



OPEN ACCESS

EDITED BY

Carlos I. Arbizu,
Universidad Nacional Toribio Rodríguez de
Mendoza de Amazonas, Peru

REVIEWED BY

Sandra Rychel-Bielska,
Wroclaw University of Environmental and Life
Science, Poland
Jianbo He,
Nanjing Agricultural University, China

*CORRESPONDENCE

Neeraj Kumar,
✉ Neeraj.Kumar2@usda.gov

RECEIVED 30 August 2025

REVISED 03 December 2025

ACCEPTED 08 December 2025

PUBLISHED 14 January 2026

CITATION

Kumar N, Smith JR, Ray JD, Gillman JD and
Bellaloui N (2026) Identification of major QTLs
for seed vigor and growth-related traits using a
biparental population in soybean.
Front. Genet. 16:1695593.
doi: 10.3389/fgene.2025.1695593

COPYRIGHT

© 2026 Kumar, Smith, Ray, Gillman and
Bellaloui. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is permitted
which does not comply with these terms.

Identification of major QTLs for seed vigor and growth-related traits using a biparental population in soybean

Neeraj Kumar^{1*}, James R. Smith¹, Jeffery D. Ray¹,
Jason D. Gillman² and Nacer Bellaloui¹

¹USDA, Agriculture Research Service, Crop Genetics Research Unit, Stoneville, MS, United States, ²Plant Genetics Research Unit USDA-ARS, 205 Curtis Hall, University of Missouri, Columbia, MO, United States

Soybean [*Glycine max* (L.) Merr.] is one of the most widely cultivated crops globally and serves as a primary source of plant-based protein and oil for human consumption. Seed vigor is a critical trait for germination and rapid seedling establishment, especially under sub-optimal environmental conditions. Genetic control of seed vigor measured as accelerated aging (AA%) was investigated using a quantitative trait locus (QTL) mapping approach. Associations of AA with six other traits: pubescence color (PC), beginning bloom (R1), reproductive period (RP), maturity (R8), plant height (PH), and stem termination (ST) were examined. A recombinant inbred line (RIL) population (247 F₆-derived RILs) from a cross between DS25-1 and DT97-4290 was developed and genotyped using genotyping-by-sequencing (GBS), which yielded a set of 8,445 curated single-nucleotide polymorphism (SNP) markers from ~90,000 SNPs. A subset of 201 RILs was phenotyped in Stoneville, Mississippi, over 3 years (2017–2019). A molecular linkage map was constructed using the SNP marker dataset, and composite interval mapping was performed using the R/qtl package. In total, 33 QTLs associated with seven phenotypic traits were identified across 12 chromosomes, using means of individual environments and best linear unbiased prediction (BLUP). Phenotypic variation of individual QTLs ranged from 0.56% to 89.77%, and the additive effects varied from –10.52% to 15.12%. Twelve QTLs detected across multiple environments on Gm03, Gm06, Gm07, and Gm19 were classified as stable. Notably, four genomic regions demonstrated QTL co-localization: Gm06 (all traits except ST), Gm07 (AA, R1, RP, and R8), Gm03 (AA and R1), and Gm19 (PH, R8, and ST). Major QTLs were in proximity to previously known genes: the *T* locus for pubescence color (Gm06), *E1* (Gm06) and *E11* (Gm07) loci for flowering and maturity, and *Dt1* (Gm19) for stem termination. The closest SNPs associated with stable QTLs can be used to develop KASP markers for marker-assisted selection programs.

KEYWORDS

growth-related traits, high-density genetic map, QTL mapping, seed vigor, soybean

Introduction

Soybean [*Glycine max* (L.) Merr.] is a major global crop; however, seed yields in farmers' fields rarely approach the yield potential of modern soybean cultivars. Achieving high yields and superior seed quality relies on multiple factors, including selecting high-yielding varieties with strong seed vigor, high germination rates, and successful seedling

establishment, and effective management practices. These practices also involve minimizing biotic and abiotic stress throughout the crop season.

Seed vigor reflects a seed's capacity to quickly germinate and establish a stand in a wide range of environmental conditions. Poor seed emergence directly reduces crop yield by decreasing plant density per unit area due to low germination rates and weak vigor (Han et al., 2014; Börner et al., 2018). Seed vigor is governed by complex interactions between internal (endogenous) and external (exogenous) factors that influence seed development, harvest, and seed storage (Bentsink et al., 2000). The endogenous factors include chlorophyll content (as seen in chlorophyllous seeds), seed coat structure, phytohormone levels, and the integrity of proteins and nucleic acids (DNA and RNA), along with their repair mechanisms (Zhou et al., 2020). Exogenous factors begin to influence physiological maturity and continue to do so throughout seed storage. These include environmental conditions such as sub-optimal temperatures, relative humidity, and oxygen levels during seed storage (Nadarajan et al., 2023). High temperatures and humidity during seed storage can significantly reduce seed viability by disrupting critical biochemical processes. These adverse conditions enhance hydrolytic enzyme activity, increase respiration rates, and promote lipid peroxidation, thereby elevating free fatty acid levels (Copeland and McDonald, 2001; Wang et al., 2018).

In the mid-south of the United States, soybeans are grown using an early soybean production system (ESPS), where planting occurs early (April), and optimally, the crop is harvested early (August) as well. However, when harvest is delayed, the seed can be exposed to adverse weather conditions (Pinheiro et al., 2021), increasing the risk of seed deterioration from microbes that cause rot and prevent the seed from sprouting within the pod due to excess moisture. Eventually, both seed quality and composition are compromised. Such mature seed damage reduces seed vigor and can lead to substantial revenue losses when the grain is eventually marketed (Heatherly, 1999). Yet, the ESPS is used because it has been shown to provide consistent yield advantages and increased profitability under both irrigated and non-irrigated conditions (Heatherly and Spurlock, 1999). Gaining insight into the genetic basis of seed vigor can facilitate the development of cultivars with enhanced viability and seed quality.

Previously, the screening of genetically diverse exotic soybean germplasm, including 25 ancestral lines of U.S. cultivars, revealed poor germination of the ancestor lines under ESPS conditions (Smith et al., 2008; Smith et al., 2017; Gillman et al., 2020). The average germination rates for these ancestors were 50.8% in standard tests and 37.9% under accelerated aging (AA). Moreover, ancestral seeds exhibited a high incidence of hard seeds (13.2%), wrinkling (29.6%), and *Phomopsis longicolla* infection (22.5%), demonstrating that the use of the ancestor gene pool alone would likely be insufficient to improve seed vigor in modern cultivars. In contrast, seeds of an unadapted exotic landrace, PI 587982A, maintained 92% and 90% germination in standard and AA tests, respectively, with no hard, wrinkled, or diseased seeds (Smith et al., 2008). One of the parental lines used in this study, DS25-1 (PI 684675: <https://npgsweb.ars-grin.gov/gringlobal/>), was derived from PI 587982A and was crossed with DT97-4290 (Paris et al., 2006), a low-vigor line, to develop the RIL population for the current study.

In the last 2 decades, genomic approaches, such as genome-wide association studies (GWAS) and quantitative trait locus (QTL) mapping, have been successfully used to dissect various complex traits related to agronomy, growth period, growth habits, flowering, maturity, and tolerance to biotic and abiotic stresses (see review from Xia et al., 2013; Zhang et al., 2015; Liu et al., 2016; Zimmer et al., 2021; Kim et al., 2023; Liu et al., 2023). Fang et al. (2017) used GWAS to identify several candidate genes linked to various important traits related to growth period, architecture, seed yield, and nutrient composition using large collections of soybean germplasm (809 accessions). Previously, biparental RIL populations were widely used to map QTLs associated with complex quantitative traits, including seed yield, seed size, and seed composition (Yao et al., 2023; Gao et al., 2024).

