BMC Plant Biology

RESEARCH

Open Access

Multi-locus genome wide association study uncovers genetics of fresh seed dormancy in groundnut

Deekshitha Bomireddy[1,2], Vinay Sharma[1], Sunil S. Gangurde[1], D. Khaja Mohinuddin[1,3], Rakesh Kumar[4], Ramachandran Senthil[1], Kuldeep Singh[1], Mangala Reddisekhar[2], Sandip K. Bera[5] and Manish K. Pandey[1*]

Abstract

Pre-harvest sprouting (PHS) in groundnut leads to substantial yield losses and reduced seed quality, resulting in reduced market value of groundnuts. Breeding cultivars with 14–21 days of fresh seed dormancy (FSD) holds promise for precisely mitigating the yield and quality deterioration. In view of this, six multi-locus genome-wide association study (ML-GWAS) models alongside a single-locus GWAS (SL-GWAS) model were employed on a groundnut mini-core collection using multi season phenotyping and 58 K "Axiom_*Arachis*" array genotyping data. A total of 9 significant SNP-trait associations (STAs) for FSD were detected on A01, A04, A08, A09, B02, B04, B05, B07 and B09 chromosomes using six ML-GWAS models. Additionally, the SL-GWAS model identified 38 STAs across 14 chromosomes of groundnut. A single STA on chromosome B02 (*qFSD-B02-1*) was consistently identified in both ML-GWAS and SL-GWAS models. Furthermore, candidate gene mining identified nine high confidence genes viz., *Cytochrome P450 705 A*, *Dormancy/auxin associated family protein*, *WRKY family transcription factor*, *Protein kinase superfamily protein*, *serine/threonine protein phosphatase*, *myb transcription factor*, *transcriptional regulator STERILE APETALA-like*, *ethylene-responsive transcription factor 7-like and F-box protein interaction domain protein* as prime regulators involved in Abscisic acid/Gibberellic acid signaling pathways regulating dormancy/germination. In addition, three of the allele-specific markers developed from the identified STAs were validated across a diverse panel. These markers hold potential for increasing dormancy in groundnut through marker-assisted selection (MAS). Thus, this research offers insights into genetic and molecular mechanisms underlying groundnut seed dormancy in addition to providing markers and donors for breeding future varieties with 2–3 weeks of FSD.

Keywords Seed dormancy, Seed in-situ germination, Association mapping, Candidate genes, Molecular mechanism, Diagnostic markers

*Correspondence:
Manish K. Pandey
Manish.Pandey@icrisat.org
[1]Center of Excellence in Genomics & Systems Biology (CEGSB) and Centre for Pre-breeding Research (CPBR), International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad 502324, India
[2]Department of Genetics & Plant Breeding, S. V. Agricultural College, ANGRAU, Tirupati 517502, India
[3]University of Agricultural Sciences, Raichur, Karnataka 584104, India
[4]Department of Life Sciences, Central University of Karnataka, Kalaburagi 585367, India
[5]ICAR-Directorate of Groundnut Research (DGR), Junagadh, Gujarat 362001, India

Introduction

The physiological states of seed encompassing both induction of dormancy and germination initiation play a pivotal role in regulating the success of field establishment of various crop plants after sowing [1]. Dormancy and germination are primarily governed by various physiological mechanisms and environmental factors [2]. Despite their contrasting expressions, both seed dormancy and germination hold equal significance in the effective management and planning of crop cultivation. There is considerable diversity in the germination behavior among groundnut germplasm [3]. Typically, bunch types tend to be non-dormant and may undergo preharvest sprouting (PHS) when sufficient moisture is present in the field at the time of maturity or before harvest. In contrast, spreading and semi-spreading types exhibit longer seed dormancy [4]. Therefore, Spanish bunch groundnut varieties with a medium degree of dormancy are advantageous in preventing in-situ germination/PHS [5]. Conversely, longer dormancy could delay normal germination even under favorable conditions, leading to reduced germination percentage in the field. Previously, efforts have been undertaken to alleviate seed dormancy in dormant varieties by employing various growth regulators and chemicals [6].

The sustainable solution to address PHS involves developing cultivars with 14–21 days of fresh seed dormancy (FSD), capable of withstanding detrimental impact of rain between maturity and harvest. However, improving PHS resistance through phenotypic selection is a challenging task. Where complicating factors include: (i) the presence of significant genetic and environmental interactions; (ii) variation in the mechanisms governing dormancy among different plant materials [4, 7]; (iii) the involvement of multiple genes in controlling seed dormancy [8, 9]; (iv) intergenic/epistatic interactions having predominant role in governing genetic basis of seed dormancy [8, 10]; and (v) environmental conditions conducive to PHS may not be always available. Controlled environments such as rooms with sprinklers or germinators offer ideal conditions for phenotyping PHS, however, their efficacy in extensive screening of lines within breeding programs may be limited.

Genomics-assisted breeding (GAB) is a potential tool to overcome these constraints. GAB has been effectively deployed to improve leaf rust and late leaf spot resistance and high oleic acid content in groundnut [11–16]. However, the development of highly efficient linked markers is a pre-requisite for successful deployment of GAB. Identification of genomic regions and candidate genes linked to FSD can facilitate marker development, aiding in the effective transfer of FSD trait into elite groundnut cultivars. Significant efforts have been dedicated to understand the molecular mechanisms of PHS in cereals, and several genomic regions and candidate genes have been identified in rice [17, 18] wheat [19, 20], and barley; [21, 22]. In rice, *GA20-oxidase* gene was identified in the quantitative trait loci (QTL) region controlling PHS [21]. Similarly, *Mitogen-Activated Protein Kinase Kinase 3* (*MKK3*) and *Alanine Aminotransferase* (*AlaAT*) in barley [23, 24]; *mother of FT and TFL1* (*MFT*) and *Phs1* in wheat [25, 26] were identified as underlying genes regulating seed dormancy. Though PHS is a widespread constraint in groundnut, only limited efforts were made to map FSD QTLs so far, mainly using bi-parental populations [8, 27–30]. In previous reports, only few genes, viz., *zeaxanthin epoxidase*, *RING-H2 finger protein* [27] and *ethylene-responsive transcription factor* [63, 29] were identified as candidate genes involved in hormonal regulation of dormancy in groundnut. Bi-parental QTL mapping generally faces limitations due to limited recombination events taking place during the development of recombinant inbred line population, because of which biological processes governing dormancy remain not fully understood [31]. Consequently, future efforts should focus on employing highly efficient and reliable QTL mapping methods to identify additional novel QTLs associated with this trait.

Genome-wide association studies (GWAS) have become increasingly recognized as a potent methodology to identify QTLs and genes linked with complex traits based on the historic recombination in a large natural population [32–35]. GWAS can surpass the limitations of bi-parental linkage mapping, allowing for the dissection of complex traits with high mapping resolution [36]. However, so far there are no reports on genome wide association studies on FSD using diverse germplasm in groundnut. Further, methods such as mixed linear model (MLM) [37], implemented in single-locus GWAS (SL-GWAS), have been extensively utilized to investigate several genetic variants linked to complex agronomic traits. However, SL-GWAS methods face limitations in identifying minimal effect significant SNP-trait associations (STAs) influenced by the stringent Bonferroni correction and multigenic background [38]. To overcome these limitations, development of Multi-Locus GWAS models (ML-GWAS) has been introduced as a multi-faceted genome scanning approach, simultaneously estimating the effect of all the markers [39].

In the view of above gap of the knowledge and advancements in association mapping analysis, in this study, we utilized association mapping strategies by employing six ML-GWAS and one SL-GWAS (MLM) models in the groundnut mini-core collection. We aim to identify all the possible STAs using multiple methodologies to derive candidate genes regulating FSD to facilitate marker development.

Materials and methods

Plant material

The groundnut mini-core collection, consisting of 184 accessions developed at ICRISAT, Patancheru was used as association mapping panel [40]. This mini-core represents the extensive genetic diversity within the entire ICRISAT groundnut germplasm, collected from 45 countries spanning Asia, Africa, North America and South America [3]. Encompassing six botanical types - *hypogaea*, *hirsuta*, *fastigiata*, *peruviana*, *aequatoriana* and *vulgaris*, the collection provides a comprehensive resource for groundnut genetic studies and breeding efforts.

Field experiment and phenotypic evaluation

Phenotyping data on days to 50% germination was recorded at two locations namely ICRISAT, Hyderabad, India and Dry land farm of S.V. Agricultural College in Tirupati, Andhra Pradesh, India. A total of four seasons (Post-rainy 2018–2019, 2019–2020, 2022–2023 and Rainy season of 2022) of phenotyping data was generated at ICRISAT, Patancheru located at 545 m altitude, 17° 31' 48.00" N latitude and 78° 16' 12.00" E longitude. Additionally, all 184 accessions were also grown at Tirupati, located at 182.9 m altitude, 13°54' N latitude and 79°54' E longitude and phenotyped during season Rainy 2021. Standard agronomic practices for groundnut cultivation were followed during all the growing seasons. To ensure that any differences in dormancy are more likely due to genetic factors rather than differences in maturity, accessions were categorized into three groups/sets (early/medium/late maturity) and harvested accordingly. Matured pods with development of black pigmentation inside the shell were freshly harvested and selected for phenotyping using an in-vitro germination assay. Twenty good-quality and uniformly sized seeds from the two replicates of each accession were selected, treated with fungicides (Mancozeb and Carbendazim), and placed on moist germination paper within petri plates. Petri plates were kept in complete darkness and watered regularly at 24-hour intervals to maintain moisture conditions [41]. Over the period of 30 days, data was recorded on each accession on the number of days required to achieve 50% germination, referred as days to 50% germination. Detailed methodology of which has been explained in our previous publication [3].

