



A Genome Wide Association Study Revealed Key Single Nucleotide Polymorphisms/Genes Associated With Seed Germination in *Gossypium hirsutum* L.

Aijun Si^{1,2}, Zhengwen Sun¹, Zhikun Li¹, Bin Chen¹, Qishen Gu¹, Yan Zhang¹, Liqiang Wu¹, Guiyin Zhang¹, Xingfen Wang^{1*} and Zhiying Ma^{1*}

¹ State Key Laboratory of North China Crop Improvement and Regulation, North China Key Laboratory for Crop Germplasm Resources of Education Ministry, Key Laboratory for Crop Germplasm Resources of Hebei, Hebei Agricultural University, Baoding, China, ² Key Laboratory of China Northwestern Inland Region, Ministry of Agriculture, Cotton Research Institute, Xinjiang Academy of Agricultural and Reclamation Science, Shihezi, China

OPEN ACCESS

Edited by:

Ahmad M. Alqudah,
Aarhus University, Denmark

Reviewed by:

Samar Gamal Thabet,
Fayoum University, Egypt
Mian Faisal Nazir,
Cotton Institute (CAAS), China

*Correspondence:

Zhiying Ma
mzhy@hebau.edu.cn
Xingfen Wang
cotton@hebau.edu.cn

Specialty section:

This article was submitted to
Plant Physiology,
a section of the journal
Frontiers in Plant Science

Received: 29 December 2021

Accepted: 21 February 2022

Published: 16 March 2022

Citation:

Si A, Sun Z, Li Z, Chen B, Gu Q, Zhang Y, Wu L, Zhang G, Wang X and Ma Z (2022) A Genome Wide Association Study Revealed Key Single Nucleotide Polymorphisms/Genes Associated With Seed Germination in *Gossypium hirsutum* L.
Front. Plant Sci. 13:844946.
doi: 10.3389/fpls.2022.844946

Fast and uniform seed germination is essential to stabilize crop yields in agricultural production. It is important to understand the genetic basis of seed germination for improving the vigor of crop seeds. However, little is known about the genetic basis of seed vigor in cotton. In this study, we evaluated four seed germination-related traits of a core collection consisting of 419 cotton accessions, and performed a genome-wide association study (GWAS) to explore important loci associated with seed vigor using 3.66 million high-quality single nucleotide polymorphisms (SNPs). The results showed that four traits, including germination potential, germination rate, germination index, and vigor index, exhibited broad variations and high correlations. A total of 92 significantly associated SNPs located within or near 723 genes were identified for these traits, of which 13 SNPs could be detected in multiple traits. Among these candidate genes, 294 genes were expressed at seed germination stage. Further function validation of the two genes of higher expression showed that *Gh_A11G0176* encoding Hsp70-Hsp90 organizing protein negatively regulated *Arabidopsis* seed germination, while *Gh_A09G1509* encoding glutathione transferase played a positive role in regulating tobacco seed germination and seedling growth. Furthermore, *Gh_A09G1509* might promote seed germination and seedling establishment through regulating glutathione metabolism in the imbibitional seeds. Our findings provide unprecedented information for deciphering the genetic basis of seed germination and performing molecular breeding to improve field emergence through genomic selection in cotton.

Keywords: cotton, seed germination, GWAS, SNP, candidate genes

INTRODUCTION

Cotton is an important fiber crop and a valuable oilseed crop. Among the cultivated species, upland cotton (*Gossypium hirsutum* L.) has the largest cultivated areas and plays an essential part in daily life and global textile industry (Fang et al., 2017; Naoumkina et al., 2019). Seed germination is the beginning of the life cycle of seed plants, as well as an important link for high-yield and resistance breeding of cotton. Seed vigor is an important character of seed quality that reflects

seed germination rate and uniformity, as well as seedling emergence and growth in the field. Hence, seedling vigor may influence crop yield through both indirect and direct effects (Ellis, 1992; Foolad et al., 2007; Sun et al., 2007). In modern agricultural cultivation practice, rapid and uniform seed germination is the key factor to maximize crop yield potential. Seeds with high vigor germinate quickly and result in uniform and strong seedlings, which ultimately affect plant growth and development (Foolad et al., 2007). With the rapid application of mechanized cultivation and harvest, the quality and vitality of cotton seeds will be particularly important to achieve precision sowing technology.

Seed vigor is a comprehensive embodiment that involves many aspects from seed germination to seedling growth, such as seed germination rate and uniformity, seedling growth rate, and growth rhythm (Ellis, 1992). Seed germination trait is under multiple-genes control and susceptible to environmental factors (Yu et al., 2019a), and is the main component of seed vigor. In previous studies, there have been reports on seed vigor, seed germination, seed field emergence, exogenous hormones, and hormone signaling pathways that regulate seed germination in cotton and other species (Zhang et al., 2014; He et al., 2019; Xiao et al., 2019; Yu et al., 2019a,b; Song et al., 2020; Zhou et al., 2020). In cotton, cold atmospheric-pressure plasma (CAP) treatment improved seed germination and chilling tolerance (Groot et al., 2018). Matrix-localized heat shock protein GhHSP24.7 mediated seed germination *via* temperature-dependent reactive oxygen species (ROS) germination (Ma et al., 2019). In rice, *OsIPMS1* affected seed vigor associated with amino acid and energy metabolism (He et al., 2019), *osa-miR164c* and *osa-miR168a* played a key role in regulating seed vigor in transgenic plants (Zhou et al., 2020). In soybean, quantitative proteomic analyses of two low phytic acid mutants showed that the high germination rate in the TW-1-M might be strongly attributed to ROS-related and plant hormone-related genes (Yu et al., 2019a). In addition, exogenous hormones and hormone signaling pathways have also been reported to play a pivotal role in the regulation of seed germination and dormancy. For instance, exogenous melatonin promoted seed germination and osmotic regulation under arid and salt stress (Xiao et al., 2019; Bai et al., 2020; Chen et al., 2020). *AtPER1* reduced primary seed germination *via* suppressing ABA catabolism and promoting GA biosynthesis in Arabidopsis seeds (Chen et al., 2019). *OsRACK1A* positively regulated seed germination by means of changing the endogenous amounts of ABA and ROS, as well as their interplay (Zhang et al., 2014). *OsMFT2* was involved in the regulation of ABA signaling-mediated seed germination through interacting with *OsZIP23/66/72* in rice (Song et al., 2020).

Glutathione S-transferases (GSTs) are multifunctional enzymes that play important roles in biological processes such as plant development, metabolism and abiotic, and biotic stress responses by catalyzing the conjugation of electrophilic substrates with glutathione (GSH), thereby reducing their toxicity (Dixon et al., 2002; Moons, 2005; Frova, 2006; Kao et al., 2016). It has been shown that GSTs promote seed germination under abiotic stresses. Seed germination and seedling growth of *GST* overexpressing tobacco was significantly improved under stressful conditions (Roxas et al., 1997). Overexpression

of *SbGSTU* in tobacco enhanced seed germination under salt stress (Jha et al., 2010). *Atgstu17* regulated seed germination by the combined effect of GSH and ABA (Chen et al., 2012), and GSH treatment improved germination after seed dehydration (Kalemba and Ratajczak, 2018). The functional deficiency of *AtGSTU7* resulted in increased GSH content and decreased H₂O₂ content in germinating seeds, and GSH was involved in seed germination under ABA treatment, implying that *AtGSTU7* involvement in seed germination was mediated by GSH-ROS homeostasis and ABA signaling (Wu et al., 2020).

In the past few years, QTL mapping based on biparental linkage analysis has become an effective approach to identify seed germination related genes in many crops, such as rice (Cui et al., 2002; Jiang et al., 2017; Yang et al., 2019; Jiang et al., 2020), barley (Moursi et al., 2020), and *Brassica rapa* (Basnet et al., 2015), but it was usually limited by the number of markers that could be employed along the chromosomes. As an excellent complement to QTL, GWAS is an effective method to detect mark-trait association (Zhao et al., 2011; Yu et al., 2017), it has been successfully implemented in rice (Sales et al., 2017; Shi et al., 2017; Yang et al., 2019), maize (Hu et al., 2017), oat (Huang et al., 2020), *Brassica napus* (Hatzig et al., 2015; Tan et al., 2017), soybean (Zhou et al., 2015; Liu et al., 2020), and other crops during germination for the identification of single nucleotide polymorphism (SNP) loci and candidate genes for various ecological and agricultural traits. In conclusion, GWAS has been successfully applied to identify the potential candidate genes underlying important agronomic traits with high-density SNPs from diverse germplasms. However, there are few studies on the rapid and accurate identification of a large number of candidate genes for seed germination in cotton, and the seed germination mechanism is still unclear. In addition, previous studies have shown that there is a very significant correlation between seed germination potential and field seedling emergence (Chen, 2012; Xie et al., 2019), and the germination rate measured in sand bed is positively correlated with the seedling emergence rate in the field (Wang, 2007). Therefore, GWAS analysis based on genotypic and phenotypic data for large-scale accessions and SNP markers should provide a powerful strategy to detect candidate genes and unravel the molecular mechanism for seed germination that is important for cotton improvement.

In the present study, we performed a GWAS for seed germination traits based on 3,665,030 SNPs from a core collection consisting of 419 diverse germplasm resources in *G. hirsutum* L. (Ma et al., 2018). The objectives were to identify SNPs significantly associated with germination capacity and candidate genes, providing useful information for better understanding the genetic mechanism of cotton seed germination so as to facilitate molecular breeding with increased field emergence rate and precision sowing.

MATERIALS AND METHODS

Plant Materials

In this study, a core collection comprising of 419 upland cotton accessions was used to conduct GWAS. The 419

accessions had abundant phenotypic variation and were used to conduct GWAS for fiber quality, yield, and phosphorus deficiency tolerance in the previous study of our group (Ma et al., 2018; Gu et al., 2020). Among which, 317 accessions were collected from different provinces of China and the remaining accessions were derived from major cotton-growing countries, including the United States, the former Soviet Union, Pakistan, Turkey, Australia, Mexico, Brazil, Chad, Uganda, Sudan, Bulgaria, and Spain.

Identification of Seed Germination Relative Traits and Statistical Analysis

After the cotton seeds were delinted by sulfuric acid, 400 full and uniform seeds were chosen from each accession. Four replicates with 100 seeds each were used for each accession. The seeds were placed evenly into a germination chamber containing 800 g dry quartz sand, then covered the seeds with 250 g of dry quartz sand, finally added 250 mL double-distilled H₂O. The germination chambers were placed in a culture room with 25/20°C temperature and 16/8 h light/dark regime during the period of seed germination. Seeds were considered as germinated when the radicle broke through the seed coat. Seedlings were considered to be established when the root reached half of the seed length. We counted the germinated seeds from the 3rd to the 7th day. Germination potential (GP) refers to the ratio of the number of normal germination seeds to the number of tested seeds in the initial stage of seed germination, usually specified as 3 days, that is, the germination rate (%) of the initial count (Yang et al., 2019; Yuan et al., 2019). GP indicates the speed of germination and the strength of seed vigor. Germination rate (GRA) refers to the proportion of all normal germinated seeds to the number of tested seeds at the end of the germination test, usually specified as 7 days (Yang et al., 2019; Yuan et al., 2019). A high seed germination rate means that there are more viable seeds and more seedlings emergence after sowing. Germination index (GI) represents the sum of the ratio of the number of germinated seeds per day to the corresponding germinating days. GI is calculated based on the formula: $GI = \sum(Gt/Dt)$. In the formula, Dt refers to the number of days to germinate; Gt is the number of seeds germinated per day corresponding to Dt (Yuan et al., 2019). Vigor index (VI) is a comprehensive reflection of seed germination rate and growth, $VI = GI \times S$, here, S refers to the length (cm) or weight (g) of normal seedlings in a certain period (Yang et al., 2019; Yuan et al., 2019). In this study, S is calculated by mass. Statistical analysis of the GP, GRA, GI, and VI were performed with SPSS 20.0. All of the phenotypic data from 419 cotton accessions were used to calculate the frequency distribution of each trait and descriptive statistics.