QTL mapping remains a powerful tool for dissecting the genetic architecture of complex traits in crop species. In previous years, significant efforts were made to identify loci associated with high seed vigor and storability in various crops through genetic studies. In *Arabidopsis*, several loci associated with seed storability were identified (Bentsink et al., 2000; Clerkx et al., 2004). In rice, numerous QTL mapping studies using RIL populations pinpointed loci for early seed vigor (Yang et al., 2019; Zhang et al., 2017), maturity-related seed vigor (Liu et al., 2014; Lai et al., 2016), and seed longevity (Raquid et al., 2021). Integrative analyses, combining GWAS, QTL mapping, and transcriptomic data, have further identified candidate genes for seed vigor in rice (Yang et al., 2021). Similarly, studies in maize using RIL populations uncovered several QTLs related to seed maturity at harvest (Jing-bao et al., 2011), seed longevity (Han et al., 2018), tolerance to stressful germination conditions (Han et al., 2014), and artificial aging tolerance (Wu et al., 2019). In wheat, Shi et al. (2020) conducted QTL and candidate gene analysis to identify genes associated with seed vigor. Seed vigor studies have also been conducted in barley (Nagel et al., 2016; Thabet et al., 2018) and oat using QTL mapping and GWAS (Huang et al., 2020). To date, only a few studies have identified loci associated with seed vigor and storability in soybean using QTL mapping approaches. Dargahi et al. (2014) reported three loci controlling seed storability on linkage groups C1 (Gm04), F (Gm13), and L (Gm19). Zhang et al. (2019) identified 34 stable QTLs for seed storability across 11 chromosomes using two different RIL populations. Wang et al. (2021) identified two major QTLs and eight potential minor QTL on chromosomes 3, 6, 9, 11, 15, 16, 17, and 19 for traits related to accelerated aging using a biparental mapping population using seeds conserved at -20°C . Jiamtae et al. (2022) reported five QTLs associated with seed viability and seed vigor on chromosomes 2, 6, and 8, which were co-localized with loci linked to seed variability, maturity, germination, hardness, and micromorphology.

For the current study, we constructed a high-density genetic linkage map to map QTLs for seed vigor and growth-related traits using an entire RIL population developed from a cross between two parental lines (DS25-1 and DT97-4290). The genetic map, in conjunction with phenotypic data collected over 3 years (2017–2019), was used for composite interval mapping (CIM) using the R/qtl package (Broman and Sen, 2009). Several consistent QTLs across the environments were identified for seven traits (AA, PC, R1, RP, R8, ST, and PH), including both novel and previously identified QTLs and candidate genes. In

addition, we identified four genomic regions that harbor multiple stable QTLs for different traits. SNPs associated with stable QTLs for the corresponding traits can be converted into KASP markers to facilitate marker-assisted selection for improving seed vigor and overall soybean productivity in improved cultivars.

Materials and methods

Development of the RIL population

The cross was made between DS25-1 (Smith et al., 2017) and DT97-4290 (Paris et al., 2006) at Stoneville, MS, in 2012. DS25-1 is a maturity group (MG) IV line and is tolerant to heat-induced seed degradation. DT97-4290 is a high-yielding MG IV public germplasm line with moderate resistance to charcoal rot (*Macrophomina phaseolina* [Tassi] Goid.) but susceptibility to heat-induced seed degradation reduced germination (Chebrolo et al., 2016) and increased levels of seed wrinkling (Narayanan et al., 2020), purple seed stain, and mold (Smith et al., 2025). In 2013, F₂ seeds were harvested at the USDA Tropical Agriculture Research Station at Isabela, Puerto Rico, and planted at Stoneville. In the fall, 247 individual F₂ plants were randomly selected to initiate the RIL population. The F₃ and F₅ RIL generations were advanced by randomly selecting a single plant from each heterogeneous progeny row at Homestead, FL (27 Farms of Homestead, Inc.), during the winters of 2013/2014 and 2014/2015. In contrast, the F₄ and F₆ generations were advanced at Stoneville during the summers of 2014 and 2015 by the same method. Single F₆-derived F₇ seeds were harvested in 2015 for seed increase and further field evaluations.

Field evaluation

All 247 RILs were planted on 10 April 2017 in 2.7 m-long single-row plots with a row spacing of 0.91 m with two replications using a randomized complete block design (RCBD) at Stoneville at the Delta Research and Extension Center (33° 26' N, -90° 54' W) in 2017. Seeds were sown with a machine planter into rows at a depth of 2.5 cm, with a seeding rate of 25 seeds m⁻¹ of row. The soil type was a Sharkey clay soil (very fine, smectitic, thermic Chromic Epiaquert). The 2017 experiment was furrow-irrigated as needed to avoid water-deficit stress and ensure adequate seed production for future years of testing. After planting and emergence, Reflex (Syngenta, DE, United States) and Select (Valent Corporation, CA, United States) herbicides, along with multiple timely cultivations, were used for weed control. Non-irrigated (rainfed) Stoneville plantings were made on 20 April 2018 and 22 April 2019 using a subset of 201 RILs, with lines maturing earlier than Pella 86 (Fehr et al., 1987) and later than Dillon (Shipe et al., 1997) not planted in the second and third years. The same experimental procedures as noted above, except for irrigation, were used in 2018 and 2019. The checks used in 2017 were early MG III Pella 86, late MG III Maverick (Sleper et al., 1998), late MG IV parents DT97-4290 and DS25-1, MG V Osage (Chen et al., 2007), and MG VI Dillon. In 2018 and 2019, the same checks were used with the addition of the following: mid MG III 7037-12111 (an unreleased high germinability line derived from PI 587982A × N98-4445A), late MG III germplasm lines DS65-1 and LG03-4561-14, early MG IV DS34-1 (which is

referred to as 34-3-1-2-4-1), early MG IV commercial cultivars Progeny 4211 and AG4232, and late MG IV and mid MG V commercial cultivars C4926 and AG5606 (<https://legacy.soybase.org/uniformtrial/index.php?page=lines&filter>).

Trait phenotyping

Growth stages (Fehr et al., 1977) R1 (beginning bloom) and R8 (full maturity) were recorded in days after planting (DAP) for each plot. In each year, R1 was estimated weekly when 50% of the plants in a plot had flowered, whereas R8 was estimated twice weekly in each plot when 95% of the pods had reached their mature pod color. The duration of the reproductive period (RP) was estimated in days as the difference between R1 and R8. Pubescence color (PC) and stem termination (ST) type were recorded for each plot at R8. The RIL population was segregating for two pubescence colors, as one parent (DS25-1) had gray pubescence, whereas the other parent (DT97-4290) had tawny pubescence. Similarly, the population segregated for stem termination habit, as one parent (DS25-1) was a determinate (D) type, whereas the other parent (DT97-4290) was an indeterminate (I) type. Plant height (PH) for each plot was estimated in 2019, measured as the distance in cm from the ground to the tip of the plant. PH estimates were taken on three plants per plot and then averaged. The timely harvest of seed was completed for each plot shortly after full maturity by hand-harvesting each plot, followed by storing/drying the cut plants at 21 °C and 60% relative humidity (RH). Harvested plants were threshed using an Almaco bundle thresher (Almaco, Nevada, IA). Split seeds and debris (sticks, dirt, etc.) were removed from each seed lot at the time of threshing, whereas damaged seed (wrinkled, shriveled, discolored, green, purple, moldy, insect feeding) was retained for later grading. The seed of each line was stored at 21 °C and 60% RH until all lines were harvested, threshed, and cleaned of debris. Accelerated aging (%; AA) tests were conducted in 2018 and 2019 at the State Seed Testing Laboratory (Mississippi State, MS) following standard protocols of the Association of Official Seed Analysts (2002). For each seed lot, ~42 g of random seed was sampled, and 200 seeds were divided into four replicates of 50. Damaged or abnormal seeds were removed, and seed moisture content was adjusted to 10%–12%. Each replicate was placed as a single layer on a screen tray in an accelerated-aging box containing ~40 mL distilled water to maintain nearly 100% relative humidity. Boxes were sealed and incubated at 41 °C–45 °C for 72 h to simulate rapid aging. After aging, boxes were cooled for 1 h before opening to prevent moisture loss. Seeds were then tested for germination on moist paper or in moist sand using the rolled towel or between-paper method at 25 °C ± 2 °C for 7 days. The percentage of normal seedlings, determined by AOSA criteria, was recorded as the accelerated aging value.

$$AA(\%) = \frac{\text{Number of normal seedlings after aging}}{\text{Total number of seeds tested}} \times 100.$$

Statistical analyses

To study the relationship among the phenotypic traits, a Pearson's correlation coefficient matrix was generated using *corr_plot* and *plot.corr_coef* functions of the *metan* package in R software

(R Foundation, 2024). The variance components for genotypes (i.e., RILs) were estimated for quantitative traits (AA, R1, RP, R8, and PH) using the *lme4()* function in R (Bates et al., 2015). We used $y_{ijkl} = \mu + E_j + G_i + (GE)_{ij} + R_{k(j)} + \varepsilon_{ijkl}$ where environment (year) was treated as a fixed effect, and genotype (G_i), genotype $\times$ environment interaction ($GE)_{ij}$, and replication within year $R_{k(j)}$ were fitted as random effects. The mixed linear models were optimized using restricted maximum likelihood and were used to determine variance components for each random effect. Best linear unbiased predictions (BLUPs) were also calculated for the above traits using the random effect of each genotype within and across environments using the *lme4* package in R software (Bates et al., 2015). Broad-sense heritability was calculated on an entry-mean basis according to Holland et al. (2003), with the variance components estimated using a model where all variables were treated as random.

