Theoretical and Applied Genetics (2026) 139:100
https://doi.org/10.1007/s00122-026-05174-0

ORIGINAL ARTICLE

Understanding the genetics of quality traits in groundnut: GWAS highlights drought-responsive markers and candidate genes

Ashutosh Purohit[1,2] · Devraj Lenka[2] · Ankush Purushottam Wankhade[1] · Anilkumar Vemula[1] · Sunita Choudhary[1] · Dnyaneshwar Bandu Deshmukh[1] · Sunil S. Gangurde[1] · Gopi Potupureddi[1] · Manish K. Pandey[1] · Janila Pasupuleti[1]

Received: 31 January 2025 / Accepted: 23 January 2026

Abstract

Key message **The study identified novel, drought-responsive markers and potential candidate genes for kernel quality traits under drought stress which are important genomic resources to breed resilient and nutritious groundnut varieties.**

Abstract Nutritional quality traits are crucial in crop breeding, especially with climate change affecting quality and yield. Increasing dry spells in semiarid regions require understanding drought stress effects on oil, protein, and fatty acid contents of groundnut kernels, and the molecular mechanisms governing these traits. The study showed a reduction in oil content and increase in protein content in a MAGIC population of groundnut under drought stress. MAGIC population exhibited high diversity with 13,937 polymorphic markers and are valuable for the genetic dissection of complex traits as compared to a RIL population with 800 polymorphic markers. Genome-wide association study (GWAS) employing multi-locus models such as BLINK, and MLMM identified 45 significant MTAs for oil, protein, oleic acid, palmitic acid, stearic acid, and linoleic acid under normal and drought stress across all chromosomes except A06, B01, B02, B03, B04, B05, and B06. Fourteen SNPs are identified with > 10% PVE linked to six quality traits, seven pleiotropic SNPs associated with multiple traits, and nine SNPs in both models across at least in two environments. Potential candidate genes for quality traits under drought stress are *RING finger proteins* and *protein kinases* for oil content; *CASP-like protein* and *geranylgeranyl diphosphate reductase* (*GLDR*) for protein content, and *fatty acid desaturase 2*, *MYB transcription factor*, and *AMMECR1* family for oleic and palmitic acid content having role in various biosynthetic and lipid pathways. The MTAs identified in this study hold potential for development into assays for genomic-assisted selection to enhance nutritional quality and drought resilience in groundnut.

Introduction

Groundnut (*Arachis hypogaea* L.) is a widely cultivated oilseed crop in the arid and semiarid areas of Asia and Africa, where the annual precipitation is between 400 and 800 mm. The global area for groundnut cultivation covers 30.53 million hectares, yielding a total production of 54.23 million tons, resulting in a productivity of 1776.2 kg/ha (FAOSTAT, 2022). India, China, Nigeria, and Sudan have large cultivated area (5.70 M ha, 4.45 M ha, 3.40 M ha, and 3 M ha, respectively) (FAOSTAT, 2022). Asia and Africa collectively account for 90% of the world's groundnut production. Groundnut oil is a key agricultural product in the export market which has generated a revenue of 709.4 M USD globally in the year 2022 (FAOSTAT, 2022). Groundnut kernels have high-quality fat and protein content (Shen et al. 2016) and high nutritive value with 40–60% edible oil, 20–30% protein, and 10–20% carbohydrates (Janila et al. 2016; Yol et al. 2017) which is advantageous in processing and confectionary industry as compared to other legumes. In addition to these major components, groundnut kernels are composed of several minerals like Ca, Mg, K, Zn, Se, Fe, Cu, Mn, and total folate, vitamins such as vitamin B complex (B1, B3, B5, B6, and B9), E, K, dietary

Communicated by Elena Bitocchi.

✉ Janila Pasupuleti
janila.pasupuleti@icrisat.org

[1] International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad 502324, India

[2] Odisha University of Agriculture and Technology (OUAT), Bhubaneshwar 751003, India

Published online: 20 March 2026

fiber, tocopherol, resveratrol, beta-sitosterol, and coenzyme (Janila et al. 2013).

Groundnut is mostly cultivated in marginal lands as a rainfed crop, where drought stress during critical crop growth stages occurs frequently. Drought stress affects filling of kernels and thereby reduces shelling percentage, and kernel size and mass which are yield contributing traits. According to previous studies, drought stress can reduce yield by 33–70% (Pereira et al. 2016; Carvalho et al. 2017; Manjonda et al. 2018; Nassar et al. 2018; Abady et al. 2021a, b). Groundnut quality attributes are also impacted by drought stress in several ways. A notable decrease in oil content, along with an increase in protein and oleic acid contents, has been observed during drought conditions at the end of the growing season (Dwivedi et al. 1996; Pokhrel et al. 2025). Altered fatty acid composition due to drought conditions leads to reduced oil stability (Hashim et al. 1993). Groundnut kernel protein content increases under moderate drought conditions, but decreases when the stress is more severe, mainly due to changes in nitrogen metabolism (Songsri et al. 2009). While most research has primarily focused on yield effects, studies examined drought's impact on quality factors such as oil content, fatty acids, and protein, and the results are often unclear (Conkerton et al. 1989; Chakraborty et al. 2013; Solanki et al. 2024). Studies on these nutritional quality traits, which are key market traits, are limited compared to studies on drought tolerance mechanisms, and the varied responses of groundnut genotypes to drought are not well understood.

Global groundnut breeding programs have prioritized genetic enhancement for drought tolerance, leading to a deeper understanding of drought tolerance mechanisms. Significant advancements have been made in bringing drought-tolerant groundnut cultivars to market. Several drought-tolerant varieties have been released, including ICR 48, ICGV 00350, ICGV 87846,55–437, 55–21, 55–33, GC 8–35, SRV 1–3, and SRV 1–96 (Mayeux et al. 2003; Vindhiyavarman et al. 2014). ICGV 91114 is a short-duration, high-yield, drought-tolerant variety developed by ICRISAT that gained popularity in India (Pande et al. 2005). In the recent past, ICGV 15083, another ICRISAT-bred high oleic acid variety, which is increasingly gaining popularity among the community, has been identified to be drought-tolerant (Abady et al. 2021a, b). Historically, screening for drought tolerance has focused on selecting those with minimal yield reduction using an empirical approach (Blum 2010). To identify genotypes that combine resilience and superior agronomic performance, a selection strategy that utilizes a scoring scale-based method for assessing the Productivity Capacity Index (PCI) and Resilience Capacity Index (RCI), which aids in selecting high-yield, drought-tolerant genotypes was proposed (Thiry et al. 2016). However, under drought stress, yield protection remained a major focus (Janila et al. 2016), and effects on nutritional quality were not well addressed by the breeding program; consequently, protection of yield along with maintaining nutritional quality remained a challenge. The wet chemistry method to determine quality traits is tedious, labor-intensive, destructive and require significant resources, making it difficult for evaluating many genotypes in a breeding program. To address this challenge, proximate analysis through near-infrared spectroscopy (NIRS) has emerged as an essential tool in modern agriculture (Gullifa et al. 2023) for determining grain composition. NIRS has proved to be faster, efficient, and nondestructive alternative compared to traditional wet chemistry methods for estimating oil, protein, and fatty acid content in groundnuts (Kassie et al. 2024). NIRS offers the advantage of simultaneously analyzing multiple traits, thereby saving both time and cost compared to wet chemistry. However, accurate estimation of grain composition requires the use of robust and reliable reference calibrations or equations developed from spectral data and wet chemistry. The NIRS equations standardized at ICRISAT have shown a strong correlation with biochemical measurements, as demonstrated by high external validation coefficients of determination (r^2). For oleic acid, the r^2 value reached 0.96, indicating exceptional accuracy in predicting oleic acid content (Deshmukh et al. 2021).

There exists significant potential for the analysis of quantitative traits in multi-parental populations, such as multi-parental advanced generation intercross (MAGIC). To overcome the drawback of a biparental population, Mackay and Powell (2007) first suggested the use of a MAGIC population for crops. The MAGIC population can be utilized to boost mapping resolutions, recombination rates, and analysis of numerous alleles (Cavanagh et al. 2008). A range of crops including wheat (*Triticum aestivum* L.) (Stadlmeier et al. 2018), rice (*Oryza sativa* L.) (Meng et al. 2016), and maize (*Zea mays* L.) (Dell'Acqua et al. 2015) has access to several MAGIC populations. MAGIC populations are important genetic assets for using linkage and association mapping techniques to pinpoint quantitative trait loci (QTL) for complex traits (Nawade et al. 2024). The extensive number of recombinations that occur due to multiple rounds of intercrossing significantly enhances the QTL mapping and genome-wide association mapping, allowing for the identification of marker–trait associations with high precision (Hoque et al. 2020a, b; Arrones Olmo 2024). The MAGIC population has been developed in groundnut and utilized to study complex traits such as drought tolerance (Sharma et al. 2024) and resistance to late leaf spot (Wankhade et al. 2023) to identify genomic regions.

Genome-wide association studies (GWAS) are utilized to uncover statistically significant associations between single-nucleotide polymorphisms (SNPs) and particular traits throughout the genome. A study on QTL mapping was conducted using SSR markers in a RIL population derived from the cross of TAG 24 and GPBD 4. This investigation uncovered a significant QTL associated with oil content on chromosome A03 and another one related to protein content on chromosome A01, in addition to several QTLs with minor effects (Sarvamangala et al. 2011). A GWAS by Shaibu et al. (2019) involving 170 germplasm lines from ICRISAT found significant MTAs for fatty acids, minerals, and proximate compositions, which need validation. Additionally, a study of a US peanut minicore collection using a 58 K SNP array (Pandey et al. 2017) identified SNPs related to oleic acid, linoleic acid, protein content, and oil content across several chromosomes (Otyama et al. 2022). Research in crops like Indian mustard has revealed marker–trait associations for oil content under heat stress (Pal et al. 2024), while wheat studies have linked markers to yield components and physiological traits under drought and heat stress (Jha et al. 2020; Lal et al. 2022). We hypothesize that drought affects groundnut kernel quality traits such as oil, protein, and fatty acids. Therefore, conducting a GWAS in a genetically diverse MAGIC population, genotyped with high-density 48 k SNP array, can help in identifying stable loci associated with these traits, which will be useful for preserving quality traits under drought conditions.

Materials and methods

Plant material

A MAGIC population of 620 entries, including 8 founder parents and 12 checks, was screened under managed stress conditions at ICRISAT, India (17.3850° N, 78.4867° E and 545 m above mean sea level), during the post-rainy (PR) seasons of 2018–2019. In 2021–2022, 574 entries, including the checks and founders, were screened again. The founder parents, ICGV 91114, ICGV 06040, ICGV 00440, ICGV 00308, ICGV 05155, ICGV 88145, GPBD 4, and 55–437, include drought stress-tolerant lines (ICGV 91114, ICGV 00308, and 55–437) and parents with varying oil content (ICGV 00440 at 45% and ICGV 05155 at 55%) (Wankhade et al. 2023). The protein content of ICGV 91114 ranges from 26 to 28%, while ICGV 06040 contains approximately 25.2%, and ICGV 05155 has a protein content of 22% (Janila et al. 2015). Significant phenotypic variation is observed among the founders in the MAGIC population, with protein content varying between 22 and 28%. The palmitic acid content of ICGV 00440 is between 11.3 and 13.2% (Bera et al. 2018), which is considered a high level of palmitic acid. There is no substantial variation in stearic acid content among the founder parents. Two-way crosses involving the foundational parent plants were crossed during 2012–2013. This was succeeded by four-way and eight-way crosses in later seasons. Five groups of F_1s were progressed from F_2 to F_8 utilizing the single seed descent (SSD) methodology. For this study, 120 recombinant inbred lines (RILs) from each of the five groups were randomly chosen to create the MAGIC population (Fig. 1).

Managed stress environment (MSE) experiment

The study utilized an alpha lattice design with two replicates each under well-watered (WW) and water-stressed (WS) scenarios. The soil of experimental field was neutral (pH 7.43) with normal electrical conductivity (0.24 dS m^{-1}). Organic carbon content was moderate (0.40%). Available phosphorus, potassium, and calcium are at medium levels and magnesium is high, indicating overall adequate macronutrient status. Sulfur, zinc, boron, iron, copper, and manganese are within sufficient ranges, suggesting no major micronutrient constraints. The basal fertilizer dose of 60 kg phosphorus pentoxide (P_2O_5) was applied. Gypsum was applied @ 400 kg/ha at peak flowering stage of the crop as per recommended package of practice. Weather data for the duration of experimentation are provided in Table S4. Each experimental plot consisted of two rows measuring 2 m, which were spaced 30 cm apart, and plants were spaced 10 cm apart within those rows. Standard agricultural practices were followed to ensure a good crop stand. The experiment was irrigated using sprinklers for 20–25 days after planting (DAP), followed by drip irrigation until harvest, allowing better water control. Mid-season drought stress was induced in the WS plots by withholding 2–3 irrigations at 1000 growing degree days (GDD), and a single irrigation was provided during the period to prevent wilting. Drought stress was imposed at approximately 1000 GDD (base temperature 10 °C), which corresponds to the early pod development and pod filling stage in groundnut. This phenological stage is highly sensitive to water deficit and is critical for determining pod yield and kernel traits. The use of GDD instead of days after sowing helps in the synchronization of stress imposition at a particular crop stage at different locations and seasons (Reddy et al. 2003). Decisions to relieve stress were made based on weather information, soil moisture

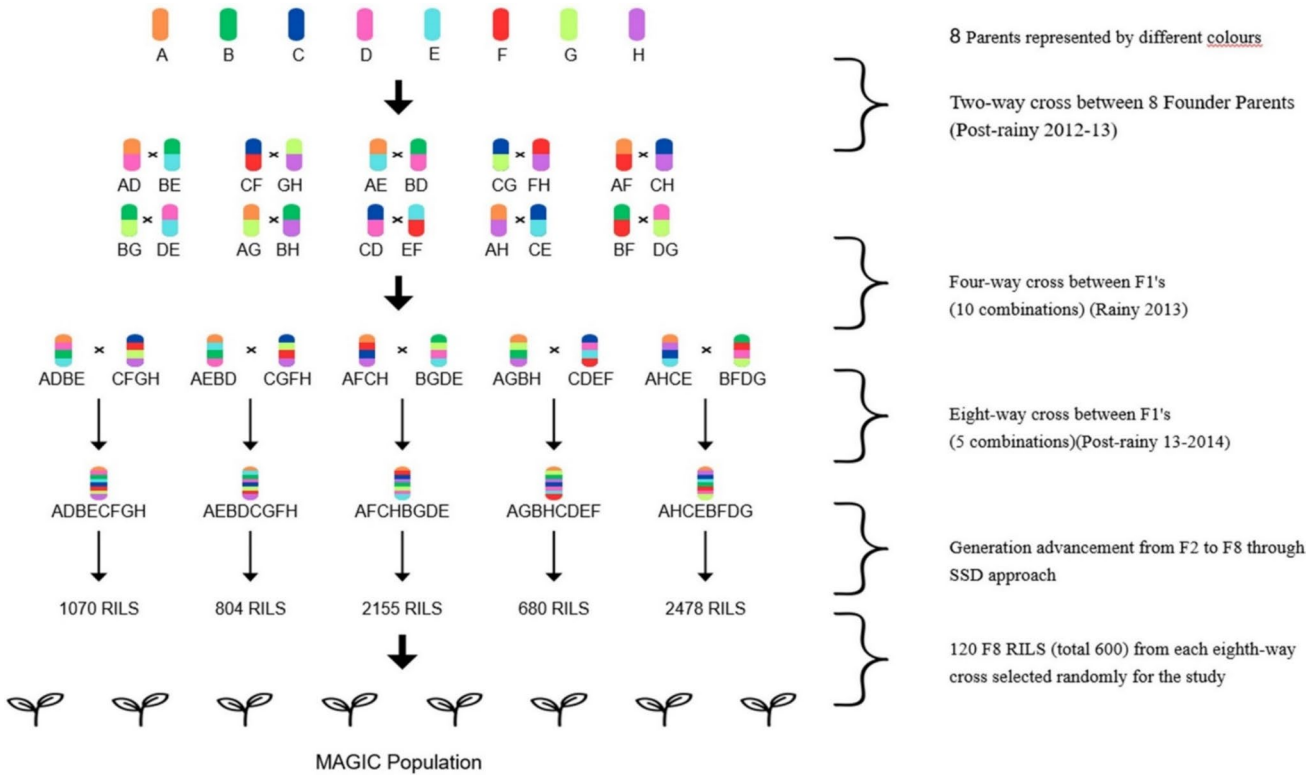

Fig. 1 Schematic diagram for MAGIC population development, **A** ICGV 91114, **B** ICGV 06040, **C** 55–437, **D** ICGV 00440, **E** ICGV 00308, **F** GPBD, **G** 05155, **H** ICGV 88145

content, visible signs of wilting, and the permanent wilting point (PWP) of 8.94% for alfisol, as established using a pressure plate extractor (Klute 2018). Soil moisture levels were monitored through neutron probes (Neutron Probe Smart503, ICT international) and Time Domain Reflectometry probes (TRIME®-FM). After harvest, pods were dried and shelled, and a full cup of kernels from each plot was analyzed in NIRS for oil, protein, oleic acid, linoleic acid, palmitic acid, and stearic acid.