Genome-Wide Association Study and the Identification of Candidate Genes

Association analysis was performed by the genome-wide efficient mixed model association (GEMMA) package using the following equation: $y = X\alpha + S\beta + K\mu + e$. Here, y represents the phenotype; α and β are fixed effects, representing marker and non-marker effects, respectively; μ represents unknown random

effects; and X , S , and K are the incidence matrices for α , β , and μ , respectively; and e is the vector of random residual effects. The top three PCs were used to build up the S matrix for population-structure correction. The matrix of simple matching coefficients was used to build up the K matrix. The analyses were implemented in the GEMMA software package (Zhou and Stephens, 2012). According to the Bonferroni correction principle, $-\log_{10}(P) > 6.59$ ($P = 1/n$, n is the number of SNPs in this study) is too stringent that we could not find the significant SNPs for four traits with this threshold. Thus, $-\log_{10}(P) > 5.0$ was used to identify significant SNP markers with seed germination related traits (Song et al., 2019; Abdelraheem et al., 2021). We identified candidate genes within 300-kb flanking significant-associated SNP loci as putative candidates based on the decay of LD (Ma et al., 2021) and conducted gene annotation based on *G. hirsutum* TM-1 genome (Zhang et al., 2015). LD block identification was performed for associated SNPs using Haploview 4.2 software (Bashir et al., 2018). Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were carried out for all candidate genes (Xie et al., 2011).

Expression Profile of Candidate Genes

To screen the possible candidate genes involving in the seed germination, the expression level of these genes were analyzed based on the transcriptomic data from the seeds soaking in water for 0, 5, and 10 h. Transcriptome data of gene expression in different tissues and different germination periods was obtained from a previous study (Zhang et al., 2015) and the fragments per kilobase of transcript per million mapped reads (FPKM) values of putative candidate genes were extracted for comparison between different tissues and different germination periods. The heat maps were generated with HemI version 1.0.¹

In order to verify the expression trend of the candidate genes, two cotton varieties with fast germination (Jinmian2) and slow germination (Qunkemian) were chosen for expression analysis with qRT-PCR. The seeds of selected varieties were spread evenly on the sterilized gauze in the germination chamber, sprayed enough sterile water to keep seeds moist during the whole germination period. Imbibitional seeds were taken at 0, 6, 12, 24, 36, 48, 60, and 72 h, respectively. The samples were immediately frozen in liquid nitrogen and stored at -80°C for RNA extraction with three replicates. Total RNA was extracted with an EASYspin Plus Plant RNA purification kit (Aidlab, Beijing, China). Total cDNA was synthesized with the PrimerScriptTM RT Reagent Kit together with gDNA Eraser (TaKaRa, Dalian, China). Quantitative real-time PCR was performed with Auge GreenTM qPCR Master Mix (US EVERBRIGHT RINC) on an ABI 7500 Real-Time PCR machine. The qRT-PCR mixtures consisted of 10 μl of AugeGreenTM Master Mix (US EVERBRIGHT RINC), 2.0 μL of ROX reference dye, 2.0 μL of cDNA, 1.0 μL of primers, and ddH₂O supplemented the volume to 20 μL . The reactions were amplified at 95°C for 30 s, followed by 40 cycles of 95°C for 5 s, 55°C for 30 s, and 72°C for 30 s. All the reactions were performed as three technical replicates. Relative gene expression

¹<http://hemi.biocuckoo.org/>

levels were calculated with the $2^{-\Delta\Delta CT}$ method. The primers used are listed in **Supplementary Table 1**.

Gene Cloning and Plant Transformation

The open reading frame (ORF) of *Gh_A11G0176* was obtained through PCR using cDNAs synthesized from RNA, the amplified products were subsequently cloned into the pGreen vector which was driven by the cauliflower mosaic virus (CaMV) 35S promoter. Using the floral dip method (Clough and Bent, 1998), the recombinant vector was transformed into *Arabidopsis thaliana* Columbia type by *Agrobacterium tumefaciens* GV3101. Transgenic plants were obtained by screening successive generations on Basta. PCR was utilized to identify homozygous T₃ transgenic lines, which were subsequently employed in further investigations. For the seed germination assay, thirty plump seeds were surface-sterilized for 5 min in 30% NaClO and 1 min in 75% ethanol, washed at least five times with sterile water, plated on Murashige and Skoog (MS) solid medium (with 1% sucrose) and stratified at 4°C for 48 h, then grown at 22°C under long-day conditions (16/8 h light/dark). For each germination assay, biological triplicates were performed, and germination ability was observed every 12 h until all seeds germinated after 24 h.

In addition, another target gene *Gh_A09G1509* was selected to validate with overexpressing transgenic tobacco which was previously acquired by our research group (Li et al., 2019). Approximately 50 plump seeds each from the wild-type (WT) and over-expression (OE) tobacco were surfaced sterilized and plated on MS media (pH 5.7–5.9) at 28°C in the incubator under long-day conditions (16/8 h light/dark), with independent biological triplicates. Germination ability was observed every 24 h until all seeds germinated after 2 days, and the method was the same as above.

Measurement for Hormone Content

Gibberellin (GA) and abscisic acid (ABA) were extracted from 0.1 g germinating seeds using 1 mL precooling reagent one (methanol: water: acetic acid = 80: 20: 1) overnight at 4°C and centrifuged at 8,000 g at 4°C for 10 min. The residue was collected and added 0.5 mL reagent one (methanol: water: acetic acid = 80: 20: 1) for 2 h, then the supernatant was collected after centrifugation. Two supernatants were combined for drying with nitrogen at 40°C until no organic phase remained. The supernatants were added 0.5 mL reagent two (petroleum ether) to extract and decolorize three times at 60–90°C, the upper ether phase was discarded, and the lower aqueous was adjusted to pH 2.8 with reagent three (saturated citric acid aqueous solution). Next, the mixture was extracted three times with equal volume of reagent four (ethyl acetate). After the organic phase with nitrogen was blown dry, the extraction was diluted to 0.5 mL with reagent five (methanol) through vortex oscillation, and filtered by 0.22- μ m membrane filter. A high-performance liquid chromatography (HPLC) device was used to analyze the final filtrate solution. The content of GA and ABA were determined at 210 and 254 nm, respectively, and was expressed as μ g/g fresh weight (FW).

H₂O₂ Extraction and Analysis

The H₂O₂ level was determined using commercial assay kits according to the manufacturer's instructions (Suzhou Keming Bioengineering Company, China). Approximately 0.1 g FW of each sample was quickly put into pre-cooled acetone (4°C) and homogenized on ice bath. The reaction solutions were then mixed into the homogenate. The mixture was centrifuged at 8,000 g at 4°C for 10 min, and the absorbance of the supernatant was measured at 415 nm right away. The H₂O₂ content was expressed as μ mol/g FW.

Protein, Glucose, and Amylase Activity Assays

Protein, glucose, and amylase activity were measured using commercial assay kits following the manufacturer's instructions (Suzhou Keming Bioengineering Company, China). The levels of protein and glucose were expressed as mg/g FW. One unit (U) of amylase is defined as 1 mg of reducing sugar produced by enzyme in 1 g FW of the sample in 1 min at 40°C. The activity of α -amylase were expressed as U/g FW.

Extraction and Measurement of Glutathione and Enzymes Related to Glutathione Metabolism

Glutathione, oxidized glutathione (CSSG), glutathione peroxidase (GPX), glutathione reductase (GRE), and GST were measured using commercial assay kits following the manufacturer's instructions (Suzhou Keming Bioengineering Company, China). The levels of GSH and CSSG were expressed as μ mol/g FW and nmol/g FW, respectively.

Glutathione reductase activity was determined following the rate of NADPH oxidation at 340 nm. The GPX activity was calculated by measuring the rate of disappearance of NADPH at 340 nm. One unit (U) of GRE activity and GPX activity was defined as each gram of sample catalyzed the oxidation of 1 nmol NADPH per minute. GST activity was calculated by measuring the increase in absorbance at 340 nm. One unit (U) of GST activity was defined as each gram of sample catalyzed the combination of 1 nmol/L CDNB and GSH per minute. The activities of GRE activity, GPX activity, and GST activity were expressed as U/g FW.

RESULTS

Phenotypic Variation of Seed Germination Related Traits

We analyzed the phenotypic variation of seed germination relevant traits including GP, GRA, GI, and VI. The results showed that all the traits displayed broad variations. The GP ranged from 4.00 to 60.00% with an average of 25.41%, the GRA ranged from 35.00 to 94.00% with an average of 75.04%, the GI ranged from 7.52 to 20.41 with an average of 15.55, and the VI ranged from 3.44 to 17.21 with an average of 9.85. The coefficient of variation (CV) of GP, GRA, GI, and VI were 47.82, 11.75, 13.21, and 18.98%, respectively (**Table 1**). High correlations were

observed among these seed germination traits (**Figure 1**). GP was significantly ($P < 0.001$) and positively correlated with GRA ($r = 0.232^{**}$), GI ($r = 0.599^{**}$) and VI ($r = 0.449^{**}$), and GRA was significantly positively correlated with GI ($r = 0.910^{**}$) and VI ($r = 0.677^{**}$), which may facilitate the identification of pleiotropic gene in response to seed germination. Moreover, the phenotypic distribution of these traits displayed continuous variation (**Figure 1**), indicating that seed germination related traits were quantitatively inherited.

Identification of Significantly Associated Single Nucleotide Polymorphisms and Candidate Genes Related to Seed Germination Traits

Genome-wide association study results showed that 92 significantly associated SNPs were identified, of which, 10 and 35, 24 and 36 were associated with GP, GRA, GI, and VI, respectively, and distributed on 20 chromosomes (**Figures 2A–D**). There were 69, 10, and 13 SNPs in the A-subgenome, D-subgenome, and scaffolds, respectively (**Supplementary Table 2**). Among them, the maximum number of associated SNPs was detected on chromosome A09 (20), and no significant SNPs were detected on chromosome D01, D03, D06, D08, D09, or D10. In addition, 13 associated SNPs located on chromosomes A01, A05, A08, A09, A10, and D13 were observed in multiple traits (**Supplementary Table 3**). Among these, four SNPs on chromosome A01 were found to be significantly associated with GI and GRA, two SNPs on chromosome A08 with GI and GRA, four SNPs on chromosome A09 with GRA and VI, and three SNPs on chromosome A05, A10, and D13 with GRA and GI.

