of anther exsertion and pollen shedding is also affecting the pollination of maize hybrids. In maize, Kohls et al. (2011) identified 7 QTLs for partial restoration of anther quality (rates of stunted, dehiscent and shedding of the anthers) of CMS-C. However, the genetic interaction of the two traits to the restoration ability of CMS-S is not clear in maize. Anther exsertion and pollen shedding are genetically complex and molecular mechanisms have been studied using mutants. In Arabidopsis, Yang et al. (2007) reported that the anther endothelium in *ms35*, resulted from a mutation of the gene *MYB26*, could not thicken normally. Recently, a mutation of the gene encoding a dioxygenase for auxin oxidation caused abnormal anther exsertion, dehiscence and pollen fertility (Zhao et al. 2013). So far, research on the genetic basis of anther exsertion and pollen shedding in maize has not been reported.

Gnome-wide association study (GWAS) has become a promising method in genetic research for many plant species with the development of high-throughput genotyping techniques. High diversity and rapid linkage disequilibrium decay in maize make it an excellent model crop for GWAS (Yan et al. 2011). Besides genetic architecture dissection, GWAS would also provide information of candidate genes for fine-mapping and gene cloning (Li et al. 2013). In this study, GWAS was firstly employed to dissect the genetic architecture of fertility restoration ability (pollen fertility, anther exsertion and pollen shedding) to CMS-S using a testcrossing association mapping panel. The aims of this study are (1) to reveal and compare the genetic basis of the restoration ability of pollen fertility, anther exsertion and pollen shedding to maize CMS-S on genome-wide level and (2) to deduce possible mechanisms of fertility restoration to CMS-S and prospect its usage in hybrid breeding. Results in this study would provide theoretical and practical starting points to improve the use of CMS-S in maize hybrid breeding.

Materials and methods

Plant materials

Plants of a CMS-S line (S-Mo17^{rf3rf3}) were pollinated with an association panel with 513 diverse maize inbreds in 2012 in the experimental station of Huazhong Agricultural University, Wuhan, China. Totally, 360 testcrossing hybrids with sufficient seeds were used for GWAS in this study. Among the association panel, 112, 116, 258 and 27 lines belong to stiff stalk, non-stiff stalk, tropical-subtropical and admixed groups, respectively. The lines were genotyped with 556,809 SNPs (Yang et al. 2014a). The CMS line with the background of the inbred Mo17, S-Mo17^{rf3rf3}, its maintainer, N-Mo17^{rf3rf3} and restorer, S-Mo17^{Rf3Rf3} were developed by the maize research group of Huazhong Agricultural University. The testcrossing association mapping panel, S-Mo17^{rf3rf3} and N-Mo17^{rf3rf3} (as a control) were sowed in three environments, Hainan (HN, 18°N, 109°E) in the winter of 2012, Wuhan (WH, 30°N, 114°E) in the spring of 2013, and Haerbin (HB, 46°N, 127°E) in the summer of 2013. Among the three environments, the temperature was lowest (average maximum daily temperature was 27.3 centigrade) in HN, and highest (average maximum daily temperature was 30.5 centigrade) in WH during the month around flowering time. The field experiments were conducted as a random complete block design with two replicates. Eight to ten plants were kept in each one-row plot. Field management was applied following local agronomic practice.

Phenotyping

Pollen fertility (or fertile pollen rate) of 3–5 plants in each plot for the GWAS panel was investigated. Spikelets close to the newly flowered spikelets on the middle of the tassel (for the individuals which anthers unable exsertion,

Table 1 Performance of pollen fertility, anther exsertion and pollen shedding in the testcrossing association mapping panel under different environments

Traits	Environments	Mean	SD	Range	Skew	Kurt
Pollen fertility (%)	HN	20.82	20.90	0.00–80.46	0.87	−0.38
	HB	20.22	22.11	0.00–77.08	0.91	−0.52
	WH	18.17	21.73	0.00–78.91	1.03	−0.36
Anther exsertion	HB	3.42	1.56	1.00–5.00	−0.30	−1.57
	WH	3.51	1.65	1.00–5.00	−0.37	−1.63
Pollen shedding	HB	2.36	0.09	1.00–5.00	0.74	−1.10
	WH	2.30	1.78	1.00–5.00	0.76	−1.30

HN Hainan, *HB* Haerbin, *WH* Wuhan, *SD* standard deviation, *Skew* skewness, *Kurt* kurtosis

spikelets were harvested from the middle part of the tassels just as the ears were silking) were sampled and stored in a fix solution [75 % (v/v) ethanol in acetic acid] in plastic tubes. Pollens were stained with an iodine–potassium iodine solution [containing 0.5 % (w/v) iodine and 1 % (w/v) iodine potassium] and examined under a light microscope (Olympus IX71, Japan) at 100× magnification. Photos of the pollens in three representative microscope fields (about 200 pollens in each field) of each sample were taken, then the numbers of fertile pollens and sterile pollens on each photo were counted, respectively. Rounded grains in dark blue were scored as fertile, and shrunken grains without staining were scored as aborted. Pollen fertility (%) was calculated as the percentage of fertile pollens, and means of the two replicates in each environment were used for GWAS.

Anther exsertion degrees with scales ranged from 1 to 5 were scored in WH and HB for the entire GWAS panel. The five scales are described as follows:

1. no anther exserted at all;
2. about one-third of the anthers exserted;
3. about half of the anthers exserted;
4. about two-third of the anthers exserted;
5. anthers exserted absolutely.

Pollen shedding (or anther dehiscence) degrees for the entire GWAS panel in WH and HB were also observed in field by shaking the tassels when normal maize plants begin shedding pollens in morning. A scale of 1 (no pollen shedding), 2 (about one-third pollens shedding), 3 (about half pollens shedding), 4 (about two-third pollens shedding) and 5 (anthers dehisced absolutely and no pollen left in the pollen sacs) was used to evaluate the pollen shedding ability.