DNA extraction and genotyping using 58 K SNP array

Total genomic DNA was isolated from the tender leaves of 25–30 day old plants of groundnut mini-core collection using Nucleospin Plant II kit (Macherey-Nagel, Düren, Germany). The purity and concentration of the isolated genomic DNA samples were assessed by electrophoresis on a 0.8% agarose gel and Thermo Fisher Scientific Nanodrop 8000 Spectrophotometer, respectively. Genotyping was carried out with 'Axiom_*Arachis*' SNP array of 58,233 SNP markers derived from DNA resequencing of 41 wild diploid ancestors and tetraploid accessions of groundnut [42, 43]. The DNA samples from mini-core accessions were genotyped on the Affymetrix GeneTitan platform following previously described methods [32] and resulting data for each accession in .CEL file format was generated and stored. Subsequent SNP calling and data analysis performed using Axiom™ Analysis Suite version 1.0 (Thermo Fisher Scientific, USA) to implement quality control (QC) measures and select samples that successfully passed the QC test. Of the 58,233 SNPs retrieved from Axiom™ analysis suit, high-quality SNPs were filtered out with minor allele frequency (MAF) of ≥ 0.05 and maximum missing sites fixed to $<20\%$ per SNP using Tassel v5.0 software [44]. After stringent filtration a total of 10,064 high-quality SNPs were subsequently employed for association mapping studies.

Genome-wide association analysis for fresh seed dormancy

Multi-season phenotyping data, along with genotyping data on 10,064 SNP of the mini-core set was used to perform genome-wide association analysis using multi-locus model. Six ML-GWAS methods, namely Multi-Locus Random-SNP-Effect Mixed Linear Model (mrMLM) [38], Fast Multi-Locus Random-SNP-Effect Mixed Linear Model (FASTmrMLM) [45], Fast Multi-Locus Random-SNP-Effect Efficient Mixed-Model Association (FASTmrEMMA) [46], Polygenic Local-Architecture Random Mixed Effect Bayesian (pLARmEB) [47], Polygenic Kruskal-Wallis Mixed Linear Model with Bayesian (pKWmEB) [48] and Iterative Sure Independence Screening Enhanced by Expectation Maximization and Bayesian LASSO (ISIS EM-BLASSO [49] in the mrMLM R package (https://cran.r-project.org/web/packages/mrMLM/index.html) were implemented. Default parameter values were utilized, and Logarithm of the Odds (LOD) score of 3 was set to identify robust STAs. All six methods utilized PCA and kinship matrices.

A single-locus mixed linear model (MLM) was also performed using the Tassel v5.0 software [44]. For correcting population structure and reducing the false-positive rate, we employed the first three principal components (PCs) and a kinship matrix as covariates. The association threshold to determine significant marker-trait associations was computed with Bonferroni correction by calculating a *p-value* of 4.9682×10^{-6}, derived from the negative log transformation of α/n (α is the overall significance threshold (0.05) and n represents total number of SNPs (10,064) used for GWAS analysis).

Identification of candidate genes corresponding to significant STAs

STAs were further used for finding of candidate genes in the genomic region of 100 kb upstream and 100 kb downstream from the identified SNP position using diploid genome assemblies on Peanutbase (www.peanutbase.org). Based on the previous reports available in the literature, genes that were reported to regulate ABA/GA mediated processes were then identified as candidate genes regulating seed dormancy/germination.

Development and validation of allele specific markers

For designing the allele specific markers, Batch primer 3 software (https://probes.pw.usda.gov/batchprimer3/) was used. A flanking sequence of 1200 bp (600 bp up stream and 600 bp downstream) to the identified SNP positions were used to design the primers. Default parameters were used to design the primers with product size range of 400–600 bp and 50% GC content. These markers were validated on a diverse set of genotypes.

Results

Phenotyping for fresh seed dormancy on mini-core collection

In this study, we used number of days required for an accession to achieve 50% germination as a measure of dormancy. In the GWAS panel, days to 50% germination was in the range of 1 to 30 days. Accessions of Virginia Bunch and Virginia Runner (var. *hypogaea*) showed a longer duration (16–30 days) of dormancy compared to Spanish Bunch (var. *vulgaris*) and Valencia (var. *fastigiata*) (1–25 days). The mean performance and phenotypic distribution of these accessions (from two replications), screened for days to 50% germination across five seasons are represented in Fig. 1 & Table S1. For more detailed information on the phenotyping data, refer to our previous publication, Bomireddy et al. [3].

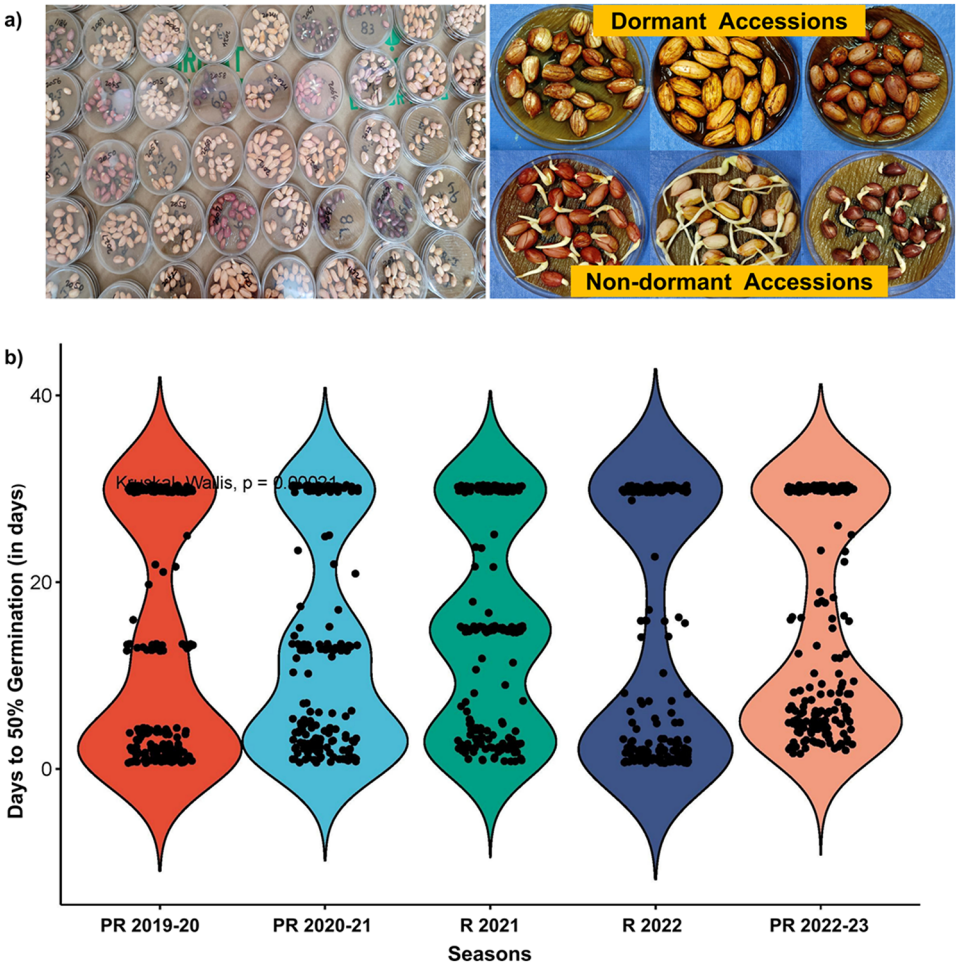

Fig. 1 Phenotypic variability for fresh seed dormancy in the mini-core accessions. (**a**) Phenotypic variability, dormant and non-dormant mini-core accessions during in-vitro germination assay (**b**) Violin plots representing days to 50% germination of the mini-core accessions during Post-rainy 2019–2020, Post-rainy 2020–2021, Rainy 2021, Rainy 2022 and Post-rainy 2022-23 seasons

Multi-locus GWAS identified significant STAs associated with fresh seed dormancy

A total of 9 STAs (LOD ≥ 3) significantly associated with FSD were identified in six ML-GWAS approaches (Fig. 2). Of these, seven STAs (*qFSD_A04-1*, *qFSD_A08-1*, *qFSD_A09-1*, *qFSD_B02-1*, *qFSD_B05-1*, *qFSD_B07-1*, *qFSD_B09-1*) were consistently identified in at least two ML-GWAS methods (Fig. 3; Table 1). Notably, STAs on chromosomes A09 (*qFSD_A09-1*) and B07 (*qFSD_B07-1*) were consistently identified in all six methods, with higher LOD score ranging from 11.87 to 61.35 and 4.35–36.85, respectively. The STA *qFSD_B05-1* on B05 was identified in four methods with LOD score ranging from 11.80 to 59.71. Additionally, two STAs (*qFSD_A09-1*, *qFSD_B02-1*) were identified by at least three ML-GWAS methods, with the LOD score ranging from 3.33