Grain quality analyses using near-infrared spectroscopy (NIRS)

Kernel oil content, crude protein (nitrogen content × 6.25), oleic acid, palmitic acid, linolenic acid, and stearic acid (% of fat) were assessed. A representative whole kernel sample of 25–30 g per genotype from each replication in both well-watered and water-stressed condition was taken for quality analysis using NIRS, calibrated for these traits using conventional wet laboratory data (Sundaram et al. 2010). NIRS produces multiple spectral data for each sample and average spectral data is used. A NIRS instrument (FOSS DS 2500) with the software package Mosaic Solo and calibration was build using Calibrator Pro. Calibration was conducted on a dataset consisting of 315 lines of groundnut, from which 205 lines chosen for calibration and the remaining lines were utilized for validation, resulting in a Global H value of 1.3. The Kennard–Stone algorithm was employed to select samples for both the calibration and validation sets. The validation process involved blind predictions of laboratory measurements using the NIRS equations developed during the calibration phase. The relationships between the blind predictions and the actual measurements were assessed using the R^2 value and standard error of calibration (SEP), as detailed in Deshmukh et al. (2021). The current equation used for oil ($R^2 = 0.89$), protein ($R^2 = 0.83$), and palmitic ($R^2 = 0.80$) has a high values of coefficient of determination (R^2) in external validation. Calibration has been further improved by using ANN model and optimized for DS2500 NIRS model for oleic acid ($R^2 = 0.97$; SEC = 1.8) and linoleic acid ($R^2 = 0.88$; SEC = 2.04).

Statistical analysis

In the current study, 620 entries × 2 replications during 2018–2019 and 574 entries × 2 replications during 2021–2022, observed 4 plots per entry (2 replications each year) under well-watered and water-stressed condition. Treatment-wise

combined analysis of variance across years was conducted using SAS (v9.4) (SAS Institute Inc., 2015) mixed model approaches to evaluate the significance of main and interaction effects of year and genotype. The analysis treated year and replication as fixed factors, while block and genotype were considered random factors. Individual year variances were estimated and incorporated into the error distribution using the residual maximum likelihood (REML) method. The analysis yielded best linear unbiased predictors (BLUPs) for genotype and the year × genotype interaction, along with best linear unbiased estimates (BLUEs) for year. Broad-sense heritability (repeatability) (Falconer, 1989) was determined from the combined analysis using the formula:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \left(\sigma_{gxy}^2 / Y\right) + \left(\sigma_e^2 / r * Y\right)}$$

where $\sigma_g^2, \sigma_{gxy}^2, and \sigma_e^2$ are genetic variance component, genotype × year variance component, and residual variance, respectively; and Y and r are the number of years at which the trial was conducted and replications, respectively.

Isolation of genomic DNA and genotyping

To extract the genomic DNA of 600 MAGIC lines and 8 founding parents, a NucleoSpin® 96 Plant II Kit (Machery Nagel Düren, Germany; https://guest.link/UM6) was deployed. A NanoDrop 8000 Spectrophotometer (Thermo Fisher Scientific, Inc., Waltham, MA, USA) was used to quantify genomic DNA, and a 0.8% agarose gel was used for quality control. A 48 K Affymetrix SNP array ("Axiom_Arachis 2.0") was used for genotyping of extracted DNA samples. The output files (Cell Intensity File) from the Affymetrix were analyzed in Axiom Analysis Suite (AAS) v 5.2 (Thermo Fisher Scientific, Inc.). Axiom Analysis Suite integrates SNP genotyping, Indel detections, multi-allele analysis, off-target variants (OTVs) calling as well as copy number detection in a graphical interface.

Filtering of genotypic data and estimation of BLUPs for phenotypic data

Before importing the raw data file into Axiom Analysis Suite (AAS, Thermo Fisher Scientific), path for library files was set. AAS executes genotyping analysis as per best practices workflow, where it runs genotyping algorithms, and allows to view cluster graphs and export of data. Best practices workflow controls the quality with a dish QC value ≥ 0.82 and QC call rate of ≥ 97%. All markers were visually verified to inspect the quality of the cluster pattern. After filtering out 47,837 SNPs with a missing rate of > 10%, MAF < 0.05 and heterozygous call rate of 0.3, remaining 13,937 SNPs were used for GWAS analysis. Genotype imputation was not performed at any stage of the analysis, and missing genotypes remaining after quality filtering were retained as missing.

Comparison of recombination events between biparental and multi-parental population

The six recombinant inbred lines derived from a cross between Chico and ICGV 02251, and six MAGIC lines derived from an intercross between eight founders were used to compare the recombination breakpoints obtained in biparental mapping populations (BPP) and multi-parental mapping populations (MPP). The recombination breakpoints were analyzed using parent-specific SNPs for the parents in BPP and MPP. The parent-specific alleles were traced in the RILs and MAGIC lines. These were then visualized individually using circos plots (Krzywinski et al. 2009a, b).

Population structure and kinship matrix

ADMIXTURE v.1.3.0 (Alexander et al. 2009) was utilized to assess the population structure, with the results illustrated using R/pophelper (Francis 2017). Analyses were performed for K values ranging from 2 to 8, employing fivefold cross-validation to initialize the primary algorithm (QuasiNewton) for enhanced convergence speed. To ensure reliability and accuracy, 20 iterations were carried out for each K value, with the ideal number of clusters identified by the K value that showed the least cross-validation error. The "GAPIT 3.0" R package facilitated a kinship analysis through pairwise SNP comparisons. The resulting square symmetric matrix was subsequently employed to create a kinship plot, which illustrates the relationships among genotypes based on their similarity indices.

Linkage disequilibrium (LD), LD decay, PCA, and clustering of genotypes

"TASSEL v5.0" (Bradbury et al. 2007) was utilized to analyze the linkage disequilibrium (LD) among SNP pairs within a 50 Kb range and the R Studio script was used to create the LD decay graphs. LD decay was estimated by plotting pairwise r^2 values against physical distance across the genome. The "GAPIT 3.0" package in R was used to compute PCA, and a 3D plot with the determined PCs was produced (Wang et al. 2021). The neighbor joining approach in "TASSEL v5.0," which makes use of a dissimilarity matrix, was used to cluster the genotypes.

Genome-wide association study and identification of candidate genes

A genome-wide association study (GWAS) was conducted on nutritional traits: oil content (OIL), protein content (PRO), oleic acid content (OA), palmitic acid content (PAL), linoleic acid content (LA), and stearic acid content (SA) under six conditions (E1 = water-stressed 2018–2019, E2 = well-watered 2018–2019, E3 = water-stressed 2021–2022, E4 = well-watered 2021–2022, E5 = Pooled water-stressed, E6 = Pooled well-watered). Predicted means (BLUPs) were calculated using GAPIT 3 (Wang and Zhang 2021) in R, utilizing 13,937 highly polymorphic SNP markers from a 48 k SNP array. A total of 561 genotypes including 8 founder parents were utilized for GWAS after quality control for each of the above-mentioned six conditions. The analysis employed the BLINK and MLMM to control for population structure and kinship. A Bonferroni-corrected threshold (Bland and Altman 1995) of 5.445 was set to reduce Type I and II errors. QQ plots, Manhattan plots, and association tables assessed the results, with high-confidence MTAs defined as those explaining over 10% of phenotypic variance. The diploid ancestors of the cultivated peanut, *A. duranensis* and *A. ipaensis*, whose genomes are accessible in PeanutBase (<https://peanutbase.org/home>), were used to identify genes within LD decay region from the significant SNPs and perform functional annotations on them. Their corresponding tetraploid IDs were BLAST searched in peanut base.

Results

Evaluation of MAGIC population for quality traits under managed stress environment (MSE)

All the kernel quality traits, oil, protein, oleic, linoleic, palmitic, and stearic acids, varied significantly among genotypes across seasons and conditions. Overall, oil content was higher in well-watered conditions, while protein content tended to increase under water stress. Pooled data across two years of evaluation, heritability of nutritional traits, and other parameters are summarized in Table 1. There were significant differences between genotypes, and genotype × environment interactions were significant for all traits studied across the water regime, seasons, and pooled data (Table 2; Table S1). All traits showed a medium to high range of heritability across two water regimes, seasons, and pooled data. Across two years of evaluation, clear trends emerged in the nutritional quality of MAGIC lines under contrasting moisture regimes. The correlation matrix derived from the combined data indicated several notable positive correlations: protein with palmitic and oleic acid, oil with linoleic and stearic acid, stearic acid with linoleic acid, and palmitic acid with linoleic acid. In contrast, protein displayed significant negative correlations with linoleic acid, stearic acid, and oil. Oleic acid showed negative correlations with linoleic acid, palmitic acid, stearic acid, and oil as well. Moreover, significant negative correlations were found between oil and palmitic acid, as well as between stearic acid and palmitic acid. Additional negative correlations were noted between oleic acid and linoleic acid, oleic acid and palmitic acid, oleic acid and protein content, and oleic acid and oil content (Fig. S1).

Population structure, kinship matrix, linkage disequilibrium (LD), and SNP density analysis

Diversity analysis of this population in our previous studies (Wankhade et al. 2023) found that the MAGIC population was genetically diverse. Population structure analysis suggested the presence of three subpopulations. Clustering of genotypes formed three clusters, based on an unrooted NJ (neighbor joining) tree. Kinship analysis revealed varying degrees of relatedness among the genotypes, indicating broad genetic diversity and suitability of the population for genome-wide association analysis. The LD decay curve indicated that r^2 declined below 0.2 at approximately 3.52 Mb (Fig. 2). The maximum number of SNPs were in chromosome number B09 followed by B10 and B08, whereas the minimum SNPs were in A08. The number of SNPs from the A genome and B genome used for further analysis are 5381 and 8556, respectively. The SNP density plot suggests a smaller number of markers toward the centromere and more toward telomeres (Fig. S2).

Comparison of a RIL and MAGIC population for the number of recombination breakpoints

MAGIC populations exhibit high diversity and are valuable for the genetic dissection of complex traits compared to a biparental RIL population (Fig. 3). In the MAGIC population used for this study, a total of 1854 parent-specific SNPs were identified from 48 K SNP array data. Of these, a large number of 599 parent-specific SNPs were identified for ICGV 05155, a high oil-containing groundnut cultivar; 412 parent-specific SNPs identified in ICGV 91114, a drought-tolerant groundnut cultivar; 365 parent-specific SNPs were identified in ICGV00308; and 119 parent-specific SNPs were identified in GPBD4, a rust and LLS-resistant groundnut variety having resistant alleles from *A. cardenacii*. For

Table 1 Summary of kernel quality trait variation, heritability, and superior MAGIC lines under contrasting moisture regimes

Trait	Heritability	Well-watered range (%)	Water-stressed range (%)	Top MAGIC lines vs. check (Well-watered)	Top MAGIC lines vs. check (Water-stressed)
Oil	0.68	43.6 (ICGR171128) – 54.4 (ICGR171383)	44.0 (ICGR171545) – 51.0 (ICGR171211)	ICGR171383 (54.40), ICGR171412 (53.10), ICGR171461 (53.00), ICGR171470 (52.86), ICGR171445 (52.68) > ICGV06420 (50.36)	ICGR171211 (51.00), ICGR171454 (49.22), ICGR171466 (48.84), ICGR171470 (48.61), ICGR171461 (48.23) > ICGV03042 (47.04)
Protein	0.24	23.0 (ICGR171470) – 30.6 (ICGR171370)	27.3 (ICGR171358) – 31.0 (ICGR171549)	ICGR171370 (30.6), ICGR171175 (28.65), ICGR171277 (28.60), ICGR171022 (28.59), ICGR171402 (28.53) > JL 24 (27.67)	ICGR171549 (31.00), ICGR171246 (30.36), ICGR171578 (30.33), ICGR171045 (30.30), ICGR171582 (30.28) > GG20 (29.25)
Oleic acid	0.54	37.6 (ICGR 171461) – 50.7 (ICGR171478)	39.2 (ICGR 171002) – 55.8 (ICGR171211)	ICGR171478 (50.74), ICGR171388 (49.98), ICGR171503 (48.39), ICGR171473 (48.26), ICGR171128 (48.16) > 55–437 (47.08)	ICGR171211 (55.81), ICGR171429 (52.86), ICGR171315 (51.80), ICGR171144 (51.75), ICGR171454(51.33) > ICGS76 (49.64)
Linoleic acid	0.57	29.1 (ICGR171388) – 42.0 (ICGR171470)	26.3 (ICGR171315) – 36.9 (ICGR171470)	ICGR171470 (41.95), ICGR171461 (41.81), ICGR171393 (40.61), ICGR171383 (40.51), ICGR171457 (40.35) > ICGV03043 (40.16)	ICGR171470 (36.94), ICGR171002 (36.77), ICGR171386 (35.60), ICGR171085 (35.39), ICGR171458 (35.26) > ICGV03042 (34.06)
Palmitic acid	0.59	10.3 (ICGR171370) – 13.1 (ICGR171365)	9.6 (ICGR171211) – 12.3 (ICGR171002)	ICGR171365 (13.18), ICGR171279 (13.17), ICGR171584 (13.09), ICGR171326 (13.09), ICGR171233 (13.08) > ICGV05155 (12.99)	ICGR171002 (12.36), ICGR171527 (12.17), ICGR171066 (12.13), ICGR171088 (12.13), ICGR171386 (12.12) > ICGV03043 (12.10)
Stearic acid	0.64	2.2 (ICGR171572) – 3.5 (ICGR171461)	2.2 (ICGR171081) – 3.1 (ICGR171245)	ICGR171461 (3.53), ICGR171383 (3.33), ICGR171454 (3.24), ICGR171470 (3.12), ICGR171477 (3.03) > ICGV06040 (2.78)	ICGR171245 (3.11), ICGR171354 (3.07), ICGR171395 (3.02), ICGR171461 (3.01), ICGR1713545 (2.99) > GG20 (2.83)

the ICGV 88145 and ICGV 06040 genotypes, 151 parent-specific SNPs were identified. The least number of SNPs, 22 and 35 were specific to 55–437 and ICGV 00440, respectively. In RIL population, a total of 2156 parent-specific SNPs were identified from 58 K SNP array data (Gangurde et al. 2023a, b). Of these SNPs, 752 SNPs were specifically identified in high seed weight donor parent ICGV 02251, and 1404 parent-specific SNPs were identified for Chico that has small kernels. Recombination events were observed in the MAGIC lines on all the chromosomes except chromosome B08. The extensive recombination breakpoints across the chromosomes resulted in diverse MAGIC lines. Conversely, in the biparental population of Chico × ICGV 02251, recombination occurred only between the two parents.