Statistical analysis and association mapping

ANOVA and correlation analyses were performed by the procedures of generalized linear modeling (GLM) and correlation (CORR) using the SAS program (Release 9.1.3; SAS Institute, Cary, NC). Heritability (h^2) of the traits was estimated as follows: $h^2 = \delta_g^2/(\delta_g^2 + \delta_{ge}^2/n + \delta_e^2/nr)$, where δ_g^2 represents genotypic variance, δ_{ge}^2 is genotype × environment variance, δ_e^2 is error variance, n is number of environments, and r is the number of replications.

Association analysis for the traits was conducted by MLM with software package TASSEL V.3.0 (Bradbury et al. 2007). The Bonferroni-corrected threshold at $\alpha = 1$ [$n = 556{,}809$, $P_{1/n} < 1.8 \times 10^{-6}$, $-\log_{10}(P) = 5.7$] was used to declare a significant SNP-trait association for all the traits. The positions of the SNPs at the identified loci for the corresponding traits were based on the public maize genome data set B73 RefGen_v2.

Results

Trait performance

Pollen fertility in the GWAS panel investigated in the three environments is given in Table 1. The average pollen fertility was highest in HN, and lowest in WH. Percentage of fertile pollens varied largely under all environments (ranged from 0 to 80.46 %), and all fit to normal distributions. This indicated that fertility restoration to CMS-S is governed by multiple loci in this association panel.

Anther exsertion and pollen shedding ability observed in WH and HB are also presented in Table 1. Means of the two traits did not change much in the two environments, but they varied largely among the genotypes in both environments, and fit or closely fit to normal distributions. This indicated that multiple genes are also involved in the restoration ability of anther exsertion and pollen shedding.

Analysis of variance (ANOVA) showed that significant difference exists among genotypes ($p < 0.01$) for all the traits, while no significant difference was detected between blocks (replicates, $p > 0.57$) for the three traits. No significant difference was detected between environments ($p > 0.14$) for anther exertion and pollen shedding; however, significant difference was detected across the environments ($p < 0.01$) for pollen fertility. Pollen fertility

Table 2 Coefficients of pairwise correlations of the three traits in different environments

	HB AE	WH AE	HB PS	WH PS	HB PF	WH PF
WH AE	0.78	1				
HB PS	0.62	0.54	1			
WH PS	0.50	0.53	0.70	1		
HB PF	0.70	0.67	0.83	0.75	1	
WH PF	0.63	0.65	0.76	0.80	0.92	1
HN PF	0.66	0.67	0.75	0.78	0.91	0.93

The coefficients are all significant at the level of $p < 0.01$

AE anther exsertion, *PS* pollen shedding, *PF* pollen fertility

in WH is significantly lower than that in HB and HN at the level of $p < 0.01$, suggesting that high temperature in the summer of WH might be unsuitable for pollen development. No significant genotype by environment interaction was detected for the three traits. Heritability of the traits was high, and h^2 of pollen fertility (0.92) was higher than that of anther exsertion (0.88) and pollen shedding (0.80).

Correlation between the traits

Positive, strong correlations ($r > 0.91$) were detected for pollen fertility investigated under the three environments, and correlations of anther exsertion and pollen shedding under different environments were also significant ($r = 0.70$ for pollen shedding, $r = 0.78$ for anther exsertion) (Table 2). The three traits were highly correlated with each other (r ranged from 0.50 to 0.83). Of them, the correlation coefficients between pollen fertility and pollen shedding were higher ($0.75 < r < 0.83$), and those between anther exsertion and pollen shedding were relatively lower ($0.50 < r < 0.62$). This implies that the genetic basis of pollen fertility and pollen shedding is more similar in comparison with that between pollen fertility and anther exsertion. Although the correlations among the traits were significant, inconsistent performance of the three traits was observed, especially between the traits of pollen fertility and anther exsertion.

Genome-wide association mapping for pollen fertility

Genotypes with 556,809 high quality SNPs with minor allele frequency greater than 0.05 of the association mapping panel have been obtained in a previous study (Yang et al. 2014a). Nineteen SNP-trait associations were identified for pollen fertility at the threshold $-\log_{10}(P) > 5.7$ ($P < 1.8 \times 10^{-6}$, $\alpha = 1$) (Fig. 1). These SNPs were distributed on chromosomes 2, 5, 6 and 10. The significant SNP (chr2: 227,045,358) identified across all the environments alone explained 25.23–28.26 % of phenotypic variation (Table 3). This locus was located in bin 2.09 where *Rf3* was mapped (Kamps and Chase 1997; Tie et al. 2006; Zhang et al. 2006; Xu et al. 2009), thus it was referred to as *Rf3* thereafter.

Since *Rf3* was the major locus for pollen fertility in this GWAS panel, conditional association analysis for pollen fertility was then performed using the *Rf3* locus as a covariate to detect fertility restoration loci independent of *Rf3*. Seven loci were revealed by both GWAS and conditional analysis, and 11 new loci were detected by conditional analysis. These 18 loci were considered independent of *Rf3*. Information of the 18 loci and the *Rf3* locus is given in Table 3. Of them, five and nine loci were identified simultaneously under three and two environments, respectively. Except for *Rf3*, phenotypic variation explained by each locus was less than 11.49 %.