To explore the potential candidate genes related to seed germination, we selected a 300-kb region flanking significant SNPs according to the linkage decay value (Ma et al., 2018). A total of 723 candidate genes were identified, of which, 569 and 154 were located on the A-subgenome and D-subgenome, respectively (**Supplementary Table 4**). Their expression were analyzed on the basis of the transcriptome data from cottonFGD (**Supplementary Figures 1, 2**). For the A-subgenome, the expressions patterns of all the genes could be divided into five types according to the expression specificity from different tissues and different germination periods. Type I were mainly expressed in seed, and their expression gradually decreased with the extension of seed imbibition time, such as *Gh_A11G0176*, *Gh_A11G0179*, *Gh_A11G0184*, and *Gh_A09G1510*. Type II displayed higher expression in root than in seed. Type III was specifically expressed in cotyledon. Type IV was less expressed in seed, but their expressions gradually increased with the extension of seed imbibition time, such as *Gh_A09G1508* and *Gh_A09G1509*. Type V was mainly expressed in cotyledon and seed, and the expression gradually decreased with the extension of seed imbibition time, such as *Gh_A11G0177* (**Supplementary Figure 1**). Similarly, the gene expressions in the D-subgenome were also divided into five types the same as to those in the A-subgenome (**Supplementary Figure 2**). KEGG pathway analysis was performed to display the top 20 significantly enriched pathway, and thereinto, glutathione

metabolism (KO00480) and circadian rhythm (KO04712) are related to the multifunction of GST and seed germination (**Figure 2E**). GO enrichment analysis was conducted to further infer the functions of candidate genes (**Figure 2F**). At $P < 0.05$ and gene number $> 3,723$ candidate genes were classified into three major categories: biological process (BP), molecular function (MF), and cell component (CC). In the biological process, transport (GO: 0006810), glutathione metabolic process (GO: 0006749), and toxin catabolic process (GO: 0009407) were the most functional terms associated with seed germination. In the molecular function, protein binding (GO: 0005515), dioxygenase activity (GO: 0051213), and glutathione transferase activity (GO: 0004364) were the significantly enriched items. In the cellular component (CC) category, chloroplast envelope (GO: 0009941), and cytoplasm (GO: 0005737) were the two most prevalent functional terms. GRE was mainly distributed in chloroplast but also in cytoplasm, which related to the stress tolerance and seed germination (Ding et al., 2009; Gill and Tuteja, 2011).

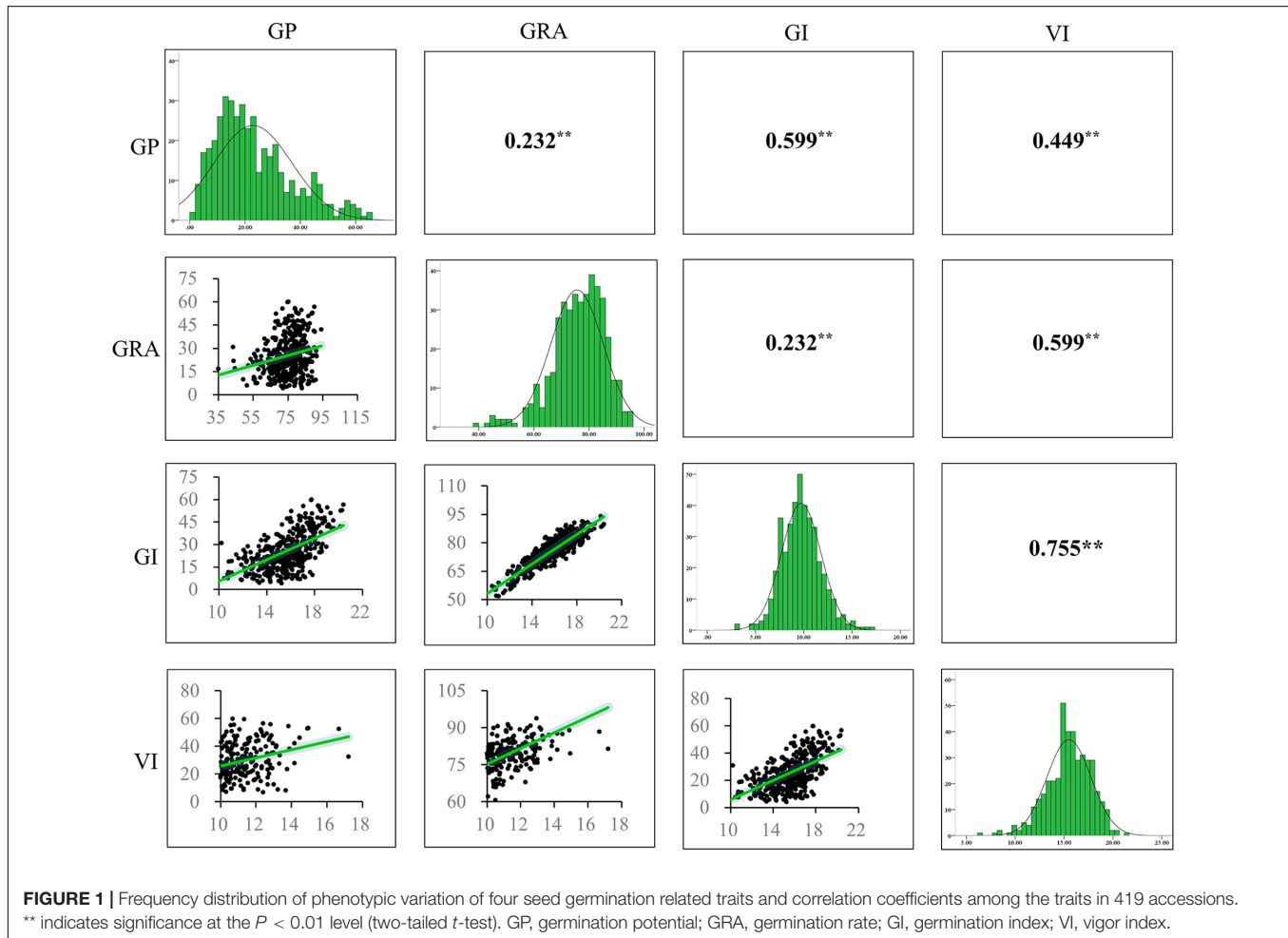
Functional Analysis of Candidate Genes

On chromosome A11, we focused on the locus mapped from 1.4 to 2.0 Mb, where a locus (A11:1681419) was significantly associated with GRA (**Figure 3A**), and the two genotypes of the locus showed significant difference (**Figure 3B**). A total of 135 candidate genes were identified within the LD region. We focused on the expression of 23 candidate genes in the 50-kb region flanking the significant SNP (**Figure 3C**). These genes were involved in heat shock protein 70 (Hsp 70) family protein, translation elongation factor EFG/EF2 protein, thioredoxin family protein, AT motif nuclear localization protein, and other proteins that performed molecular functions. It can be seen from the heat map that the expression levels of *Gh_A11G0176*, *Gh_A11G0177*, *Gh_A11G0179*, and *Gh_A11G0184* gradually decreased with the extension of seed imbibition time (**Figure 3C**). Their expression patterns were analyzed *via* qRT-PCR using a fast germinating and a slow germinating variety at 0, 6, 12, 18, 24, 36, 48, and 72 h after imbibition (HAI). Compared with the fast germinating variety, *Gh_A11G0176* and *Gh_A11G0177* displayed a higher expression in the slow germinating variety (**Figures 3D,E**). However, *Gh_A11G0179* and *Gh_A11G0184* exhibited higher expression in the fast germinating varieties (**Figures 3F,G**). To understand the function of target gene *Gh_A11G0176* encoding Hsp70-Hsp90 organizing protein 3, we overexpressed the gene in Arabidopsis and obtained the homozygous lines. The OE Arabidopsis lines and WT were put on MS medium for germination assays. The results showed that the germination rates of OE were markedly reduced compared with WT during germination (**Figures 3H,I**), that is, the OE of *Gh_A11G0176* caused delayed germination. At 24 HAI, the germination rate of WT (78%) was approximately 4.3 times that of OE-3 (18%). At 36 HAI, most of the OE and WT have already germinated, and the germination rates of WT and OE were almost the same.

On chromosome A09, we focused on the locus mapped from 68.0 to 68.5 Mb with significant signals, of which SNP A09:68240653 was significantly associated with GRA and VI

TABLE 1 | Phenotypic variation statistics for seed germination traits.

| Variable | Minimum | Maximum | Mean | Std. deviation | CV% | Skewness | Kurtosis |
|-------------------------------|---------|---------|-------|----------------|-------|----------|----------|
| GP, germination potential (%) | 4.00 | 60.00 | 25.41 | 12.15 | 47.82 | 0.597 | -0.205 |
| GRA, germination rate (%) | 35.00 | 94.00 | 75.04 | 8.82 | 11.75 | -0.865 | 1.372 |
| GI, germination index | 7.52 | 20.41 | 15.55 | 2.05 | 13.18 | -0.527 | 0.467 |
| VI, vigor index | 3.44 | 17.21 | 9.85 | 1.87 | 18.98 | 0.247 | 0.664 |