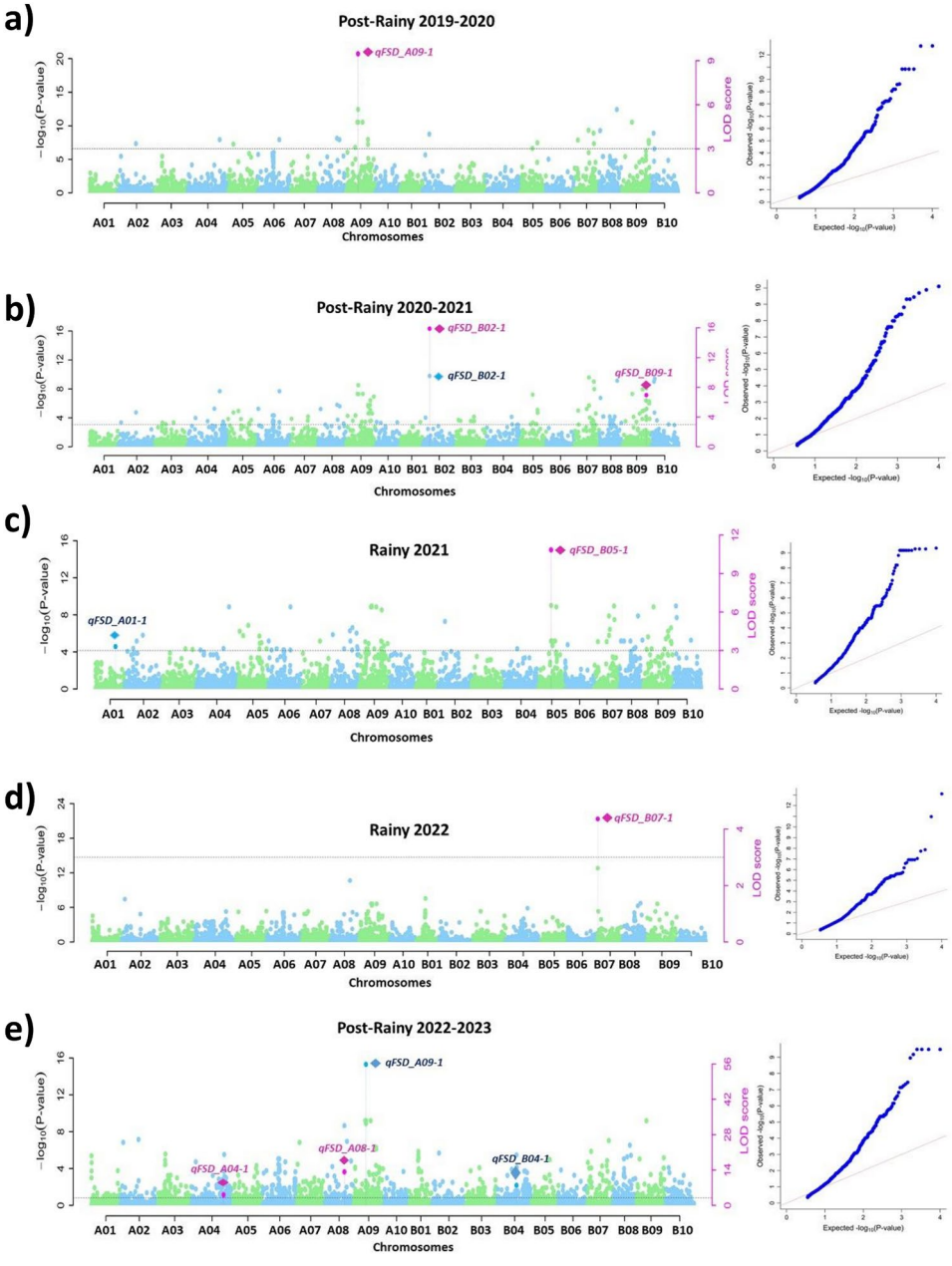

Fig. 2 Multi-Locus Genome-wide association studies of fresh seed dormancy measured as days to 50% germination using a diverse groundnut mini-core collection of 184 accessions. Manhattan plots and Quantile–Quantile (QQ) plots for fresh seed dormancy during (**a**) Post-Rainy 2019–2020 (**b**) Post-Rainy 2020–2021; (**c**) Rainy 2021 (**d**) Rainy 2022 (**e**) Post-Rainy 2022–2023. Manhattan plots were created for the genomic regions identified across five seasons, by at least one model. They displayed SNP associations, with peaks showing potential genomic regions linked to fresh seed dormancy. Names of the associated genomic regions/QTLs that were used for candidate gene prediction are mentioned near the pointing peak. Q-Q plots displayed that observed associations are aligned with expected distributions, highlighting true/significant associations

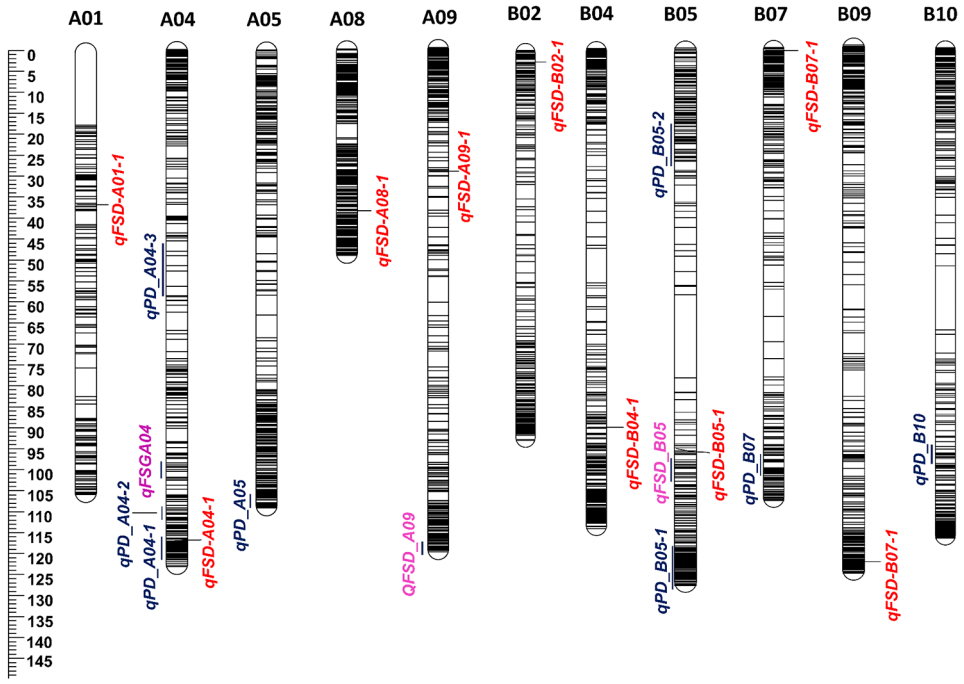

Fig. 3 Distribution of fresh seed dormancy QTLs measured as Days to 50% germination (D_50%) identified using ML-GWAS in groundnut mini-core collection. QTLs on the right are the ones identified from ML-GWAS. QTLs on the left with different colours represents the genomic regions identified for seed dormancy from different studies (Light Purple: Kumar et al., [27] Dark Purple: Zhang et al., [9] Blue: Wang et al., [29]

Table 1 The significant STAs for fresh seed dormancy identified using multi-locus GWAS models

SNP	Chr	Position (bp)	QTL region	Seasons	Model name	LOD score	R^2 (%)
AX_147210899	A01	36,775,841	*qFSD_A01-1*	Rainy 2021	6	3.3	2.88
AX_147221160	A04	119,365,410	*qFSD_A04-1*	Post-rainy 2022–2023	4,6	3.98–4.30	5.81–6.03
AX_147231175	A08	38,560,701	*qFSD_A08-1*	Post-rainy 2022–2023	2,4,5	3.33–48.76	70.98–74.50
AX_147233202	A09	29,726,644	*qFSD_A09-1*	Post-rainy 2020–2021, Post-rainy 2022–2023	1,2,3,4,5,6	11.87–61.35	71.75–84.54
AX_147240363	B02	3,495,023	*qFSD_B02-1*	Post-rainy 2020–2021	1,2,4	10.13–17.16	74.84–85.07
AX_147247942	B04	105,437,967	*qFSD_B04-1*	Post-rainy 2022_2023	1	7.92	70.72
AX_147250106	B05	112,344,292	*qFSD_B05-1*	Post-rainy 2020–2021	2,3,4,6	11.80-59.71	53.44–85.07
AX_147254360	B07	995,898	*qFSD_B07-1*	Rainy 2022	1,2,3,4,5,6	4.35–36.85	38.58–84.56
AX_147262364	B09	143,731,363	*qFSD_B09-1*	Post-rainy 2020–2021	1,6	6.26–7.62	15.22–18.74

Chr: Chromosome; Model – 1: mrMLM; 2: FASTmrMLM; 3: FASTmrEMMA; 4: pLARmEB; 5: pKWmEB; 6: ISIS EM-BLASSO; LOD: Logorithm of Odds: R^2: Coefficient of Determination/Phenotypic Variation Explained

to 48.76 and 10.13–17.16, respectively. Similarly, the STA on chromosome A09 (*qFSD_A09-1*) was consistently detected during the seasons, Post-rainy 2019–2020 and 2022–2023 through ISIS EM-BLASSO approach.

Moreover, MLM or single locus model in Tassel v5.0 identified 38 significant STAs across 14 chromosomes, with R^2 values in the range of 3.1–8.9% (Figure S1; Table S2). Notably, STAs on chromosome A09 and B09 were consistently identified in all the five seasons, with R^2 value of 4.4 and 8.1%, respectively. Additionally, a single STA on chromosome B02 (*qFSD_B02-1*) identified in both ML-GWAS and SL-GWAS approaches.