Combining effects of elite alleles of the 19 loci for pollen fertility

In the present association mapping panel, a total of 27 hybrids had the favorable allele of *Rf3*. Pollen fertility in these hybrids ranged from 31.76 to 69.18 % with an average of 54.96 % (Fig. 2a). On average, pollen fertility in the hybrids with no favorable allele or at least one favorable allele (except *Rf3*) was 8.48 and 27.36 %, respectively. Combining effects were observed between the favorable alleles of *Rf3* and the other significant loci. Among the 27 hybrids, 10, 9 and 8 hybrids with only *Rf3*, *Rf3* plus one and *Rf3* plus 2–4 favorable alleles of other loci had an average pollen fertility of 49.24, 57.34 and 62.70 %, respectively.

Combining effects of the favorable alleles of the significant loci except *Rf3* were also analyzed (Fig. 2b). Among the hybrids without the favorable allele of *Rf3*, 65 and 22 hybrids had one and two favorable alleles, respectively. Only one to three hybrids with favorable alleles numbered from 4 to 17, thus we divided the hybrids with 4–10 and 11–17 favorable alleles into two groups. Hybrids with

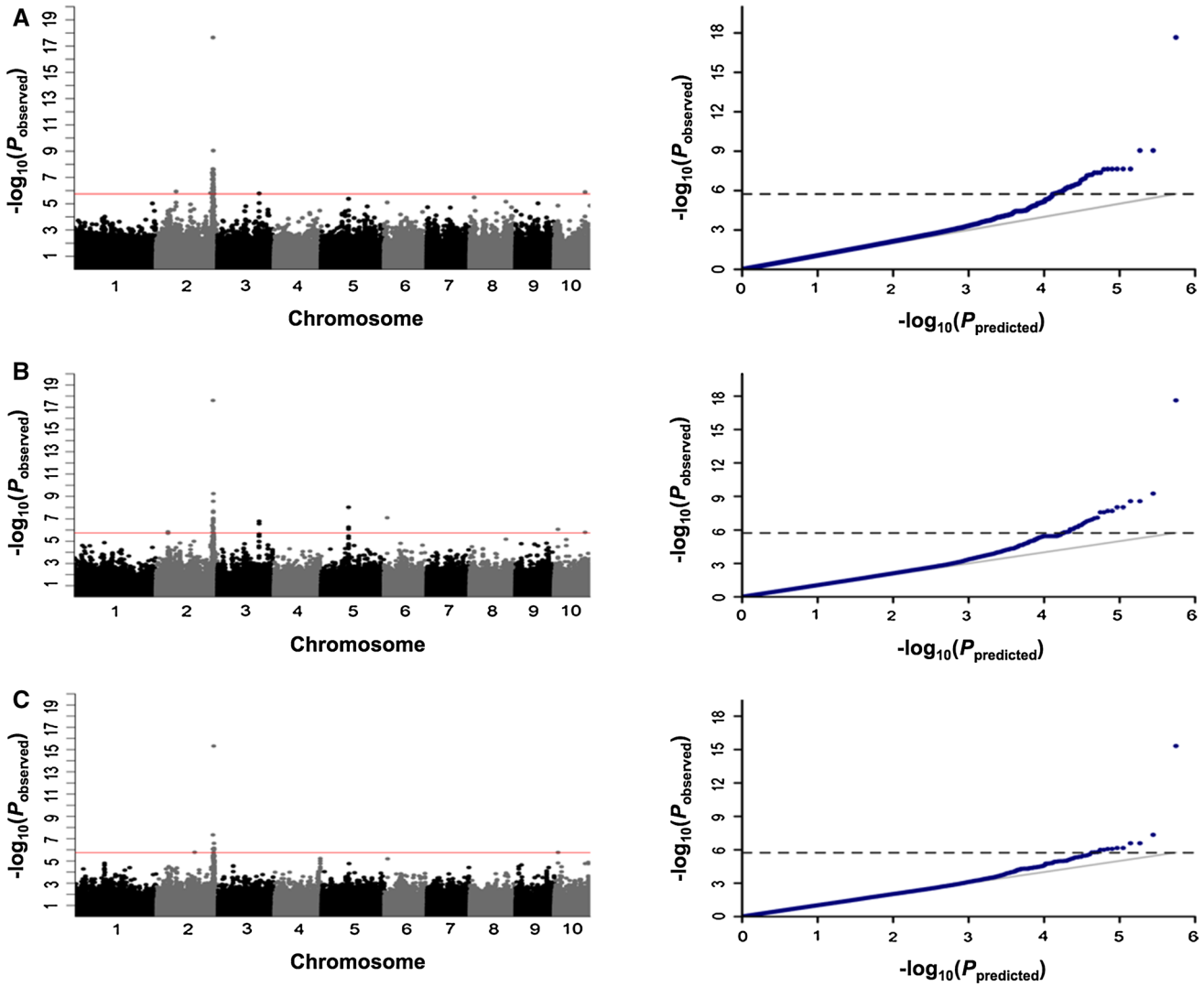

Fig. 1 Manhattan and quartile–quartile plots of GWAS results for pollen fertility. **a–c** The results in the environments of Hainan, Wuhan and Haerbin, respectively

1, 2, 3, 4–10 and 11–17 favorable alleles had an average pollen fertility of 17.18, 29.86, 42.46, 44.31 and 56.22 %, respectively.

Combining analysis also revealed that the total of 19 loci for pollen fertility explained 60.93 % of phenotypic variation. In addition, the effects of *Rf3* alone and the other 18 significant loci (in total) were also compared, they explained about 27.06 and 37.23 % of phenotypic variation, respectively.

These results indicated that *Rf3* alone could completely restore the pollen fertility of CMS-S in the lines with *Rf3* allele, but the *Rf3* frequency was only about 10 % of the lines in the association panel. Although the average effect of the other loci was relatively small, pyramiding of favorable alleles from *Rf3* and the other loci could increase pollen fertility of CMS-S.