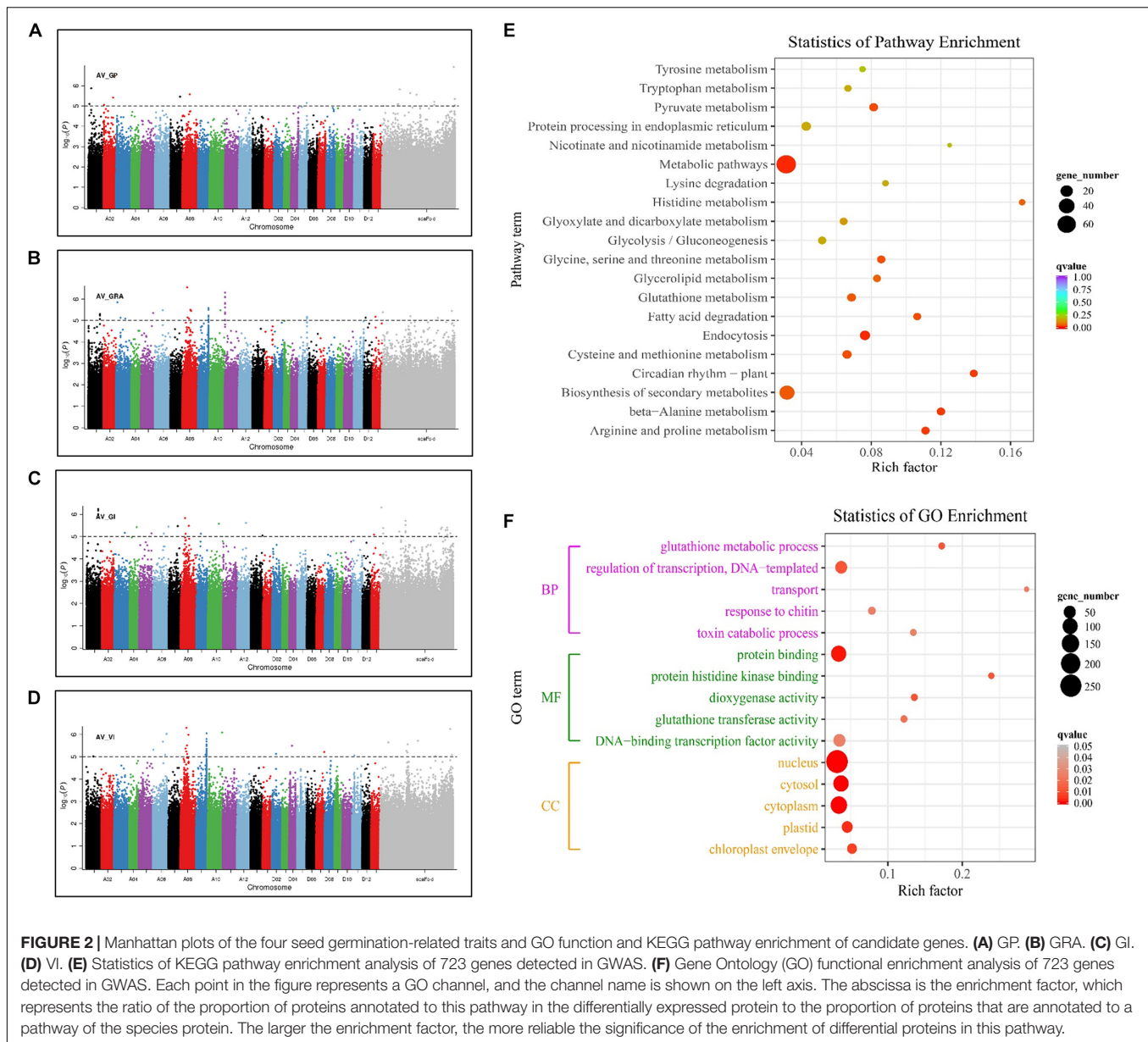


(Figures 4A,B). Ninety-five candidate genes were identified (Supplementary Table 4). The accessions carrying the alternate genotype germinated faster than those with the reference (Figure 4C). We analyzed the expression of 23 candidate genes within the 50-kb region flanking the significant SNP and found that glutathione transferase encoding genes, *Gh_A09G1508* and *Gh_A09G1509*, displayed an increased trend with the extension of seed imbibition time excluding *Gh_A09G1510* (Figure 4D). We further analyzed the expression of three genes using a fast germinating and a slow germinating variety at 0, 6, 12, 18, 24, 36, 48 and 72 HAI by qRT-PCR. The results showed that the expression of *Gh_A09G1508* and *Gh_A09G1509* was higher in the fast germinating variety, however, *Gh_A09G1510* was higher in the slow germinating variety (Figures 4E–G). Moreover, the OE of *Gh_A09G1509* in tobacco resulted in

faster seed germination and seedling growth as well as longer hypocotyls at 3 days after imbibition (DAI) compared with WT. The cotyledons of transgenic tobacco unfolded earlier than WT at 4 DAI (Figures 4H,I), indicating that *Gh_A09G1509* promoted the plant seed germination.

Effects of *Gh_A09G1509* on Endogenous Phytohormones, H_2O_2 , Starch Mobilization, and Soluble Sugar Content During Germination

Abscisic acid and gibberellin are the key endogenous substances that work antagonistically in the regulation of seed germination. Furthermore, H_2O_2 in seeds as a signal can promote germination and seedling growth (Barba-Espin et al., 2010;

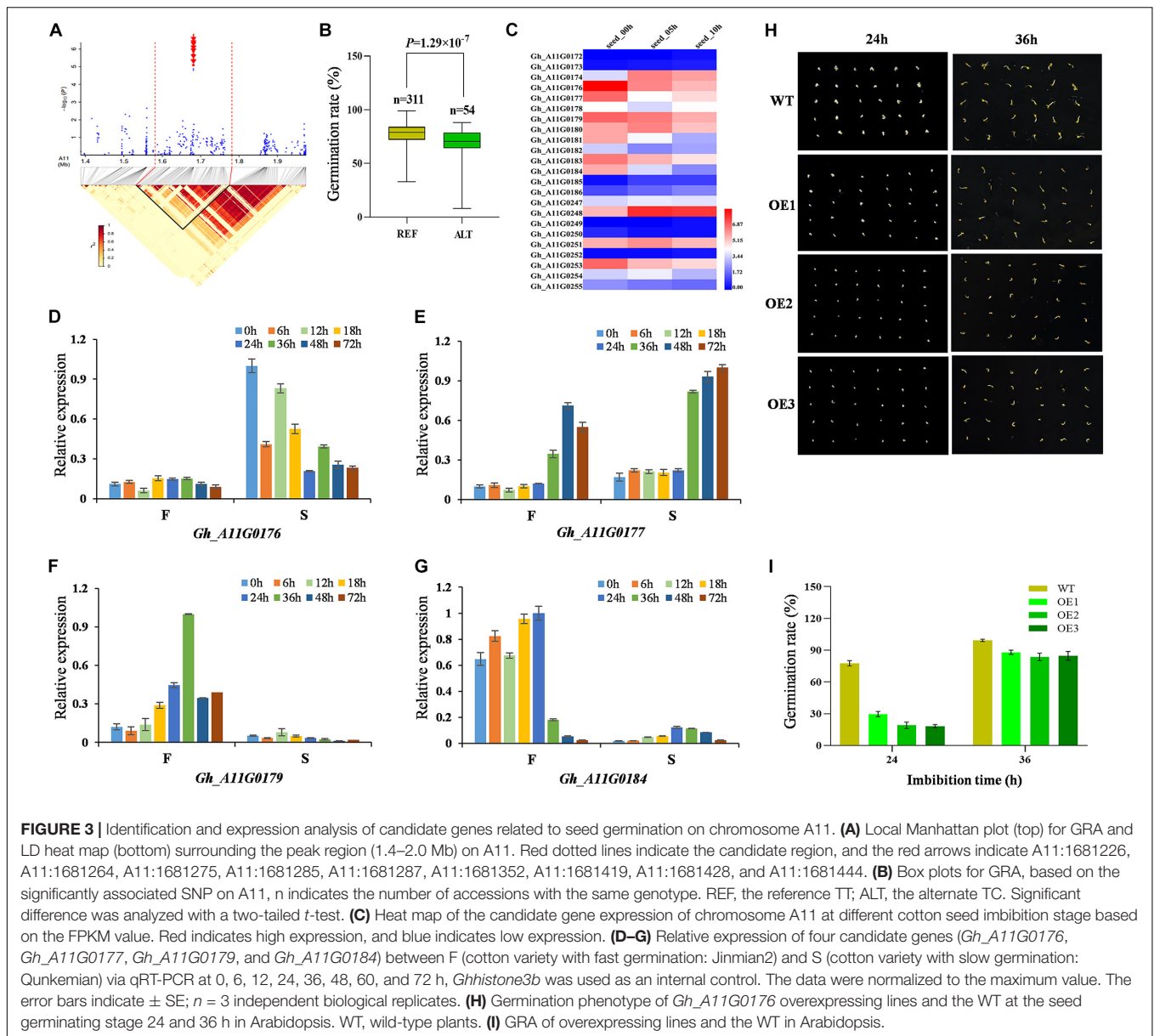


Katsuya-Gaviria et al., 2020), which was induced by GA but suppressed by ABA (Ishibashi et al., 2012). Therefore, endogenous GA₃, ABA, and H₂O₂ contents were measured during seed germination. Compared with WT, the OE transgenic tobacco showed significantly higher GA₃ contents (Figure 5A). For ABA contents, it was significantly lower in the transgenic tobacco at 48 HAI (Figure 5B). At 48 HAI, the hypocotyl of the tobacco broken through the seed coat and reached to protrusion. As similar to GA₃, the endogenous H₂O₂ was significantly higher at 48 HAI in transgenic tobacco than WT (Figure 5C). Imbibition and starch hydrolysis are the critical steps during seed germination. We further compared the changes of α-amylase and glucose contents between the transgenic tobacco and the WT. The content of α-amylase was significantly higher in the OE lines at 48 HAI (Figure 5D). Furthermore, we observed

that the content of glucose appeared an apparent increase at 48 HAI (Figure 5E). These results indicated that *Gh_A09G1509* regulated seed vigor through adjusting the relative contents of endogenous phytohormones and altering starch hydrolysis and glucose contents in germinating seeds.

Changes of Substances Related to Glutathione Metabolism During Germination

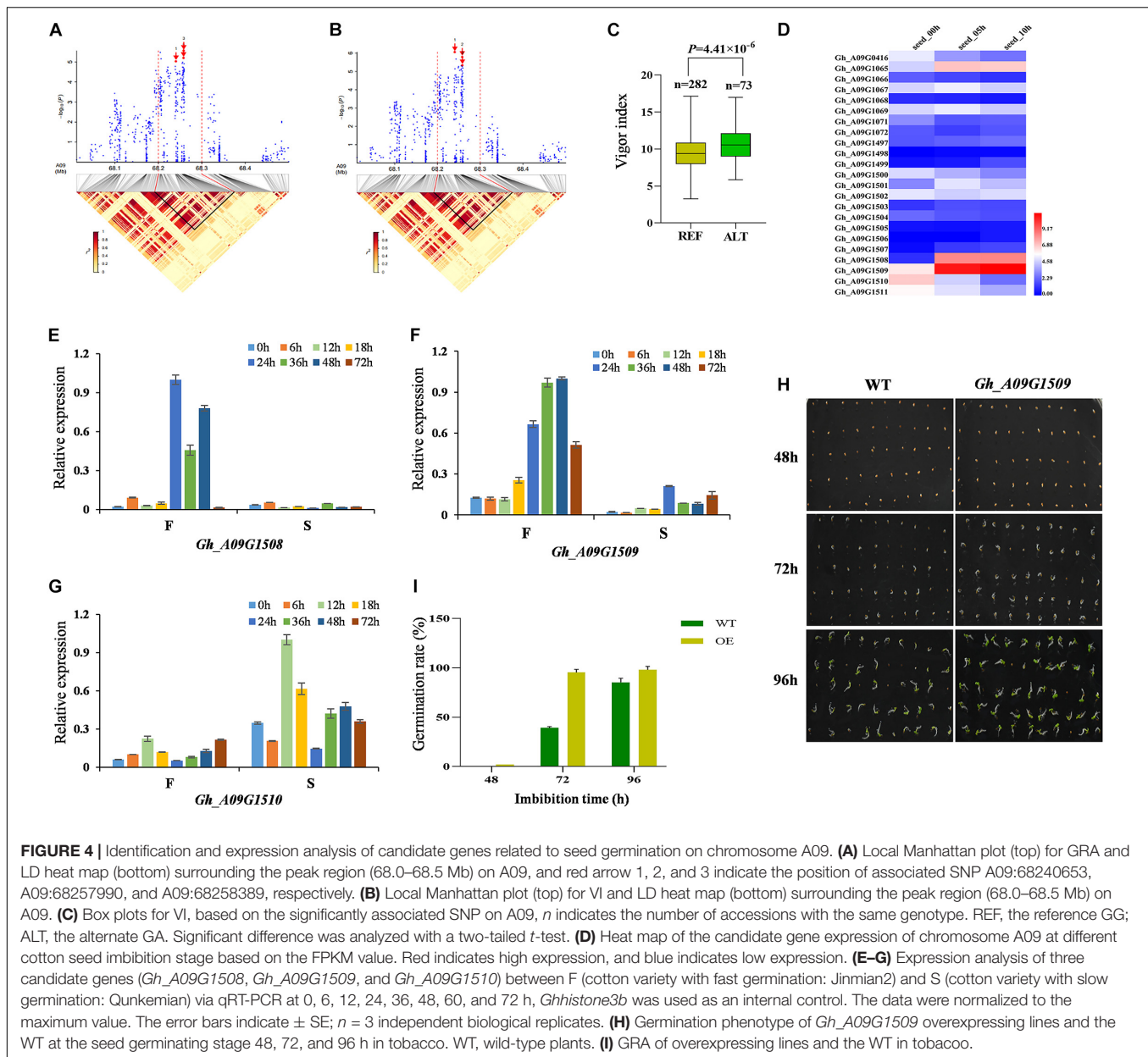
Doubled glutathione improved the germination capacity (Kalemba and Ratajczak, 2018), and GRE and GPX regulated the balance between the reduced and oxidized forms of glutathione in plants (Zhao et al., 2021), which is crucial for cellular redox state homeostasis and plant development



(Hernandez et al., 2015). Since *Gh_A09G1509* belonged to *GST* gene family, the related substances of glutathione metabolism were measured during germination. There was no significant difference in GRE between the OE tobacco plants and the WT (Figure 5F), but the activities of GPX and GST decreased during germination, and the activities of GPX and GST in the OE plants were significantly lower than those in WT plants at 48 HAI (Figures 5G,H), when the hypocotyl broken through the seed coat. Compared with WT, the content of GSH and GSSG were also significantly higher in transgenic tobacco at 48 HAI (Figures 5I,J). Similarly, the ratio of GSH/GSSG in the OE tobacco plants was significantly higher during the whole germination period, especially at 48 HAI (Figure 5K). The results suggested that *Gh_A09G1509* might facilitate glutathione metabolism during seed germination.