Marker–trait associations (MTAs) for quality traits

A genome-wide association study (GWAS) was conducted over two years to examine six quality traits under two water regimes, using a high-density SNP array with 13,937 filtered SNPs. To minimize false discoveries, we utilized a false discovery rate (FDR) of 0.05. A Bonferroni correction set a significant threshold of -log10(p) value of 3.587E-6. We analyzed the data using two models: BLINK for rapid discovery and MLMM for multi-locus confirmation which provides a robust GWAS strategy that balances computational speed, statistical power, and control of false positives, especially for complex traits. Key findings from

Table 2 Treatment-wise combined analysis of variance across two years (2018/19 and 2021/22) for quality traits under water-stressed and well-watered conditions

Tratis		OIL		PRO		PAL		SA		OA		LA	
	Treatment Condition	WS	WW	WS	WW	WS	WW	WS	WW	WS	WW	WS	WW
Covariance Parameter	Component	Random effect											
Block (Rep Year)	Variance	0.079	0.157	0.156	0.048	0.001	0	0.004	0.00422	0.016	0	0.095	0
	probChisq	0.0409	0.0014	< .0001	0.0447	0.7184		< .0001	< .0001	0.8231		0.1035	
Genotype	Variance	0.585	2.074	0.244	0	0.038	0.1002	0.1002	0.02609	1.483	3.11	1.261	2.374
	probChisq	< .0001	< .0001	0.0062		0.0011	< .0001	< .0001	< .0001	< .0001	< .0001	< .0001	< .0001
Year × Genotype	Variance	0.579	0.569	0.323	1.03	0.056	0.0429	0.014	0.00538	2.166	0.78	1.231	0.802
	probChisq	0.0003	0.0022	0.0098	< .0001	0.0002	0.0450	0.0003	< .0001	< .0001	0.1062	< .0001	0.0188
Experiment		Residuals											
Year_2018/19	Variance	4.935	5.337	3.483	6.038	0.517	0.441	0.126	0.103	9.193	11.72	7.650	8.3
Year_2021/22	Estimate	2.917	3.402	2.623	1.199	0.259	0.65	0.062	0.0508	6.150	12.38	4.173	7.859
Fixed term		Fixed (Type-III) effect											
Year	F statistic	54.380	20.5	107.690	50.39	1156.010	497.72	10.670	0.81	212.260	54.61	200.300	436.86
	prob F	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.001	0.37	< 0.001	< 0.001	< 0.001	< 0.001
Rep (Year)	F statistic	0.940	1.59	5.290	3.63	0.230	114.07	4.190	3.94	0.950	78.76	2.680	12.66
	prob F	0.393	0.209	0.006	0.03	0.798	< 0.001	0.018	0.022	0.391	< 0.001	0.074	< 0.001
	CV %	4.31	4.43	5.97	6.91	5.56	6.11	11.64	10.91	6.04	7.91	7.73	8.09
	Repeatability	0.32	0.60	0.21	0.00	0.23	0.39	0.23	0.54	0.33	0.48	0.38	0.50

CV%: percentage of coefficient of variance, OIL: oil content, PRO: protein content, PAL: palmitic acid content, SA: stearic acid content, OA: oleic acid content, LA: linoleic acid content

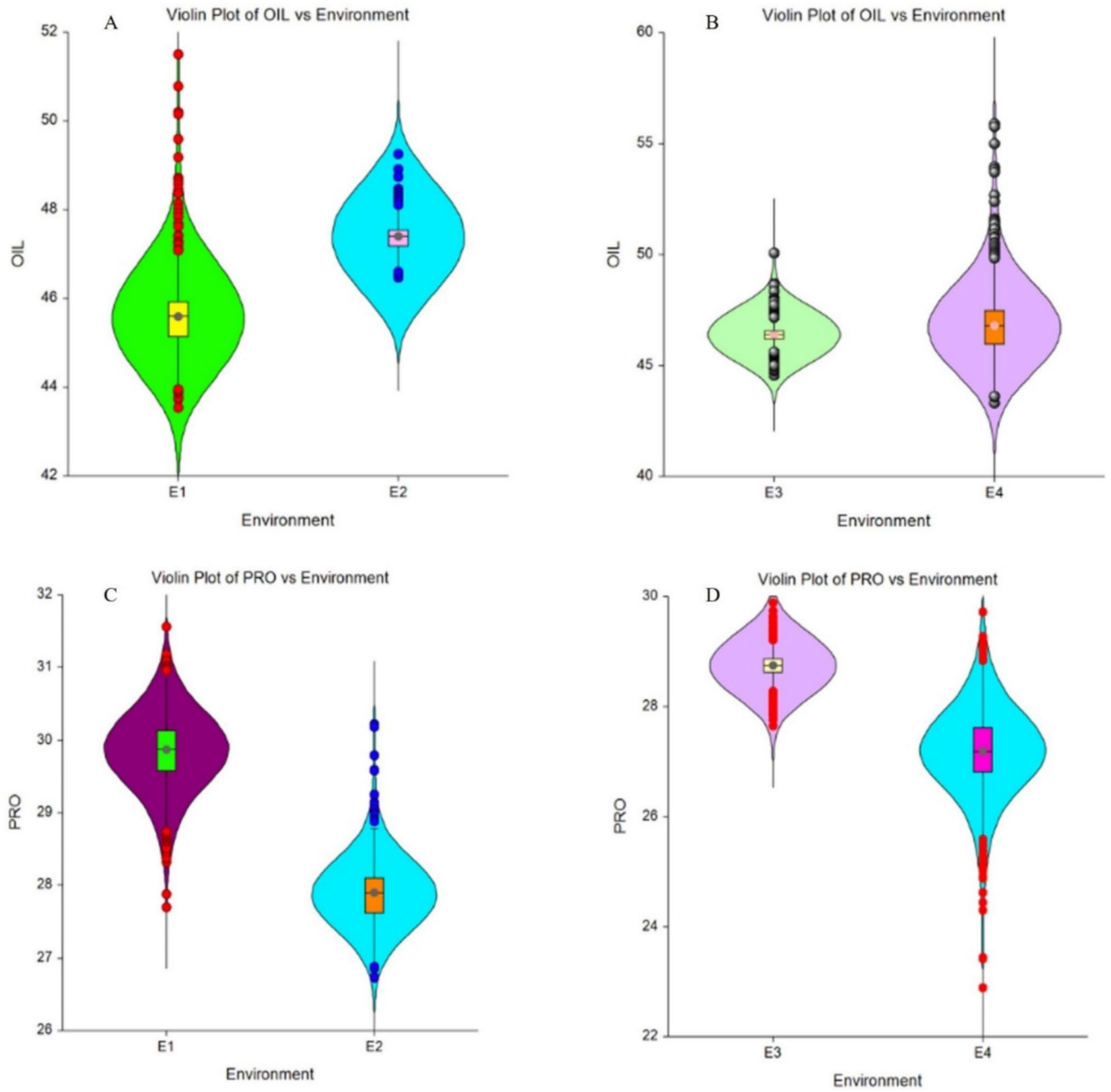

Fig. 2 Variation of oil and protein content in different water regime, **A** water-stressed condition 2018–2019**, B** well-watered condition 2018–2019, **C** water-stressed condition 2021–2022, **D** well-watered condition 2021–2022

the GWAS revealed 45 significant SNPs associated with six quality traits across all chromosomes except A06, B01, B02, B03, B04, B05, and B06 (Fig. 4; Fig. S3; Table S2), with BLINK outperforming MLMM in SNP detection. BLINK detected 40 MTAs, while MLMM detected 13 MTAs, with partial overlap between models. A total of 14 major SNPs were identified as significantly associated with six traits in this study, each contributing over 10% to phenotypic variance (Table 3).

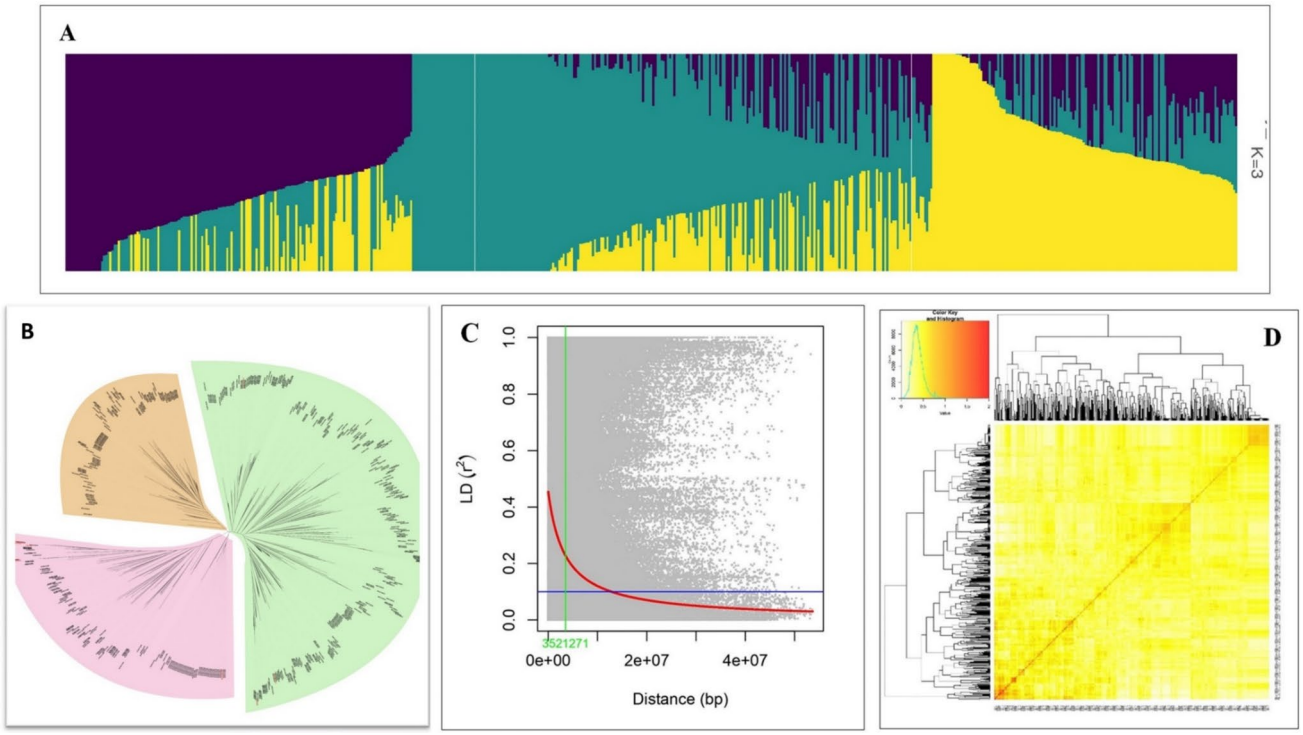

Fig. 3 **A** Population structure of MAGIC, **B** unrooted NJ tree, **C** LD decay plot, **D** kinship matrix

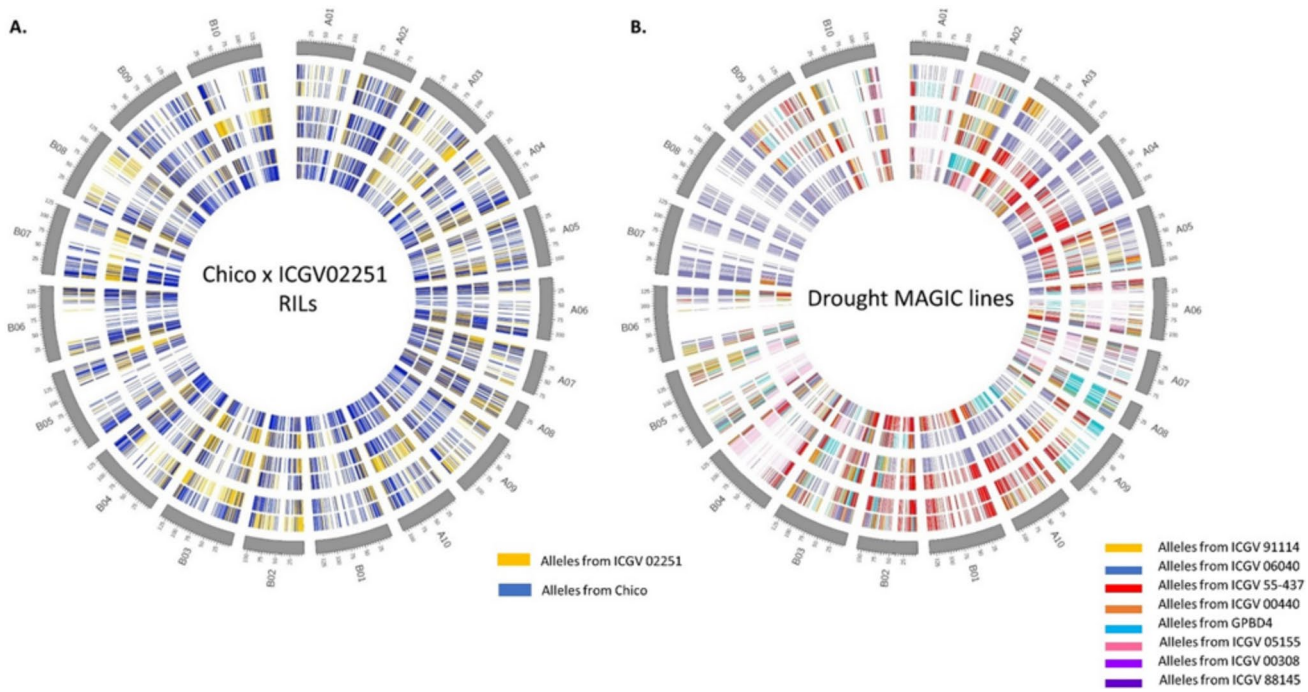

Fig. 4 Comparison of recombination breakpoints between set of recombinant inbred lines from biparental and multi-parental populations. **A** Recombination breakpoints in the 6 recombinant inbred lines (RIL32, RIL87, RIL134. RIL208, RIL290 and RIL364) derived from the cross Chico × ICGV 02251. Where, two color schemes were used to illustrate the alleles from two parents. **B** Recombination breakpoints in the 6 MAGIC lines (ICGR171586, ICGR171157, ICGR171594, ICGR171567, ICGR171226, ICGR171493) derived from the cross between 8 founder parents

Table 3 List of high-confidence SNPs (PVE > 10%) associated with quality traits

Trait	SNP ID	Chr	Environment	Stable across ≥ 2 environments		Detected in (BLINK/MLMM/Both)	Position (bp)	*p* value	MAF	Effect	r2 (%)
				Yes	No						
OIL	AX_176817775	B06	E1	–	No	MLMM	129,495,569	2.61E-13	0.13	-0.81	19.73
OIL	AX_177637369	B10	E1, E5	Yes	-	MLMM	127,616,318	4.50E-20	0.17	0.96	26.48
OIL	AX_176821854	A05	E4	–	No	MLMM	7,556,626	8.62E-07	0.08	0.62	10.09
OIL, PRO	AX_177638445	A10	E1, E4	Yes	–	Both	101,203,312	2.87E-06	0.07	0.64	17.03
PRO	AX_176811969	A01	E1	–	No	BLINK	99,527,079	5.73E-11	0.05	0.30	12.55
PRO	AX_147209601	A01	E2	–	No	MLMM	9,873,427	3.47E-06	0.17	0.20	15.31
PRO	AX_176794206	A01	E5	–	No	BLINK	106,097,712	2.28E-08	0.28	-0.09	20.34
OA, LA, PAL	AX_147234396	A09	E1, E2, E3, E4, E5	Yes	–	Both	114,692,575	4.68E-16	0.11	-1.09	44.49 (E1)
PAL	AX_147234411	A09	E1, E2, E6	Yes	–	Both	114,883,471	2.68E-10	0.13	0.12	61.81 (E1)
PAL	AX_176819108	B09	E2	–	No	BLINK	142,395,397	4.04E-13	0.13	-0.07	12.24
PAL	AX_176801823	A03	E4	–	No	MLMM	130,885,295	5.18E-10	0.06	-0.16	27.05
PAL	AX_176823131	A02	E6	–	No	Both	497,336	7.97E-11	0.07	-0.14	27.10
SA	AX_176823393	B06	E2, E6	Yes	–	BLINK	128,311,575	4.11E-08	0.33	0.02	12.54
SA	AX_176802683	A09	E6	–	No	MLMM	116,676,384	1.34E-06	0.07	-0.06	10.13

OIL: oil content, PRO: protein content, PAL: palmitic acid content, SA: stearic acid content, OA: oleic acid content, LA: linoleic acid content, E1: water-stressed 2018–2019, E2: well water 2018–2019, E3: water-stressed 2021–2022, E4: well water 2021–2022, E5: pooled water-stressed, E6: pooled well-watered, MAF: minor allele frequency

Oil content

BLINK identified 11 MTAs across three environments on six A genome chromosomes and five MTAs across two environments on five B genome chromosomes, while MLMM detected fewer overlapping signals. In WS_2018 (E1), BLINK detected four MTAs and MLMM three MTAs, sharing AX_147221664 (Chr. A05) and AX_177637369 (Chr. B10). In WW_2018 (E2), BLINK exclusively detected four MTAs, while MLMM identified two MTAs in WW_2021 (E4). Under combined WS (E5), BLINK detected five MTAs. Across environments, AX_147221664 and AX_177637369 were repeatedly detected, indicating environmental stability. These stable loci also showed relatively high PVE under both well-watered and drought conditions (Table 3).

Protein content

BLINK detected eight MTAs across four environments on five chromosomes, while MLMM identified one MTA in WW_2018 (E2) on chromosome A01. In WS_2018 (E1), BLINK identified three MTAs. In WW_2018 (E2), BLINK detected three MTAs, sharing AX_147209601 (Chr. A01) with MLMM. BLINK detected one MTA in WW_2021 (E4) and combined WS (E5). The shared SNP AX_147209601 showed moderate–high PVE in WW_2018, but no protein MTAs were consistently detected across multiple environments, indicating strong environmental specificity.

Oleic acid

BLINK identified eight environment-specific MTAs, with one major SNP detected in all environments except E3, while MLMM detected fewer but overlapping MTAs. BLINK detected five MTAs in WS_2018 (E1), and MLMM detected one. In WW_2018 (E2), MLMM detected one MTA, while in WW_2021 (E4) BLINK detected two MTAs and MLMM detected one. BLINK detected three MTAs in combined WS (E5), and MLMM one in combined WW (E6). SNP AX_147234396 (Chr. A09) was detected by both BLINK and MLMM in WS_2018, WW_2021, and combined WS, and across all individual and combined environments, explaining > 10% PVE (up to ~ 44% under WW and higher under stress), demonstrating robust effects.