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_{G \times R}^2}{R} + \frac{\sigma_{G \times Y}^2}{Y} + \frac{\sigma_E^2}{RY}}$$

where G is genotype, R is replicate, Y is year, and E is error.

Production of genotyping-by-sequencing (GBS) markers and data processing

Leaf samples from each of the 247 RILs and both parental lines were collected from experimental field plots grown at Stoneville, MS, in 2015. The collected leaf sample was lyophilized using a Model 2400 freeze dryer (The Freeze Dry Company, Nisswa, MN 56468, United States). Genomic DNA was extracted employing a DNA isolation kit (Promega, Madison, WI 53711, United States) in a Maxwell 16™ automated DNA isolation machine. The genotyping was carried out using GBS markers technology (Elshire et al., 2011) by LGC Genomics GmbH (Berlin, Germany) following their established protocols (<https://www.biosearchtech.com/>) for normalized GBS (nGBS) with MseI and Illumina NextSeq (2 $\times$ 150 bp) sequencing. DNA sequences were processed by LGC Genomics bioinformaticists for quality, and the GBS tags were aligned to the “Williams 82” soybean reference genome assembly 1 (<https://www.soybase.org/>). The aligned sequences were separated for each plant sample using the in-line barcode sequences, and then the reads were processed to remove the barcode and adapter sequences. For sequence alignment, the Burrows–Wheeler Aligner (BWA v.0.7.12; <http://bio-bwa.sourceforge.net/>) program was used (Li and Durbin, 2009), and SNPs were called using Freebayes v.1.92-16 (Garrison and Marth, 2012).

Quality control of GBS-SNP markers

Prior to genetic mapping, monomorphic SNPs that had excessive heterozygous or missing values (>15%) or did not follow a 1:1 segregation ratio (chi-square P value ≤ 0.01) were removed. Missing data for SNPs with <15% missing data were imputed using an LD-kNNi method, which is based on a k-nearest-neighbor genotype (Money et al., 2015). This analysis was implemented using TASSEL software (<https://www.maizegenetics.net/tassel>). This resulted in a set of 8,925 (8,890 + 35 scaffold) high-

quality-filtered SNPs, of which 8,890 polymorphic SNPs were used to construct a genetic linkage map.

Construction of the genetic linkage map

Prior to linkage map construction, the genotypes of all 247 RILs were converted into ABH allele format, where allele “A” is from DS25-1, “B” is from DT97-4290, and “H” is heterozygous. All polymorphic SNPs (8,890) were used for the construction of the linkage map using R/qtl (Broman et al., 2003). The ABH data format was used to form linkage groups with a logarithm of odds (LOD) threshold ≥ 3.0 and a maximum recombination frequency of 0.35. In total, 8,445 SNPs were assigned to 22 linkage groups (LGs), corresponding to 20 soybean chromosomes. The LGs were assigned to different chromosomes based on the genomic position of SNPs identified during SNP calling. The markers within the LGs were ordered using the “ordermarkers” and “ripple” functions in R/qtl to minimize the length of linkage maps. The genetic distance between the markers was calculated in centiMorgans (cM) using the Kosambi mapping function (Kosambi, 1944).

QTL analysis and candidate gene identification

QTL analysis was performed using the genetic linkage map described above, along with phenotypic data from seven traits (AA, PC, R1, RP, R8, PH, and ST). Mean values from individual environments of each phenotypic trait were used for QTL analysis. The phenotypic observations recorded for discrete traits (PC and ST) were converted into a numerical variable (binary coding) prior to data analysis. For PC, gray pubescence color was coded as “0,” tawny as “2,” and a heterozygous individual was coded as “1.” For the stem termination habit, the indeterminate type was coded as “0,” the determinate type as “2,” and a heterozygous individual was recorded as “1.” The individual environment includes Stoneville 2017 (I), Stoneville 2018 (II), and Stoneville 2019 (III). Across-environment QTL analysis was carried out using BLUP values for the continuous traits (AA, R1, RP, R8, and PH), which were denoted as IV. We performed composite interval mapping (CIM) via the “cim” function for QTL detection, following the methodology outlined in R/qtl (Broman and Sen, 2009). CIM analysis was performed using multiple QTL regression models using R software (Broman and Sen, 2009). A minimum logarithm of odds (LOD) score ≥ 2.5 was used to determine the presence of a putative QTL, whereas the threshold LOD scores were computed using 1,000 permutations. A QTL detected in at least 50% or more environments was classified as a stable QTL (Cao et al., 2019; Bazzar et al., 2020). Additionally, a QTL was classified as “Major” if it appeared with a LOD score ≥ 4.0 and phenotypic variation explained (PVE) $\geq 10\%$. Candidate genes were identified in SoyBase (<https://www.soybase.org/>) using the Williams 82.a1 (wm82.a1) reference genome (Schmutz et al., 2010). The intervals of stable QTLs were used to search for the candidate genes of each trait. Identified candidate genes were analyzed using Gene Ontology (GO) to reveal their molecular

TABLE 1 Descriptive statistics of the parents (DS25-1 and DT97-4290) and RIL population for plant height (PH) for a single year (2019), for accelerated aging (AA%) for 2 years (2018–2019), and beginning bloom (R1), reproductive period (RP), and maturity (R8) for 3 years (2017–2019).

Trait ^a	Parental lines of RILs		RIL population		
	DS25-1	DT97-4290	Mean	Range	CV
AA	85	45	54 (1.6)	6–89	31
R1 (DAP)	72	55	66 (0.7)	53–84	14
RP (days)	95	103	97 (0.5)	79–116	7
R8 (DAP)	168	158	163 (1.0)	130–187	9
PH (cm)	65	81	85 (1.9)	23–157	33

^aCV, coefficient of variation; DAP, days after planting. Value in parentheses is standard error.

function via SoyBase (<https://www.soybase.org/>). For the corresponding gene positions from the Glyma.wm82.a1 assembly to Glyma.wm82.a4, we used “gene model correspondence lookup” available at SoyBase (<https://www.soybase.org/>).

Results

Phenotypic diversity between parents and the RIL population

The parental lines (DS25-1 and DT97-4290) and the RIL population exhibited a wide range of phenotypic diversity across the quantitative traits (Table 1). DS25-1 displayed a significantly higher seed vigor score (accelerated aging, AA), at 85%, than DT97-4290, which had a score of 45% across years (2018–2019). DT97-4290 was an early-blooming line, reaching the R1 stage (beginning bloom) at 55 DAP, while DS25-1 reached R1, 17 days later, at 72 DAP. Similarly, DT97-4290 reached full maturity (R8) at 158 DAP, which was 10 days earlier than DS25-1 (168 DAP). However, the reproductive period (RP) was shorter in DS25-1 (96 days) than in DT97-4290 (103 days). In terms of plant height (PH), DT97-4290 (81 cm) was significantly taller than DS25-1 (66 cm). The RIL population showed a broad range of variation across all traits (Table 1), with transgressive segregation observed for all quantitative traits (AA, R1, RP, R8, and PH). The AA of the RILs ranged from 6% to 89%, with an average of 54%, which was significantly less than that of the parents. The RILs' PH ranged from 23 cm to 157 cm, with an average (84 cm) that was greater than the parental means. However, the means for three traits (R1, RP, and R8) in the RIL population were very close to the parental means. Variance component analysis showed that genotypic variance in the RILs, genotype $\times$ environment ($G \times E$), replication, and year had significant contributions to total phenotypic variance for each trait (AA, R1, RP, and R8) in this population (Supplementary Table S1). The environmental effects on phenotypic trait values were primarily attributed to year and replication for R1, RP, and R8, while AA showed a highly significant genotype $\times$ year interaction. The RIL population exhibited continuous variation in a near-normal distribution for RP and PH (Figure 1), while AA showed a slight negative skew. The distribution for R1 and R8 was bimodal.