DISCUSSION

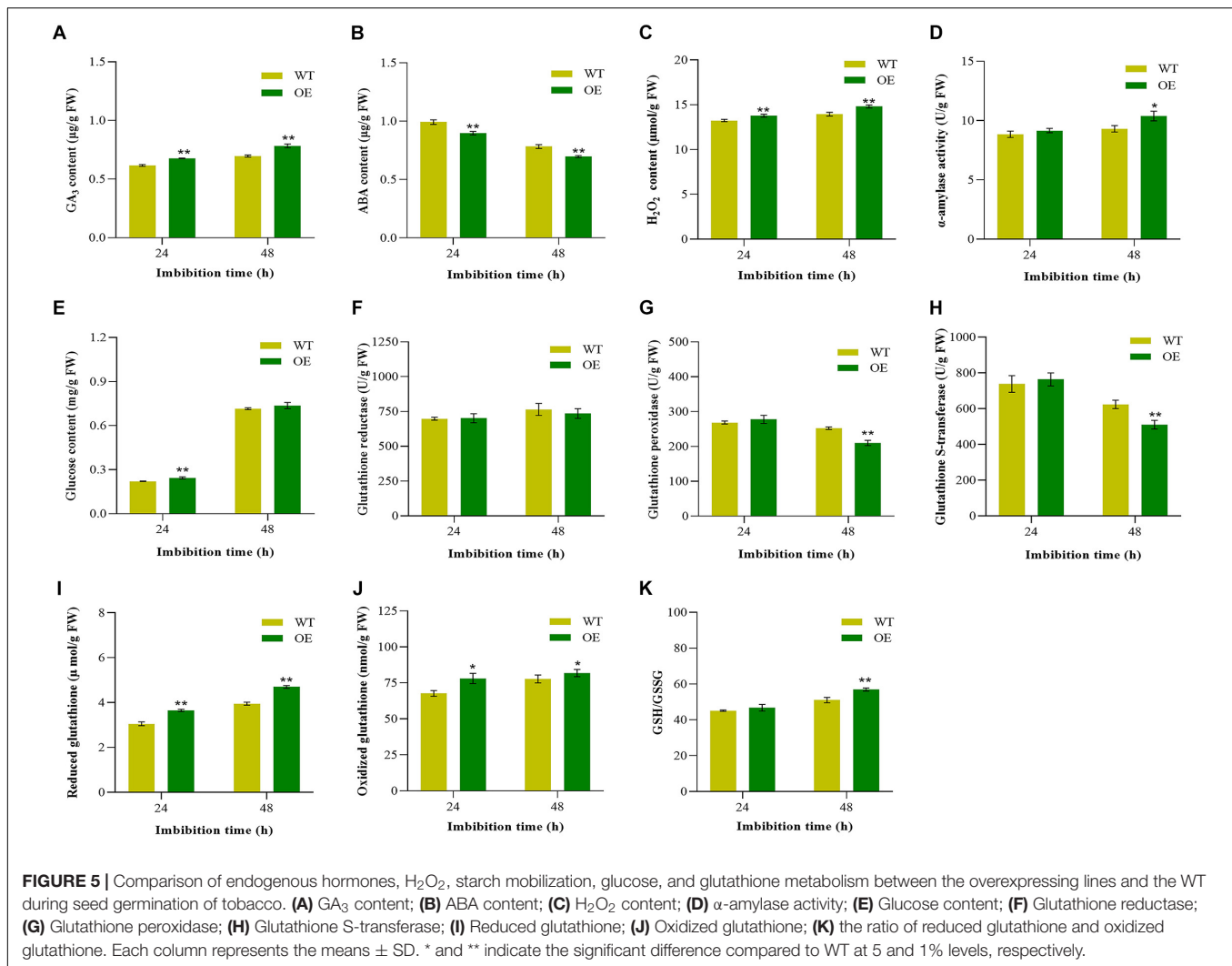
Rapid and uniform seed germination under various conditions is an agronomically important trait for high yield of crops (Foolad et al., 2007). Seeds with high vigor have obvious production advantages and potential, thus, it is critical to identify and utilize seed vigor related genes for improving seed vigor (He et al., 2019). The development of seedlings with enhanced vigor could allow for earlier planting, extended agricultural growing seasons for crops and the expansion of crop production in marginal locations (Roxas et al., 1997). In the present study, a total of 92 significantly associated SNPs located within or near 723 genes were identified for seed germination traits on the basis of a core collection containing 419 cotton accessions with 3.66 million high-quality SNPs. More associated SNPs were



located in the A-subgenome than D-subgenome. The candidate genes were involved in various metabolic pathways, including toxin metabolism, toxin decomposition, glutathione metabolism, and circadian rhythms with plants. Some of these candidate genes related to germination have been identified in rice and other plant species (Guo et al., 2011; Kaur et al., 2015; Wang et al., 2015). The results provide new insights into the genetic basis of seed germination vigor and molecular tools for crop improvement in cotton.

On chromosome A09 and A11, *Gh_A09G1499*, *Gh_A09G1610*, *Gh_A11G0179*, and *Gh_A11G0252* belonged to the protein kinase superfamily which played fundamental roles in the modulation of plant growth and development,

including seed development, dormancy and germination, seedling and root growth, flowering, fruit development, and ripening and leaf senescence (Halford and Hey, 2009; Wang et al., 2020). Overexpression of the receptor protein kinase gene *ZmRLK7* in Arabidopsis reduced the plant height, organ size (e.g., petals, silique, and seeds) and seed weight (He et al., 2020). The kinase-associated protein phosphatase (KAPP) in Arabidopsis was negatively involved in ABA-mediated seed germination and early seedling growth (Lu et al., 2020). *SnRK2* regulated key traits of crop improvement and production such as seed maturation and germination via ABA-dependent or ABA-independent pathways (Mao et al., 2020).



Gh_A11G0180 and *Gh_A09G1503* encoded thioredoxin (Trx) family proteins. During seed germination, Trx acts as a signal in early germination to promote amylase and proteinase activities and initiated the hydrolysis of storage materials, thereby promoting seed germination (Smiri et al., 2010; Guo et al., 2011). Trx has been observed to accelerate germination and α-amylase synthesis in some transgenic studies (Cho et al., 1999; Wong et al., 2002; Kim et al., 2003). Trx *h* present in starchy endosperm functions in germination and early seedling development in cereals (Jiao et al., 1992; Kobrehel et al., 1992; Besse et al., 1996; Bewley, 1997; Cho et al., 1999; Wong et al., 2003). Overexpression of *Trx* in barley (*Hordeum vulgare*) endosperm accelerated germination (Wong et al., 2002), and suppressed expression of *Trx* in wheat (*Triticum aestivum*) inhibited germination (Guo et al., 2007).

Gh_A09G1506 encoding a seed storage protein (SSP) was deposited in the protein bodies of developing seeds and subsequently utilized during the germination of plant as a source of nitrogen and carbon (Kawakatsu et al., 2010). In peanut, embryonic properties could be suppressed via

repression of *SSP* genes during germination (Yang et al., 2015). Ectopic expression analysis of *PtCP5* showed decreased storage protein accumulation, delayed seed germination, and seedling development in *OX-PtCP5* transgenic *Arabidopsis* (Liu et al., 2021). Research on rice mutants and transgenic complementary mutants suggested that *OsTudor-SN* functioned in post-transcriptional regulation of storage protein expression and seed development (Chou et al., 2019).

Four candidate genes, *Gh_A11G0171*, *Gh_A11G0174*, *Gh_A11G0176*, and *Gh_A11G0181*, belonged to the heat shock protein (HSP) family that responded to abiotic stress and protected plants from adverse environmental effects (Charng et al., 2006; Sun et al., 2012). Recent studies showed that HSP played an important role in seed germination. Overexpression of *ZmHSP16.9* in transgenic tobacco increased seed germination rate (Sun et al., 2012), and overexpression of *CaHsp25.9* in *A. thaliana* resulted in increased germination and root length under abiotic stress (Feng et al., 2019). *CsHSP* transformed plants improved seed germination vigor under heat stress (Wang et al., 2017), and *GhHSP24.7* mediated seed germination via

thermal sensing. Under rugged environmental conditions, the OsHSP18.2 positively controlled the germination and cotyledon emergence (Kaur et al., 2015). In the present study, we validated the function of *Gh_A11G0176*, and the OE Arabidopsis lines showed delayed germination compared with the WT, indicating *Gh_A11G0176* might play a negative role during seed germination.

Several GST genes, including *Gh_A09G1508*, *Gh_A09G1509*, *Gh_A09G1510*, *Gh_A11G0199*, and *Gh_A11G0200* were identified in the present study. GSTs are versatile enzymes and catalyze the conjugation of electrophilic substrates to GSH and thus reduce their toxicity (Frova, 2006). In addition to glutathione transferase activity, some GSTs were found to possess glutathione peroxidase activity (Bartling et al., 1993; Xu et al., 2016) or be involved in light-dependent pathways and circadian rhythm changes (Galle et al., 2018). GSTs were activated by a variety of environmental stimuli and were found to perform a direct function in lowering oxidative damage and hazardous compounds produced during xenobiotic metabolism (Dixon et al., 2002; Moons, 2005; Frova, 2006). Seed germination and seedling growth were significantly improved in transgenic tobacco lines that overexpressed plant GST/GPX under stressful conditions (Roxas et al., 1997), while overexpression of *Gh_A09G1509* in tobacco resulted in enhanced Verticillium wilt resistance (Li et al., 2019). In the present study, *Gh_A09G1509* was observed to have a higher expression level in fast germinating varieties than in slow germinating varieties, which was consistent with the transcriptome results. Moreover, *Gh_A09G1509* OE tobacco resulted in faster seed germination and seedling growth as well as longer hypocotyls, indicating that *Gh_A09G1509* played a role in the positive regulation of promoting seed germination.