Linoleic acid

BLINK detected seven MTAs in WS_2018 (E1), while MLMM detected one. In WS_2021 (E3), MLMM detected one MTA. BLINK identified three MTAs in combined WS (E5) and four in combined WW (E6), while MLMM

detected two in combined WW. The SNP AX_147234396 (Chr. A09) was shared between BLINK and MLMM and repeatedly detected across WS_2018, WS_2021, and in combined analyses, explaining a high proportion of phenotypic variance in both models and indicating strong locus robustness and environmental stability.

Palmitic acid

BLINK detected five MTAs across three environments, while MLMM detected four on overlapping chromosomes. In WS_2018 (E1), MLMM identified one MTA. In WW_2018 (E2), BLINK identified two MTAs, MLMM one. In WS_2021 (E3), BLINK identified one MTA; in WW_2021 (E4), MLMM identified two. MLMM detected two MTAs in combined WS (E5), while both models shared two MTAs in combined WW (E6). SNP AX_147234411 (Chr A09) was detected by both models in WW_2018, WW_2021, and combined WW, explaining up to 61.81% PVE under drought, indicating a highly stable and stress-responsive major locus.

Stearic acid

For stearic acid, BLINK detected three MTAs on chromosomes A09 and B06, while MLMM detected two MTAs on A09 and B08, primarily under WW conditions. In WW_2018 (E2), BLINK detected one MTA, while no MTAs were detected by MLMM. Under combined WW (E6), BLINK detected three MTAs and MLMM detected two, with AX_176802683 (Chr A09) shared between models. Additionally, AX_176823393 (Chr B06) was detected in both WW_2018 and combined WW analyses, demonstrating cross-environment stability.

Under well-watered conditions, major SNPs associated with oil, protein, and fatty acids explained 10–44% of phenotypic variance, while under drought they accounted for up to 62% (Table 3). Notably, AX_147234396 was consistently linked with oleic, linoleic, and palmitic acids across environments, and AX_147234411 showed a strong effect on palmitic acid under stress (61.81% PVE). Nine SNPs were stable across models and minimum of two environments, including AX_177638445 and AX_177637369 (oil), AX_147209601 (protein), AX_147234396 (oleic/linoleic), AX_147234411 and AX_176823131 (palmitic), and AX_176823393 (stearic).

A list of drought-specific, well-watered-specific and common SNPs identified for oil and protein content are provided in Table 4. Seven single-nucleotide polymorphisms (SNPs) with pleiotropic effects on multiple quality traits were identified. Notably, three SNPs, AX_177638445, AX_147209601, and AX_176814496, were associated with both oil and protein content, with AX_177638445 positively affecting oil and negatively affecting protein. AX_176811969 had a positive effect on protein content and a negative effect on LA content. Additionally, AX_177639855, AX_147234396, and AX_176803029 were significantly associated with oleic, linoleic, and palmitic acids, whereas AX_177639855 and AX_147234396 had negative effects on oleic acid and positive effects on linoleic acid, and AX_176803029 exhibited the opposite effect. Some SNPs were detected only by BLINK due to its superiority in identifying SNPs with a small PVE.

Potential candidate gene analysis for quality traits

From the candidate gene analysis, six intragenic SNPs for oil, three for protein, one for oleic acid, two for linoleic acid, three for palmitic acid, and one for stearic acid were identified, respectively. Candidate gene analysis was conducted on significant SNPs using Gbrowse on PeanutBase. Potential candidate genes were searched in the LD window of 3.52 MB around the significant SNPs (Table S5). The candidate gene–trait links are based on functional annotations, which identify genes associated with specific biological processes reported in earlier studies.

Candidate genes for quality traits under well-watered condition

In the analysis of oil content, several intragenic SNPs (at zero physical distance) were identified, suggesting the presence of potential causal variants. Notable among these are AX_176821854 on chromosome A05, located within the *Aradu.7RM7Z* gene, which encodes a *phosphoinositide-specific phospholipase C X domain-containing protein*; AX_147209601 on chromosome A01, mapping intragenically to *Aradu.ZT8PK*, which encodes *phospholipase DP2*; and AX_147234096 on chromosome A09, residing within *Aradu.HBT8X*, which encodes a *proteasome subunit beta type*. Additionally, SNP AX_147209601 on chromosome A01 is situated approximately 4.1 kb downstream of the *MID1* gene. The stable SNP, AX_177638445 on chromosome A10, is located approximately 2.3 kb upstream of the key gene encoding a *NAC domain protein*. Notably, AX_147209601 on chromosome A01 exhibited pleiotropic effects, being associated with both oil and protein contents (Tables 5 and S3). Under well-watered conditions, protein-associated intragenic marker–trait associations (MTAs),

Table 4 List of drought-specific, well-watered-specific and common SNPs associated with oil and protein content

Trait	Environment	SNP ID	Detected in (BLINK/ MLMM/Both)	Stable across ≥ 2 environments	Chromosome	Position (bp)	p value	MAF	Effect
OIL	Drought-specific	AX_147221664	Both	Yes	A05	7,694,037	3.23E-10	0.081105	-0.39359
		AX_177637369	BLINK	Yes	B10	127,616,318	4.50E-20	0.174688	0.964495
		AX_177640152	BLINK	No	B07	5,225,224	2.31E-07	0.131907	-0.20662
		AX_176817775	MLMM	No	B06	129,495,569	2.61E-13	0.134581	-0.80815
		AX_176818308	BLINK	No	A04	120,922,052	1.71E-06	0.095365	0.156508
		AX_177641922	BLINK	No	A08	1,734,733	9.16E-07	0.431373	0.124935
		AX_177643864	BLINK	No	B09	6,008,519	1.89E-06	0.185383	0.162517
	Well-watered-specific	AX_147209601	BLINK	No	A01	9,873,427	9.20E-09	0.170232	0.112977
		AX_176814496	BLINK	No	A03	126,548,835	5.27E-09	0.397504	0.085456
		AX_147234096	BLINK	No	A09	111,195,533	4.66E-07	0.227273	0.084689
		AX_177644302	BLINK	No	B10	126,354,470	2.74E-07	0.227273	0.071717
		AX_176821854	MLMM	No	A05	7,556,626	8.62E-07	0.08125	0.621054
	Common SNPs	AX_177638445	Both	Yes	A10	101,203,312	2.76E-08	0.069519	0.356857
PROTEIN	Drought-specific	AX_176811969	BLINK	No	A01	99,527,079	5.73E-11	0.053476	0.300707
		AX_176821384	BLINK	No	A04	77,777,651	3.98E-08	0.268271	0.136844
		AX_147253555	BLINK	No	B06	124,628,344	4.24E-07	0.196078	0.127471
		AX_176794206	BLINK	No	A01	106,097,712	2.28E-08	0.282531	-0.09257
	Well-watered-specific	AX_147209601	Both	No	A01	9,873,427	8.35E-09	0.170232	0.142697
		AX_176814496	BLINK	No	A03	126,548,835	3.22E-08	0.397504	0.103095
		AX_176810566	BLINK	No	A10	5,797,502	5.87E-11	0.163102	0.166429
		AX_177638445	BLINK	No	A10	101,203,312	1.38E-06	0.069643	-0.26796

such as SNP AX_147209601 on chromosome A01 located within *Aradu.PJ6S6* encoding a *GHMP kinase family protein*. Other key protein content-associated genes encode *pentatricopeptide repeat (PPR) proteins*, *peroxidases*, and *NAC domain transcription factors*, indicating a coordinated regulation of seed storage metabolism. Furthermore, one intragenic SNP, AX_176819108, on chromosome B09, located within *Araip.90GMX* (*Sec23/Sec24 protein transport*) and *Aradu.8JX0L* (*phosphatidylinositol-4-phosphate 5-kinase*), was associated with palmitic acid under well-watered conditions. For stearic acid, the intragenic SNP AX_177643021 in the *Araip.F9QDS* locus on chromosome B08 encodes a *poly(A) RNA polymerase cid11-like isoform X2*. The associated genes predominantly encode proteins such as *sn1-specific diacylglycerol lipase beta-like protein*, pollen protein *Ole E I-like*, *Nse4*, a component of the *Smc5/6 DNA repair complex*, *transcription factor bHLH128-like*, *zinc finger CCCH domain-containing protein 14*, *clathrin interactor EPSIN 2-like isoform X2*, *Cytochrome P450 superfamily protein*, *serine/threonine protein phosphatase 7 long form homolog*, and various disease resistance proteins. Most candidate genes related to stearic acid encode proteins, including *phosphatidylinositol-4-phosphate 5-kinase* and disease resistance proteins.

Candidate genes for quality traits under drought stress

Intragenic SNPs associated with oil content include AX_147221664 on chromosome A05, located within the *Aradu.U5XP0* gene, which encodes a *serine/threonine protein phosphatase 7 long form homolog*; AX_177640152 on chromosome B07, situated within the *Araip.65E1M* gene encoding an *LA RNA-binding protein*; and AX_177643864 on chromosome B09, found within the *Araip.9T22U* gene encoding an *F-box/RNI-like superfamily protein*. Furthermore, SNP AX_147221664 on A05 is intragenic to *RING finger proteins* and is positioned approximately 1.0 kb downstream of *zinc finger MYM-type proteins*. Two intragenic SNPs associated with protein content were identified: AX_176811969 on chromosome A01, located within the *Aradu.8GS2N* gene encoding a *CASP-like protein*, and AX_147253555 on chromosome B06, located within the *Araip.WKL89* gene encoding *receptor-like kinase 1*. Another lead SNP, AX_176794206 on A01, is approximately 2.2 kb downstream of the *Aradu.646B6* gene encoding *geranylgeranyl diphosphate reductase (GLDR)* and approximately 33 kb upstream of a *ZF-HD homeobox protein*. The lead SNP, AX_147234396 on A09, is associated with palmitic and oleic acid and is located within the *Aradu.G1YNF* gene (*fatty acid desaturase 2*). For linoleic acid, two intragenic SNPs were

Table 5 Potential candidate genes for quality traits

Trait	SNP ID	Chr	Diploid Gene ID	Tetraploid Gene ID	position	Distance to lead SNP (Kb)	Alleles	Sart	End	Length	Gene description
OIL, PRO	AX_147209601	A01	*Aradu.ZT8PK*	*Arahy.Q75RKD*	Intragenic	*0*	*T/C*	9,872,100	9,874,639	2539	*phospholipase DP2*
OIL	AX_177638445	A10	*Aradu.H5KV7*	*Arahy.CK11CG*	Intergenic	~2.3 kb upstream	*A/G*	101,199,650	101,202,008	2358	*NAC domain protein*
OIL	AX_176821854	A05	*Aradu.7RM7Z*	*Arahy.4E9YQ9*	Intragenic	*0*	*T/C*	7,572,916	7,576,801	3885	*PI-PLC X domain-containing protein*
OIL	AX_147234096	A09	*Aradu.TV5GQ*	*Arahy.2B1HMX*	Intrageniic	*0*	*T/C*	111,198,554	111,199,954	1400	*Lipid transfer protein*
OIL	AX_177643864	B09	*Araip.3M49X*	*Arahy.S0UFBH*	Intragenic	*0*	*A/C*	5,987,244	5,992,001	4757	*F-box/RNI-like superfamily protein*
OIL	AX_147234096	A09	*Aradu.HBT8X*	*Arahy.1L9YJM*	Intergenic	*0*	*T/C*	111,194,806	111,195,664	858	*Proteasome subunit beta type*
OIL	AX_147221664	A05	*Aradu.48HYY*	*Arahy.U8FDA3*	Intragenic	*0*	*A/G*	7,723,058	7,723,778	720	*RING finger proteins*
OIL	AX_147221664	A05	*Aradu.6I6D2*	*Arahy.W5KZ3V*	Intergenic	*~1.0 kb downstream*	*A/G*	7,695,388	7,696,353	965	*Zinc finger MYM-type proteins*
OIL	AX_177640152	B07	*Araip.SJE5I*	*Arahy.8SNE72*	Intragenic	*0*	*T/C*	5,256,564	5,257,448	884	*FAR1*
OIL	AX_147209601	A01	*Aradu.JA2PJ*	*Arahy.85H9WH*	intergenic	*~4.1 kb downstream*	*T/C*	9,828,119	9,832,109	3990	*MID1*
OIL	AX_147234096	A09	*Aradu.LY9Z0*	*Arahy.Z2JA3J*	intergenic	*~4.1 kb downstream*	*T/C*	111,149,129	111,154,056	4927	*Inositol polyphosphate phosphatase family proteins*
PRO	AX_147209601	A01	*Aradu.PJ6S6*	*Arahy.FEY4ZS*	Intragenic	*0*	*T/C*	9,824,555	9,824,781	226	*GHMP kinase family protein*
PRO	AX_147253555	B06	*Araip.WKL89*	*Arahy.9DML6S*	intergenic	*~26 kb upstream*	*T/C*	124,654,620	124,655,652	1032	*DUF protein*
PRO	AX_176811969	A01	*Aradu.8GS2N*	*Arahy.ATDP9G*	Intragenic	*0*	*A/G*	99,568,410	99,570,325	1915	*CASP-like protein*
PRO	AX_176810566	A10	*Aradu.7XS8V*	*Arahy.FFL3CR*	intergenic	*~4.8 kb upstream*	*A/G*	5,802,354	5,804,187	1833	*Peroxidase superfamily proteins*
PRO	AX_176794206	A01	*Aradu.646B6*	*Aradu.646B6*	intergenic	*~2.2 kb downstream*	*A/G*	106,061,034	106,063,294	2260	*Geranylgeranyl diphosphate reductase (GLDR)*
PRO	AX_176794206	A01	*Aradu.7WL5U*	*Arahy.75AC6V*	intergenic	*~33 kb upstream*	*A/G*	106,131,115	106,132,100	985	*ZF-HD homeobox protein*
PRO	AX_147253555	B06	*Araip.V9CFB*	*Arahy.X13C0W*	Intragenic	*0*	*T/C*	124,628,023	124,630,972	2949	*receptor-like kinase 1*
OA	AX_147234396	A09	*Aradu.5Y9LT*	*Arahy.E5H8ET*	Intragenic	*0*	*A/G*	114,679,131	114,682,099	2968	*AMMECR1 family protein*
OA	AX_176820814	B09	*Araip.N8HQ9*	*Arahy.XU9J8D*	intergenic	*~15 kb upstream*	*T/C*	146,782,942	146,788,882	5940	*NAD (P)-binding Rossmann-fold superfamily protein*
OA	AX_176822179	B09	*Araip.55WJR*	*Arahy.XA2D1U*	intergenic	*~28 kb downstream*	*T/C*	136,702,585	136,705,774	3189	*dihydrolipoamide acetyltransferase*
OA	AX_177639855	A07	*Aradu.6B67D*	*Arahy.MZ2CHN*	intergenic	*~38 kb upstream*	*A/G*	1,371,842	1,375,748	3906	*calcineurin B-like protein (CBL)*
OA	AX_177639855	A07	*Aradu.C64A0*	*Arahy.VAAE0N*	intergenic	~3.8 kb upstream	*A/G*	1,350,124	1,353,931	3807	*receptor-like protein kinases (RLKs)*
OA	AX_176812743	A09	*Aradu.65HV5*	*Arahy.MU4VHD*	intergenic	*~43 kb upstream*	*A/G*	3,865,712	3,869,228	3516	*Myo-inositol oxygenase (MIOX)*
OA	AX_147234396	A09	*Aradu.M1UTK*	*Arahy.EL78PJ*	intergenic	*~31 kb upstream*	*A/G*	114,724,017	114,726,254	2237	*Hemerythrin class glutathione S-transferase protein family*

Table 5 (continued)

Trait	SNP ID	Chr	Diploid Gene ID	Tetraploid Gene ID	position	Distance to lead SNP (Kb)	Alleles	Sart	End	Length	Gene description
OA	AX_176822179	B09	*Araip.SW89G*	*Arahy.I0VBSZ*	intergenic	*~36.8 kb downstream*	*T/C*	136,639,164	136,640,202	1038	*DEAD-box ATP-dependent RNA helicase-like protein*
SA	AX_176802683	A09	*Aradu.21AU1*	*Arahy.ILR90A*	intergenic	*~37.4 kb upstream*	*T/C*	116,713,791	116,720,604	6813	*Diacylglycerol lipases (DGLs)*
SA	AX_177643021	B08	*Araip.F9QDS*	*Arahy.CIGQ09*	Intragenic	*0*	*A/G*	127,142,249	127,150,238	7989	*poly(A) RNA polymerase cid11-like isoform X2*
LA	AX_147222791	A05	*Aradu.PB1FR*	*Arahy.WVJ4GZ*	Intragenic	*0*	*T/C*	84,428,601	84,447,881	19,280	*Disease resistance protein (TIR-NBS-LRR class) family*
SA, PAL	AX_176823393	B06	*Araip.0S5VX*	*Arahy.VT6DM9*	Intergenic	~2.6 kb downstream	*A/G*	128,350,912	128,353,580	2668	*Dehydration-responsive protein*
SA, PAL	AX_176823393	B06	*Araip.PE5QA*	*Arahy.463B8J*	Intergenic	*~23.4 kb downstream*	*A/G*	128,286,125	128,288,099	1974	*Rhomboid-like protein*
PRO, LA	AX_176811969	A01	*Aradu.TY35U*	*Arahy.YF5V20*	Intragenic	*0*	*A/G*	99,527,017	99,529,783	2766	*Membrane transporter D1*
OA, PAL	AX_147234396	A09	*Aradu.G1YNF*	*Arahy.8TPQ4A*	Intragenic	*0*	*A/G*	114,690,777	114,693,267	2490	*fatty acid desaturase 2*
PAL	AX_147234411	A09	*Aradu.8JX0L*	*Arahy.18R96V*	Intragenic	*0*	*T/G*	114,876,941	114,883,715	6774	*Phosphatidylinositol-4-phosphate 5-kinase family protein*
PAL	AX_176819108	B09	*Araip.90GMX*	*Arahy.Q6VS78*	Intragenic	*0*	*A/C*	142,392,111	142,404,297	12,186	*Sec23/Sec24 protein transport family protein*

OIL: oil content, PRO: protein content, PAL: palmitic acid content, SA: stearic acid content, OA: oleic acid content, LA: linoleic acid content

identified: AX_147222791 on A05, encoding a membrane transporter protein, and AX_176811969 on A01, encoding a *TIR-NBS-LRR* class disease-resistant protein. Numerous loci in the significant SNP regions were linked to *BTB/POZ domain-containing proteins* and *TIR-NBS-LRR-type* resistance proteins. Many genes related to oil content also influence protein content, including the *RING finger protein 38-like* located within the *Aradu.48HYY* gene on A05, *MYB transcription factor*, *F-box family protein*, *NAC domain protein*, *phospholipase DP2*, *FAR1 DNA-binding domain protein*, and *protein kinase superfamily proteins*.