Candidate genes identified for fresh seed dormancy

Candidate gene search in 200 kb flanking region of the 9 significant STAs identified a total of 134 genes (Table S3). Of these, previous studies have extensively reported 63 genes as potential regulators in the process of seed dormancy and/or seed germination regulating *via* ABA, GA, and ethylene signaling pathways (Table S4). Among 63 genes, 5 genes unveiled missense variants in the CDS coding region, 3 genes had downstream gene variants, 1 gene had 3'_prime_UTR_variants, 2 genes had 5'_prime_UTR_variant and 1 gene had upstream gene variant, determined by *SnpEff* prediction tool (Table S5). A STA (*qFSD-A09-01*) identified in all six methods could be an important genomic region regulating FSD. Corresponding to the STA (*qFSD-A09-01*), important

genes such as *dormancy/auxin associated family protein (Aradu.2Q7VA)*, *DUF223 domain protein (Aradu.XE42X)* and *hypoxia-responsive family protein (Aradu.WAM0A)* were identified with functional relevance to dormancy/germination. In the genomic region of STA *qFSD-A04-01*, multiple copies of *protein kinase superfamily proteins (Aradu.VA4EQ)* along with *F-box interaction domain proteins (Aradu.XB89Z)*, and *serine/threonine-protein phosphatase* (*Aradu.X6CDT*) were identified to regulate seed dormancy as reported in various crops [33, 25, 29]. The genes corresponding to the STA on chromosome A08 (*qFSD-A08-01*) included *RNA methyltransferase (Aradu.3F83L)*, *Pentatricopeptide repeat (PPR) superfamily protein (Aradu.H1R88)*, and *eukaryotic aspartyl protease family protein (Aradu.HGT8J)*. Similarly, *ethylene-responsive transcription factors* (*Araip.LL89K*; *Araip.I6HJK*) from *qFSD-B04-01* are recognized as promising genes involved in regulating dormancy. In the STA *qFSD_B07-01*, *BTB/POZ domain-containing protein (Araip.JP0WQ)*, *late embryogenesis abundant (LEA) protein* (*Araip.S5KEZ*) and *receptor-like protein kinases* (*Araip.RLI4W*) from *qFSD_B09-01* were prominent genes known for their involvement in regulation of dormancy/germination.

In the 200 kb genomic region around the identified significant STAs by SL-GWAS methods, a total of 346 genes were retrieved (Table S6). Potential genes among these included *auxin transport protein (Aradu.05XZ1)*, *E3 ubiquitin-protein ligase (Aradu.7Y7XJ)*, *WRKY family transcription factor family protein (Aradu.76148)*, *serine/threonine protein phosphatase (Aradu.25KN6)*, *cytochrome P450 (Aradu.FN562)*, *abscisic acid receptor (Aradu.640E1)*, *eukaryotic aspartyl protease family protein (Araip.2P2KT) etc..*, which were identified as the prominent regulators of dormancy/germination (Table S7).

To elucidate the molecular mechanisms governing dormancy and germination, we conducted a comprehensive investigation to understand the functional involvement of the identified candidate genes in the ABA and GA biosynthesis and or signaling pathways. Among these candidate genes, *cytochrome P450 705 A (Aradu.FN562)* was distinctly discerned for its involvement in ABA catabolism. *Dormancy/auxin associated family protein* (*ARF*) (*Aradu.2Q7VA*), *WRKY family transcription factor (Aradu.76148)*, *Protein kinase superfamily protein (Aradu.VA4EQ)*, *serine/threonine protein phosphatase* (*PP2C*) (*Aradu.25KN6*) and *MYB transcription factor* (*Aradu.GFS4B*) were observed as key players participating in ABA signaling. *Transcriptional regulator of STERILE APETALA-like* (*Araip.14WNT*) was known for its involvement in GA catabolism, while *ethylene-responsive transcription factors* (*Araip.LL89K*) and *F-box protein interaction domain protein* (*Aradu.XB89Z*) emerged as a noteworthy participant in GA signaling (Fig. 4).

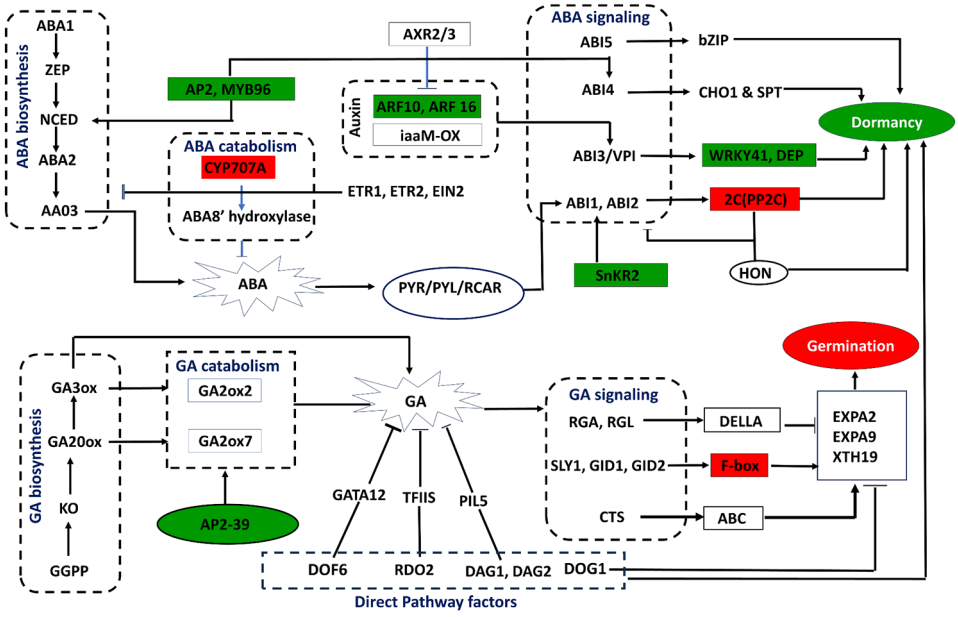

Fig. 4 Abscisic and gibberellic acid metabolism and signaling pathways in plants. Picture highlights the genes identified in this study, depicting their role and how interplay of ABA/GA genes is enabling seed dormancy or germination. *AP2* (*Araip.14WNT*)- *transcriptional regulator STERILE APETALA-like*; *MYB96* (*Aradu.GFS4B*)- *myb transcription factor*; *ARF10* (*Aradu.2Q7VA*)-*Dormancy/auxin associated family protein*; *CYP707A* (*Aradu.FN562*)- *cytochrome P450 705A5*; *ETR* (*Araip.LL89K*)- *ethylene-responsive transcription factor 7-like*; *SnKR2* (*Aradu.VA4EQ*)- *Protein kinase superfamily protein*; *WRKY41* (*Aradu.76148*)- *WRKY family transcription factor family protein*; *2 C*(*PP2C*) (*Aradu.25KN6*)- *serine/threonine-protein phosphatase 7 long form homolog*; *F-box* (*Aradu.XB89Z*)- *F-box protein interaction domain protein*. Source: Sohindji et al. [50]

Based on the phenotyping data, a representative panel comprising of mini-core accessions with varied dormancy durations were selected to assess the efficacy of the identified STAs. Allele calls of the selected accessions for nine significantly associated SNP markers (identified from ML-GWAS models) were used from 'Axiom_*Arachis*' 58K SNP array genotyping data. There was no polymorphism between non-dormant and dormant mini-core accessions for AX_147210899 and AX_147221160 markers from A01 and A04 chromosomes (Fig. 5). However, AX_147233202 from A09 chromosome differentiated between dormant and non-dormant genotypes. Accessions with all favorable dormant alleles for the other 6 significant SNPs exhibited longer dormancy durations (require ≥ 30 days for 50% germination). Conversely, an increase in the number of unfavorable non-dormant alleles corresponded to a decreased dormancy duration. Therefore, these markers can be used for development of allele specific or KASP assays for their deployment in marker-assisted selection (MAS) to improve popular groundnut cultivars with 14–21 days of dormancy.

Allele specific markers developed and validated for fresh seed dormancy

Of the nine STAs (identified from ML-GWAS models), seven were shortlisted based on clear polymorphism between non-dormant and dormant accessions. However, the remaining two markers AX-147,210,899 (A01) and AX-147,221,160 (A04) were dropped due to large proportion of heterozygous calls. Of the seven markers, AX-147,233,202 from A09 showed clear polymorphism between dormant and non-dormant lines, while the other six had one or two ambiguous calls between them. Based on the allelic combinations, four markers were selected to develop allele-specific primers. During validation, three out of the four primers successfully distinguished between dormant and non-dormant lines by showing clear bands. Using these three markers, entire

Name of the Accession	Agronomic Type	Associated Markers									DAG_50%
		AX_147210899	AX_147221160	AX_147231175	AX_147233202	AX_147240363	AX_147247942	AX_147250106	AX_147254360	AX_147262364	
ICG10092	*Valencia Bunch*	AG	AG	GG	CC	CT	CC	AA	AA	AA	1
ICG397	*Valencia Bunch*	AG	AG	GG	CC	CT	CC	AA	AA	AA	2
ICG442	*Spanish Bunch*	AG	AG	GG	CC	CT	CC	AA	AA	AA	3
ICG5221	*Valencia Bunch*	AG	AG	GG	CC	CT	CC	AA	AA	AA	3
ICG5195	*Spanish Bunch*	AG	AG	GG	CC	CT	CC	AA	AA	AA	4
ICG14985	*Spanish Bunch*	AG	AG	GG	CC	CT	CC	GG	AA	AA	8
ICG15042	*Valencia Bunch*	AG	AG	GG	CC	CT	CC	AA	AG	AA	9
ICG15419	*Peruvinia Runner*	AG	AG	GG	CC	CT	CC	AA	AG	GG	16
ICG721	*Virginia Bunch*	AG	AG	AA	TT	CC	TT	GG	GG	GG	30
ICG4156	*Virginia Runner*	AG	AG	AA	TT	CC	TT	GG	GG	GG	30
ICG11457	*Virginia Runner*	AG	AG	AA	TT	CC	TT	GG	GG	GG	29
ICG11322	*Virginia Bunch*	AG	AG	AA	TT	CC	CC	GG	GG	GG	29
ICG14705	*Virginia Bunch*	AG	NN	AA	TT	CC	TT	GG	AA	GG	25
ICG6057	*Virginia Bunch*	AG	AG	AA	TT	CC	TT	GG	GG	AA	24
ICG5827	*Virginia Runner*	GG	AG	AA	TT	CT	TT	AA	GG	GG	22
ICG11426	*Virginia Bunch*	AG	AG	GG	TT	CC	CC	GG	AA	GG	22

Fig. 5 Validation of the significantly associated markers with allele calls of the representative panel of mini-core accessions. Accessions with all favorable dormant alleles exhibited longer dormancy durations. Whereas, an increase in the number of unfavorable non-dormant alleles exhibited reduced dormancy duration

mini-core set was genotyped along with the non-dormant (ICGV 02266) and dormant (ICGV 97045) parents of a RIL population, developed at ICRISAT [8]. ICGV 02266 is a non-dormant genotype Among these markers, *GMFSD2*, *GMFSD3*, and *GMFSD4* successfully differentiated between dormant and non-dormant lines (Table 2; Fig. 6). As depicted in the gel images, Virginia runner and Virginia bunch lines predominantly exhibited dormancy, with a dormancy period of 23–30 days. In contrast, Valencia bunch and Spanish bunch lines are non-dormant, germinating within 1–2 days. However, ICG 118, a Spanish bunch line, did not germinate until 30 days and exhibited dormancy.