Heritability and the relationship between phenotypic traits

The broad-sense heritability was highest for R8 (0.96) and lowest for PH (0.52). Heritability was also high for R1 (0.88), RP (0.81), and AA (0.81) (Table 2). Pearson's correlation coefficients among seven phenotypic traits are presented based on the means of phenotypic traits, including data for PH (2019), AA (2018–2019), and across the years for R1, RP, and R8 (2017–2019; Table 2). As expected, R1 and R8 showed strong, positive correlations across the 3 years ($r = 0.87$). However, R1 exhibited a moderate ($r = 0.39$) correlation with RP. Seed vigor (AA) exhibited moderate to strong positive correlations with R1, R8, and PH, but surprisingly, it had only a weak but significant ($0.12, P \leq 0.05$) relationship with RP. PH showed a moderate to strong positive relationship with R1, RP, and R8 ($r = 0.41$ – 0.55). Stem termination (ST) exhibited a moderate positive relationship with PH ($r = 0.38$), whereas the remaining traits did not show any relationship with RP or showed weak negative correlations with ST (Table 2). In contrast, pubescence color (PC) showed a significant moderate-to-strong negative relationship with all other phenotypic traits (Table 2) except ST, where it showed a significant weak but positive correlation ($r = 0.19$).

Construction of the genetic linkage map

A genetic linkage map was constructed using 8,445 quality-filtered SNPs derived from GBS across 247 RILs. In total, 22 linkage groups were formed, with one linkage group assigned to each of the 20 soybean chromosomes, except for Gm09 and Gm12, which each formed two separate linkage groups (Table 3). The cumulative length of the genetic linkage map was 2,519.91 cM with an average marker distance of 0.43 cM. The full genetic map is shown in Supplementary Figure S1. The maximum marker density was observed on Gm04 (1 marker/0.09 cM), while the minimum marker density was found on Gm12 (1 marker/0.97 cM). Among the chromosomes, Gm18 had the highest number of markers (999 SNPs), while Gm12 harbored the fewest (137 SNPs).

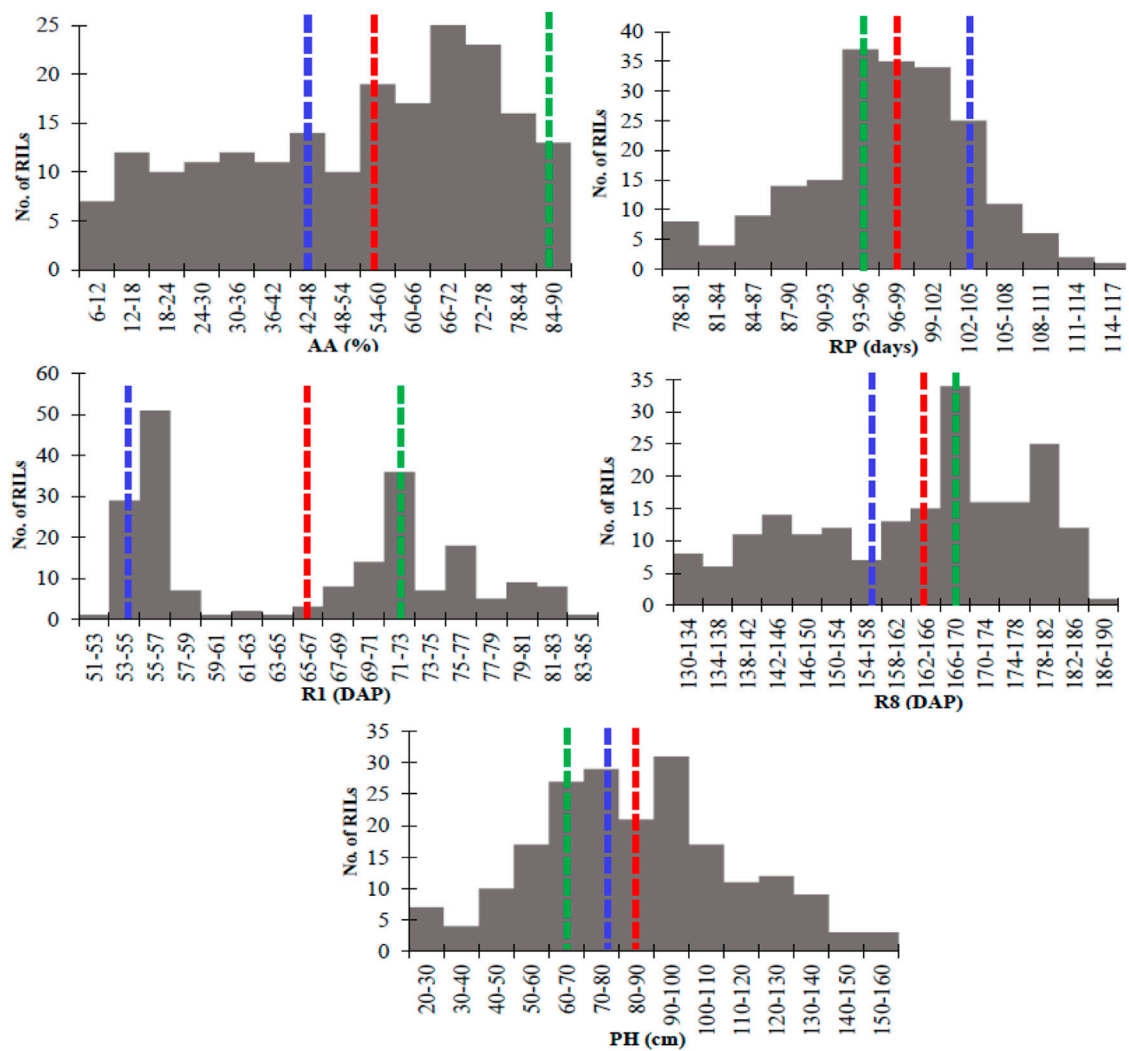


FIGURE 1 Frequency distribution of the means of different phenotypic traits of the recombinant inbred lines (RILs) derived from DS25-1 and DT97-4290. AA, accelerated aging (%); R1, beginning bloom; RP, reproductive period; R8, maturity; PH, plant height. Dashed green, blue, and red lines represent the means of the DS25-1, DT97-4290, and RIL populations, respectively. Means values of each trait were used for R1, RP, and R8 (2017–2019), AA (2018–2019), and PH (2019).

TABLE 2 Estimates of trait heritabilities and Pearson's correlation coefficients among phenotypic traits of the RIL population derived from a cross between DS25-1 and DT97-4290 across years.

Trait	PC	ST	R1	RP	R8	AA	H ²
ST	0.19**						-
R1	-0.87***	-0.16**					0.88
RP	-0.38***	0.02	0.39***				0.81
R8	-0.78***	-0.1	0.87***	0.80***			0.96
AA	-0.62***	-0.04	0.66***	0.12*	0.49***		0.81
PH	-0.49***	0.38***	0.51***	0.41***	0.55***	0.42***	0.52

Data used for pubescence color (PC), stem termination (ST), and plant height (PH), for only a single year, for accelerated aging (AA) for 2 years (2018–2019), and for beginning bloom (R1), reproductive period (RP), and maturity (R8), across the 3 years (2017–2019). H², broad-sense heritability. *, **, and *** indicate significance at the 0.05, 0.01, and 0.001 levels, respectively.

TABLE 3 Summary of genetic linkage maps using the biparental RIL mapping population derived from a cross between DS25-1 and DT97-4290.

Linkage groups ^a	Mapped SNPs	LG length (cM)	Average distance between SNPs (cM)	Max spacing between SNPs (cM)
Gm01/LG_D1a	393	126.56	0.32	9.40
Gm02/LG_D1b	292	177.08	0.61	35.45
Gm03/LG_N	345	148.72	0.43	26.18
Gm04/LG_C1	714	61.29	0.09	20.05
Gm05/LG_A1	179	132.74	0.75	16.70
Gm06/LG_C2	382	170.13	0.45	19.81
Gm07/LG_M	236	111.8	0.48	13.75
Gm08/LG_A2	188	124.58	0.67	21.3
<i>Gm09/LG_K.1</i>	597	55.88	0.09	13.75
<i>Gm09/LG_K.2</i>	56	34.65	0.63	11.96
Gm10/LG_O	173	60.66	0.35	20.77
Gm11/LG_B1	139	106.14	0.77	29.59
<i>Gm12/LG_H.1</i>	57	54.35	0.97	12.62
<i>Gm12/LG_H.2</i>	80	45.44	0.6	14.60
Gm13/LG_F	570	185.85	0.33	23.79
Gm14/LG_B2	797	128.16	0.16	8.62
Gm15/LG_E	257	133.53	0.52	24.33
Gm16/LG_J	573	128.14	0.22	35.44
Gm17/LG_D2	484	141.18	0.29	26.55
Gm18/LG_G	999	145.41	0.14	11.44
Gm19/LG_L	589	126.74	0.22	23.19
Gm20/LG_I	345	120.89	0.35	9.01
	8,445	2519.91	0.43	35.45

^acM, centiMorgan. Linkage groups (LGs) are italicized for chromosomes 09 and 12, where two LGs were made.