Seed germination is a complex trait that not only affected by temperature and environment, but also by endogenous hormones, such as ABA and GA. The induction and maintenance of dormancy are favorably regulated by ABA, while germination is enhanced by GA. The time of germination depends on the balance and the physiological interaction between ABA and GA (Savage and Metzger, 2006). Researches in rice showed that *OsIPMS1* could promote the synthesis of GA₃ biosynthesis-related amino acids in germinated seeds, which resulted in an increase in the amount of soluble sugars available for glycolysis during seed germination. And then, the tricarboxylic acid cycle (TCA) will be boosted, resulting in increased glycolysis and TCA cycle metabolites, which contributed to quick seed germination and strong seedling growth (He et al., 2019). *OsMFT2* positively regulated ABA response genes through interacting with *OsbZIP23/66/72* and negatively regulated seed germination in rice (Song et al., 2020). In cotton, appropriate melatonin may promote seed germination by regulating the endogenous phytohormones GA₃ and ABA (Xiao et al., 2019). In the present study, it was demonstrated that *Gh_A09G1509* improved seed germination by increasing GA₃ content and decreasing ABA content, in consistent with the literature described on these hormones (Kucera et al., 2005; Tuan et al., 2018). ROS triggered protein carbonization to release

dormancy, more and more evidence subsequently showed that ROS homeostasis was essential for germination (Oracz et al., 2007). Studies barley seeds found that dormant seeds had low ROS content but high ABA content. Seed dormancy and germination may be influenced by changes in the equilibrium between ABA and ROS (Ishibashi et al., 2017). The activated C kinase 1 (RACK1) receptor OsRACK1A positively regulated seed germination by regulating endogenous levels of ABA and ROS, as well as their interplay (Zhang et al., 2014). Additionally, the H₂O₂ accumulation might change the hormone balance by increasing GA_s and decreasing ABA and ethylene, which was crucial for seed dormancy and germination (Barba-Espin et al., 2010, 2011; Jeevan et al., 2015). In the present study, we observed that the H₂O₂ levels in overexpressed plants were significantly higher compared to WT during the whole seed germination stage. At the later stage of germination (48 h), GA₃ induced an increase in the activity of α -amylase which catalyzed the hydrolysis of starch into glucose. Soluble sugars such as glucose, serve as the primary energy source for seed germination (Ding et al., 2012). We speculated that *Gh_A09G1509* may regulate seed germination through GA and reactive oxygen signaling pathways. The interaction between GA and H₂O₂ promoted the hydrolysis of starch, which increased the glucose content and thus promoted seed germination.

Glutathione is not only an essential metabolite of plant life, but also plays an important role in protein biosynthesis of plant cells, the ratio of GSH to CSSG is the important indicator reflecting the activity of glutathione in plants (Cairns et al., 2006). At the early stage of seed germination, a high ratio of GSH to CSSG is necessary for the synthesis of proteins required for growth and development (Fahey et al., 1980). During the germination process, GPX can catalyze GSH to GSSG and reduce toxic peroxides to non-toxic hydroxyl compounds, and GST can catalyze the combination of GSH and toxic substances or peroxides to inactivate them, thereby protecting cells from oxidative damage (Roxas et al., 2000; Kalemba and Ratajczak, 2018), and promoting seed germination (Wu et al., 2020). In silver maple, GSH treatment caused less dehydration and increased germination (Kalemba and Ratajczak, 2018). In the present study, the activity of GRE gradually increased, while the activities of GPX and GST gradually decreased with the germinating of tobacco seeds. Moreover, overexpression of GRE in transgenic plants leads to elevated levels of GSH. Thereby, we speculated that *Gh_A09G1509* might promote seed germination and seedling establishment by glutathione metabolism.

CONCLUSION

In this study, a total of 92 significantly associated SNPs and 294 expressed genes at seed germination stage were screened out. *Gh_A11G0176* might play a negative role while *Gh_A09G1509* play a positive role in regulating the germination of cotton seeds. *Gh_A09G1509* might regulate seed vigor and seedling establishment mainly *via* glutathione metabolism and H₂O₂ level in germinating seeds. This provides a valuable reference for

understanding the molecular mechanism and facilitating crop improvement of seed germination in cotton.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

ZM conceptualized and designed the research. AS, ZL, BC, QG, YZ, LW, and GZ performed the experiments. AS, ZS, and BC analyzed the data. AS wrote the manuscript. ZM and XW revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the Key Research and Development Program of Hebei Province (21326314D), China

REFERENCES

- Abdelraheem, A., Thyssen, G. N., Fang, D. D., Jenkins, J. N., McCarty, J. C., Wedegaertner, T., et al. (2021). GWAS reveals consistent QTL for drought and salt tolerance in a MAGIC population of 550 lines derived from intermating of 11 Upland cotton (*Gossypium hirsutum*) parents. *Mol. Genet. Genomics* 296, 119–129. doi: 10.1007/s00438-020-01733-2
- Bai, Y., Xiao, S., Zhang, C., Zhang, Y., Sun, C., Zhang, X., et al. (2020). Melatonin improves the germination rate of cotton seeds under drought stress by opening pores in the seed coat. *Peer J.* 8:e9450. doi: 10.7717/peerj.9450
- Barba-Espin, G., Diaz-Vivancos, P., Clemente-Moreno, M. J., Albacete, A., Faize, L., Faize, M., et al. (2010). Interaction between hydrogen peroxide and plant hormones during germination and the early growth of pea seedlings. *Plant Cell Environ.* 33, 981–994. doi: 10.1111/j.1365-3040.2010.02120.x
- Barba-Espin, G., Diaz-Vivancos, P., Job, D., Belghazi, M., Job, C., and Hernandez, J. A. (2011). Understanding the role of H₂O₂ during pea seed germination: a combined proteomic and hormone profiling approach. *Plant Cell Environ.* 34, 1907–1919. doi: 10.1111/j.1365-3040.2011.02386.x
- Bartling, D., Radzio, R., Steiner, U., and Weiler, E. W. (1993). A glutathione S-transferase with glutathione-peroxidase activity from *Arabidopsis thaliana*. Molecular cloning and functional characterization. *Eur. J. Biochem.* 216, 579–586. doi: 10.1111/j.1432-1033.1993.tb18177.x
- Bashir, K., Sarwar, R., Saeed, S., Mahjabeen, I., and Kayani, M. A. (2018). Interaction among susceptibility genotypes of *PARP1* SNPs in thyroid carcinoma. *PLoS One* 13:e0199007. doi: 10.1371/journal.pone.0199007
- Basnet, R. K., Duwal, A., Tiwari, D. N., Xiao, D., Monakhos, S., Bucher, J., et al. (2015). Quantitative trait locus analysis of seed germination and seedling vigor in *Brassica rapa* reveals QTL hotspots and epistatic interactions. *Front. Plant Sci.* 6:1032. doi: 10.3389/fpls.2015.01032
- Besse, I., Wong, J. H., Kobrehel, K., and Buchanan, B. B. (1996). Thiocalcin: a thioredoxin-linked, substrate-specific protease dependent on calcium. *Proc. Natl. Acad. Sci. U.S.A.* 93, 3169–3175. doi: 10.1073/pnas.93.8.3169
- Bewley, J. D. (1997). Seed germination and dormancy. *Plant Cell* 9, 1055–1066. doi: 10.1105/tpc.9.7.1055
- Cairns, N. G., Pasternak, M., Wachter, A., Cobbett, C. S., and Meyer, A. J. (2006). Maturation of *Arabidopsis* seeds is dependent on glutathione biosynthesis within the embryo. *Plant Physiol.* 141, 446–455. doi: 10.1104/pp.106.077982

Agricultural Research System (CARS-15-03), and Hebei Top Talent Project (031601801).

ACKNOWLEDGMENTS

We greatly appreciate the active contributions of all participated students and reviewers for their precious work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.844946/full#supplementary-material>

Supplementary Figure 1 | Heat map of the expression of candidate genes in different tissues, including cotyledon, root, and seed at different time periods on the A-subgenome.

Supplementary Figure 2 | Heat map of the expression of candidate genes in different tissues, including cotyledon, root, and seed at different time periods on the D-subgenome.

- Chang, Y. Y., Liu, H. C., Liu, N. Y., Hsu, F. C., and Ko, S. S. (2006). *Arabidopsis* Hsa32, a novel heat shock protein, is essential for acquired thermotolerance during long recovery after acclimation. *Plant Physiol.* 140, 1297–1305. doi: 10.1104/pp.105.074898
- Chen, D. (2012). Discussion on the importance of seed germination potential to crop emergence rate in field. *Chinese Seed Industry* 3:2. doi: 10.19462/j.cnki.1671-895x.2012.03.025
- Chen, H., Ruan, J., Chu, P., Fu, W., Liang, Z., Li, Y., et al. (2019). *AtPER1* enhances primary seed dormancy and reduces seed germination by suppressing the ABA catabolism and GA biosynthesis in *Arabidopsis* seeds. *Plant J.* 101, 310–323. doi: 10.1111/tj.14542
- Chen, J. H., Jiang, H. W., Hsieh, E. J., Chen, H. Y., Chien, C. T., Hsieh, H. L., et al. (2012). Drought and salt stress tolerance of an *Arabidopsis* glutathione S-transferase U17 knockout mutant are attributed to the combined effect of glutathione and abscisic acid. *Plant Physiol.* 158, 340–351. doi: 10.1104/pp.111.181875
- Chen, L., Liu, L., Lu, B., Ma, T., Jiang, D., Li, J., et al. (2020). Exogenous melatonin promotes seed germination and osmotic regulation under salt stress in cotton (*Gossypium hirsutum* L.). *PLoS One* 15:e0228241. doi: 10.1371/journal.pone.0228241
- Cho, M. J., Wong, J. H., Marx, C., Jiang, W., Lemaux, P. G., and Buchanan, B. B. (1999). Overexpression of thioredoxin *h* leads to enhanced activity of starch debranching enzyme (pullulanase) in barley grain. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14641–14646. doi: 10.1073/pnas.96.25.14641
- Chou, H. L., Tian, L., Fukuda, M., Kumamaru, T., and Okita, T. W. (2019). The role of RNA-binding protein *OsTudor-SN* in post-transcriptional regulation of seed storage proteins and endosperm development. *Plant Cell Physiol.* 60, 2193–2205. doi: 10.1093/pcp/pcz113
- Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x
- Cui, H., Peng, B., Xing, Z., Xu, G., Yu, B., and Zhang, Q. (2002). Molecular dissection of seedling-vigor and associated physiological traits in rice. *Theor. Appl. Genet.* 105, 745–753. doi: 10.1007/s00122-002-0908-2
- Ding, G., Che, P., Ilarslan, H., Wurtele, E. S., and Nikolau, B. J. (2012). Genetic dissection of methylcrotonyl CoA carboxylase indicates a complex role for mitochondrial leucine catabolism during seed development and germination. *Plant J.* 70, 562–577. doi: 10.1111/j.1365-313X.2011.04893.x