Discussion

Groundnut or peanut is a major oilseed and grain legume mainly cultivated under rainfed regions of tropical, subtropical and temperate countries worldwide. Unlike the Virginia genotypes, the widely grown Spanish cultivars have lost the dormancy trait during domestication and selective breeding, and resulted in introduction of PHS trait in cultivated groundnut. Developing commercial Spanish cultivars with 14–21 days of dormancy can prevent yield losses due to PHS. Utilizing GAB offers a distinct advantage over conventional breeding by facilitating efficient tracking of alleles among segregating lines through the use of trait-linked markers [15]. In this context, the present investigation employed multi-model genome wide association analysis on mini-core collection using genotyping data generated from "Axiom_*Arachis*" 58 K SNP array and multi-environment phenotyping data to identify the genomic regions and candidate genes regulating dormancy/germination.

Conventional single-locus methods like Generalized Linear Model (GLM) and Mixed Linear Model (MLM) have been frequently deployed for identifying genetic variants in several crops [51]. However, these methods have limitations as they neglect combined effects of multiple loci and face issues with multiple test corrections to determine critical values [52]. ML-GWAS methods, however, addresses these challenges [53]. Comparative studies indicated that ML-GWAS has higher statistical power and lower false-positive errors as compared to SL-GWAS methods [54, 55]. Investigators typically integrate the strengths of various ML-GWAS algorithms to identify target loci/QTL associated with complex traits, as each algorithm possesses unique characteristics and QTL detection capabilities [56].

A total of 9 significant STAs using ML-GWAS were identified for FSD trait using mini-core collection association panel. Previously, a QTL-Seq study reported two genomic regions on B05 and A09 chromosomes for FSD and also developed a potential marker on chromosome B05, *GMFSD1* [27]. In this study we have developed three more allele specific markers namely, GMFSD2 (A09), GMFSD3 (B05) and GMFSD4 (B09). High-density genetic mapping for FSD identified two dormancy QTLs on chromosomes A04 and A05 [29]. Similarly, a major stable QTL associated with fresh seed germination was identified on chromosome A04 [9]. Moreover, our previous FSD study used a 5 K SNP assay based bi-parental genetic mapping and identified five major QTLs on Ah01, Ah06 Ah11, Ah16 and Ah17 chromosomes and two minor QTLs on Ah04 and Ah15 chromosomes [8]. Additionally, *qFSD_A04-1* (119 Mb) on chromosome A04 was found to be located in the close proximity of *qPD_A04-2* [29], while the location of *qFSD_B05-1* (112 Mb) on chromosome B05 was physically close to the genomic region identified by Kumar et al. [27]. Thus, the identification of significant STAs around previously reported genomic regions using multi-locus GWAS underscores the method's reliability. In addition to ML-GWAS, our study has also identified 38 significant STAs on 14 chromosomes of cultivated groundnut by SL-GWAS revealing all the possible genomic regions associated with FSD.

Differences in the mapping results from various studies can be attributed to the factors such as seed development stage, population composition or the pedigree of the parents used in the population development and the prevailing environment during crop growth period [57, 58]. Comparable results were also documented in association mapping studies on wheat and rice seed dormancy [59, 31]. Because of genome similarity between the homeologous chromosomes of diploid progenitor

Table 2 Primers sequences for three allele specific markers validated on diverse germplasm lines of groundnut

SN	Probe ID	Chr	Pos	Marker name	Dormant Alleles	F/*R*	Primer sequence	Melting temperature (Tm)	Product size (bp)
1.	AX_147233202	A09	29,726,644	*GMFSD2*	T	F	AACTTGAACTTTCCTGGGAT	48	472
						R	TCCTGACTTCCCTGATGTTG	52	
2.	AX_147250106	B05	112,344,292	*GMFSD3*	G	F	TATTTGGTCTGCTCCGCTCT	52	286
						R	TCTACAAACTTCTCTCCGGTCC	55	
3.	AX_147262364	B09	143,731,363	*GMFSD4*	G	F	AACCAAGGGAAGGATCAACC	52	265
						R	TCAAGACTGTTCCCGAATGAC	52	

Chr: Chromosome; **Pos**: Position

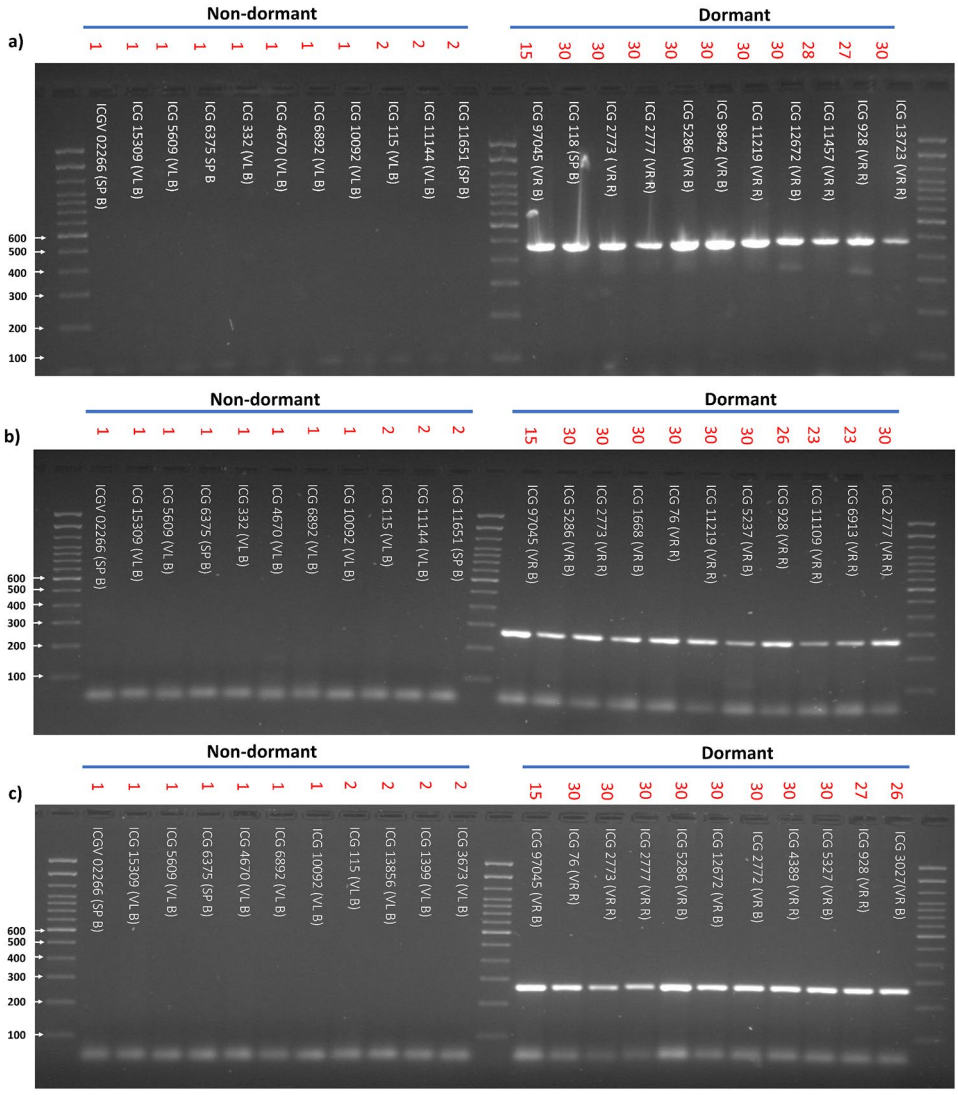

Fig. 6 Three allele specific markers were validated in diverse mini-core collection. (**a**) GMFSD2 (AX_147233202) on A09; (**b**) GMFSD3 (AX_147250106) on B05 and (**c**) GMFSD4 (AX_147262364) on B09 chromosomes. These three markers differentiated non-dormant lines of Spanish Bunch (SP B); Valencia Bunch (VL B) from dormant lines of Virginia Runner (VR R); Virginia Bunch (VR B)

genomes (A and B genome), homeologous associations on both sub-genomes of groundnut were identified. Nested association mapping for seed and pod weight in groundnut identified associations on homeologous chromosomes A05/B05, A06/B06 [12, 32]. Similarly, in other allopolyploids such as wheat, seed dormancy QTLs were detected on homeologous 3 A/3B [60], 4 A/4B and 5 A/5B [61] chromosomes.