Allele distribution pattern and in silico gene expression analysis

Allele distribution pattern suggested that high-oil and high-protein MLs mostly carried favorable alleles, while low oil and protein MLs predominantly carried unfavorable alleles (Fig. 5). The phenotypic data of the top ten MAGIC lines (MLs) with high and low oil and protein content from the 2021–2022 post-rainy season were analyzed to assess SNP utility and efficiency. The tissue-specific expression of few key candidate genes as mentioned in Fig. 6 was examined utilizing the *A. hypogaea* gene expression atlas (*AhGEA*) for the *fastigiata* subspecies (Sinha et al. 2020) which provides evidence for the relationship between gene expression levels and trait variation. Also, gene expression during drought stress expression resource (Brasileiro et al. 2015) was studied. Among the potential candidate genes identified in this investigation, few exhibited differential expressions in at least one tissue during critical developmental stages, as observed across 20 tissues in the gene expression atlas (Fig. 6). Utilizing a drought-specific expression atlas, it was observed that several genes associated with lipid metabolism, including *FAD2A*, *GLDR*, *phospholipases*, and *lipid transfer proteins*, exhibited moderate to strong upregulation under drought stress, particularly within leaf tissues and seeds. Key transcription factors, such as members of the *NAC*, *MYB*, and *ZF-HD families*, were consistently induced under water-deficit conditions, indicating their regulatory role in drought-responsive pathways. Furthermore, genes involved in signal transduction, including *receptor-like kinases*, *calcineurin B-like proteins*, and *RING finger proteins*, demonstrated drought-responsive expression patterns, underscoring their potential involvement in stress adaptation mechanisms (Supplementary Table S6).

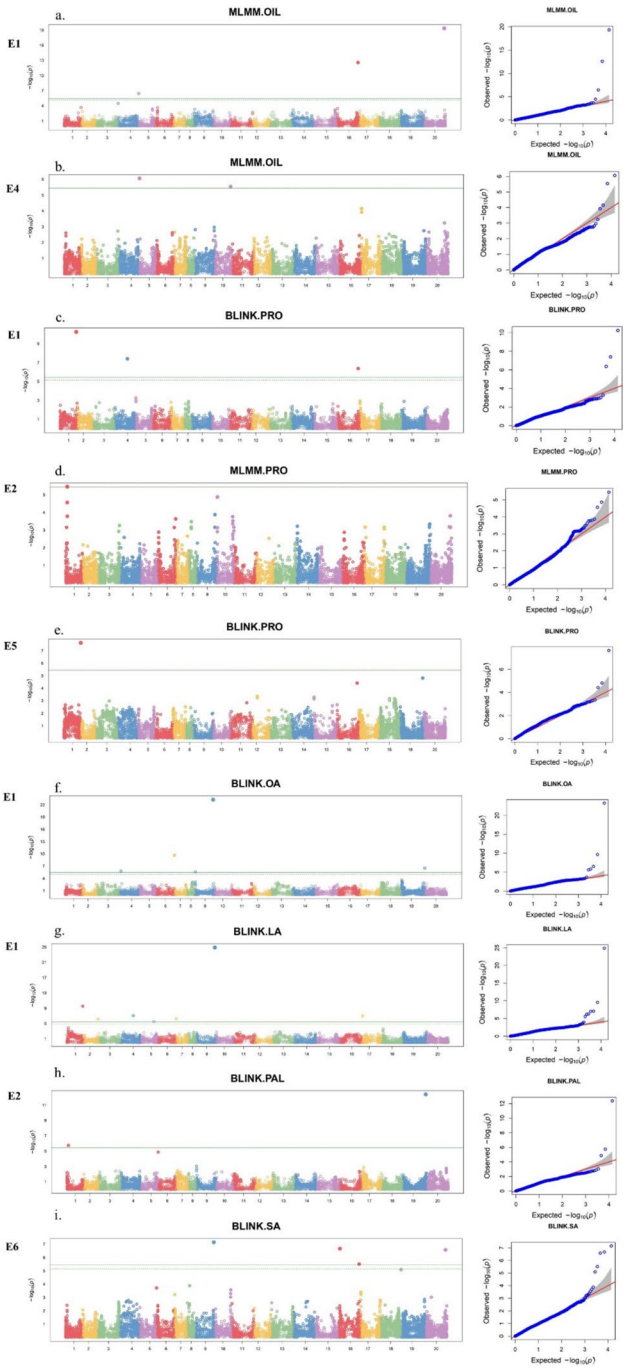

Fig. 5 Manhattan plots (left panels) and corresponding quantile–quantile (Q–Q) plots (right panels) depicting genome-wide association results for **A** oil content in environment E1 using the MLMM model, **B** oil content in E4 using MLMM, **C** protein content in E1 using BLINK, **D** protein content in E2 using MLMM, **E** protein content in E5 using BLINK, **F** oleic acid content in E1 using BLINK, **G** linoleic acid content in E1 using BLINK, **H** palmitic acid content in E2 using BLINK, and **I** stearic acid content in E6 using BLINK

Discussion

Ensuring the high nutritive value of grain and fodder in changing climatic conditions, particularly under drought stress, is becoming a crucial breeding objective for most crop breeding programs, including groundnut breeding. A significant reduction in oil content and other fatty acids due to water stress highlights the need for a deeper

Genotype	oil_ws_2018	AX_147221664	AX_177638445	AX_177640152	AX_177637369	AX_147209601	AX_176814496	AX_147234096	AX_177644302	AX_176821854	AX_176818308	AX_177641922	AX_177643864
ICGR171470	51.50	A	G	C	C	T	G	T	T	T	G	G	A
ICGR171461	50.78	A	G	C	C	T	G	T	T	T	G	G	A
ICGR171454	50.20	A	G	T	C	T	A	C	T	T	G	T	C
ICGR171370	50.15	A	G	C	C	T	G	T	T	T	G	G	A
ICGR171395	49.59	A	G	C	C	T	G	T	T	T	G	G	A
ICGR171466	49.18	G	G	T	C	C	A	C	T	C	G	T	C
ICGR171391	48.72	A	G	T	C	T	G	T	T	T	G	G	A
ICGR171383	48.69	G	G	T	C	T	A	T	T	C	G	G	C
ICGR171445	48.64	A	A	C	-	C	A	C	T	T	A	T	C
ICGR171509	48.57	G	A	T	A	-	G	T	T	C	A	T	C
ICGR171132	44.03	G	A	T	A	C	A	C	C	C	A	T	C
ICGR171502	44.02	G	A	T	A	C	A	-	C	C	A	T	C
ICGR171340	43.95	G	A	T	A	C	A	C	T	C	A	G	A
ICGR171131	43.92	G	A	T	A	C	A	C	T	C	A	G	C
ICGR171260	43.90	G	A	T	A	C	G	C	T	C	A	G	C
ICGR171410	43.90	G	G	T	A	C	A	C	T	C	A	T	C
ICGR171203	43.87	G	A	T	A	C	A	T	C	C	A	T	A
ICGR171311	43.77	G	A	T	A	C	G	C	T	C	A	T	C
ICGR171361	43.73	G	A	T	A	C	A	C	C	-	A	T	C
ICGR171549	43.54	G	A	T	A	C	G	C	C	C	A	G	C

Genotype	pro_ww_2021	AX_176821384	AX_147253555	AX_147209601	AX_176810566	AX_177638445	AX_176794206
ICGR171498	29.72	A	T	C	A	A	A
ICGR171277	29.27	A	T	C	A	A	A
ICGR171022	29.15	A	T	C	A	A	A
ICGR171128	29.13	A	T	C	A	A	A
ICGR171076	29.05	A	T	C	A	A	A
ICGR171282	28.99	A	T	C	A	A	G
ICGR171483	28.93	A	T	C	A	A	A
ICGR171295	28.88	A	T	C	A	A	A
ICGR171486	28.85	A	T	C	A	A	A
ICGR171220	28.84	A	T	C	A	A	A
ICGR171471	25.15	C	C	T	G	A	G
ICGR171323	25.15	C	T	T	G	A	G
ICGV5155	25.08	C	T	T	A	A	G
ICGR171466	25.02	A	T	C	G	G	-
ICGR171102	24.91	A	T	T	A	A	A
ICGR171261	24.62	A	T	T	A	A	G
ICGR171321	24.44	C	C	T	G	A	A
ICGR171383	24.30	C	C	T	G	G	A
ICGR171412	23.44	C	C	T	G	G	G
ICGR171461	23.42	C	C	T	G	G	G
ICGR171470	22.89	C	C	T	G	G	G

Fig. 6 Allele distribution pattern of favorable and unfavorable alleles of significant SNPs associated with high and low oil content **A** and protein content **B** under water-stressed condition during post-rainy 2018–2019

understanding of how environmental factors affect quality traits, as well as the importance of identifying genotypes that remain stable in these drought-stressed conditions. Due to the limited variation in oil, protein, and fatty acid content within breeding populations, it is crucial to explore diverse germplasms and multi-parental populations. In the MAGIC population used for the study, a wide range of variability has been observed as mentioned in Table 1. All these traits exhibited moderate to high heritability values in the current study. In groundnut, medium heritability is generally defined as 30–60% and high heritability as > 60%, for interpreting genetic control of quality traits such as oil and protein content. In previous studies, heritability (h^2) values of 0.52 (oil content) and 0.64 (protein content) were reported, along with low genetic advance and nonadditive gene effects (Noubissie et al. 2012). The wide phenotypic variation and moderate to high heritability observed for oil, protein, and fatty acid traits indicate exploitable genetic variation, supporting effective selection for nutritional quality under drought-prone environments and highlighting suitability of these traits for GWAS and QTL mapping (Abou-Elwafa and Shehzad 2020). The significance of block effects for specific traits across years highlights the strong influence of environmental heterogeneity, reinforcing the importance of multi-environment testing. Furthermore, a strong negative correlation between oil and protein, as well as between oleic and linoleic acids, aligns with findings from previous studies. (Xu et al. 2025). The consistent negative associations between oil and protein content, and between oleic and linoleic acids, highlight inherent biochemical trade-offs that must be considered in breeding programs. However, contrasting correlations between protein and linoleic acid across years under similar environments, negative in 2018–2019 and positive in 2021–2022, suggest strong genotype × environment interactions influencing trait relationships, particularly under stress. These findings indicate that selection for quality traits under drought conditions should emphasize stability and adaptability across environments rather than relying on single-year evaluations (Pal et al. 2024). Variations in stress intensity and phenological stage at stress imposition may further explain the observed year-to-year differences. Overall, the findings demonstrate the value of multi-parental populations for unraveling complex trait relationships and provide a strong foundation for integrating nutritional quality traits into drought-resilient groundnut breeding programs.

Implications of drought on quality traits

The current study found that drought stress conditions resulted in a decrease in oil content and an increase in protein content (Fig. 7). Few studies have documented that mid-season

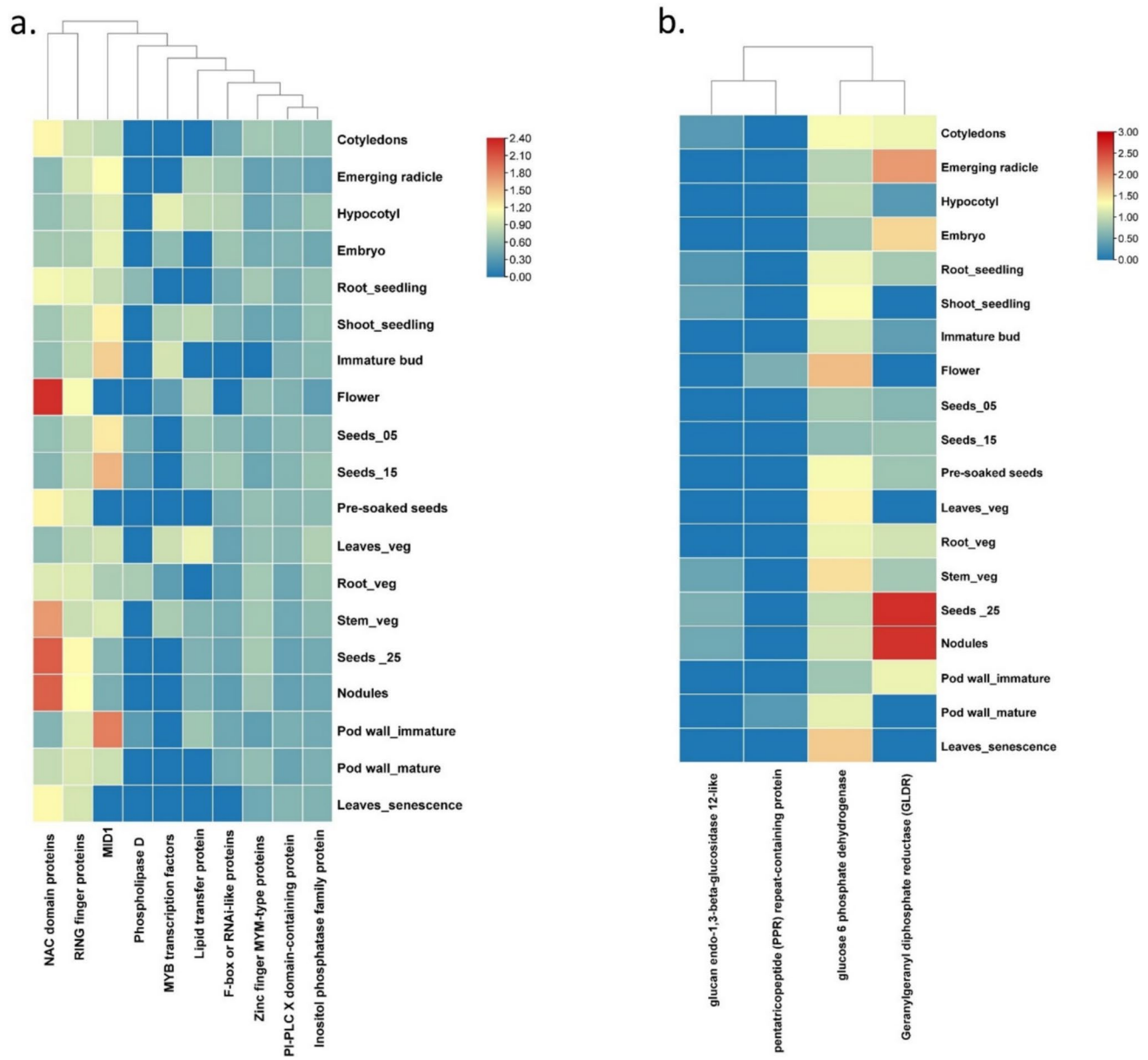

Fig. 7 Tissue-specific expression of potential candidate genes associated with **A** oil and **B** protein content

drought affects oil, protein, and fatty acid content in groundnut kernels (Chaiyadee et al 2013; Aydinsakir et al. 2016). Recent research also supports the findings of the current study (Dwivedi et al. 1996; Ai et al. 2025). This shows that even mid-season drought can lower oil content but increase protein content in groundnut kernels (Chakraborty et al. 2013). Ambiguities in the impact of drought on quality, may be attributed to the genotype-by-treatment interactions, as the response to drought stress varies with genotypes. Therefore, the results of the current study, which utilized a large set of genotypes and a multi-parental population, could provide clarity regarding these ambiguities. Drought stress impairs photosynthesis (Aguirrezabal et al. 2009), affects key enzymes that regulate lipid and fatty acid biosynthesis and metabolism (Fabian et al. 2019), decreases the movement of assimilates from source to sink (Vadez et al. 2013), impairs seed development, and induces oxidative stress. All these factors may be responsible toward reduction of oil content can lead to reduction in oil content of groundnut kernels. Under drought stress, increase in protein content could be primarily attributed to the changes in nitrogen metabolism and physiological adaptations. Shifting metabolic substrates away from lipid biosynthesis under drought stress provides more substrates for protein synthesis, leading to increased protein content in groundnut seeds. Under drought stress, there is an induction of stress-related protein synthesis (Tunnacliffe and Wise 2007), increased nitrogen

uptake and assimilation, and a reduction in seed size, which leads to higher protein content in the kernel.