- Ding, S., Lu, Q., Zhang, Y., Yang, Z., Wen, X., Zhang, L., et al. (2009). Enhanced sensitivity to oxidative stress in transgenic tobacco plants with decreased glutathione reductase activity leads to a decrease in ascorbate pool and ascorbate redox state. *Plant Mol. Biol.* 69, 577–592. doi: 10.1007/s11103-008-9440-3
- Dixon, D. P., Laphorn, A., and Edwards, R. (2002). Plant glutathione transferases. *Genome Biol.* 3, 3004.3001–3004.3010. doi: 10.1186/gb-2002-3-3-reviews3004
- Ellis, R. H. (1992). Seed and seedling vigour in relation to crop growth and yield. *Plant Growth Regul.* 11, 249–255. doi: 10.1007/BF00024563
- Fahey, R. C., Di Stefano, D. L., Meier, G. P., and Bryan, R. N. (1980). Role of hydration state and thiol-disulfide status in the control of thermal stability and protein synthesis in wheat embryo. *Plant Physiol.* 65:1062. doi: 10.1104/pp.65.6.1062
- Fang, L., Wang, Q., Hu, Y., Jia, Y., Chen, J., Liu, B., et al. (2017). Genomic analyses in cotton identify signatures of selection and loci associated with fiber quality and yield traits. *Nat. Genet.* 49, 1089–1098. doi: 10.1038/ng.3887
- Feng, X. H., Zhang, H. X., Ali, M., Gai, W. X., Cheng, G. X., Yu, Q. H., et al. (2019). A small heat shock protein *CaHsp25.9* positively regulates heat, salt, and drought stress tolerance in pepper (*Capsicum annuum* L.). *Plant Physiol. Biochem.* 142, 151–162. doi: 10.1016/j.plaphy.2019.07.001
- Foolad, M. R., Subbiah, P., and Zhang, L. (2007). Common QTL affect the rate of tomato seed germination under different stress and nonstress conditions. *Int. J. Plant Genomics* 2007:97386. doi: 10.1155/2007/97386
- Frova, C. (2006). Glutathione transferases in the genomics era: new insights and perspectives. *Biomol. Eng.* 23, 149–169. doi: 10.1016/j.bioeng.2006.05.020
- Galle, A., Czekus, Z., Bela, K., Horvath, E., Csiszar, J., and Poor, P. (2018). Diurnal changes in tomato glutathione transferase activity and expression. *Acta Biol. Hung.* 69, 505–509. doi: 10.1556/018.69.2018.4.11
- Gill, S. S., and Tuteja, N. (2011). Cadmium stress tolerance in crop plants: probing the role of sulfur. *Plant Signal Behav.* 6, 215–222. doi: 10.4161/psb.6.2.14880
- Groot, G., Hundt, A., Murphy, A. B., Bange, M. P., and Mai-Prochnow, A. (2018). Cold plasma treatment for cotton seed germination improvement. *Sci. Rep.* 8:14372. doi: 10.1038/s41598-018-32692-9
- Gu, Q., Ke, H., Sun, Z., Wu, L., Zhang, G., Zhang, Y., et al. (2020). Genetic variation associated with the shoot biomass of upland cotton seedlings under contrasting phosphate supplies. *Mol. Breed.* 40:80. doi: 10.1007/s11032-020-01156-2
- Guo, H. X., Yin, J., Ren, J. P., Wang, Z. Y., and Chen, H. L. (2007). Changes in proteins within germinating seeds of transgenic wheat with an antisense construct directed against the thioredoxin. *J. Plant Physiol. Mol. Biol.* 33, 18–24.
- Guo, H., Zhang, H., Li, Y., Ren, J., Wang, X., Niu, H., et al. (2011). Identification of changes in wheat (*Triticum aestivum* L.) seeds proteome in response to anti-*trx s* gene. *PLoS One* 6:e22255. doi: 10.1371/journal.pone.0022255
- Halford, N. G., and Hey, S. J. (2009). Snf1-related protein kinases (SnRKs) act within an intricate network that links metabolic and stress signalling in plants. *Biochem. J.* 419, 247–259. doi: 10.1042/BJ20082408
- Hatzig, S. V., Frisch, M., Breuer, F., Nesi, N., Ducourneau, S., Wagner, M.-H., et al. (2015). Genome-wide association mapping unravels the genetic control of seed germination and vigor in *Brassica napus*. *Front. Plant Sci.* 6:221. doi: 10.3389/fpls.2015.00221
- He, C., Wang, J., Dong, R., Guan, H., Liu, T., Liu, C., et al. (2020). Overexpression of an antisense RNA of maize receptor-like kinase gene *ZmRLK7* enlarges the organ and seed size of transgenic *Arabidopsis* plants. *Front. Plant Sci.* 11:579120. doi: 10.3389/fpls.2020.579120
- He, Y., Cheng, J., He, Y., Yang, B., Cheng, Y., Yang, C., et al. (2019). Influence of isopropylmalate synthase *OsIPMS1* on seed vigour associated with amino acid and energy metabolism in rice. *Plant Biotechnol. J.* 17, 322–337. doi: 10.1111/pbi.12979
- Hernandez, L. E., Sobrino-Plata, J., Montero-Palmero, M. B., Carrasco-Gil, S., Flores-Caceres, M. L., Ortega-Villasante, C., et al. (2015). Contribution of glutathione to the control of cellular redox homeostasis under toxic metal and metalloid stress. *J. Exp. Bot.* 66, 2901–2911. doi: 10.1093/jxb/erv063
- Hu, G., Li, Z., Lu, Y., Li, C., Gong, S., Yan, S., et al. (2017). Genome-wide association study identified multiple genetic loci on chilling resistance during germination in maize. *Sci. Rep.* 7:10840. doi: 10.1038/s41598-017-11318-6
- Huang, C. T., Esvelt Klos, K., and Huang, Y. F. (2020). Genome-wide association study reveals the genetic architecture of seed vigor in oats. *G3 Genes Genom. Genet.* 10, 4489–4503. doi: 10.1534/g3.120.401602
- Ishibashi, Y., Aoki, N., Kasa, S., Sakamoto, M., Kai, K., Tomokiyo, R., et al. (2017). The interrelationship between abscisic acid and reactive oxygen species plays a key role in barley seed dormancy and germination. *Front. Plant Sci.* 8:275. doi: 10.3389/fpls.2017.00275
- Ishibashi, Y., Tawaratsumida, T., Kondo, K., Kasa, S., Sakamoto, M., Aoki, N., et al. (2012). Reactive oxygen species are involved in gibberellin/abscisic acid signaling in barley aleurone cells. *Plant Physiol.* 158, 1705–1714. doi: 10.1104/pp.111.192740
- Jeevan, K., Rajendra, P. S., Rintu, B., and Chakradhar, T. (2015). Seed birth to death: dual functions of reactive oxygen species in seed physiology. *Ann Bot.* 116, 663–668. doi: 10.1093/aob/mcv098
- Jha, B., Sharma, A., and Mishra, A. (2010). Expression of *SbGSTU* (tau class glutathione S-transferase) gene isolated from *Salicornia brachiata* in tobacco for salt tolerance. *Mol. Biol. Rep.* 38, 4823–4832. doi: 10.1007/s11033-010-0625-x
- Jiang, N., Shi, S., Shi, H., Khanzada, H., Wassan, G. M., Zhu, C., et al. (2017). Mapping QTL for seed germinability under low temperature using a new high-density genetic map of rice. *Front. Plant Sci.* 8:1223. doi: 10.3389/fpls.2017.01223
- Jiang, S., Yang, C., Xu, Q., Wang, L., Yang, X., Song, X., et al. (2020). Genetic dissection of germinability under low temperature by building a resequencing linkage map in japonica rice. *Int. J. Mol. Sci.* 21:1284. doi: 10.3390/ijms21041284
- Jiao, J. A., Yee, B. C., Kobrehel, K., and Buchanan, B. B. (1992). Effect of thioredoxin-linked reduction on the activity and stability of the Kunitz and Bowman-Birk soybean inhibitor proteins. *J. Agric. Food Chem.* 40, 2333–2336. doi: 10.1021/jf00024a002
- Kalemba, E. M., and Ratajczak, E. (2018). The effect of a doubled glutathione level on parameters affecting the germinability of recalcitrant *Acer saccharinum* seeds during drying. *J. Plant Physiol.* 223, 72–83. doi: 10.1016/j.jplph.2018.02.010
- Kao, C. W., Bakshi, M., Sherameti, I., Dong, S., Reichelt, M., Oelmüller, R., et al. (2016). A Chinese cabbage (*Brassica campestris* subsp. *Chinensis*) tau-type glutathione-S-transferase stimulates *Arabidopsis* development and primes against abiotic and biotic stress. *Plant Mol. Biol.* 92, 643–659. doi: 10.1007/s11103-016-0531-2
- Katsuya-Gaviria, K., Caro, E., Carrillo-Barral, N., and Iglesias-Fernandez, R. (2020). Reactive oxygen species (ROS) and nucleic acid modifications during seed dormancy. *Plants (Basel)* 9:679. doi: 10.3390/plants9060679
- Kaur, H., Petla, B. P., Kamble, N. U., Singh, A., Rao, V., Salvi, P., et al. (2015). Differentially expressed seed aging responsive heat shock protein OsHSP18.2 implicates in seed vigor, longevity and improves germination and seedling establishment under abiotic stress. *Front. Plant Sci.* 6:713. doi: 10.3389/fpls.2015.00713
- Kawakatsu, T., Hirose, S., Yasuda, H., and Takaiwa, F. (2010). Reducing rice seed storage protein accumulation leads to changes in nutrient quality and storage organelle formation. *Plant Physiol.* 154, 1842–1854. doi: 10.1104/pp.110.164343
- Kim, Y. B., Garbisu, C., Pickering, I. J., Prince, R. C., George, G. N., Cho, M. J., et al. (2003). Thioredoxin *h* overexpressed in barley seeds enhances selenite resistance and uptake during germination and early seedling development. *Planta* 218, 186–191. doi: 10.1007/s00425-003-1102-8
- Kobrehel, K., Kobrehel, K., Wong, J. H., Balogh, A., Kiss, F., Yee, B. C., et al. (1992). Specific reduction of wheat storage proteins by thioredoxin *h*. *Plant Physiol.* 99, 919–924. doi: 10.1104/pp.99.3.919
- Kucera, B., Cohn, M. A., and Leubner-Metzger, G. (2005). Plant hormone interactions during seed dormancy release and germination. *Seed Sci. Res.* 15, 281–307. doi: 10.1079/ssr2005218
- Li, Z. K., Chen, B., Li, X. X., Wang, J. P., Zhang, Y., Wang, X. F., et al. (2019). A newly identified cluster of glutathione S-transferase genes provides *Verticillium wilt* resistance in cotton. *Plant J.* 98, 213–227. doi: 10.1111/tpj.14206
- Liu, X., Mo, L., Guo, X., Zhang, Q., Li, H., Liu, D., et al. (2021). How cysteine protease gene *PtCP5* affects seed germination by mobilizing storage proteins in *populus trichocarpa*. *Int. J. Mol. Sci.* 22:12637. doi: 10.3390/ijms222312637
- Liu, Z., Li, H., Gou, Z., Zhang, Y., Wang, X., Ren, H., et al. (2020). Genome-wide association study of soybean seed germination under drought stress. *Mol. Genet. Genomics* 295, 661–673. doi: 10.1007/s00438-020-01646-0
- Lu, K., Zhang, Y. D., Zhao, C. F., Zhou, L. H., Zhao, Q. Y., Chen, T., et al. (2020). The *Arabidopsis* kinase-associated protein phosphatase KAPP, interacting with protein kinases SnRK2.2/2.3/2.6, negatively regulates abscisic acid signaling. *Plant Mol. Biol.* 102, 199–212. doi: 10.1007/s11103-019-00941-8