QTL for PC, PH, ST, and AA

To validate the RIL population, pubescence color (PC), a well-characterized qualitative trait in soybeans, was used for QTL mapping. Our results confirmed the presence of a major QTL at ~18.7 Mb on Gm06, which exhibited a very high LOD score (110.3) and phenotypic variation (89.77%) and had minimal additive effects (−0.75) (Supplementary Figure S2A). Additionally, a minor QTL was also identified on Gm05 (~41.7 Mb), with a LOD score of 2.80 and accounting for <1% of the phenotypic variation (Table 4).

We identified a major QTL (*qST.1*) for stem termination (ST), with a high LOD score (27.8), on Gm19 (~45.8 Mb). This QTL accounted for 65.63% of the phenotypic variation (Table 4; Supplementary Figure S2B) and had minimal additive effects (−0.75). For plant height (PH), two significant QTLs, *qPH.1* and *qPH.2*, were identified on chromosomes Gm06 (20.3 Mb) and Gm19 (45.8 Mb), with LOD scores of 16.2 and 7.2, respectively (Supplementary Figure S2C). A major QTL

(*qPH.1*) for PH accounted for 28.49% of the phenotypic variation, and another locus (*qPH.2*) accounted for 12.12% of the phenotypic variation. Additive effects were 15.52 and −10.09 (Table 4). Interestingly, QTLs for ST (*qST.1*) and PH (*qPH.2*) showed a large overlapping peak on Gm19 (~45.7 Mb), with a common closest SNP (S19_45757122).

In total, eight significant QTLs for AA were identified (Table 4), distributed on six different chromosomes (Gm03, Gm05, Gm06, Gm11, Gm13, and Gm14). A major AA QTL (*qAA.4*) was identified with a LOD score of 24.6, which explained the largest proportion of phenotypic variation (32.63%) (Supplementary Figure S2D). It was also identified in two environments, suggesting greater stability. The other QTLs associated with AA are minor, with LOD scores ranging from 2.6 to 4.4 and accounting for phenotypic variation ranging from 2.70% to 5.64%. Of these eight QTLs, a locus (*qAA.3*) detected in two environments (III and IV) on Gm05, and another locus (*qAA.4*) detected on Gm06 across the environments (II, III, and IV), were both considered stable QTLs (Supplementary

TABLE 4 QTL detected for pubescence color (PC), plant height (PH), and stem termination (ST) from the means of two replications of an individual environment, whereas for accelerated aging (AA%), 2 years of means (2018–2019) and estimated BLUPs were used for composite interval mapping using RIL population derived from DS25-1 and DT97-4290.

Trait ^a	QTL name	Chr/LG	QPP (Mb)	QPP (bp)	QGP (cM)	LOD	PVE (%)	AD effects	Env.	Parent allele
PC	qPC.1	Gm05/LG_A1	41.7	41,675,693	131.7	2.8	0.43	0.06	II	P1
	qPC.2	Gm06/LG_C2	18.7	18,666,488	117.1	110.3	89.77	0.90	II	P1
AA	qAA.1	Gm03/LG_N	29.1	29,076,783	56.2	2.8	3.05	−4.57	III	P2
	qAA.2	Gm03/LG_N	37.4	37,433,048	98.0	2.8	3.64	−3.16	IV	P2
	qAA.3	Gm05/LG_A1	41.7	41,675,693	131.7	3.1	3.32	4.81	III	P1
		Gm05/LG_A1	41.7	41,675,693	131.7	2.7	3.02	2.94	IV	P1
	qAA.4	Gm06/LG_C2	19.6	19,595,738	120.5	11.5	21.60	11.73	II	P1
		Gm06/LG_C2	19.6	19,616,246	120.5	24.6	32.63	14.91	III	P1
		Gm06/LG_C2	19.6	19,595,695	120.5	22.2	36.45	9.76	IV	P1
	qAA.5	Gm11/LG_B1	4.9	4,908,170	34.4	2.7	4.11	5.22	II	P1
	qAA.6	Gm11/LG_B1	7.2	7,224,445	48.2	2.6	2.70	4.30	III	P1
	qAA.7	Gm14/LG_B2	3.7	3,745,833	18.1	3.0	3.35	−4.69	III	P2
	qAA.8	Gm13/LG_F	27.9	27,940,207	68.3	4.4	5.64	4.03	IV	P1
PH	qPH.1	Gm06/LG_C2	20.3	20,291,873	121.5	16.2	28.49	15.52	III	P1
	qPH.2	Gm19/LG_L	45.8	45,757,122	90.9	7.2	12.12	−10.09	III	P2
ST	qST.1	Gm19/LG_L	45.8	45,757,122	90.9	27.8	65.63	−0.75	II	P2

^aChr/LG, chromosome/linkage group; QPP, QTL, physical position (megabase); QGP, QTL, genetic position (centiMorgan); LOD, logarithm of the odds; PVE, phenotypic variation explained; AD, additive effects; Env., environment; I, Stoneville 2017; II, Stoneville 2018; III, Stoneville 2019; IV, BLUP, dataset; P1, DS25-1, P2, DT97-4290.

Table S2). The remaining QTLs detected on Gm03, Gm11, and Gm14 were significant only in single environments (Table 4).

QTL for reproductive traits (R1, RP, and R8)

For R1, six QTLs (qR1.1 to qR1.6) were identified on five chromosomes (Gm01, Gm03, Gm06, Gm07, and Gm15), and the LOD scores ranged from 2.5 to 88.7 (Table 5; Supplementary Figure S2E). A major QTL identified on chromosome 6 (qR1.3) had the highest LOD score (88.7) and accounted for most of the phenotypic variation (82.58%). Comparatively, the other five QTLs were minor, with LOD scores ranging from 2.5 to 7.5 that explained a small portion of the phenotypic variation (0.83%–3.53%). Of these QTLs, three (qR1.3, qR1.5, and qR1.2) were consistently detected on Gm06 (~18.0–20.3 Mb), Gm07 (3.9 Mb), and Gm03 (37.4–38.9 Mb) in a minimum of two to four environments (I, II, III, and IV) and had above-the-threshold LOD scores. They were considered stable QTLs. However, two additional significant QTLs were detected only in a single environment, each on Gm07 (7.5 Mb) and Gm15 (~11.3 Mb).

Altogether eight QTLs were identified for RP (Table 5; Supplementary Figure S2F) on seven different chromosomes (Gm01, Gm03, Gm06, Gm07, Gm17, Gm18, and Gm19), and the LOD score range was relatively narrow (2.5–12.3), compared to other growth-period-related traits (R1 and R8). However, qRP.3 and qRP.4 had much higher LOD scores, 12.3 and 12.1, respectively, and PVE values, 20.85% and 19.85%, respectively, than the other RP QTLs (maximum LOD 2.9 and PVE 5.52%, Table 5). Of these eight, four QTLs, on Gm01 (qRP.1), Gm06 (qRP.3), Gm07 (qRP.4), and Gm18 (qRP.7), were detected consistently, had above-threshold LOD scores, and were considered stable (Supplementary Table S2). The other four QTLs (qRP.2, qRP.5, qRP.6, and qRP.8) were considered minor.

For maturity (R8), six significant QTLs were identified (Table 5; Supplementary Figure S2G), which were distributed on six different chromosomes (Gm03, Gm05, Gm06, Gm07, Gm13, and Gm19) and exhibited a wide range of LOD scores (2.6–52.0). Two major QTLs were identified (Gm06: qR8.3 and Gm07: qR8.4) based on the high LOD scores (52.0 and 21.7, respectively) and PVE (64.48% and 16.58% respectively). Of these

TABLE 5 QTL detected for beginning bloom (R1), reproductive period (RP), and maturity (R8), from the means of each year (2017–2019) and the BLUP dataset following composite interval mapping using the RIL population derived from DS25-1 and DT97-4290.