Genome-wide association mapping for anther exsertion and pollen shedding

A total of three SNP-trait associations were identified for anther exsertion with the value of $-\log_{10}(P)$ larger than 5.7 ($P < 1.8 \times 10^{-6}$, $\alpha = 1$) (Table 4). These SNPs were distributed on different chromosomes. None of them was identified simultaneously under two environments. Individual locus explained less than 9.26 % of phenotypic variation. One locus, *Rf3*, was also associated with pollen fertility.

For the trait of pollen shedding, eight loci distributed on three chromosomes were detected at the threshold of $-\log_{10}(P) = 5.7$ ($P < 1.8 \times 10^{-6}$, $\alpha = 1$), and individual locus explained 9.20–15.75 % of phenotypic variation (Table 4). Only one locus, *Rf3*, was identified simultaneously under two environments, and *Rf3* alone explained

Table 3 *Rf3* and the 18 significant loci for pollen fertility revealed by conditional association analysis in the three environments

Chr	Position[a]	Alleles[b]	MAF	HN		WH		HB	
				$-\log_{10}(P)$	PVE (%)	$-\log_{10}(P)$	PVE (%)	$-\log_{10}(P)$	PVE (%)
2	16,219,902	T/C	0.12			6.14	5.57		
2	22,611,619	T/G	0.07			7.19	6.83	6.35	6.85
2	149,204,832	T/C	0.06	6.21	6.29	5.75	5.37		
2	**224,905,041**	T/G	0.08	6.58	6.63	5.62	5.18		
2	225,851,740	G/C	0.05	7.86	7.93	5.71	5.09	5.66	5.79
2	226,205,610	T/C	0.06	7.68	7.77	8.58	8.28		
2	226,336,663	T/A	0.07	6.00	5.94	6.84	6.48		
2	**227,045,358**	T/C	0.06	17.65	28.26	17.61	26.52	15.31	25.23
2	**228,065,331**	C/G	0.05	10.29	10.91	10.46	10.36	8.63	9.58
2	**228,895,908**	C/G	0.07	6.03	5.91	9.05	7.69		
2	**229,571,211**	T/A	0.06	9.67	10.25	8.81	8.56	8.07	8.85
2	**229,674,212**	C/T	0.08	10.30	11.49	8.27	8.32	7.51	8.53
3	160,790,386	A/C	0.06	6.09	5.92	6.00	5.38		
3	**198,788,317**	G/T	0.07	6.10	5.94	8.19	7.83		
5	205,544,441	T/G	0.05	5.91	5.72				
8	12,196,297	A/C	0.13	6.07	5.72				
8	14,702,495	T/C	0.10	6.34	6.55				
8	164,028,314	T/A	0.06	8.38	8.69	8.13	7.74		
10	**143,717,120**	G/A	0.05	6.11	6.03				

MAF minor allele frequency, *PVE* phenotypic variation explained by the locus

[a] *Rf3* locus (Chr2: 227,045,358) and the seven loci revealed by both GWAS and conditional analysis are shown in bold

[b] The favorable alleles are presented on the left of the slash sign

about 15 % of phenotypic variation. Three loci were also associated with pollen fertility.

Candidate genes associated with fertility restoration ability to CMS-S

A total of 83 annotated candidate genes were predicted based on the extension regions from 50 Kb upstream and downstream (LD of the panel) of the 19 significant SNPs. Detailed information of these genes is summarized in Table S1. For each locus, one to eight (on average of 3.2) candidate genes were predicted. These genes can be classified into five categories according to the motifs and putative functions of their encoding proteins. The first type (2/83) includes two PPRs, *GRMZM2G158308* and *GRMZM2G450166*, and PPR-mediated fertility restoration is one of the known mechanisms of CMS/Rf system in plants (Wang et al. 2006; Hu et al. 2012; Tang et al. 2014).

Two genes encoding proteins with AP2 domain (*GRMZM2G081892_T01*, *GRMZM2G310368*), four F-box genes (*GRMZM2G135087_T01*, *GRMZM2G150383_T01*, *GRMZM2G001180_T01*, *GRMZM2G107945_T01*), two RNA processing related genes (*GRMZM2G053239* and *GRMZM2G070831*), two genes related to protein folding or chaperone (*GRMZM2G135354_T01*, *GRMZM2G519073*), a gene encoding a 14-3-3 protein (*GRMZM2G408768*) and seven genes encoding transcription factors with zinc fingers, leucine zipper, WAKY or HLH domains (*GRMZM2G445163_T02*, *GRMZM2G007028*, *GRMZM2G007028_T01*, *GRMZM2G440125_T01*, *GRMZM2G318412_T01*, *GRMZM2G468056*, *GRMZM2G114873*) are regarded as the second type (18/83); they are involved in DNA/RNA binding, RNA processing, protein folding or protein–protein interactions, playing roles in regulation of gene expression.

The third group (4/83) is characterized as signal transduction, including an auxin-carrier (*GRMZM2G025742*), an auxin-responsive gene (*GRMZM2G429254*), a GTP binding protein gene (*GRMZM2G302233*), and a gene encoding Ca/CaM-dependent phosphorase (*GRMZM2G133854*).