As discussed earlier, identification of candidate genes underlying the QTLs/STAs provide insights for better understanding of the trait. As ABA, GA and ethylene have been demonstrated to be associated with seed dormancy and germination regulation in many crops, identifying genes involved in the regulation of their metabolic pathways is of major interest. The involvement of ABA signaling and, its interaction with GAs/ethylene tends to modulate seed dormancy and germination initiation. Therefore, the genes retrieved in this study from both ML-GWAS and SL-GWAS models were thoroughly reviewed in previous literature for assessing their functional role in ABA and GA signaling pathways. *Cytochrome P450 705 A* (*Aradu.FN562*) was identified as an important participant in ABA catabolism. A *cytochrome P450 superfamily protein* (*CYP707A*) in Arabidopsis encodes *ABA 8'-hydroxylases*, an enzyme involved in ABA 8'-hydroxylation pre-dominant for ABA catabolism. Expression profiling indicated that *cyp707a2* mutant displayed six times higher ABA levels, resulting in hyper seed dormancy compared to wild types [62]. It indicated that *CYP707A2* negatively regulates seed dormancy by declining ABA levels during seed imbibition. Supporting this, *cytochrome P450 superfamily protein*

gene copies displayed high transcript abundance in ICGV 91114 (non-dormant) gene expression atlas, indicating their positive role in regulating germination [8].

WRKY family transcription factor family protein (*Aradu.76148*) identified in this study is known to be involved in ABA signaling. Lack of *WRKY transcription factor 41* (*WRKY41*) in imbibed seeds of Arabidopsis resulted in decreased *ABI3* (play crucial role in seed dormancy) expression, while overexpressing transgenic *WRKY41* lines increased *ABI3* expression [63]. Examination of the double mutant *wrky41 aba2* revealed that the regulation of *ABI3* expression and seed dormancy is a combined effect between *WRKY41* and ABA. Therefore, *WRKY41* acts as a key regulator of *ABI3* expression, thereby influencing seed dormancy. The identified protein kinase superfamily protein (*Aradu.VA4EQ*) in this study is known for its positive role in ABA signaling. In Arabidopsis, *SNF1-RELATED PROTEIN KINASE 2.2* (*SnRK2.2*) and *SnRK2.3*, were reported to regulate ABA responses in seed dormancy and germination by mediating ABA signaling [64]. Double mutants of *snrk2.2* and *snrk2.3* exhibited decreased expression of several ABA-induced genes, demonstrating their positive role in ABA signaling. Similarly, in Arabidopsis, redundant ABA-activated *SnRK2s* were identified as prominent regulators of seed maturation and dormancy [65]. *Dormancy/auxin associated family protein* (*Aradu.2Q7VA*) or *Auxin responsive factors* (*ARF*) identified in this study were reported to stimulate ABA signaling to induce seed dormancy [66]. Defects in auxin signaling of *MIR160*-overexpressing plants and auxin receptor mutants significantly reduce dormancy, while increase in auxin biosynthesis prolong dormancy.

Serine/threonine-protein phosphatase 7 long form homolog (Aradu.25KN6) and *myb transcription factor* (*Aradu.GFS4B*) identified in this study are known for their role in ABA signaling. In Arabidopsis, loss of *ABSCISIC ACID-INSENSITIVE1* (*ABI1*) which encodes 2 C class of *serine/threonine phosphatases* (*PP2C*), leads to enhanced ABA responsiveness, indicating its negative role in ABA signaling [67]. Further, overexpressing *HON* (*Protein Phosphatase 2 C family group*) lines revealed that it suppresses dormancy by impeding ABA signaling [68]. However, some *PP2Cs* like *HIGHLY ABA-INDUCED PP2C GENE1* (*HAI1*) were reported to promote ABA signaling [69]. *MYB transcription factor 96* (*MYB96*) was presumed to fine tune seed dormancy as it enhances ABA biosynthetic NCED genes and down regulate GA biosynthetic *GA20ox1*, *GA3ox1* genes in Arabidopsis [70]. The *myb96-1* mutant seeds exhibited germination earlier than the wild *MYB96-1*, while activation-tagging of *myb96-1D* seeds delayed the germination process. Differential gene expression analysis identified *MYB60* as the promising candidate regulating dormancy with higher transcript abundance observed in the dormant var. 'Tifrunner' gene expression atlas [8].

The identified *Transcriptional regulator of STERILE APETALA-like (Araip.14WNT)* is particularly known to be involved in GA catabolism. *APETALA 2* (*AP2*)-domain-containing transcription factors (ATFs), including *OsAP2-39* in rice and *ABI4* in Arabidopsis, were reported to play a prominent role in ABA and GA antagonistic crosstalk [71, 72]. The *AP2/ethylene-responsive element binding factor (AP2/ERF)* family constitutes a substantial group of plant transcription factors, playing diverse roles at various plant developmental stages. In rice, transcriptome analysis of *OsAP2-39* overexpression lines unveiled upregulation of *OsNCED-I* (ABA biosynthesis gene), increasing endogenous ABA levels, and enhanced GA-inactivating gene *OsEUI*, resulting in reduced GA content [72]. *OsAP2-39* directly governs the expression of *OsNCED-I* and *OsEUI*, elucidating a novel mechanism regulating ABA/GA balance and ultimately influencing rice growth.

Ethylene-responsive transcription factors (*Araip.LL89K*) and *F-box protein interaction domain protein* (*Aradu.XB89Z*) identified in this study are the significant contributors for GA signaling. Ethylene (ET) was reported to break seed dormancy by antagonizing ABA biosynthesis and signaling, thereby promoting seed germination [73]. In Arabidopsis, *Delay of Germination1* (*DOG1*) interacts with *ethylene responsive factor 12* (*ERF12*) and co-regulate seed dormancy [74]. High transcript levels of ethylene-responsive transcription factors in seed and embryo of ICGV 91114 (*Arachis hypogaea* sub spp. *fastigiata*) gene expression atlas revealed its functional relevance in seed germination [8, 75]. In plants, F-box proteins regulate various physiological processes in various ways. Song et al. [76] had isolated *OsFbx352* gene (encoding for F-box domain protein) from rice to characterize its role in germination. Overexpression of *OsFbx352*, (a F-box protein in rice) demonstrated lower ABA contents by reduced expression of ABA synthesis genes (*OsNced2*, *OsNced3*) and increased ABA catabolism gene expression (*OsAba-ox2*, *OsAba-ox3*) leading to seed germination. Whereas, knockdown of *OsFbx352* led to higher ABA contents and suppressed seed germination, revealing its regulatory role in modulating ABA metabolism. In groundnut, high gene expression value of *F-box/RNI-like superfamily protein* (*Arahy.LZ56CD*) in all the six selected tissues of seed and pod of non-dormant ICGV 91114 also indicated its positive regulatory role in germination [8]. In Arabidopsis, transgenic plants overexpressing for *AtTLP9* (*Tubby-F-box like protein*) also demonstrated hypersensitiveness to ABA [77]. As all these genes were explored for their involvement in ABA/GA/Ethylene signaling pathways, and identified as

promising candidates in the hormonal regulation of dormancy and germination.

The allele calls of the representative panel of mini-core accessions for the identified significant STAs revealed that presence of all favorable dormant alleles in an accession result in increased dormancy duration. Conversely, an increase in the number of unfavorable non-dormant alleles corresponded to a decreased dormancy duration. Based on this, allele-specific markers were developed for these STAs, where three of them could successfully differentiate dormant and non-dormant genotypes. These markers can be validated on much larger panel of breeding material to confirm its efficacy to use in MAS for cultivar development. Additionally, these STAs show potential for developing Kompetitive allele specific PCR (KASP) assays, and accessions with all the favorable dormant alleles can be used as donors in marker-assisted back-cross (MABC) programs aiming for dormancy of ≥30 days. However, if the goal is to achieve FSD of 14–21 days, accessions with a combination of dormant and non-dormant alleles for these markers can be utilized. Thus, these assays could play a crucial role in future molecular breeding programs, offering targeted and efficient means to select lines with 2–3 weeks of dormancy.

Conclusion

With a diverse panel of 184 groundnut mini-core accessions, this study employed a combined approach of ML-GWAS and SL-GWAS models and identified 9 and 38 STAs for FSD, respectively. Examining the vicinity of these STAs revealed potential candidate genes; *Cytochrome P450 705 A*, *Dormancy/auxin associated family protein*, *WRKY family transcription factor*, *Protein kinase superfamily protein*, *serine/threonine protein phosphatase*, *myb transcription factor*, *transcriptional regulator STERILE APETALA-like*, *ethylene-responsive transcription factor 7-like and F-box protein interaction domain protein* to be involved in ABA/GA associated pathways. This investigation underscores dormancy as a complex trait controlled by multiple genes, highlighting the importance of understanding gene interactions across multiple biological pathways. Furthermore, examination of allelic calls in the mini-core accessions for the identified STAs revealed the intricate regulation of dormancy and germination through the expression and suppression of multiple genes in diverse combinations, affecting dormancy duration. This suggests that successful molecular breeding strategies must incorporate multiple genes from various biological pathways. Further characterization of the candidate genes identified in this study is recommended through overexpression studies and the CRISPR/Cas9 approach, which may provide more insights on precise function of these candidate genes in FSD. Additionally, haplotype analysis for the identified candidate genes can aid in identifying superior haplotypes for FSD trait, which could be utilized in haplotype-based breeding.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-024-05897-6.

Supplementary Material 1

Acknowledgements

The authors are thankful to ICRISAT Gene-Bank for providing seed material. DB acknowledges Acharya N.G. Ranga Agricultural University for collaborating with ICRISAT and opportunity given as a PhD student to pursue this investigation at ICRISAT. VS acknowledges the Council of Scientific and Industrial Research (CSIR), Govt. of India for the award of SRF-Direct fellowship (File No: 09/0800(18433)/2024-EMR-I) for PhD.