Trait ^a	QTL name	Chr/LG	QPP (Mb)	QPP (bp)	QGP (cM)	LOD	PVE (%)	AD. effects	Env.	Parent allele
R1	qR1.1	Gm01/LG_D1a	1.3	1,335,864	6.7	2.5	1.01	0.94	IV	P1
	qR1.2	Gm03/LG_N	38.9	38,913,468	106.1	3.6	1.93	−1.22	II	P2
		Gm03/LG_N	39.3	39,272,951	107.1	2.7	0.57	0.91	IV	P2
	qR1.3	Gm06/LG_C2	19.0	18,959,143	117.8	48.7	66.83	7.17	II	P1
		Gm06/LG_C2	19.9	19,864,606	121.1	62.2	75.53	9.63	III	P1
		Gm06/LG_C2	20.3	20,291,912	121.5	88.7	82.58	10.57	I	P1
		Gm06/LG_C2	20.3	20,291,912	121.5	78.2	82.34	8.95	IV	P1
	qR1.4	Gm07/LG_M	4.0	3,978,484	18.8	7.5	3.53	−1.59	II	P2
		Gm07/LG_M	4.0	3,978,484	18.8	4.9	3.41	−2.02	III	P2
		Gm07/LG_M	4.0	3,978,484	18.8	4.9	2.24	−1.42	IV	P2
	qR1.5	Gm07/LG_M	7.5	7,540,731	40.5	5.2	0.83	−1.21	I	P2
RP	qRP.1	Gm01/LG_D1a	56.3	56,264,066	126.3	2.6	3.03	−1.96	I	P2
		Gm01/LG_D1a	56.3	56,264,066	126.3	2.6	2.74	−1.28	IV	P2
	qRP.2	Gm03/LG_N	37.8	37,825,537	99.6	2.9	3.96	−2.20	I	P2
	qRP.3	Gm06/LG_C2	18.0	18,029,734	116.7	10.5	20.10	5.08	I	P1
		Gm06/LG_C2	19.6	19,595,695	120.5	12.3	20.85	4.48	II	P1
		Gm06/LG_C2	18.0	18,029,734	116.7	9.8	17.70	3.25	IV	P1
	qRP.4	Gm07/LG_M	4.0	3,978,484	18.8	8.1	12.24	−3.95	I	P2
		Gm07/LG_M	4.0	3,978,484	18.8	12.1	19.85	−4.31	II	P2
		Gm07/LG_M	4.0	3,978,484	18.8	7.4	12.26	−2.64	III	P2
		Gm07/LG_M	4.0	3,978,484	18.8	12.2	18.83	−3.35	IV	P2
	qRP.5	Gm17/LG_D2	0.8	0,773,001	1.1	2.5	2.66	1.24	IV	P2
	qRP.6	Gm18/LG_G	8.0	7,955,333	54.4	2.5	4.55	−1.61	IV	P2
	qRP.7	Gm18/LG_G	18.1	18,093,932	67.5	2.5	5.52	−1.80	III	P2
	qRP.8	Gm19/LG_L	43.1	43,084,686	67.7	2.5	4.69	−2.42	I	P2
R8	qR8.1	Gm03/LG_N	37.8	37,825,537	99.6	4.0	2.08	−2.88	I	P2
	qR8.2	Gm05/LG_A1	38.1	38,062,600	106.3	2.9	1.46	2.33	I	P1
		Gm05/LG_A1	38.1	38,062,600	106.3	2.7	2.32	2.44	II	P1
		Gm05/LG_A1	38.1	38,062,600	106.3	3.1	2.40	2.20	III	P1
		Gm05/LG_A1	38.2	38,273,815	107.1	2.8	3.19	2.18	IV	P1
	qR8.3	Gm06/LG_C2	19.6	19,616,246	120.5	45.2	56.12	10.22	III	P1
		Gm06/LG_C2	19.7	19,740,115	121.1	52.0	64.48	15.43	I	P1
		Gm06/LG_C2	19.7	19,740,115	121.1	51.1	61.38	12.16	II	P1
		Gm06/LG_C2	19.7	19,740,115	121.1	46.1	49.94	11.97	IV	P1

(Continued on following page)

TABLE 5 (Continued) QTL detected for beginning bloom (R1), reproductive period (RP), and maturity (R8), from the means of each year (2017–2019) and the BLUP dataset following composite interval mapping using the RIL population derived from DS25-1 and DT97-4290.

Trait ^a	QTL name	Chr/LG	QPP (Mb)	QPP (bp)	QGP (cM)	LOD	PVE (%)	AD. effects	Env.	Parent allele
	qR8.4	Gm07/LG_M	4.0	3,978,484	18.8	9.3	7.51	−5.04	I	P2
		Gm07/LG_M	4.0	3,978,484	18.8	21.7	16.78	−6.36	II	P2
		Gm07/LG_M	4.0	3,978,484	18.8	15.2	12.35	−4.73	III	P2
		Gm07/LG_M	4.0	3,978,484	18.8	13.4	13.53	−4.91	IV	P2
	qR8.5	Gm13/LG_F	34.7	34,676,079	94.2	2.6	1.57	−1.98	II	P2
	qR8.6	Gm19/LG_L	45.8	45,757,122	90.9	3.0	1.46	−2.47	I	P2

^aChr/LG, chromosome/linkage group; QPP, QTL, physical position (megabase); QGP, QTL, genetic position (centiMorgan); LOD, logarithm of the odds; PVE, phenotypic variation explained; AD effects, additive effects; Env., environment; I, Stoneville 2017; II, Stoneville 2018; III, Stoneville 2019; IV, BLUP, dataset; P1, DS25-1, P2, DT97-4290.

six loci, two (Gm06: qR8.3 and Gm07: qR8.4) were detected in multiple environments and thus considered as stable and major QTLs. The remaining four QTLs (qRP.1, qRP.2, qRP.5, and qRP.6) were not detected in multiple environments and were thus considered minor (Table 5).

Co-localization of QTLs among the traits and candidate gene identification

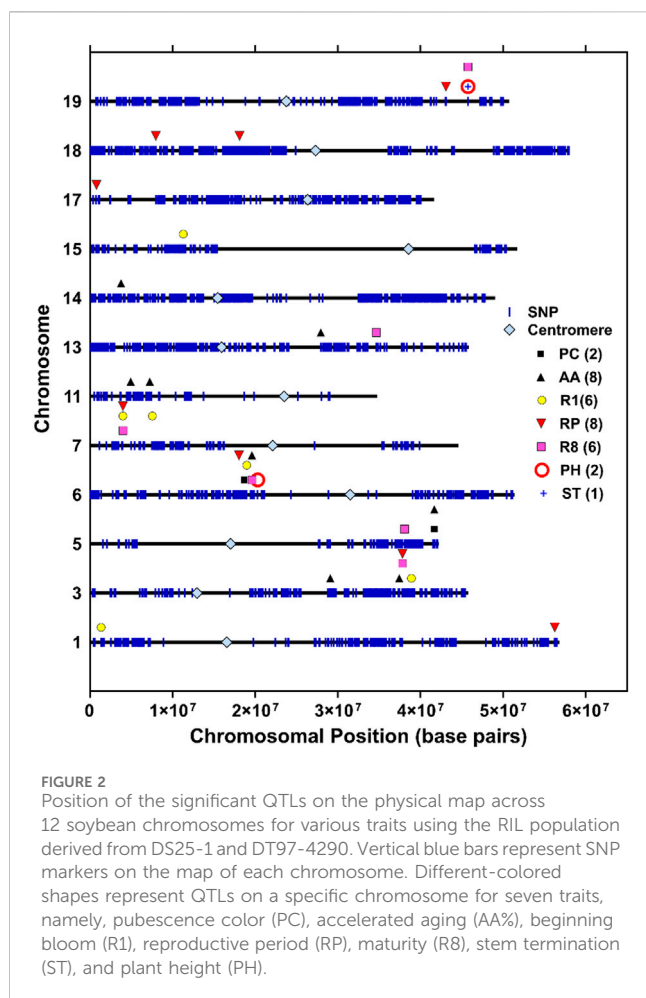
We compared QTL results based on their positions and intervals to identify overlapping genomic regions among the seven phenotypic traits (Tables 4, 5; Figure 2). At least four major genomic regions located on chromosomes Gm03, Gm06, Gm07, and Gm19 were identified, exhibiting overlapping loci among multiple phenotypic traits. A genomic region on Gm03 (37.4–38.9 Mb) exhibited co-localized QTLs across the four traits (AA, R1, RP, and R8) between multiple environments. Another genomic region on Gm06 (18.0–20.3 Mb) exhibited co-localization across the six traits (AA, PC, R1, RP, R8, and PH). All these putative QTL loci fell within a region of 2.3 Mb, and most were within a region of 0.9 Mb. Similarly, a genomic region on Gm07 (3.9 Mb) exhibited co-localized QTLs across the growth-period-related traits (R1, RP, and R8). An additional major genomic region on Gm19 (43.1–45.8 Mb) showed co-localized QTLs for PH, ST, and R8 (Tables 4, 5; Figure 2).