The majority of the predicted genes (48/83) involved in biosynthesis, metabolism, cell division and growth are in the fourth type, such as *GRMZM2G057514* encoding a cytochrome P450 protein, *GRMZM2G372068* encoding an UDP-glucosyl transferase, *GRMZM2G435993_T01*, encoding a 1,3-beta-glucan synthase, *GRMZM2G105570* encoding an ABC transporter, which found to be involved in the

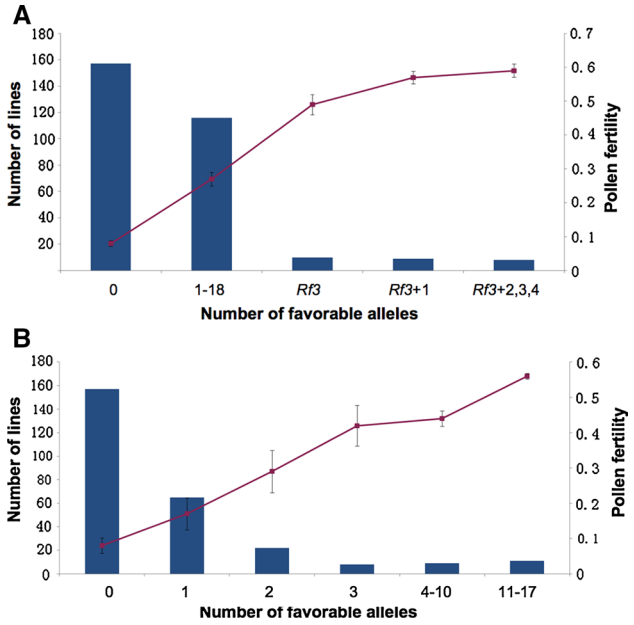

Fig. 2 Combining effects of favorable alleles of the 19 loci for pollen fertility. **a** Combining effects of *Rf3* and the other loci. *0* represents the lines had on elite allele at all, *1–18* means the lines contained 1–18 elite alleles excluding *Rf3*, *Rf3*, *Rf3* + 1, and *Rf3* + 2, 3, 4 represent the lines with favorable alleles from *Rf3* alone, *Rf3* and one other significant loci, *Rf3* and 2–4 other loci; **b** combining effects of the other significant loci except *Rf3*

biosynthesis pathways of starch and fatty alcohol, which are important in pollen and anther development. The last type includes 11 members (such as *AC204502.4_FGT006*, *GRMZM2G140803_T01*) encoding putative proteins with unknown functions.

Discussion

Besides *Rf3*, multiple loci with relatively smaller effects involved in fertility restoration to CMS-S in maize had been detected

In this study, a total of 30 significant loci were identified for pollen fertility, anther exsertion and pollen shedding. Individual locus explained up to 28.26 % of phenotypic variation for these traits. This well agreed with the fact that these traits fit or closely fit to normal distributions. Although *Rf3* alone could completely restore the pollen fertility of CMS-S, it only presented in less than 10 % of the association panel. In addition, combining the favorable alleles of the other significant loci would increase the pollen fertility of CMS-S. It indicates that multiple genes with different effects might be involved in the fertility restoration ability to CMS-S in maize, and a number of loci with smaller effects explained that pollen fertility in majority of the association panel was partially restored. In maize, Kohls et al. (2011) also identified several major and minor QTLs for partial fertility restoration to C type CMS. Among the significant loci revealed in the association panel, 11 clustered on the chromosome 2L where *Rf3* is located. A previous study also reported that the *Rf3* region of 2L potentially harbors a complex of linked *Rf* genes of CMS-S (Gabay-Laughnan et al. 2004).

In addition, of the three and eight loci for anther exsertion and pollen shedding, only one and three loci were also associated with pollen fertility, respectively. This indicated that the genetic architecture of pollen fertility is not the same as anther exsertion and pollen shedding, and that

Table 4 Significant loci for anther exsertion and pollen shedding in the two environments

Traits	Chr	Position[a]	Allele[b]	MAF[c]	HB		WH	
					$-\log_{10}(P)$	PVE (%)[d]	$-\log_{10}(P)$	PVE (%)
Anther exsertion	**2**	**227,045,358**	T/C	0.06			6.43	8.48
	3	166,436,207	T/C	0.46			6.44	9.26
	8	170,293,742	G/T	0.24	6.29	8.52		
Pollen shedding	2	225,781,253	T/C	0.16			5.76	9.23
	2	**227,045,358**	T/C	0.06	8.54	14.49	9.08	15.75
	2	227,531,495	A/G	0.34			6.79	10.82
	2	**228,895,908**	C/G	0.07			7.33	12.11
	2	**229,674,212**	C/T	0.08			6.25	9.84
	3	157,918,370	G/A	0.20			5.94	10.54
	3	188,669,740	A/C	0.11	6.13	10.00		
	9	152,065,296	G/C	0.09			5.79	9.20

[a] The loci shown in bold represent that they were also associated with pollen fertility. See the footnotes for b, c and d in Table 3

between pollen fertility and pollen shedding is relatively similar. This also agreed with the fact that correlation coefficients between pollen fertility and pollen shedding are higher than that between pollen fertility and anther exsertion. Although the effects of genotype by environment for the three traits were not significant, half (15) of the significant loci were detected only in one environment, especially for the traits of anther exsertion and pollen shedding. This might be caused by the polygenic, small effect genetic pattern of the traits and high threshold used to declare significant SNP-trait associations for these traits. As we reduced the threshold to $-\log_{10}(P) = 5.0$ ($P < 9.0 \times 10^{-6}$, $\alpha = 5$), most of the loci were detected in at least two environments. In addition, variations of flowering time in the population under different environments would also influence the detection of some loci with minor effects, although the correlations between flowering time and the traits were not significant ($-0.08 < r < 0.06$). We also noticed that more significant loci were identified in HN (the environment with the lowest temperature among the three environments); this implied that some loci might be thermosensitive as indicated for *Rf9* (Gabay-Laughnan et al. 2009).