Population structure and LD decay

Experimental populations like MAGIC and NAM are structured, which could influence the significance of genetic associations. In order to accurately capture genetic recombination, a larger number of lines is required in the population. Compared to GWAS using natural populations, multi-parental populations can reduce false discoveries due to population stratification and increase the frequency of rare alleles in wider populations (Kover et al. 2009). Higher recombination in MAGIC population breaks down linkage blocks more effectively than in biparental populations, leading to haplotype diversity and new combinations of gene variants. This disruption of linkage disequilibrium promotes random association between genetic loci and increases allelic diversity. In MAGIC populations, this leads to the formation of fine-scale genetic mosaics from multiple parental segments. Additionally, the lines traced back to known founders. MAGIC populations help mitigate hidden population structure, thereby reducing the chance of false positives in GWAS, which is common issue when using unrelated populations. GWAS analysis is dependent on understanding LD, particularly its decay, for effective association mapping. In the present study involving the MAGIC population, a LD block of 3.52 MB was observed. Self-pollinating crops, such as groundnut, typically exhibit larger LD blocks and demonstrate slower LD decay due to limited recombination events (Vos et al. 2017). The higher LD value observed in the MAGIC population can be ascribed to the limited recombination events and the specific set of founder parents used. Comparing LD values across different populations, such as inbread lines with diverse genetic structures and recombination histories, may not yield meaningful insights. However, numerous studies have reported extensive LD in various groundnut breeding populations (Otyama et al. 2019; Zhang et al. 2020; Zhou et al. 2017). For instance, Otyama et al. (2019) reported an LD decay of up to 4.8 MB, which is more than the LD reported in the current study. This phenomenon is mostly attributed to the recent tetraploid history of peanuts. Several investigations have been conducted to map QTLs in the MAGIC population, which features a substantial LD block size (Kover et al. 2009; Dell'Acqua et al. 2015).

Marker–trait association (MTAs) for quality traits

In case of groundnut, QTL mapping studies have been conducted to identify QTLs associated with oil, protein, and fatty acids (Pandey et al. 2014; Sarvamangala et al. 2011). However, only a few GWAS studies have been conducted in groundnuts for these nutritional traits. Environmental factors greatly impact quantitative traits, such as nutrient content, and ultimately influence the expression of genes and QTLs associated with these traits (Xue et al. 2017). Nonetheless, the current study utilized a high-density and highly accurate "Axiom Arachis2" 48 K SNP array genotyping platform (Peng et al. 2020) in a MAGIC population under both drought stress and well-watered conditions to identify drought-responsive SNPs under drought condition. Genome-wide association studies (GWAS) are amenable to generate false positives due to factors such as population structure, sample size, and environmental influences. To avoid the discovery of false positives, we applied stringent thresholds and corrections, including Bonferroni and false discovery rate (FDR) adjustments. Additionally, we considered associations that were consistently observed across multiple models or demonstrated biological relevance. This is the first study in groundnut to identify drought-responsive MTAs for quality traits under drought stress to the best of our knowledge.

Previous research by Sarvamangala et al. (2011) identified major QTLs for oil, protein, and fatty acids on various linkage groups using SSR markers in a RIL population. They identified significant QTLs for oil on linkage groups A01, A03, A04, and A08; for protein on A01, A03, A04, A06, A07, and A08; and for oleic and linoleic acids on A01, A07, and A08. The present investigation also revealed notable SNPs for oil content on A01, A03, A04, and A08; for protein content on A01, A03, and A04; for oleic acid on A07; and for linoleic acid on A01 and A07. Furthermore, indications of association between quality traits and genomic regions on other chromosomes were observed. Zhang et al. (2020) conducted a GWAS and QTL-seq, correlating SNPs with oil, protein, and fatty acid content on chromosomes A01, A02, A03, A04, A09, B01, B02, and B10. Consistent with their findings, significant SNPs for oil content were identified on A01, A03, A04, A09, and B10, and for protein content on A04. Another GWAS highlighted regions on A08 linked to oil content, situated 36.64 MB from SNP AX_177641922 in the present study (Guo et al. 2024). Shaibu et al. (2019) discovered SNPs associated with fatty acids using DArTSeq on 170 genotypes at ICRISAT, Nigeria. The current research identified comparable SNPs on A09, A07, and B10, with several SNPs such as AX_147234396 on A09 (392.589 kb downstream), on A09 (98.45 Mb upstream), on A07 (3.47 Mb

downstream), and on B10 (79.48 Mb upstream), located near previously reported SNPs for oleic acids. For linoleic acid, on B07 (362.10 kb downstream), on A07 (3.47 and 3.61 Mb upstream). Moreover, SNPs on chromosomes A02, A09, A10, B07, B09, and B10 were within 100 Mb of previously identified SNPs for linoleic acid content. Regarding stearic acid, two SNPs on chromosome B06 were 19.05 Mb and 44.29 Mb upstream of previously identified SNPs. A GWAS conducted by Otyama et al. (2019) identified genomic regions governing fatty acid composition, some of which aligned with the results of the present study. SNP AX_147234396 on chromosome A09 was linked to oleic, linoleic, and palmitic acid content in both studies. In the current research, SNP AX_176809593, associated with linoleic acid, is 50.76 kb from a previously identified region on A10. Four additional SNPs linked to linoleic acid were also found within a 100 MB range of earlier findings. SNP AX_147234411 on A09, associated with palmitic acid, was also noted in previous studies alongside four other SNPs in a similar 100 MB area. Furthermore, five SNPs related to stearic acid in the present study were within 100 MB of significant SNPs identified earlier. A GWAS study by Zhou et al. (2022) uncovered seven SNPs linked to oleic acid content within 100 MB of the SNPs identified in the current study. For palmitic acid, one SNP on chromosome A01 was 64.20 MB from a previously reported SNP. The present study also identified additional SNPs associated with stearic and linoleic acids. For certain traits, such as protein content with low repeatability, the identified marker–trait associations (MTAs) may be subject to environmental influences. Nonetheless, some associations may retain biological significance, particularly when they are consistently observed across multiple seasons. These MTAs need to be validated through independent populations or functional studies prior to their integration into the breeding pipeline. In addition to that, drought response complexity limits consistently detected SNPs across drought stress and well-watered conditions, likely due to environmental specificity and G×E interaction. Drought stress triggers condition-specific gene expression, leading to SNPs showing significance under specific conditions. Quality traits are quantitatively inherited and affected by stress-induced changes, while differential pathway expression under drought can mask well-watered condition associations. Although consistent SNPs across environments are valuable, stress-specific SNPs represent crucial adaptive response loci for environment-specific breeding.

Genome-wide association studies (GWAS) have been pivotal in identifying marker–trait associations (MTAs) for key quality traits in groundnut. Early analyses using the US minicore collection linked FAD2 gene variants with oleic and linoleic acid content (Chu et al. 2007). Following this, marker-assisted breeding led to the development of "Tifguard High O/L," combining nematode resistance with the high oleic trait (Chu et al. 2011). Similarly, the root-knot nematode-resistant variety NemaTAM was successfully developed and released in the USA using marker-assisted selection (Simpson et al. 2003). Quality traits such as oil content, protein content, and fatty acid composition can be prioritized in breeding programs by focusing on MTAs identified in the current study with high phenotypic variance explained (PVE) and stable expression across environments. Validated MTAs can be converted into kompetitive allele-specific PCR (KASP) markers for use in early-generation selection and backcrossing to efficiently introgress favorable alleles into elite cultivars. Furthermore, integrating these MTAs into genomic selection models holds promise for enhancing prediction accuracy and accelerating the development of improved groundnut varieties adapted to both normal and drought-prone conditions.

Potential candidate genes for quality traits

An extensive literature review was conducted to identify plausible candidate genes that encode proteins with functions related to lipid metabolism, fatty acid pathways, and roles in drought tolerance mechanisms. Candidate genes associated with oil content were found to govern several proteins involved in lipid biosynthesis and lipid metabolism. Furthermore, numerous proteins regulated by these genes are important in drought response.

Potential candidate genes for quality traits under well-watered conditions

The gene *Aradu.ZT8PK* encodes *Phospholipase DP2 (PLD),* which is essential for lipid metabolism (Colley et al. 1997), and also affects macronutrient availability in *Arabidopsis* (Lu et al. 2016). In foxtail millet (*Setaria italica*), *PLD alpha 1* is reported to be involved in drought tolerance and stress signaling, particularly in the ABA pathway (Peng et al. 2010). The gene *Aradu.7RM7Z* encodes a *PI-PLC X domain protein*, which generates *glycerophosphodiester phosphodiesterase*, a significant enzyme in lipid metabolism. Additionally, *lipid transfer proteins (LTPs)* facilitate membrane biogenesis and fatty acid regulation by promoting phospholipid transfer between membranes (Kader 1996). *NAC domain proteins* (*Aradu.H5KV7*) play a substantial role in oil content and drought response. In *Medicago falcate*, *NAC transcription factors* have been reported to regulate drought tolerance by modulating genes related to lipid transport and oxidation–reduction (Duan et al. 2017). Proteins such as *phospholipase D2* (*Aradu.PE7A6*), *NAC domain protein* (*Aradu.H5KV7*), *lipid transfer protein* (*Aradu.VWQ48*), and GDSL-like lipase/acyl hydrolase superfamily protein (*Aradu.TMP0B*) are critical for lipid metabolism, supporting the connection between oil and protein content in groundnuts. The *GHMP kinase family* protein (*Aradu.PJ6S6*) contributes to isoprene and amino acid

biosynthesis and carbohydrate metabolism, influencing protein content (Bork et al. 1993). Furthermore, proteomic and lipidomic analyses revealed that increased expression of lipase *GDSL domain-containing protein* enhances total fatty acids and 18:1 composition, while reducing 18:2 composition (Xu et al. 2022). Genes identified to be regulating oleic, palmitic, linoleic, and stearic acid content have functions related to fatty acid biosynthesis. A previously known gene *Aradu.G1YNF* was identified in the current study which encodes *fatty acid desaturase 2* (*FAD2*), that influences the composition of three key fatty acids: oleic, palmitic, and linoleic acids. Mutant alleles of this gene were found in chromosomes A09 and B09, affecting peanut oil quality. *FAD2* catalyzes the conversion of oleic acid to linoleic acid (Ray et al. 1993). SNP AX_147234396, significantly linked to oleic (OA), linoleic (LA), and palmitic acid (PAL), was found to be intragenic, located within the *FAD2A* gene, based on the Tifrunner v2.0 genome. *FAD2A* encodes oleoyl-PC desaturase, which converts oleic to linoleic acid. Mutations in *FAD2A* are known to elevate OA and reduce LA (Pandey et al. 2014). The co-localization of this SNP with *FAD2A*, along with its high PVE (up to 41.66%), highlights its strong functional relevance for oil quality improvement. *Aradu.7K9KV* encodes the ferredoxin 3 protein, which is involved in desaturating lipid-bound oleate in spinach (Schmidt et al. 1990). A recently identified gene, *Aradu.5Y9LT*, encodes an *AMMECR1* family protein associated with protein content in soybeans (Vuong et al. 2023). *Araip.*N8HQ9 encodes an *NAD(P)-binding Rossmann-fold superfamily protein* linked *to fatty acyl-CoA reductase*, and is characterized by specific catalytic motifs (Fujimoto et al. 2001; Wang et al. 2016). Additionally, *Aradu.K16RE* encodes proteins in the *oxysterol-binding family*, which are reported to be involved in signaling and lipid metabolism (Raychaudhuri et al. 2010). Finally, *Araip.55WJR* encodes a component of the *pyruvate dehydrogenase complex* that aids in mitochondrial fatty acid biosynthesis (Dieckmann 2024).

Potential candidate genes for quality traits under drought stress

Several genes associated with protein content are also implicated in plant stress response. The *DUF protein* (*Araip. WKL89*) and *glucose-6-phosphate dehydrogenase*, essential for the oxidative pentose phosphate pathway (OPPP) in tomatoes, contribute to drought tolerance by enhancing proline synthesis, Hsp70 chaperone accumulation, and ascorbate peroxidase activity (Landi et al. 2016). The gene *Aradu.8GS2N* encodes a *CASP-like protein* that is reported to regulate the stress response through Casparian strip formation (Wang et al. 2020). *Geranylgeranyl diphosphate reductase (GLDR)* is vital for chlorophyll biosynthesis in rice and for the stress response in olives (Bruno et al. 2009, 2017). The *ZF-HD homeobox protein* is also significant in the response to water-deficit stress in pigeon pea (Kumar et al. 2015). *F-box proteins*, also known as *RNAi-like proteins*, are involved in protein ubiquitination and contribute to the regulation of antioxidant capacity and reactive oxygen species (ROS) under stress (An et al. 2019). *RING finger proteins* are critical for plant growth and abiotic stress responses, functioning as ABA receptors (Han et al. 2022). *Zinc finger MYM-type proteins* are also reported to regulate abiotic stress responses (Han et al. 2020), whereas *MYB transcription factors* are important for hormone signaling and stress tolerance (Zhang et al. 2018). *FAR1* is reported to be essential for the drought stress response, ABA signaling, and starch metabolism (Graf and Smith, 2011), and *MID1* is expressed in vascular tissues, enhancing drought tolerance and yield when overexpressed (Guo et al. 2016). *Inositol polyphosphate phosphatase proteins* affect drought responses in tomatoes by repressing drought-inducible genes (Na et al. 2020). *MADS-box proteins* are reported to regulate abiotic stress responses in plants (Yang et al. 2024). Genes encoding proteins associated with oil content, such as *inositol phosphatase*, the *PPR superfamily*, *RING proteins*, *MYB transcription factors*, *MID1 family proteins*, *and F-box proteins*, are also linked to protein content and contribute to abiotic stress responses, particularly drought. Several genes linked to oleic, stearic, linoleic, and palmitic acid contents are involved in drought stress mechanisms. Notably, *calcineurin B-like protein (CBL)* (Batistic et al. 2009), *FBA* in *Arabidopsis* associated with drought tolerance in roots, *receptor-like protein kinases (RLKs),* and heat shock proteins, which also respond to abiotic stress (Rahman et al. 2022). The gene *Aradu.65HV5*, which encodes *myo-inositol oxygenase (MIOX),* enhances drought tolerance by reducing oxidative damage and contributes to ascorbic acid biosynthesis. Additionally, genes related to the hemerythrin class *glutathione S-transferase* and *DEAD-box ATP-dependent RNA helicase* are linked to stress tolerance (Galle et al. 2009; Sun et al. 2010; Rezaei et al. 2013; Dang et al. 2024). For stearic acid, *Aradu.21AU1* encodes *diacylglycerol lipases (DGLs),* which are involved in producing bioactive monoacylglycerols (Yuan et al. 2016). Other proteins associated with stearic and palmitic acid, such as dehydration-responsive proteins and *bHLH transcription factors*, also play significant roles in drought tolerance (Shinozaki 1993; Wang et al. 2018; Duan et al. 2024; John et al. 2024).