The above four genomic regions were considered important, and the regions were analyzed for candidate gene predictions using the Genome Browser at SoyBase (<https://www.soybase.org/>). We searched for candidate genes within the genomic region approximately 300 kb to the left and right sides of the closest SNP identified for each major QTL (Tables 4, 5) to the corresponding trait and used them for Gene Ontology (GO) analyses. In this analysis, we identified over 362 genes within the selected regions on chromosomes Gm03, Gm06, Gm07, and Gm19 (Supplementary Table S3). Of these, we found key candidate genes near major QTLs identified for different traits. We identified a candidate gene (*Glyma.06G202600*) nearby to our major QTL (qPC.2) located on Gm06 (~18.6 Mb) for pubescence color that encodes glucan endo-1,3-beta-glucosidase-like protein 2-like. In soybeans, glucan endo-1,3-beta-glucosidases are involved in various biological processes, including plant defense, cell remodeling, cell division, and development (Perrot et al., 2022).

Glyma.06G202600 was 68 kb downstream of a previously detected gene (*Glyma.06G202300*), which encodes cytochrome P450 family proteins (Toda and Yang, 2002; Zhou et al., 2024). Another gene was predicted on chromosome Gm06 (*Glyma.06G207800*), which was previously recognized and cloned as the E1 locus (Bernard, 1972; Xia et al., 2012; Zimmer et al., 2021). Another major genomic region on Gm07 (3.9 Mb) exhibited multiple stable QTLs for R1 (qR1.4), RP (qRP.4), and R8 (qR8.4). Our candidate gene analysis results confirmed the location of the E11 locus (Wang et al., 2019; Zimmer et al., 2021), which was ~97 kb downstream of our QTL region on Gm07. Additionally, we detected a major genomic region on Gm19 (~45.8 Mb), which showed multiple co-localized QTLs for PH, ST, and R8. Because we observed a major peak in this region spanning ~5.0 Mb (43.5–47.2 Mb), we identified several candidate genes within the QTL peak. One of them has been reported as the stem termination *Dt1* gene (*Glyma.19G194300*), which encodes a protein (Terminal Flower 1) in soybean (Liu et al., 2010; Tian et al., 2010).

Discussion

High seed vigor is a fundamental determinant of seed yield potential, facilitating uniform germination and rapid plant stand establishment. These factors contribute directly to optimal canopy development and resource utilization, thereby maximizing grain yield. The significance of seed vigor is further amplified under abiotic and biotic stress conditions, including sub-optimal temperatures, water-deficit, pathogen pressure, and extreme humidity. Overall, the genetic control of seed vigor in soybeans has not been well studied. The parental lines of this population (DS25-1 and DT97-4290) vary in terms of the characteristics of seed germination, accelerated aging, flowering time, maturity, plant height, stem termination, heat tolerance, mature seed damage tolerance, and seed composition. DS25-1 was derived from a cross between PI 587982A (an unadapted exotic source of abiotic stress tolerance) and DT98-9102, a high-yielding improved germplasm that was publicly released by USDA-ARS but not entered into the USDA germplasm collection (USDA-ARS, 2020). DS25-1 was released by ARS in 2017 (Smith et al., 2017) and has repeatedly demonstrated very high seed germination, even under elevated temperatures (Chebroly et al., 2016; Gillman et al., 2020;



Smith et al., 2025). Because both parental lines of this RIL population showed numerous contrasting breeding-relevant characteristics, our group utilized the population for QTL mapping to understand the genetic control of seed vigor and other important growth-related traits (Tables 1, 4, and 5).

Trait variations, distributions, heritabilities, and correlations

In this study, a broad range of phenotypic variation was observed between the parental lines and their derived progenies (RILs) for various qualitative and quantitative traits, including AA, PC, R1, RP, R8, PH, and ST (Table 1). This contrasting phenotypic variation across multiple traits allowed us to utilize this population for QTL mapping to identify genomic regions associated with the above traits. The distributions of all the phenotypic traits exhibited continuous variation in a near-normal distribution except for PC and ST, which is expected, given the polygenic nature of AA, R1, RP, R8, and PH (Figure 1). PC and ST are known to be simply inherited. In addition, the RIL population exhibited a broader range of means than their parents, manifesting transgressive segregations for seed vigor and other growth-related traits (AA, R1, RP, R8, and PH), and indicating that comprehensive genetic recombination of alleles had occurred

among the progeny of these parental lines (Mackay et al., 2021). The transgressive segregation for corresponding traits can be utilized for trait improvement in soybean breeding programs through pyramiding the desirable allelic combinations contributed from the parents (positive and negative) in the selected progenies.

In this study, most quantitative traits (AA, R1, RP, and R8) exhibited a high broad-sense heritability (0.81–0.96, Table 2), indicating that these traits are heritable and will respond to phenotypic selection. Interestingly, days to maturity (R8, 0.96) exhibited the highest heritability among the traits (AA, R1, RP, and PH) over 3 years of testing, indicating that maturity is a relatively stable trait in this population, wherein perhaps few maturity loci are segregating. Previous studies have also reported high heritabilities for flowering and maturity (Zhang et al., 2015; Liu et al., 2016). In contrast, the moderate heritability for PH (0.52) suggests that this trait is likely influenced by a combination of genetic and environmental factors. One such environmental factor in the current study was planting date, which varied from 10 April 2017 to 22 April 2019. Planting date affects the photoperiod experienced, which affects plant height. In addition, the 2017 experiment was furrow-irrigated, whereas the 2018 and 2019 experiments were rainfed, which may also have affected plant height, although PH was only recorded in 2019.

Relationships between phenotypic traits can play a critical role in breeding programs by suggesting which selection strategy may be most effective for specific trait improvement. Our results demonstrated a significant positive correlation between seed vigor, as measured by AA and major growth-related traits (R1, R8, and PH) across the years (Table 2). Previous studies have reported positive relationships among growth-period traits (Kantolic and Slafer, 2005; Liu et al., 2016) as well as positive correlations with seed yield (Kantolic and Slafer, 2005), which can facilitate simultaneous phenotypic selection for desired trait combinations.

Candidate genes and consistent QTLs for various traits

Pubescence color (PC) is a relatively well-characterized morphological trait that is used to distinguish soybean genotypes. We identified a major QTL (qPC.2) on Gm06 (~18.6 Mb) that lies 68 kb downstream of a previously characterized gene (*Glyma.06G202300*) for PC that encodes cytochrome P450 family proteins (Zhou et al., 2024), which are highly expressed in pods and seeds. Previous studies reported that soybean hilum color and pubescence color are pleiotropic and controlled by multiple loci (Bernard and Weiss, 2004), which include a major locus (T/t) previously mapped on Gm06 (Cober and Ablett, 1998) and some minor loci reported on Gm03, Gm04, and Gm05 (Zhou et al., 2024). However, the molecular mechanisms of hilum and pubescence color formation remain unclear. A major QTL (qPC.2) identified for PC overlapped with AA, R1, RP, and R8 (Tables 4, 5). Previous QTL mapping and GWAS studies reported common loci between PC and other traits, such as maturity, growth habit, leaf shape, and cold tolerance (Toda and Yang, 2002; Tsegaw and Zegeye, 2023). Additionally, we detected a minor new locus on Gm05 (~41.7 Mb, Table 4) for PC, which overlapped with a locus for AA in this study.

QTL analysis identified eight QTLs for AA (Table 4) distributed across six chromosomes (Gm03, Gm05, Gm06, Gm11, Gm13, and Gm14), indicating that AA is polygenic in nature and regulated by multiple genes. Among these eight, two loci (qAA.3 and qAA.4) were detected on Gm05 and Gm06, respectively, in two environments, and both showed positive additive effects. Interestingly, the positive alleles for both these QTLs were contributed from DS25-1, which showed a high AA score (85%) compared to DT97-4290 (45%) and high additive effects (up to 15%) that can be used for stacking these two alleles for improving seed vigor. The location of qAA.3 overlapped with a previously reported QTL on Gm05 (Gillman et al., 2020). It should be noted that DS25-1 likely contains desirable alleles for seed vigor from PI 587982A, which also likely contributed to the improved germination for the parental line (04030-1-4-1-1) used for developing the RIL population tested in the study by Gillman et al. (2020). We identified a candidate gene (*Glyma.06G206400*) close to our major locus (Gm06, qAA.4) that is a DNA-binding transcription factor, which generally plays a critical role in seed development.