Multiple mechanisms might be involved in fertility restoration to CMS-S

Based on the B73 sequences within the 100 kb extension regions of the 26 significant SNPs (four of them were associated with more than one trait) revealed by GWAS for the three traits, totally 83 candidate genes were predicted. The first group contains two genes encoding proteins with PPR domains. The SNP (Chr2: 227,045,358) significantly associated with *Rf3* is located 536 bp down-stream of *GRMZM2G450166*. In B73, *GRMZM2G450166* encodes a P type PPR protein predicted targeting to mitochondrial ($p = 0.9338$) with online software Mitoprot II –v1.101 (http://ihg.gsf.de/ihg/mitoprot.html). Most of the *Rf*-related PPRs cloned belong to P type, and targeted to mitochondrial (Wang et al. 2006; Hu et al. 2012). Gene expression analysis (Davidson et al. 2011) showed that it expressed in all the tissues investigated with the highest level in pollens (http://qteller.com/qteller3/). In addition, with a large fine-mapping population, *Rf3* was narrowed in a 28 kb region harboring *GRMZM2G450166* (data not shown). Thus, it is considered as a candidate gene of *Rf3*. However, the inbred B73 does not carry a functional *Rf3* allele, to verify the function of this gene we need to acquire the full-length gene sequence from restorers of CMS-S. The three candidate PPR genes identified by Xu et al. (2009) are about 900 kb away from the significant locus of *Rf3*, and no significant SNP was detected near the three candidate genes. The other PPR gene might play roles in fertility restoration to CMS-S in restorer lines without *Rf3*. PPR proteins targeting to mitochondrial acted to reduce the accumulation of CMS-related RNAs in the company of other factors with functions of RNA binding or processing (Hu et al. 2012). In the second group, two RNA processing-related genes (*GRMZM2G053239* and *GRMZM2G070831*) were identified; *GRMZM2G135354_T01* and *GRMZM2G519073* encode chaperone or prefoldin, which were found to be responsible for the transportation and refolding of proteins from the cytoplasm into the mitochondrial matrix (Koll et al. 1992). These genes might be also involved in PPR-mediated fertility restoration to CMS-S.

Plant hormones, including auxin, regulate many aspects of plant growth and development, and genes related to auxin metabolism (Zhao et al. 2013), auxin-responsive (Yadav et al. 2011) and the regulatory network of GA/Auxin (Kay et al. 2013) played important roles in male fertility. A gene in the fourth group, *GRMZM2G140822_T01*, encodes a 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein, which catalyzes the irreversible oxidation of IAA to OxIAA. Zhao et al. (2013) found that a member of this family, DAO, was essential for rice reproductive development including anther dehiscence and pollen fertility. In addition, in the third group, a gene encoding an auxin-carrier (*GRMZM2G025742*), an auxin-responsive gene (*GRMZM2G429254*), a GTP-binding protein gene (*GRMZM2G302233*), and a gene encoding Ca/CaM-dependent phosphorase (*GRMZM2G133854*) are all related to auxin transporting or signal transduction. These genes might be involved in fertility restoration to CMS-S by the pathway of auxin transportation or signal transduction.

It is interesting to notice that some predicted genes in the fourth group encode transcription factors, a 14-3-3 protein and proteins with F-box domain. These genes are involved in protein–protein interactions, and might play roles in regulating the development of pollens or anthers. For instance, a helix–loop–helix transcription factor (Jeong et al. 2014) and a F-box gene (Li et al. 2015) were found to be essential for pollen and anther development recently; 14-3-3 protein was also found to be involved in starch accumulation regulation in maize pollens (Datta et al. 2002). In addition, a number of functional genes related to starch or cellulose synthesis were predicted in the fourth group, which might play important roles in the development of anthers and pollens as well. For instance, *GRMZM2G097841*, *GRMZM2G097854*, *GRMZM2G431504* and *GRMZM2G001401* encode reductase proteins; this kind of proteins was found to be associated with male sterile in Arabidopsis (Aarts et al. 1997); *GRMZM2G105570* encodes an ABC transporter. *GRMZM2G057514_T01* encodes a cytochrome P450 protein. Rice CYP703A3, a cytochrome P450 hydroxylase, was found to be essential for anther cuticle and pollen exine development (Yang et al. 2014b), and an ABC

transporter, OsABCG15, was required for the transport of lipidic precursors for anther cuticle and pollen exine development (Niu et al. 2013; Yang et al. 2014a, b). A total of eight genes encoding enzymes related to carbohydrate or lipid metabolisms (such as *GRMZM2G372068_T01* encoding a UDP-glucosyl transferase, *GRMZM2G135195_T01* encoding a galacturonosyltransferase) are in this category; these enzymes were essential for pollen and anther development, such as β-1,3-galactosyltransferase in Ms8 (Wang et al. 2013).

In conclusion, fertility restoration to CMS-S could be related to at least three mechanisms in this association mapping panel: (1) *Rf3* and other PPR genes governed interactions with the mitochondrial CMS-related RNAs; (2) Auxin synthesis, responsive related pathways; and (3) the pathways related to pollen and anther development. *Rf3* alone only explained the fertility restoration in less than 10 % of the inbreds used in this study, and other *Rf* genes restored the fertility of CMS-S with different mechanisms in the other inbreds, especially for that with partial fertility restoration ability. Mechanisms of fertility restoration to CMS not related to PPR genes have also been reported in maize (Cui et al. 1996; Gabay-Laughnan et al. 2009) and rice (Itabashi et al. 2011; Hu et al. 2013).

Utility of CMS-S in maize hybrid breeding

In hybrid breeding harnessing CMS, it is important to select sterile inbreds with absolutely sterile pollen, and restorer lines with high restoration ability to CMS. In addition, the pattern of fertility restoration to CMS-S in maize is gametophytic, thus theoretically only 50 % of the pollens are fertile in the hybrids with only one restoration gene, *Rf3*. More non-allelic restoration genes pyramiding in the restorers would increase pollen fertility in hybrids, thus it would be safer in maize production. For instance, Huang et al. (2012) found two non-allelic nuclear genes restored fertility of the gametophytic HL-CMS hybrids to 75 %, and this increased the tolerance of abiotic stress of the rice hybrids.