Author contributions

MKP conceived the idea and supervised and finalized the manuscript. KS and RS contributed to providing seed material and in-seed multiplication of the mini-core collection. DB, VS, SSG, RK phenotyped the mini-core collection. DB and VS performed the analysis and drafting the manuscript. DB, VS, SSG contributed in improvising figures. KMD and SSG designed primers and performed validation. MR, VS, SSG, SKB, MKP contributed to reviewing and improving the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

The authors are thankful to the Indian Council of Agricultural Research (ICAR) through ICAR-ICRISAT collaborative project, Department of Biotechnology (DBT), Government of India through Indo-UK (Newton-Bhabha) project, and Bill & Melinda Gates Foundation (BMGF), USA through Tropical Legumes III project.

Data availability

Genotyping data generated using 58 K Axiom_*Arachis* SNP array is provided in the Table S8.

Declarations

Ethics approval and consent to participate

This study did not involve any human or animal research participant data. Clinical trial number not applicable. The plant samples and soil samples were collected from the fields with permission from field owners.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 26 August 2024 / Accepted: 29 November 2024
Published online: 27 December 2024

References

1. Nautiyal PC, Sivasubramaniam K, Dadlani M. Seed dormancy and regulation of germination. Seed Science and Technology; 2023. pp. 39–66.
2. Koornneef M, Bentsink L, Hilhorst H. Seed dormancy and germination. Curr Opin Plant Biol. 2002;5(1):33–6
3. Bomireddy D, Sharma V, Senthil R, Reddisekhar M, Shah P, Singh K, et al. Identification of donors for fresh seed dormancy and marker validation in a diverse groundnut mini-core collection. Agronomy. 2024;14(1):112.
4. Naganagoudar YB, Kenchanagoudar PV, Rathod S, Keerthi CM, Nadaf HL, Channappagoudar BB. Inheritance of fresh seed dormancy in recombinant inbred lines (RIL) developed for mapping population TAG 24 x GPBD 4 in groundnut (*Arachis hypogea* L). Legume Research-An Int J. 2016;39(5):844–6.

5. Nautiyal PC, Bandyopadhyay A, Zala PV. In situ sprouting and regulation of fresh-seed dormancy in Spanish type groundnut (*Arachis hypogaea* L). Field Crops Res. 2001;70(3):233–41.
6. Rajan D, Kumar R, Devyani K, Ramya MJ. Role of ethrel and storage in dormancy breaking in groundnut (*Arachis hypogaea* L). Int J Curr Microbiol App Sci. 2020;9(2):2550–64
7. Yaw AJ, Richard A, Safo-Kantanka O, Adu-Dapaah HK, Ohemeng-Dapaah S, Agyeman A. Inheritance of fresh seed dormancy in groundnut. Afr J Biotechnol. 2008;7(4):421–425.
8. Bomireddy D, Gangurde SS, Variath MT, Janila P, Manohar SS, Sharma V, et al. Discovery of major quantitative trait loci and candidate genes for fresh seed dormancy in groundnut. Agronomy. 2022;12(2):404.
9. Zhang M, Zeng Q, Liu H, Qi F, Sun Z, Miao L, et al. Identification of a stable major QTL for fresh-seed germination on chromosome Arahy.04 in cultivated peanut (*Arachis hypogaea* L). Crop J. 2022;10(6):1767–73
10. Khalfaoui JB. Inheritance of seed dormancy in a cross between two Spanish peanut cultivars. Peanut Sci. 1991;18(2):65–7
11. Deshmukh D, Marathi B, Sudini HK, Variath MT, Chaudhari S, Manohar SS, et al. Combining high oleic acid trait and resistance to late leaf spot and rust diseases in groundnut (*Arachis hypogaea* L). Front Genet. 2020;11:514.
12. Janila P, Pandey MK, Shasidhar Y, Variath MT, Sriswathi M, Khera P, et al. Molecular breeding for introgression of fatty acid desaturase mutant alleles (ahFAD2A and ahFAD2B) enhances oil quality in high and low oil containing peanut genotypes. Plant Sci. 2016;242:203–13
13. Pandey MK, Gangurde SS, Shasidhar Y, Sharma V, Kale SM, Khan AW, et al. High-throughput diagnostic markers for foliar fungal disease resistance and high oleic acid content in groundnut. BMC Plant Biology. 2024;24(1):262
14. Shasidhar Y, Variath MT, Vishwakarma MK, Manohar SS, Gangurde SS, Sriswathi M, et al. Improvement of three popular Indian groundnut varieties for foliar disease resistance and high oleic acid using SSR markers and SNP array in marker-assisted backcrossing. Crop J. 2020;8(1):1–15
15. Varshney RK, Pandey MK, Janila P, Nigam SN, Sudini H, Gowda MVC, et al. Marke r-assisted introgression of a QTL region to improve rust resistance in three elite and popular varieties of peanut (*Arachis hypogaea* L). Theor Appl Genet. 2014;127:1771–81
16. Yeri SB, Bhat RS. Development of late leaf spot and rust resistant backcross lines in JL24 variety of groundnut (*Arachis hypogaea* L). Electron J Plant Breed. 2016;7:37–41
17. Chen D, Zou W, Zhang M, Liu J, Chen L, Peng T, et al. Genome-wide association study for seed dormancy using re-sequenced germpl asm under multiple conditions in rice. Int J Mol Sci. 2023;24(7):6117
18. Lee GA, Jeon YA, Lee HS, Hyun DY, Lee JR, Lee MC, et al. New genetic loci associated with preharvest sprouting and its evaluation based on the model equation in rice. Front Plant Sci. 2017;8:1393
19. Guo G, Xu S, Chen H, Hao Y, Mao H. QTL mapping for wheat seed dormancy in a yangmai16/zhongmai895 double haploid population. Plants. 2023;12(4):759
20. Ogbonnaya FC, Imtiaz M, Ye G, Hearnden PR, Hernandez E, Eastwood RF, et al. Genetic and QTL analyses of seed dormancy and preharvest sprouting resistance in the wheat germplasm CN10955. Theoretical and Applied Genetics. 2008;116:891–902
21. Li C, Ni P, Francki M, Hunter A, Zhang Y, Schibeci D, et al. Genes controlling seed dormancy and pre-harvest sprouting in a rice-wheat-barley comparison. Volume 4. Functional & Integrative Genomics. 2004:84–93
22. Nakamura S, Pourkheirandish M, Morishige H, Sameri M, Sato K, Komatsuda T. Quantitative trait loci and maternal effects affecting the strong grain dormancy of wild barley (*Hordeum vulgare ssp. spontaneum*). Front Plant Sci. 2017;8:301809
23. Nakamura S, Pourkheirandish M, Morishige H, Kubo Y, Nakamura M, Ichimura K, et al. Mitogen-activated protein kinase kinase 3 regulates seed dormancy in barley. Curr Biol. 2016;26(6):775–81
24. Sato K, Yamane M, Yamaji N, Kanamori H, Tagiri A, Schwerdt JG, et al. Alanine aminotransferase controls seed dormancy in barley. Nat Communications. 2016;7(1):p11625
25. Nakamura S, Abe F, Kawahigashi H, Nakazono K, Tagiri A, Matsumoto T, et al. A wheat homolog of MOTHER OF FT AND TFL1 acts in the regulation of germination. Plant Cell. 2011;23(9):3215–29
26. Torada A, Koike M, Ogawa T, Takenouchi Y, Tadamura K, Wu J, et al. A causal gene for seed dormancy on wheat chromosome 4A encodes a MAP kinase kinase. Curr Biol. 2016;26(6):782–7
27. Kumar R, Janila P, Vishwakarma MK, Khan AW, Manohar SS, Gangurde SS, et al. Whole-genome resequencing-based QTL-seq identified candidate genes and molecular markers for fresh seed dormancy in groundnut. Plant Biotechnol J. 2020;18(4):992–1003
28. Vishwakarma MK, Pandey MK, Shasidhar Y, Manohar SS, Nagesh P, Janila P, et al. Identification of two major quantitative trait locus for fresh seed dormancy using the diversity arrays technology and diversity arrays technology-seq based genetic map in Spanish-type peanuts. Plant Breeding. 2016;135(3):367–75
29. Wang ML, Wang H, Zhao C, Tonnis B, Tallury S, Wang X, et al. Identification of QTLs for seed dormancy in cultivated peanut using a recombinant inbred line mapping population. Plant Mol Biology Report. 2022;40(1):208–17
30. Zhang X, Zhang X, Wang L, Liu Q, Liang Y, Zhang J, et al. Fine mapping of a QTL and identificat ion of candidate genes associated with cold tolerance during germination in peanut (*Arachis hypogaea* L.) on chromosome B09 using whole genome re-sequencing. Front Plant Sci. 2023;14:1153293
31. Lu Q, Niu X, Zhang M, Wang C, Xu Q, Feng Y, et al. Genome-wide association study of seed dormancy and the genomic consequences of improvement footprints in rice (*Oryza sativa* L). Front Plant Sci. 2018;8:2213
32. Gangurde SS, Wang H, Yaduru S, Pandey MK, Fountain JC, Chu Y, et al. Nested-association mapping (NAM) based genetic dissection uncovers candidate genes for seed and pod weights in peanut (*Arachis hypogaea* L). Plant Biotechnol J. 2020;18:1457–71
33. Gangurde SS, Thompson E, Yaduru S, Wang H, Fountain JC, Chu Y, et al. Linkage Mapping and Genome-Wide Association Study Identified Two Peanut Late Leaf Spot Resistance Loci, PLLSR-1 and PLLSR-2, Using Nested Association Mapping. Phytopathology. 2024;114(6):1346–55.
34. Pandey MK, Upadhyaya HD, Rathore A, Vadez V, Sheshshayee MS, Sriswathi M, et al. Genome-wide association studies for 50 agronomic traits in peanut using the 'reference set' comprising 300 genotypes from 48 countries of the semi-arid tropics of the world. PLoS One. 2014;9(8):105228
35. Zhong H, Liu S, Sun T, Kong W, Deng X, Peng Z, et al. Multi-locus genome-wide association studies for five yield-related traits in rice. BMC Plant Biology. 2021;21(1):364
36. Bhandari A, Sandhu N, Bartholome J, Cao-Hamadoun TV, Ahmadi N, Kumari N, et al. Genome-wide association study for yield and yield related traits under reproductive stage drought in a diverse indica-aus rice panel. Rice. 2020;13:1–22
37. Zhang YM, Mao Y, Xie C, Smith H, Luo L, Xu S. Mapping quantitative trait loci using naturally occurring genetic variance among commercial inbred lines of maize (*Zea mays* L). Genetics. 2005;169(4):2267–75
38. Wang SB, Feng JY, Ren WL, Huang B, Zhou L, Wen YJ, et al. Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. Sci Rep. 2016;6(1):19444
39. Cui Y, Zhang F, Zhou Y. The application of multi-locus GWAS for the detection of salt-tolerance loci in rice. Front Plant Sci. 2018;9:1464
40. Upadhyaya HD, Bramel PJ, Ortiz R, Singh S. Developing a mini core of peanut for utilization of genetic resources. Crop Sci. 2002;42(6):2150–6
41. Upadhyaya HD, Nigam SN. Inheritance of fresh seed dormancy in peanut. Crop Sci. 1999;39(1):98–101
42. Clevenger J, Chu Y, Chavarro C, Agarwal G, Bertioli DJ, Leal-Bertioli SC, et al. Genome-wide SNP genotyping resolves signatures of selection and tetrasomic recombination in peanut. Mol Plant. 2017;10(2):309–22
43. Pandey MK, Agarwal G, Kale SM, Clevenger J, Nayak SN, Sriswathi M, et al. Development and evaluation of a high density genotyping 'Axiom_ Arachis' array with 58 K SNPs for accelerating genetics and breeding in groundnut. Sci Rep. 2017;7(1):40577
44. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, et al. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics. 2007;23(19):2633–5
45. Tamba CL, Zhang YM. A fast mrMLM algorithm for multi-locus genome-wide association studies. Biorxiv. 2018;341784
46. Wen YJ, Zhang H, Ni YL, Huang B, Zhang J, Feng JY, et al. Methodological implementation of mixed linear models in multi-locus genome-wide association studies. Brief Bioinform. 2018;19(4):700–12
47. Zhang J, Feng JY, Ni YL, Wen YJ, Niu Y, Tamba CL, et al. PLARmEB:integration of least angle regression with empirical Bayes for multilocus genome-wide association studies. Heredity. 2017;118(6):517–24
48. Ren WL, Wen YJ, Dunwell JM, Zhang YM. pKWmEB: integration of Kruskal–Wallis test with empirical Bayes under polygenic back ground control for multi-locus genome-wide association study. Heredity. 2018;120(3):208–18
49. Tamba CL, Ni YL, Zhang YM. Iterative sure independence screening EM-Bayesian LASSO algorithm for multi-locus genome-wide association studies. PLoS Comput Biol. 2017;13(1):1005357