- Ma, W., Guan, X., Li, J., Pan, R., Wang, L., Liu, F., et al. (2019). Mitochondrial small heat shock protein mediates seed germination via thermal sensing. *Proc. Natl. Acad. Sci. U.S.A.* 116, 4716–4721. doi: 10.1073/pnas.1815790116
- Ma, Z., He, S., Wang, X., Sun, J., Zhang, Y., Zhang, G., et al. (2018). Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield. *Nat. Genet.* 50, 803–813. doi: 10.1038/s41588-018-0119-7
- Ma, Z., Zhang, Y., Wu, L., Zhang, G., Sun, Z., Li, Z., et al. (2021). High-quality genome assembly and resequencing of modern cotton cultivars provide resources for crop improvement. *Nat. Genet.* 53, 1385–1391. doi: 10.1038/s41588-021-00910-2
- Mao, X., Li, Y., Rehman, S. U., Miao, L., Zhang, Y., Chen, X., et al. (2020). The sucrose non-fermenting 1-related protein kinase 2 (SnRK2) genes are multifaceted players in plant growth, development and response to environmental stimuli. *Plant Cell Physiol.* 61, 225–242. doi: 10.1093/pcp/pcz230
- Moons, A. (2005). Regulatory and functional Interactions of plant growth regulators and plant glutathione S-transferases (GSTs). *Vitam. Horm.* 72, 155–202. doi: 10.1016/S0083-6729(05)72005-7
- Moursi, Y. S., Thabet, S. G., Amro, A., Dawood, M. F. A., Baenziger, P. S., and Sallam, A. (2020). Detailed genetic analysis for identifying QTLs associated with drought tolerance at seed germination and seedling stages in barley. *Plants* 9:1425. doi: 10.3390/plants9111425
- Naoumkina, M., Thyssen, G. N., Fang, D. D., Jenkins, J. N., McCarty, J. C., and Florane, C. B. (2019). Genetic and transcriptomic dissection of the fiber length trait from a cotton (*Gossypium hirsutum* L.) MAGIC population. *BMC Genomics* 20:112. doi: 10.1186/s12864-019-5427-5
- Oracz, K., El-Maarouf Bouteau, H., Farrant, J. M., Cooper, K., Belghazi, M., Job, C., et al. (2007). ROS production and protein oxidation as a novel mechanism for seed dormancy alleviation. *Plant J.* 50, 452–465. doi: 10.1111/j.1365-313X.2007.03063.x
- Roxas, V. P., Lodhi, S. A., Garrett, D. K., Mahan, J. R., and Allen, R. D. (2000). Stress tolerance in transgenic tobacco seedlings that overexpress glutathione S-transferase glutathione peroxidase. *Plant Cell Physiol.* 41, 1229. doi: 10.1093/pcp/pcd051
- Roxas, V. P., Smith, R. K., Allen, E. R., and Allen, R. D. (1997). Overexpression of glutathione S-transferase/glutathione peroxidase enhances the growth of transgenic tobacco seedlings during stress. *Nat. Biotechnol.* 15, 988–991. doi: 10.1038/nbt1097-988
- Sales, E., Viruel, J., Domingo, C., and Marques, L. (2017). Genome wide association analysis of cold tolerance at germination in temperate japonica rice (*Oryza sativa* L.) varieties. *PLoS One* 12:e0183416. doi: 10.1371/journal.pone.0183416
- Savage, W. F., and Metzger, G. L. (2006). Seed dormancy and the control of germination. *New Phytol.* 171, 501–523. doi: 10.1111/j.1469-8137.2006.01787.x
- Shi, Y., Gao, L., Wu, Z., Zhang, X., Wang, M., Zhang, C., et al. (2017). Genome-wide association study of salt tolerance at the seed germination stage in rice. *BMC Plant Biol.* 17:92. doi: 10.1186/s12870-017-1044-0
- Smiri, M., Chaoui, A., Rouhier, N., Kamel, C., Gelhaye, E., Jacquot, J., et al. (2010). Cadmium induced mitochondrial redox changes in germinating pea seed. *Biomaterials* 23, 973–984. doi: 10.1007/s10534-010-9344-y
- Song, C., Li, W., Pei, X., Liu, Y., Ren, Z., He, K., et al. (2019). Dissection of the genetic variation and candidate genes of lint percentage by a genome-wide association study in upland cotton. *Theor. Appl. Genet.* 132, 1991–2002. doi: 10.1007/s00122-019-03333-0
- Song, S., Wang, G., Wu, H., Fan, X., Liang, L., Zhao, H., et al. (2020). *OsMFT2* is involved in the regulation of ABA signaling-mediated seed germination through interacting with *OsZIP23/66/72* in rice. *Plant J.* 103, 532–546. doi: 10.1111/tpj.14748
- Sun, L., Liu, Y., Kong, X., Zhang, D., Pan, J., Zhou, Y., et al. (2012). *ZmHSP16.9*, a cytosolic class I small heat shock protein in maize (*Zea mays*), confers heat tolerance in transgenic tobacco. *Plant Cell Rep.* 31, 1473–1484. doi: 10.1007/s00299-012-1262-8
- Sun, Q., Wang, J. H., and Sun, B. Q. (2007). Advances on seed vigor physiological and genetic mechanisms. *Agric. Sci. China* 6, 1060–1066. doi: 10.1016/s1671-2927(07)60147-3
- Tan, M., Liao, F., Hou, L., Wang, J., Wei, L., Jian, H., et al. (2017). Genome-wide association analysis of seed germination percentage and germination index in *Brassica napus* L. under salt and drought stresses. *Euphytica* 213:40. doi: 10.1007/s10681-016-1832-x
- Tuan, P. A., Kumar, R., Rehal, P. K., Toora, P. K., and Ayele, B. T. (2018). Molecular mechanisms underlying abscisic acid/gibberellin balance in the control of seed dormancy and germination in cereals. *Front. Plant Sci.* 9:668. doi: 10.3389/fpls.2018.00668
- Wang, M., Zou, Z., Li, Q., Xin, H., Zhu, X., Chen, X., et al. (2017). Heterologous expression of three *Camellia sinensis* small heat shock protein genes confers temperature stress tolerance in yeast and *Arabidopsis thaliana*. *Plant Cell Rep.* 36, 1125–1135. doi: 10.1007/s00299-017-2143-y
- Wang, P., Hsu, C. C., Du, Y., Zhu, P., Zhao, C., Fu, X., et al. (2020). Mapping proteome-wide targets of protein kinases in plant stress responses. *Proc. Natl. Acad. Sci. U.S.A.* 117, 3270–3280. doi: 10.1073/pnas.1919901117
- Wang, X., Zhou, W., Lu, Z., Ouyang, Y., Chol Su, O., and Yao, J. (2015). A lipid transfer protein, OsLTPL36, is essential for seed development and seed quality in rice. *Plant Sci.* 239, 200–208. doi: 10.1016/j.plantsci.2015.07.016
- Wang, Y. (2007). Correlation between germination rate and field emergence rate of hybrid rice seeds in paper bed and sand bed. *South. Agric.* 1, 9–10, 28. doi: 10.19415/j.cnki.1673-890x.2007.05.002
- Wong, J. H., Balmer, Y., Cai, N., Tanaka, C. K., Vensel, W. H., Hurkman, W. J., et al. (2003). Unraveling thioredoxin-linked metabolic processes of cereal starchy endosperm using proteomics. *FEBS Lett.* 547, 151–156. doi: 10.1016/s0014-5793(03)00696-3
- Wong, J. H., Kim, Y. B., Ren, P. H., Cai, N., Cho, M. J., and Hedden, P. (2002). Transgenic barley grain overexpressing thioredoxin shows evidence that the starchy endosperm communicates with the embryo and the aleurone. *Proc. Natl. Acad. Sci. U.S.A.* 99, 16325–16330. doi: 10.1073/pnas.212641999
- Wu, J., Zhang, N., Liu, Z., Liu, S., Liu, C., Lin, J., et al. (2020). The *AtGSTU7* gene influences glutathione-dependent seed germination under ABA and osmotic stress in *Arabidopsis*. *Biochem. Biophys. Res. Commun.* 528, 538–544. doi: 10.1016/j.bbrc.2020.05.153
- Xiao, S., Liu, L., Wang, H., Li, D., Bai, Z., Zhang, Y., et al. (2019). Exogenous melatonin accelerates seed germination in cotton (*Gossypium hirsutum* L.). *PLoS One* 14:e0216575. doi: 10.1371/journal.pone.0216575
- Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., et al. (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 39, W316–W322. doi: 10.1093/nar/gkr483
- Xie, S., Zhao, Y., and Wang, Y. (2019). Analysis on the importance of seed germination potential to crop emergence rate in field. *Mod. Agric. Sci. Technol.* 18:11.
- Xu, J., Tian, Y. S., Xing, X. J., Peng, R. H., Zhu, B., Gao, J. J., et al. (2016). Over-expression of *AtGSTU19* provides tolerance to salt, drought and methyl viologen stresses in *Arabidopsis*. *Physiol. Plant.* 156, 164–175. doi: 10.1111/ppl.12347
- Yang, J., Yang, G., Yang, M., Su, L., Xia, A., Li, D., et al. (2019). Quantitative trait locus analysis of seed germination and early seedling growth in rice. *Front. Plant Sci.* 10:1582. doi: 10.3389/fpls.2019.01582
- Yang, P., Zhang, F., Luo, X., Zhou, Y., Xie, J., and Mock, H. P. (2015). Histone deacetylation modification participates in the repression of peanut (*Arachis hypogaea* L.) seed storage protein gene *Ara h 2.02* during germination. *Plant Biol.* 17, 522–527. doi: 10.1111/plb.12268
- Yu, J., Zao, W., He, Q., Kim, T. S., and Park, Y. J. (2017). Genome-wide association study and gene set analysis for understanding candidate genes involved in salt tolerance at the rice seedling stage. *Mol. Genet. Genomics* 292, 1391–1403. doi: 10.1007/s00438-017-1354-9
- Yu, X., Jin, H., Fu, X., Yang, Q., and Yuan, F. (2019a). Quantitative proteomic analyses of two soybean low phytic acid mutants to identify the genes associated with seed field emergence. *BMC Plant Biol.* 19:569. doi: 10.1186/s12870-019-2201-4
- Yu, X., Liu, H., Sang, N., Li, Y., Zhang, T., Sun, J., et al. (2019b). Identification of cotton mother of *FT* and *TFL1* homologs, *GhMFT1* and *GhMFT2*, involved in seed germination. *PLoS One* 14:e0215771. doi: 10.1371/journal.pone.0215771
- Yuan, Y., Xing, H., Zeng, W., Xu, J., Mao, L., Wang, L., et al. (2019). Genome-wide association and differential expression analysis of salt tolerance in *Gossypium*

- hirsutum* L at the germination stage. *BMC Plant Biol.* 19:394. doi: 10.1186/s12870-019-1989-2
- Zhang, D., Chen, L., Li, D., Lv, B., Chen, Y., Chen, J., et al. (2014). *OsRACK1* is involved in abscisic acid- and H₂O₂-mediated signaling to regulate seed germination in rice (*Oryza sativa*, L.). *PLoS One* 9:e97120. doi: 10.1371/journal.pone.0097120
- Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J., et al. (2015). Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat. Biotechnol.* 33, 531–537. doi: 10.1038/nbt.3207
- Zhao, J., Li, W., Sun, S., Peng, L., Huang, Z., He, Y., et al. (2021). The rice small auxin-up RNA gene *OsSAUR33* regulates seed vigor via sugar pathway during early seed germination. *Int. J. Mol. Sci.* 22:1562. doi: 10.3390/ijms22041562
- Zhao, K., Tung, C. W., Eizenga, G. C., Wright, M. H., Ali, M. L., Price, A. H., et al. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* 2:467. doi: 10.1038/ncomms1467
- Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* 44, 821–824. doi: 10.1038/ng.2310
- Zhou, Y., Zhou, S., Wang, L., Wu, D., Cheng, H., Du, X., et al. (2020). *miR164c* and *miR168a* regulate seed vigor in rice. *J. Integr. Plant Biol.* 62, 470–486. doi: 10.1111/jipb.12792
- Zhou, Z., Jiang, Y., Wang, Z., Gou, Z., Lyu, J., Li, W., et al. (2015). Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat. Biotechnol.* 33, 408–414. doi: 10.1038/nbt.3096

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Si, Sun, Li, Chen, Gu, Zhang, Wu, Zhang, Wang and Ma. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.