The polygenic pattern of fertility restoration to CMS-S in maize makes the usage of this type of CMS in hybrid breeding very difficult. Only 60 out of the 360 testcrossing hybrids used in this study have less than 1 % fertile pollens, and majority of the inbreds partially restored the fertility of the CMS-S line. This will increase the difficulty of breeding new CMS-S lines. Pollen fertility in 24 testcrossing hybrids (possessed more favorable alleles, Fig. 2) is higher than 60 %, indicating that it is possible to increase pollen fertility of CMS-S hybrids by pyramiding more non-allelic restoration genes. On the other hand, stable CMS-S lines could be bred by pyramiding the unfavorable alleles at the loci identified in this study. So the genetic loci revealed in this study would provide useful information to breed male sterile and restorer lines of CMS-S by markers-assisted selection.

We also noticed that the genetic architecture between pollen fertility and fertility observed in field (anther exsertion and pollen shedding) is different in a certain degree, especially for that between pollen fertility and anther exsertion. In our testcrossing association mapping panel, hybrids with good performance of field fertility but low pollen fertility and hybrids with high pollen fertility but unable dehiscence well were all observed. Thus, in hybrids breeding, pollen fertility, anther exsertion and pollen shedding should be under consideration together.

Author contribution statement Conceived and designed the experiments: B.Y. Performed the experiments: Y.F., Q.Z., H.S., Y.W., H.W. L.J. Contributed materials/analysis tools: J.Y., Y.Z. Wrote the paper: Y.F., B.Y.

Acknowledgments This research was supported by a grant from the National Natural Science Foundation of China (31371635).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards The authors declare that the experiments comply with the current laws of the country in which they were performed.

References

Aarts MG, Hodge R, Kalantidis K, Florack D, Wilson ZA, Mulligan BJ, Stiekema WJ, Scott R, Pereira A (1997) The Arabidopsis MALE STERILITY 2 protein shares similarity with reductases in elongation/condensation complexes. Plant J 12:615–623

Bentolila S, Alfonso A, Hanson M (2002) A pentatricopeptide repeat containing gene restores male sterility to male-sterile plants. PNAS USA 99:10887–10892

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635

Brown GG, Formanova N, Jin H, Wargachuk R, Dendy C, Patil P, Laforest M, Zhang J, Cheung WY, Landry BS (2003) The radish *Rfo* restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. Plant J 35:262–272

Cui XQ, Wise RP, Schnable PS (1996) The *rf2* nuclear restorer gene of male-sterile T-cytoplasm maize. Science 272:1334–1336

Datta R, Chamusco KC, Chourey P (2002) Starch biosynthesis during pollen maturation is associated with altered patterns of gene expression in maize. Plant Physiol 130:1645–1656

Davidson RM, Hansey CN, Gowda M, Childs KL, Lin H, Waillancout B, Sekhon RS, de Leon N, Kaeppler SM, Jiang N, Buell CR (2011) Utility of RNA sequencing for analysis of maize reproductive transcriptomes. Plant Genome 4:191–203

Fujii S, Toriyama K (2009) Suppressed expression of retrograde-regulated male sterility restores pollen fertility in cytoplasmic male sterile rice plants. PNAS USA 106:9513–9518

Gabay-Laughnan S, Chase CD, Ortega VM, Zhao L (2004) Molecular-genetic characterization of CMS-S restorer-of-fertility alleles identified in mexican maize and teosinte. Genetics 166:959–970

Gabay-Laughnan S, Kuzmin EV, Monroe J, Roark L, Newton KJ (2009) Characterization of a novel thermosensitive restorer of fertility for cytoplasmic male sterility in maize. Genetics 182:91–103

Hu J, Wang K, Huang W, Liu G, Gao Y, Wang J, Huang Q, Ji Y, Qin X, Wan L, Zhu R, Li S, Yang D, Zhu Y (2012) The rice pentatricopeptide repeat protein RF5 restores fertility in Hong-Lian cytoplasmic male-sterile lines via a complex with the glycine-rich protein GRP162. Plant Cell 24:109–122

Hu J, Huang W, Huang Q, Qin X, Dan Z, Yao G, Zhu R, Zhu Y (2013) The mechanism of ORFH79 suppression with the artificial restorer fertility gene Mt-GRP162. New Phytol 199:52–58

Huang W, Hu J, Yu C, Huang Q, Wan L, Wang L, Qin X, Ji Y, Zhu R, Li S, Zhu Y (2012) Two non-allelic nuclear genes restore fertility in a gametophytic pattern and enhance abiotic stress tolerance in the hybrid rice plant. Theor Appl Genet 124:799–8077

Itabashi E, Iwata N, Fujii S, Kazama T, Toriyama K (2011) The fertility restorer gene, *Rf2*, for lead rice-type cytoplasmic male sterility of rice encodes a mitochondrial glycine-rich protein. Plant J 65:359–367

Jeong HJ, Kang JH, Zhao M, Kwon JK, Choi HS, Bae JH, Lee HA, Joung YH, Choi D, Kang BC (2014) Tomato *Male sterile 10*^{35} is essential for pollen development and meiosis in anthers. J Exp Bot 65:6693–6709

Kamps TL, Chase CD (1997) RFLP mapping of the maize gametophytic restorer of fertility locus (*Rf3*) and aberrant pollen transmission of the nonrestoring *Rf3* allele. Theor Appl Genet 95:525–531

Kay P, Groszmann M, Ross JJ, Parish RW, Swain SM (2013) Modifications of a conserved regulatory network involving INDEHISCENT controls multiple aspects of reproductive tissue development in Arabidopsis. New Phytol 197:73–87

Kazama T, Toriyama K (2003) A pentatricopeptide repeat-containing gene that promotes the processing of aberrant *apt6* RNA of cytoplasmic male-sterile rice. FEBS Lett 544:99–102