50. Sohindji FS, Sogbohossou DE, Zohoungbogbo HP, Houdegbe CA, Achigan-Dako EG. Understanding molecular mechanisms of seed dormancy for improved germination in traditional leafy vegetables: An overview. Agronomy. 2020;10(1):57
51. He L, Xiao J, Rashid KY, Yao Z, Li P, Jia G, et al. Genome-wide association studies for pasmo resistance in flax (*Linum usitatissimum* L). Front Plant Sci. 2019;9:1982
52. Odesola KA, Olawuyi OJ, Paliwal R, Oyatomi OA, Abberton MT. Genome-wide association analysis of phenotypic traits in Bambara groundnut under drought-stressed and non-stressed conditions based on DArTseq SNP. Front Plant Sci. 2023;14:1104417
53. Liu X, Huang M, Fan B, Buckler ES, Zhang Z. Iterative usage of fixed and random effect models for powerful and efficient genome- wide association studies. PLoS Genet. 2016;12(2):1005767
54. Hu X, Zuo J, Wang J, Liu L, Sun G, Li C, et al. Multi-locus genome-wide association studies for 14 main agronomic traits in barley. Front Plant Sci. 2018;9:1683
55. Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, et al. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat Genetics. 2012;44(7):825–30
56. Liu S, Zhong H, Meng X, Sun T, Li Y, Pinson SR, et al. Genome-wide association studies of ionomic and agronomic traits in USDA mini core collection of rice and comparative analyses of different mapping methods. BMC Plant Biol. 2020;20:1–18
57. Cheng J, Wang L, Du W, Lai Y, Huang X, Wang Z, et al. Dynamic quantitative trait locus analysis of seed dormancy at three development stages in rice. Mol Breed. 2014;34:501–10
58. Magwa RA, Zhao H, Xing Y. Genome-wide association mapping revealed a diverse genetic basis of seed dormancy across subpopulations in rice (*Oryza sativa* L). BMC Genet. 2016;17:1–13
59. Lin M, Zhang D, Liu S, Zhang G, Yu J, Fritz AK, et al. Genome-wide association analysis on pre-harvest sprouting resistance and grain color in US winter wheat. BMC Genomics. 2016;17:1–16
60. Shao M, Bai G, Rife TW, Poland J, Lin M, Liu S, et al. QTL mapping of pre-harvest sprouting resistance in a white wheat cultivar Danby. Theor Appl Genet. 2018;131:1683–97
61. Lin M, Cai S, Wang S, Liu S, Zhang G, Bai G. Genotyping-by-sequencing (GBS) identified SNP tightly linked to QTL for pre-harvest sprouting resistance. Theor Appl Genet. 2015;128:1385–95
62. Kushiro T, Okamoto M, Nakabayashi K, Yamagishi K, Kitamura S, Asami T, et al. The Arabidopsis cytochrome P450 CYP707A encodes ABA 8′-hydroxylases: key enzymes in ABA catabolism. EMBO J. 2004;23(7):1647–56
63. Ding ZJ, Yan JY, Li GX, Wu ZC, Zhang SQ, Zheng SJ. WRKY 41 controls Arabidopsis seed dormancy via direct regulation of ABI 3 transcript levels not downstream of ABA. Plant J. 2014;79(5):810–23
64. Fujii H, Verslues PE, Zhu JK. Identification of two protein kinases required for abscisic acid regulation of seed germination, root growth, and gene expression in Arabidopsis. Plant Cell. 2007;19(2):485–94
65. Nakashima K, Fujita Y, Kanamori N, Katagiri T, Umezawa T, Kidokoro S, et al. Three Arabidopsis SnRK2 protein kinases, SRK2D/SnRK2. 2, SRK2E/SnRK2. 6/OST1 and SRK2I/SnRK2. 3, involved in ABA signaling are essential for the control of seed development and dormancy. Plant Cell Physiol. 2009;50(7):1345–63
66. Liu X, Zhang H, Zhao Y, Feng Z, Li Q, Yang HQ, et al. Auxin controls seed dormancy through stimulation of abscisic acid signaling by inducing ARF-mediated ABI3 activation in Arabidopsis. Proc Natl Acad Sci. 2013;110(38):15485–90
67. Gosti F, Beaudoin N, Serizet C, Webb AA, Vartanian N, Giraudat J. ABI1 protein phosphatase 2 C is a negative regulator of abscisic acid signaling. Plant Cell. 1999;11(10):1897–909
68. Kim W, Lee Y, Park J, Lee N, Choi G. HONSU, a protein phosphatase 2 C, regulates seed dormancy by inhibiting ABA signaling in A rabidopsis. Plant Cell Physiol. 2013;54(4):555–72
69. Saez A, Apostolova N, Gonzalez-Guzman M, Gonzalez-Garcia MP, Nicolas C, Lorenzo O, et al. Gain-of-function and loss-of-function phenotypes of the protein phosphatase 2 CHAB1 reveal its role as a negative regulator of abscisic acid signalling. Plant J. 2004;37(3):354–69
70. Lee HG, Lee K, Seo PJ. The Arabidopsis MYB96 transcription factor plays a role in seed dormancy. Plant Mol Biol. 2015;87:371–81
71. Shu K, Liu XD, Xie Q, He ZH. Two faces of one seed: hormonal regulation of dormancy and germination. Mol Plant. 2016;9(1):34–45
72. Yaish MW, El-Kereamy A, Zhu T, Beatty PH, Good AG, Bi YM, et al. The APETALA-2-like transcription factor OsAP2-39 controls key interactions between abscisic acid and gibberellin in rice. PLoS Genet. 2010;6(9):1001098
73. Corbineau F, Xia Q, Bailly C, El-Maarouf-Bouteau H. Ethylene, a key factor in the regulation of seed dormancy. Front Plant Sci. 2014;5:113486
74. Li F, Wen W, Liu J, Zhang Y, Cao S, He Z, et al. Genetic architecture of grain yield in bread wheat based on genome-wide association studies. BMC Plant Biol. 2019;19:1–19
75. Sinha P, Bajaj P, Pazhamala LT, Nayak SN, Pandey MK, Chitikineni A, et al. Arachis hypogaea gene expression atlas for fastigiate subspecies of cultivated groundnut to accelerate functional and translational genomics applications. Plant Biotechnol J. 2020;18(11):2187–200
76. Song S, Dai X, Zhang WH. A rice F-box gene, Os Fbx352, is involved in glucose-delayed seed germination in rice. J Exp Bot. 2012;63(15):5559–68
77. Lai CP, Lee CL, Chen PH, Wu SH, Yang CC, Shaw JF. Molecular analyses of the Arabidopsis TUBBY-like protein gene family. Plant Physiol. 2004;134(4):1586–97

Publisher's note