Conclusion

The current study revealed the utility of MAGIC population for association mapping studies as the number of recombination breakpoints identified are large compared to a RIL population. The study identified novel significant SNPs such as AX_176817775 and AX_177637369 associated with oil,

AX_176811969 and AX_176794206 associated with protein, AX_147234411 associated with palmitic acid under drought stress acid along with a previously reported SNPs AX_147234396 on A09 associated with oleic, linoleic and palmitic acid. Fourteen high-confidence SNPs with > 10% PVE for the above-mentioned quality traits can be further validated and developed in to markers to be utilized in breeding programs. Candidate genes such as *RING finger proteins* and *protein kinases* for oil content; *CASP-like protein* and *geranylgeranyl diphosphate reductase (GLDR)* for protein content, and *fatty acid desaturase 2*, *MYB transcription factor* and *AMMECR1 family* for oleic and palmitic acid content identified in the study provides new insight in to the molecular mechanisms governing nutritional traits under water-deficit stress condition.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00122-026-05174-0.

Acknowledgements The authors are thankful to Gopi Potupureddi and K. Sivasakthi for their support in conducting the experiments.

Author contributions AP conducted investigation, write original draft, data curation, validation, visualization, formal analysis, and software. DL supervised the investigation, writing and eidting. APW developed MAGIC population, conducted investigation and reviewing the MS. AV conducted data analysis, software, and writing and reviewing. SC supported phenotyping methodology and writing and reviewing. DBD performed supervision. SG participated in visualization and writing and reviewing. GP supported the field phenotyping and layout of experiments, data collection. MKP generated genomic data, contributed to methodology, and writing and reviewing. JP conceptualized, designed methodology, project administration, funding acquisition, resources, supervision, and writing, reviewing and editing.

Funding The funds were received from CGIAR's CRP-Grain Legumes and Dryland Cereals (GLDC), and Department ofAgricutlure, Odisha state, India.

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Abady S, Shimelis H, Janila P, Deshmukh D, Wankhade A, Chaudhari S, Manohar SS (2021a) Combining ability analysis of groundnut (*Arachis hypogaea* L.) genotypes for yield and related traits under drought-stressed and non-stressed conditions. Euphytica 217(11):200. https://doi.org/10.1007/s10681-021-02932-7

Abady S, Shimelis H, Janila P, Yaduru S, Shayanowako AIT, Deshmukh D, Chaudhari S, Manohar SS (2021b) Assessment of the genetic diversity and population structure of groundnut germplasm collections using phenotypic traits and SNP markers: implications for drought tolerance breeding. PLoS ONE. https://doi.org/10.1371/journal.pone.0259883

Abou-Elwafa SF, Shehzad T (2020) Genetic diversity, GWAS and prediction for drought and terminal heat stress tolerance in bread wheat (*Triticum aestivum* L.). Genet Resour Crop Evol 68(2):711–728. https://doi.org/10.1007/s10722-020-01018-y

Aguirrezabal LAN (2009) Seed composition of peanut as affected by drought stress during seed filling. Field Crops Research

Ai X, Wang S, Dai R, Ma X, Li C, Zhong C, Wang J, Liu X, Zhao X, Zhang H, Yu H (2025) Dynamic changes of seed development, oil accumulation and fatty acid composition in peanut under soil water deficit. Plant Physiol Biochem 219:109336

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19(9):1655–1664. https://doi.org/10.1101/gr.094052.109

Arrones-Olmo A, Antar O, Pereira-Dias L, Solana-Garcia A, Ferrante P, Aprea G, Vilanova Navarro S (2024) A novel tomato inter-specific (Solanum lycopersicum var. cerasiforme and S. pimpinellifolium) MAGIC population facilitates trait association and candidate gene discovery in untapped exotic germplasm. Horticulture Research 11(7). https://doi.org/10.1093/hr/uhae154

Aydinsakir K, Dinc N, Buyuktas D, Bastug R, Toker R (2016) Assessment of different irrigation levels on peanut crop yield and quality components under Mediterranean conditions. J Irrig Drain Eng. https://doi.org/10.1061/(ASCE)IR.1943-4774.0001062

Batistic O, Kudla J (2009) Plant calcineurin B-like proteins and their interacting protein kinases. Biochim Biophys Acta 1793(6):985–992. https://doi.org/10.1016/j.bbamcr.2008.10.006

Bera SK, Manohar SS, Variath MT, Chaudhari S, Yaduru S, Thankappan R, Pasupuleti J (2018) Assessing variability for disease resistance and nutritional quality traits in an interspecific collection of groundnut (*Arachis hypogaea*). Plant Breed 137(6):883–894. https://doi.org/10.1111/pbr.12647

Bland JM, Altman DG (1995) Multiple significance tests: the Bonferroni method. BMJ 310(6973):170. https://doi.org/10.1136/bmj.310.6973.170

Blum A (2010) Plant breeding for water-limited environments. Springer Science Business Media, Berlin

Bork P, Sander C, Valencia A (1993) Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. Protein Sci 2:31–40. https://doi.org/10.1002/pro.5560020104

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Brasileiro AC, Morgante CV, Araujo AC, Leal-Bertioli SC, Silva AK, Martins AC, Vinson CC, Santos CM, Bonfim O, Togawa RC, Saraiva MA (2015) Transcriptome profiling of wild Arachis from waterlimited environments uncovers drought tolerance candidate genes. Plant Mol Biol Rep 33(6):1876–1892

Bruno L, Chiappetta A, Muzzalupo I, Gagliardi C, Iaria D, Bruno A, Bitonti MB (2009) Role of geranylgeranyl reductase gene in organ development and stress response in olive (*Olea europaea*) plants. Funct Plant Biol 36(4):370–381. https://doi.org/10.1071/FP08219

Carvalho MJ, Vorasoot N, Puppala N, Muitia A, Jogloy S (2017) Effects of terminal drought on growth, yield and yield components in valencia peanut genotypes. SABRAO J Breed Genet 49:270–279

Cavanagh C, Morell M, Mackay I, Powell W (2008) From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. Curr Opin Plant Biol 11(2):215–221. https://doi.org/10.1016/j.pbi.2008.01.002

Chaiyadee S, Jogloy S, Songsri P, Singkham N, Vorasoot N, Sawatsitang P, Holbrook CC, Patanothai A (2013) Soil moisture affects

fatty acids and oil quality parameters in peanut. Int J Plant Prod 7(1):81–96
Chakraborty K, Bishi SK, Singh AL, Kalariya KA, Kumar L (2013) Moisture deficit stress affects yield and quality in groundnut seeds. Indian J Plant Physiol 18(2):136–141. https://doi.org/10.1007/s40502-013-0020-4
Chu Y, Ramos L, Holbrook CC, Ozias-Akins P (2007) Frequency of a loss-of-function mutation in oleoyl-PC desaturase (ahFAD2A) in the mini-core of the US peanut germplasm collection. Crop Sci 47(6):2372–2378. https://doi.org/10.2135/cropsci2007.02.0117
Chu Y, Wu CL, Holbrook CC, Tillman BL, Person G, Ozias-Akins P (2011) Marker-assisted selection to pyramid nematode resistance and the high oleic trait in peanut. The plant genome, 4(2)
Colley WC, Sung TC, Roll R, Jenco J, Hammond SM, Altshuller Y, Bar-Sagi D, Morris AJ, Frohman MA (1997) Phospholipase D2, a distinct phospholipase D isoform with novel regulatory properties that provokes cytoskeletal reorganization. Curr Biol 7(3):191–201. https://doi.org/10.1016/S0960-9822(97)70090-3
Conkerton EJ, Ross LF, Daigle DJ, Kvien CS, McCombs C (1989) The effect of drought stress on peanut seed composition II. Oil, protein and minerals: 593–602
Dang P, Patel J, Sorensen R, Lamb M, Chen CY (2024) Genome-wide association analysis identified quantitative trait loci (QTLs) underlying drought-related traits in cultivated peanut (*Arachis hypogaea* L.). Genes 15(7):868. https://doi.org/10.3390/genes15070868
Dell'Acqua M, Gatti DM, Pea G, Cattonaro F, Coppens F, Magris G, Hlaing AL, Aung HH, Nelissen H, Baute J, Frascaroli E, Churchill GA, Inze D, Morgante M, Pe ME (2015) Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in *Zea mays*. Genome Biol 16(1):167. https://doi.org/10.1186/s13059-015-0716-z
Deshmukh DB, Variath MT, Patne N, Kona P, Marathi B, Pasupuleti J (2021) Inheritance of oleic acid trait and high throughput non-destructive phenotyping for nutritional traits in groundnut kernels. Indian J Genet Plant Breed 81(04):72–80. https://doi.org/10.31742/IJGPB.81.4.9
Dieckmann CL (2024) A hub for regulation of mitochondrial metabolism: fatty acid and lipoic acid biosynthesis. IUBMB Life 76(6):332–344. https://doi.org/10.1002/iub.2802
Duan M, Zhang R, Zhu F, Zhang Z, Gou L, Wen J, Dong J, Wang T (2017) A lipid-anchored NAC transcription factor is translocated into the nucleus and activates glyoxalase I expression during drought stress. Plant Cell 29(7):1748–1772. https://doi.org/10.1105/tpc.17.00044
Duan X, Xu Y, Zhang K, Liu Z, Yu Y (2024) Identification and expression analysis of soybean (*Glycine max* L.) dynamin genes reveal their involvements in plant development and stress response. Trop Plant Biol. https://doi.org/10.1007/s12042-024-09361-5
Dwivedi SL, Nigam SN, Rao RCN, Singh U, Rao KVS (1996) Effect of drought on oil, fatty acids and protein contents of groundnut (*Arachis hypogaea* L.) seeds. Field Crops Res 48:125–133. https://doi.org/10.1016/S0378-4290(96)01027-1
Fábián A. et al (2019) Effects of drought on the grain quality traits of maize and wheat: A review. Field Crops Research
FAOSTAT (2022). https://www.fao.org/faostat/en/
Fujimoto K, Hara M, Yamada H, Sakurai M, Inaba A, Tomomura A, Katoh S (2001) Role of the conserved Ser–Tyr–Lystriad of the SDR family in sepiapterin reductase. Chem-Biol Interact 130:825–832. https://doi.org/10.1016/S0009-2797(00)00238-6
Francis R (2017) Pophelper: an R package and web app to analyse and visualize population structure. Mol Ecol Resour 17(1):27–32. https://doi.org/10.1111/1755-0998.12509
Gangurde SS, Khan AW, Janila P, Variath MT, Manohar SS, Singam P, Chitikineni A, Varshney RK, Pandey MK (2023a) Whole-genome sequencing-based discovery of candidate genes and diagnostic markers for seed weight in groundnut. Plant Genome 16(4):e20265. https://doi.org/10.1002/tpg2.20265
Gangurde SS, Pasupuleti J, Parmar S, Variath MT, Bomireddy D, Manohar SS, Varshney RK, Singam P, Guo B, Pandey MK (2023b) Genetic mapping identifies genomic regions and candidate genes for seed weight and shelling percentage in groundnut. Front Genet 14:1128182. https://doi.org/10.3389/fgene.2023.1128182
Gullifa G, Barone L, Papa E, Giuffrida A, Materazzi S, Risoluti R (2023) Portable NIR spectroscopy: the route to green analytical chemistry. Front Chem 11:1214825. https://doi.org/10.3389/fchem.2023.1214825
Guo C, Yao L, You C, Wang S, Cui J, Ge X, Ma H (2016) MID1 plays an important role in response to drought stress during reproductive development. Plant J 88(2):280–293. https://doi.org/10.1111/tpj.13250
Guo M, Deng L, Gu J, Miao J, Yin J, Li Y, Fang Y, Huang B, Sun Z, Qi F, Dong W (2024) Genome-wide association studyand development of molecular markers for yield and quality traits in peanut (Arachis hypogaea L.). BMC Plant Biol 24(1):244. https://doi.org/10.1186/s12870-024-04937-5
Galle A, CsiszAr J, Secenji M, Guoth A, Cseuz L, Tari I, Györgyey J, Erdei L (2009) Glutathione transferase activity andexpression patterns during grain fi lling in fl ag leaves of wheat genotypes differing in drought tolerance: response towater deficit. J plant physiol 166(17):1878–1891. https://doi.org/10.1016/j.jplph.2009.05.016
Han G, Lu C, Guo J, Qiao Z, Sui N, Qiu N, Wang B (2020) C2H2 zinc finger proteins: master regulators of abiotic stress responses in plants. Front Plant Sci 11:115. https://doi.org/10.3389/fpls.2020.00115
Han G, Qiao Z, Li Y, Yang Z, Wang C, Zhang Y, Liu L, Wang B (2022) RING zinc finger proteins in plant abiotic stress tolerance. Front Plant Sci 13:877011. https://doi.org/10.3389/fpls.2022.877011
Hashim IB, Koehler PE, Eitenmiller RR, Kvien CK (1993) Fatty acid composition and tocopherol content of drought stressed Florunner peanuts. Peanut Sci 20(1):21–24. https://doi.org/10.3146/i0095-3679-20-1-6
Hoque A, Fiedler JD, Rahman M (2020a) Genetic diversity analysis of a flax (*Linum usitatissimum L.*) global collection. BMC Genom 21:1–13. https://doi.org/10.1186/s12864-020-06922-2
Hoque A, Fiedler JD, Rahman M (2020b) Genetic diversity analysis of a flax (*Linum usitatissimum* L.) global collection. BMC Genomics 21(1):557. https://doi.org/10.1186/s12864-020-06922-2
Janila P, Nigam SN, Pandey MK, Nagesh P, Varshney RK (2013) Groundnut improvement: use of genetic and genomic tools. Front Plant Sci 4:23. https://doi.org/10.3389/fpls.2013.00023
Janila P, Nigam SN, Abhishek R, Kumar VA, Manohar SS, Venuprasad R (2015) Iron and zinc concentrations in peanut (*Arachis hypogaea* L.) seeds and their relationship with other nutritional and yield parameters. J Agric Sci 153(6):975–994. https://doi.org/10.1017/S0021859614000525
Janila P, Variath MT, Pandey MK, Desmae H, Motagi BN, Okori P, Manohar SS, Rathnakumar AL, Radhakrishnan T, Liao B, Varshney RK (2016) Genomic tools in groundnut breeding program: status and perspectives. Front Plant Sci 7:289. https://doi.org/10.3389/fpls.2016.00289
Jha UC, Bohra A, Nayyar H (2020) Advances in "omics" approaches to tackle drought stress in grain legumes. Plant Breed 139(1):1–27. https://doi.org/10.1111/pbr.12761
John A, Kramer M, Lehmann M, Kunz HH, Aarabi F, Alseekh S, Neuhaus HE (2024) Degradation of FATTY ACID EXPORT PROTEIN1 by RHOMBOID-LIKE PROTEASE11 contributes to cold tolerance in *Arabidopsis*. Plant Cell 36(5):1937–1962. https://doi.org/10.1093/plcell/koae011

Jung S, Swift D, Sengoku E, Patel M, Teule F, Powell G, Moore K, Abbott A (2000) The high oleate trait in the cultivated peanut (*Arachis hypogaea* L.) isolation and characterization of two genes encoding microsomal oleoyl-PC desaturases. Mol Gen Genet 263:796–805. https://doi.org/10.1007/s004380000244

Kader JC (1996) Lipid-transfer proteins in plants. Annu Rev Plant Biol 47(1):627–654. https://doi.org/10.1146/annurev.arplant.47.1.627

Kassie FC, Chaix G, Ngalle HB, Seye M, Fall C, Tossim HA, Nguepjop JR (2024) Application of near-infrared spectroscopy for fast germplasm analysis and classification in multi-environment using intact-seed peanut (*Arachis hypogaea* L.). Oil Crop Sci 9(2):132–141. https://doi.org/10.1016/j.ocsci.2024.03.003

Klute A (2018) Water Retention: Laboratory Methods. In Klute A (ed), SSSA Book Series (pp. 635–662). Soil Science Society of America, American Society of Agronomy. https://doi.org/10.2136/sssabookser5.1.2ed.c26

Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD, Mott R (2009) A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. PLoS Genet 5(7):e1000551. https://doi.org/10.1371/journal.pgen.1000551

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009a) Circos: an information aesthetic for comparative genomics. Genome Res 19(9):1639–1645. https://doi.org/10.1101/gr.092759.109

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009b) Circos: an information aesthetic for comparative genomics. Genome Res 19(9):1639–1645

Kumar R, Yadav S, Shrinivas D, Kumar SA, Shitole V, Naik GR (2015) Transcriptome of pigeonpea roots under water deficit analyzed by suppression subtractive hybridization. J Agr Sci Tech 17(5):1333–1345

Lal MK, Tiwari RK, Gahlaut V, Mangal V, Kumar A, Singh MP, Zinta G (2022) Physiological and molecular insights on wheat responses to heat stress. Plant Cell Rep 41(3):501–518. https://doi.org/10.1007/s00299-021-02784-4