Kohls S, Stamp P, Knaak C, Messmer R (2011) QTL involved in the partial restoration of male fertility of C-type cytoplasmic male sterility in maize. Theor Appl Genet 123:327–338

Koizuka N, Imai R, Fujimoto H, Hayakawa T, Kimura Y, Kohno-Murase J, Sakai T, Kawasaki S, Imamura J (2003) Genetic characterization of a pentatricopeptide repeat protein gene, orf687, that restores fertility in the cytoplasmic male-sterile Kosena radish. Plant J 34:407–415

Koll H, Guiard B, Rassow J, Ostermann J, Horwich AL, Neupert W, Hartl FU (1992) Antifolding activity of hsp60 couples protein import into the mitochondrial matrix with export to the intermembrane space. Cell 68:1163–1175

Komori T, Ohta S, Murai N, Takakura Y, Kuraya Y, Suzuki S, Hiei Y, Imaseki H, Nitta N (2004) Map-based cloning of a fertility restorer gene, *Rf-1*, in rice (*Oryza sativa* L.). Plant J 37:315–325

Laughnan JR, Gabay SJ (1973) Mutations leading to nuclear restoration of fertility in S male-sterile cytoplasm in maize. Theor Appl Genet 43:109–116

Laughnan JR, Gabay SJ (1978) Nuclear and cytoplasmic mutations to fertility in S male-sterile maize. In: Walden DB (ed) Maize breeding and genetics. Wiley, New York, pp 427–446

Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, Han Y, Chai Y, Guo T, Yang N, Liu J, Warburton ML, Cheng Y, Hao X, Zhang P, Zhao J, Liu Y, Wang G, Li J, Yan J (2013) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. Nat Genet 45:43–50

Li L, Li Y, Song S, Deng H, Li N, Fu X, Chen G, Yuan L (2015) An anther development F-box (ADF) protein regulated by tapetum degeneration retardation (TDR) controls rice anther development. Planta 241:157–166

Matsuhira H, Kagami H, Kurata M, Kitazaki K, Matsunaga M, Hamaguchi Y, Hagihara E, Ueda M, Harada M, Muramatu A, Yui-Kurino R, Taguchi K, Tamagake H, Mikami T, Kubo T (2012) Unusual and typical features of a novel restorer-of-fertility gene of sugar beet (*Beta vulgaris* L.). Genetics 192:1347–1358

Niu B, He F, He M, Ren D, Chen L, Liu Y (2013) The ATP-binging cassette transporter OsABCG15 is required for anther development and pollen fertility in rice. J Integr Plant Biol 55:710–720

Schmitz-Linneweber C, Small I (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci 13:663–670

Tang H, Luo D, Zhou D, Zhang Q, Tian D, Zheng X, Chen L, Liu Y (2014) The rice restorer *Rf4* for wild-abortive cytoplasmic male sterility encodes a mitochondrial-localized PPR protein that functions in reduction of WA352 transcripts. Mol Plant 7:1497–1500

Tie S, Xia J, Qiu F, Zheng Y (2006) Genome-wide analysis of maize cytoplasmic male sterility-S based on QTL mapping. Plant Mol Biol Rep 24:71–81

Wang Z, Zou Y, Zhang Q, Li X, Chen L, Wu H, Su D, Guo J, Chen Y, Luo D, Long Y, Zhong Y, Liu YG (2006) Cytoplasmic male sterility of rice with Boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. Plant Cell 18:676–687

Wang D, Skibbe DS, Walbot V (2013) Maize male sterile 8 (Ms8), a putative β-1,3-galactosyltransferase modulates cell division, expression, and differentiation during early maize anther development. Plant Reprod 26:329–338

Xu X, Liu Z, Zhang D, Liu Y, Song W, Li J, Dai J (2009) Isolation and analysis of rice Rf1-orthologus PPR genes co-segregating with *Rf3* in maize. Plant Mol Biol Rep 27:511–517

Yadav SR, Khanday I, Majhi BB, Veluthambi K, Vijayraghavan U (2011) Auxin-responsive OsMGH3, a common downstream target of OsMADS1 and OsMADS6, controls rice floret fertility. Plant Cell Physiol 52:2123–2135

Yan J, Warburton M, Crouch J (2011) Association mapping for enhancing maize genetic improvement. Crop Sci 51:433–449

Yang C, Xu Z, Song J, Conner K, Vizcay Barrena G, Willson ZA (2007) Arabidopsis MYB26/MALE STERILE35 regulates secondary thichening in the endothecium and is essential for anther dehiscence. Plant Cell 19:534–548

Yang N, Lu Y, Yang X, Huang J, Zhou Y, Ali F, Wen W, Liu J, Yan J (2014a) Genome wide association studies using a new nonparametric model reveal the genetic architecture of 17 agronomic traits in an enlarged maize association panel. PLoS Genet 10:e1004573

Yang X, Wu D, Shi J, He Y, Pinot F, Grausem B, Yin C, Zhu L, Chen M, Luo Z, Liang W, Zhang D (2014b) Rice CYP703A3, a cytochrome P450 hydroxylase, is essential for development of anther cuticle and pollen exine. J Integr Plant Biol 56:979–994

Zhang Z, Wang Y, Zheng Y (2006) AFLP and PCR-based markers linked to *Rf3*, a fertility restorer gene for S cytoplasmic male sterility in maize. Mol Genet Genomics 276:162–169

Zhao Z, Zhang Y, Liu X, Zhang X, Liu S, Yu X, Ren Y, Zheng X, Zhou K, Jiang L, Guo X, Gai Y, Wu C, Zhai H, Wang H, Wan J (2013) A role for a dioxygenase in auxin metabolism and reproductive development in rice. Dev Cell 27:113–122