Landi S, Nurcato R, DeLillo A, Lentini M, Grillo S, Esposito S (2016) Glucose-6 phosphatedehydrogenase plays a central role in the response of tomato (*Solanum lycopersicum*) plants to short and long-term drought. Plant Physiol Biochem 105:79–89. https://doi.org/10.1016/j.plaphy.2016.04.013

Lu S, Yao S, Wang G, Guo L, Zhou Y, Hong Y, Wang X (2016) Phospholipase Dε enhances *Braasica napus* growth and seed production in response to nitrogen availability. Plant Biotechnol J 14(3):926–937. https://doi.org/10.1111/pbi.12446

Mackay I, Powell W (2007) Methods for linkage disequilibrium mapping in crops. Trends Plant Sci 12(2):57–63. https://doi.org/10.1016/j.tplants.2006.12.001

Manjonda RV, Vorasoot N, Puppala N, Muetia AM, Jogloy S (2018) Reproductive efficiency and yield responses of Valencia peanut genotypes under terminal drought conditions. Khon Kaen Agric J 46:181–192

Meng L, Guo L, Ponce K, Zhao X, Ye G (2016) Characterization of three *Indica* rice multiparent advanced generation intercross (MAGIC) populations for quantitative trait loci identification. Plant Genome 9(2):2015–10. https://doi.org/10.3835/plantgenome2015.10.0109

Na JK, Metzger JD (2020) A putative tomato inositol polyphosphate 5-phosphatase, Le5PT1, is involved in plant growth and abiotic stress responses. 3 Biotech 10(1):28. https://doi.org/10.1007/s13205-019-2023-y

Nassar S, Al-Kady A, El-Saka Z (2018) Effect of drought stress on yield and yield components of 20 peanut genotypes grown under newly reclaimed soil. Egypt J Agron 40(1):45–58. https://doi.org/10.21608/agro.2018.2275.1088

Nawade B, Chu SH, Kim S, Lee CY, Bao J, Park YJ (2024) Nested association mapping population in *japonica* rice: development, characterization, and application in genome-wide association studies. Curr Plant Biol 40:100401. https://doi.org/10.1016/j.cpb.2024.100401

Noubissie TJB, Njintang NY, Dolinassou S (2012) Heritability studies of protein and oil contents in groundnut (*Arachis hypogaea* L.) genotypes. Int J Innov Biol Sci 2(3):162–171

Otyama PI, Wilkey A, Kulkarni R, Assefa T, Chu Y, Clevenger J, O'Connor DJ, Wright GC, Dezern SW, MacDonald GE, Anglin NL (2019) Evaluation of linkage disequilibrium, population structure, and genetic diversity in the US peanut mini core collection. BMC Genomics 20:1–17. https://doi.org/10.1186/s12864-019-5824-9

Otyama PI, Chamberlin K, Ozias-Akins P, Graham MA, Cannon EKS, Cannon SB, MacDonald GE, Anglin NL (2022) Genome-wide approaches delineate the additive, epistatic, and pleiotropic nature of variants controlling fatty acid composition in peanut (*Arachis hypogaea* L.). G3 Genes|genomes|genetics 12(1):382. https://doi.org/10.1093/g3journal/jkab382

Pal L, Sandhu SK, Kaur J, Bhatia D (2024) Deciphering variations, identification of marker–trait associations and candidate genes for seed oil content under terminal heat stress in Indian mustard (*Brassica juncea* L. Czern & Coss) germplasm stock. 3 Biotech 14(5):140. https://doi.org/10.1007/s13205-024-03985-w

Pande S, Upadhyaya HD, Rao JN, Reddy PL, Rao PP (2005) Promotion of Integrated Disease Management of ICGV 91114, a dual-purpose, early maturing groundnut variety for rainfed areas. Information Bulletin No. 68

Pandey MK, Agarwal G, Kale SM, Clevenger J, Nayak SN, Sriswathi M, Chitikineni A, Chavarro C, Chen X, Upadhyaya HD, Vishwakarma MK, Leal-Bertioli S, Liang X, Bertioli DJ, Guo B, Jackson SA, Ozias-Akins P, Varshney RK (2017) Development and evaluation of a high density genotyping 'Axiom_Arachis' array with 58 K SNPs for accelerating genetics and breeding in groundnut. Sci Rep 7(1):40577. https://doi.org/10.1038/srep40577

Pandey MK, Guo B, Holbrook CC, Janila P, Zhang X, Bertioli DJ, Varshney RK (2014) Molecular markers, genetic maps and QTLs for molecular breeding in peanut. Genetics, genomics and breeding of peanuts 61–113

Peng Y, Zhang J, Cao G, Xie Y, Liu X, Lu M, Wang G (2010) Overexpression of a PLDα1 gene from *Setaria italica* enhances the sensitivity of Arabidopsis to abscisic acid and improves its drought tolerance. Plant Cell Rep 29(7):793–802. https://doi.org/10.1007/s00299-010-0865-1

Peng Z, Zhao Z, Clevenger JP, Chu Y, Paudel D, Ozias-Akins P, Wang J (2020) Comparison of SNP calling pipelines and NGS platforms to predict the genomic regions harboring candidate genes for nodulation in cultivated peanut. Front Genet 11:222. https://doi.org/10.3389/fgene.2020.00222

Pereira JW, Albuquerque MB, Melo Filho PA, Nogueira RJMC, de Lima LM, Santos RC (2016) Assessment of drought tolerance of peanut cultivars based on physiological and yield traits in a semiarid environment. Agric Water Manag 166:70–76. https://doi.org/10.1016/j.agwat.2015.12.010

Pokhrel S, Kharel P, Pandey S, Botton S, Nugraha GT, Holbrook C, Ozias-Akins P (2025) Understanding the impacts of drought on peanuts (*Arachis hypogaea L.*): exploring physio-genetic mechanisms to develop drought-resilient peanut cultivars. Front Genet 15:1492434. https://doi.org/10.3389/fgene.2024.1492434

Rahman MA, Woo JH, Song Y, Lee SH, Hasan MM, Azad MAK, Lee KW (2022) Heat shock proteins and antioxidant genes involved in heat combined with drought stress responses in perennial rye grass. Life 12(9):1426. https://doi.org/10.3390/life12091426

Ray TK, Holly SP, Knauft DA, Abbott AG, Powell GL (1993) The primary defect in developing seed from the high oleate variety of peanut (*Arachis hypogaea* L.) is the absence of Δ12–desaturase activity. Plant Sci 91:15–21. https://doi.org/10.1016/0168-9452(93)90184-2

Raychaudhuri S, Prinz WA (2010) The diverse functions of oxysterol-binding proteins. Annu Rev Cell Dev Biol 26(1):157–177

Rezaei MK, Shobbar ZS, Shahbazi M, Abedini R, Zare S (2013) Glutathione S-transferase (GST) family in barley:identification of members, enzyme activity, and gene expression pattern. J plant physiol 170 (14): 1277-1284. https://doi.org/10.1016/j.jplph.2013.04.005

Reddy TVB, Rao JP, Vijayalakshmi K (2003) Physiological and biochemical evaluation of groundnut cultivars differing in drought tolerance. Indian J Plant Physiol 8:359–363

Sarvamangala C, Gowda MVC, Varshney RK (2011) Identification of quantitative trait loci for protein content, oil content and oil quality for groundnut (*Arachis hypogaea* L.). Field Crops Res 122(1):49–59. https://doi.org/10.1016/j.fcr.2011.02.010

SAS II (2015) Base SAS 9.4 procedures guide. Statistical Procedures

Shaibu AS, Motagi B, Ayittah P (2019) Genome wide association studies for fatty acids, mineral and proximate compositions in groundnut (*Arachis hypogaea* L.) seeds. Pertanika J Trop Agric Sci 42(3):939–955

Sharma V, Mahadevaiah SS, Latha P, Gowda SA, Manohar SS, Jadhav K, Pandey MK (2024) Dissecting genomic regions and underlying candidate genes in groundnut MAGIC population for drought tolerance. BMC Plant Biol 24(1):1044. https://doi.org/10.1186/s12870-024-05749-3

Shen Y, Liu YH, Zhang XJ, Sha Q, Chen ZD (2016) Gynophore miRNA analysis at different developmental stages in *Arachis duranensis*. Genet Mol Res 15(4):8691. https://doi.org/10.4238/gmr15048691

Simpson CE, Starr JL, Church GT, Burow MD, Paterson AH (2003) Registration of'NemaTAM'peanut. Crop Sci 43(4):1561

Sinha P, Bajaj P, Pazhamala LT, Nayak SN, Pandey MK, Chitikineni A, Huai D, Khan AW, Desai A, Jiang H, Zhuang W (2020) Arachis hypogaea gene expression atlas for fastigiata subspecies of cultivated groundnut to accelerate functional and translational genomics applications. Plant Biotechnol J 18(11):2187-2200. https://doi.org/10.1111/pbi.13374

Solanki MV, Mahatma MK, Varma A, Thawait LK, Singh S, Jangir CK, Meena MD, Tomar RS, Rathod PJ, Vala A, Kandoliya UK (2024) Water deficit stress enhances the bioactive compounds of groundnut (*Arachis hypogaea L.*) kernels at the expense of primary metabolites. Food Biosci 58:103670. https://doi.org/10.1016/j.fbio.2024.103670

Songsri P, Jogloy S, Holbrook CC, Kesmala T, Vorasoot N, Akkasaeng C, Patanothai A (2009) Association of root, specific leaf area and SPAD chlorophyll meter reading to water use efficiency of peanut under different available soil water. Agric Water Manage 96(5):790–798. https://doi.org/10.1016/j.agwat.2008.10.009

Sun L, Ren H, Liu R, Li B, Wu T, Sun F, Liu H, Wang X, Dong H (2010) An h-type thioredoxin functions in tobaccodefense responses to two species of viruses and an abiotic oxidative stress. Mol plant-microbe interact 23(11):1470–1485. https://doi.org/10.1094/MPMI-01-10-0029

Stadlmeier M, Hartl L, Mohler V (2018) Usefulness of a multiparent advanced generation intercross population with a greatly reduced mating design for genetic studies in winter wheat. Front Plant Sci 9:1825. https://doi.org/10.3389/fpls.2018.01825

Sundaram J, Kandala CV, Holser RA, Butts CL, Windham WR (2010) Determination of in-shell peanut oil and fatty acid composition using near-infrared reflectance spectroscopy. J Am Oil Chem Soc 87(10):1103–1114. https://doi.org/10.1007/s11746-010-1589-7

Thiry AA, Chavez Dulanto PN, Reynolds MP, Davies WJ (2016) How can we improve crop genotypes to increase stress resilience and productivity in a future climate? A new crop screening method based on productivity and resistance to abiotic stress. J Exp Bot 67:5593–5603. https://doi.org/10.1093/jxb/erw330

Tunnacliffe A, Wise MJ (2007) The continuing conundrum of the LEA proteins. Naturwissenschaften. https://doi.org/10.1007/s00114-007-0254-y

Vadez V (2013) Drought tolerance in groundnut. Advances in Agronomy

Vindhiyavarman P, Nigam SN, Janila P, Vaidhyalingam M, Manivannan N, Saravanan S, Meenakumari B, Gopalakrishnan C, Kennedy JS (2014) A new high yielding Spanish bunch groundnut variety CO 7 (ICGV 00351) for the drought prone areas of Tamil Nadu. Electron J Plant Breed 5(2):192–196

Vos PG, Paulo MJ, Voorrips RE, Visser RG, van EHJ, van EFA (2017) Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato. Theor Appl Genet 130(1):123–135. https://doi.org/10.1007/s00122-016-2798-8

Vuong TD, Florez-Palacios L, Mozzoni L, Clubb M, Quigley C, Song Q, Nguyen HT (2023) Genomic analysis and characterization of new loci associated with seed protein and oil content in soybeans. Plant Genome 16(4):e20400. https://doi.org/10.1002/tpg2.20400

Wang M, Wang Y, Wu H, Xu J, Li T, Hegebarth D, Jetter R, Chen L, Wang Z (2016) Three TaFAR genes function in thebiosynthesis of primary alcohols and the response to abiotic stresses in Triticum aestivum. Sci Rep 6(1):25008. https://doi.org/10.1038/srep25008

Wang J, Zhang Z (2021) GAPIT version 3: boosting power and accuracy for genomic association and prediction. Genom Proteom Bioinform 19(4):629–640. https://doi.org/10.1016/j.gpb.2021.08.005

Wang P, Wang H, Wang Y, Ren F, Liu W (2018) Analysis of bHLH genes from foxtail millet (*Setaria italica*) and their potential relevance to drought stress. PLoS ONE 13(11):e0207344. https://doi.org/10.1371/journal.pone.0207344

Wang X, Zhang Y, Wang L, Pan Z, He S, Gao Q, Du X (2020) Casparian strip membrane domain proteins in *Gossypium arboreum*: genome-wide identification and negative regulation of lateral root growth. BMC Genomics 21:1–16

Wankhade AP, Chimote VP, Viswanatha KP, Yadaru S, Deshmukh DB, Gattu S, Sudini HK, Deshmukh MP, Shinde VS, Vemula AK, Pasupuleti J (2023) Genome-wide association mapping for LLS resistance in a MAGIC population of groundnut (*Arachis hypogaea* L.). Theor Appl Genet 136(3):43. https://doi.org/10.1007/s00122-023-04256-7

Xu Y, Yan F, Zong Y, Li J, Gao H, Liu Y, Wang Q (2022) Proteomic and lipidomics analyses of high fatty acid AhDGAT3 transgenic soybean reveals the key lipase gene associated with the lipid internal mechanism. Genome 65(3):153–164. https://doi.org/10.1139/gen-2021-0043

Xu J, Jiang X, Yin X, Zhao X, Chen N, Pan L, Fu C, Jiao Y, Ma J, Yuan M, Chi X (2025) Genome-wide association analysis in peanut accessions uncovers the genetic basis regulating oil and fatty acid variation. BMC Plant Biol 25(1):651

Xue W, Yan J, Zhao G, Jiang Y, Cheng J, Cattivelli L, Tondelli A (2017) A major QTL on chromosome 7HS controls the response of barley seedling to salt stress in the Nure× Tremois population. BMC Genet 18(1):79

Yamaguchi-Shinozaki K, Shinozaki K (1993) The plant hormone abscisic acid mediates the drought-induced expression but not the seed-specific expression of rd22, a gene responsive to dehydration stress in *Arabidopsis thaliana*. Mol Gen Genet 238:17–25. https://doi.org/10.1007/bf00279525

Yang J, Chen R, Liu W, Xiang X, Fan C (2024) Genome-wide characterization and phylogenetic and stress response expression

analysis of the MADS-box gene family in litchi (*Litchi chinensis* Sonn.). Int J Mol Sci 25(3):1754. https://doi.org/10.3390/ijms25031754

Yol E, Ustun R, Golukcu M, Uzun B (2017) Oil content, oil yield and fatty acid profile of groundnut germplasm in Mediterranean climates. J Am Oil Chem Soc 94(6):787–804. https://doi.org/10.1007/s11746-017-2981-3

Yuan D, Wu Z, Wang Y (2016) Evolution of the diacylglycerol lipases. Prog Lipid Res 64:85–97. https://doi.org/10.1016/j.plipres.2016.08.004

Zhang T, Zhao Y, Wang Y, Liu Z, Gao C (2018) Comprehensive analysis of MYB gene family and their expressions under abiotic stresses and hormone treatments in *Tamarix hispida*. Front Plant Sci 9:1303. https://doi.org/10.3389/fpls.2018.01303

Zhang H, Chu Y, Dang P, Tang Y, Jiang T, Clevenger JP, Ozias-Akins P, Holbrook C, Wang ML, Campbell H, Hagan A, Chen C (2020) Identification of QTLs for resistance to leaf spots in cultivated peanut (*Arachis hypogaea* L.) through GWAS analysis. Theor Appl Genet 133(7):2051–2061. https://doi.org/10.1007/s00122-020-03576-2

Zhou Q, Zhou C, Zheng W, Mason AS, Fan S, Wu C, Fu D, Huang Y (2017) Genome-wide SNP markers based on SLAF-Seq uncover breeding traces in rapeseed (*Brassica napus* L.). Front Plant Sci 8:648. https://doi.org/10.3389/fpls.2017.00648

Zhou W, Song S, Dossou SSK, Zhou R, Wei X, Wang Z, Wang L (2022) Genome-wide association analysis and transcriptome reveal novel loci and a candidate regulatory gene of fatty acid biosynthesis in sesame (*Sesamum indicum* L.). Plant Physiol Biochem 186:220–231. https://doi.org/10.1016/j.plaphy.2022.07.023