Genomic regions linked to various traits (AA, R1, RP, R8, ST, and PH)

Based on QTL mapping, we identified four genomic regions on Gm03, Gm06, Gm07, and Gm19 that showed overlapping QTLs for multiple traits (AA, PC, R1, RP, R8, and PH; Tables 4, 5). Notably, most of the stable QTLs were detected on Gm06, Gm07, and Gm19 in multiple environments with high LOD scores and were thus considered important for further candidate gene analysis. We found a gene (*Glyma.06G207800*) on chromosome Gm06 that encodes the E1 gene (Bernard, 1972; Xia et al., 2012; Zimmer et al., 2021) in the same interval where we detected QTLs for multiple traits (AA, PC, R1, R8, and PH). It might be due to the confounding effect of the E1 gene. To rule out this possibility, we also performed QTL analysis using maturity data as a cofactor, and we found the same results. Another genomic region on Gm07 (3.9 Mb) showed stable QTLs for three traits (R1, RP, and R8). The closest SNP (Gm07: 3,978,484 bp) linked to the above three QTLs was mapped to ~97 kb downstream of the reported location of the E11 locus (Wang et al., 2019; Zimmer et al., 2021).

For growth habits, soybeans are generally classified as determinate, semi-determinate, or indeterminate. Determinacy and maturity determine plant architecture, which can play a major role in seed yield potential (Liu et al., 2010; Tian et al., 2010; Ping et al., 2014). Growth habits are genetically governed by two major loci (*Dt1*: Gm19, and *Dt2*: Gm18) mapped on two separate chromosomes (Bernard, 1972; Liu et al., 2010; Ping et al., 2014). The determinate type is regulated by a recessive allele at the *Dt1* locus (Liu et al., 2010), which can also affect flowering and maturity time in soybeans (Bernard, 1972; Zhang et al., 2015; Kong et al., 2018; Ogiso-Tanaka et al., 2019). The semi-indeterminate type is controlled by a dominant epistatic effect of *Dt2* and *Dt1* (Bernard, 1972; Ping et al., 2014). The RIL population used in the current study showed segregation for determinacy. We mapped the *Dt1* (Chr19) locus and confirmed the gene (*Glyma.19G194300*) that encodes a protein (*Terminal Flower 1*) in soybean (Liu et al., 2010; Tian et al., 2010). In this study, the same locus (at ~45.8 Mb, Gm19) also overlapped with

QTLs for PH and ST. This locus accounted for a high amount of the phenotypic variation for PH (28.4%) and ST (65.6%) with negative additive effects for these traits. The desirable alleles for the corresponding traits can be used to select a superior progeny from this population after yield validation. Our findings were corroborated by previous studies (Mansur et al., 1996; Lee et al., 1996; Tian et al., 2022). Other reports have also shown genomic regions that showed common QTLs on the three chromosomes (Gm06, Gm07, and Gm19) for flowering, maturity, and plant height (Kantolic and Slafer, 2005; Fang et al., 2017). Zhang et al. (2015) reported that a major locus on Gm19 (~45 Mb) exhibited shared associations between multiple traits, including days to flower, days to maturity, and duration of flowering to maturity. Previously, QTL clusters were reported on the overlapping genomic regions associated with the growth-period traits on multiple chromosomes, including Gm06, Gm07, and Gm19 (Liu et al., 2016).

Other QTLs detected in this study

Several significant QTLs were detected for flowering and maturity on genomic regions different from those of the three overlapping regions described above (Tables 4, 5; Supplementary Table S2). We compared our results with previous GWAS and QTLs studies using SoyBase (<https://www.soybase.org/>). Two significant QTLs for R1 were identified: one on Gm01 (qR1.1) and one on Gm07 (qR1.4). These coincided with regions found in a previous GWAS study conducted on growth-period traits (Zimmer et al., 2021). However, two additional QTLs were identified for the first time in this study, one on Gm07 (qR1.5) and one on Gm15 (qR1.6). Notably, R1, RP, and R8 are reproductive traits and are correlated in soybean. In this study, QTL qRP.1 and qRP.4 overlapped with full maturity (R8) loci reported previously by Zimmer et al. (2021). Two additional RP QTLs (qRP.2 and qRP.7) detected in this study coincided with those reported from a previous GWAS study (Zhang et al., 2015) for different traits (RP and PH, respectively). In our study, two QTLs, one on Gm17 (qRP.5) and one on Gm18 (qRP.6), were identified for the first time. Maturity (R8) in soybeans is also well-characterized. Although we identified six QTLs linked to R8, only one (qR8.5, Gm13) has not been previously reported. The other five QTLs identified on Gm03, Gm05, Gm06, Gm07, and Gm19 were previously reported (Panthee et al., 2007; Zhang et al., 2015; Zimmer et al., 2021).

Conclusion

In the present study, an RIL population derived from a cross between DS25-1 and DT97-4290 was phenotypically characterized for various traits, including seed vigor and growth-related characteristics, over three consecutive years. A high-density genetic linkage map comprising 8,445 SNPs was constructed and utilized for composite interval mapping to identify QTLs associated with seven traits: AA, PC, R1, RP, R8, ST, and PH. In total, 33 QTLs were detected across 12 chromosomes (Gm01, Gm03, Gm05, Gm06, Gm07, Gm11, Gm13, Gm14, Gm15, Gm17, Gm18, and Gm19) based on phenotypic datasets from both individual environments and estimated BLUP values. Four genomic regions were found to

contain stable QTL in multiple environments, including three chromosomes (Gm06, Gm07, and Gm19) with major QTLs, highlighting the significance of this population for further studies. Within this population, several genetic loci and candidate genes were validated for important traits, including seed vigor, pubescence color, plant height, maturity, and stem termination. Additionally, novel loci associated with seed vigor and growth-related traits were also identified. The discovery of major new QTLs related to seed vigor and growth-related traits provides the basis for developing specific KASP markers for these traits, which can then facilitate marker-assisted selection for the traits in multiple soybean breeding programs.

Data availability statement

The datasets presented in this study can be found in the article/[Supplementary Material](#), further inquiries can be directed to the corresponding author. We request appropriate recognition be made when any data sets are used for producing new publications.

Author contributions

NK: Writing – original draft, Conceptualization, Writing – review and editing, Formal analysis, Data curation, Software, Visualization. JS: Conceptualization, Data curation, Writing – review and editing, Investigation, Methodology, Resources, Validation, Supervision. JR: Conceptualization, Data curation, Writing – review and editing, Investigation, Resources, Validation, Supervision. JG: Writing – review and editing, Validation, Supervision. NB: Writing – review and editing, Validation, Supervision.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This research was supported in part by the U.S. Department of Agriculture, Agricultural Research Service and the Mississippi Soybean Promotion Board. The USDA is an equal opportunity provider and employer. This work was supported by ARS project numbers 6066-21220-012-00D and 6066-21220-016-00D.

References

- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., et al. (2015). Package lme4. *Convergence* 12 (1), 2.
- Bazzer, S. K., Kaler, A. S., Ray, J. D., Smith, J. R., Fritsch, F. B., and Purcell, L. C. (2020). Identification of quantitative trait loci for carbon isotope ratio ($\delta^{13}\text{C}$) in a recombinant inbred population of soybean. *Theor. Appl. Genet.* 133, 2141–2155. doi:10.1007/s00122-020-03586-0
- Bentsink, L., Alonso-Blanco, C., Vreugdenhil, D., Tesnier, K., Groot, S. P., and Koornneef, M. (2000). Genetic analysis of seed-soluble oligosaccharides in relation to seed storability of arabidopsis. *Plant Physiol.* 124 (4), 1595–1604. doi:10.1104/pp.124.4.1595
- Bernard, R. L. (1972). Two genes affecting stem termination in soybeans. *Crop Sci.* 12, 235–239. doi:10.2135/cropsci1972.0011183X001200020028x
- Bernard, R. L., and Weiss, M. G. (2004). *Qualitative genetics, 117-149 in soybean: improvement, production and uses*. 3rd Edition (Madison: American Society of Agronomy).
- Börner, A., Nagel, M., Agacka-Moldoch, M., Gierke, P. U., Oberforster, M., Albrecht, T., et al. (2018). QTL analysis of falling number and seed longevity in wheat (*Triticum aestivum* L.). *J. Appl. Genet.* 59 (1), 35–42. doi:10.1007/s13353-017-0422-5
- Broman, K. W., and Sen, S. (2009). *A guide to QTL mapping with R/qtl* (Vol. 46). New York, NY: Springer.
- Broman, K. W., Wu, H., Sen, S., and Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19 (7), 889–890. doi:10.1093/bioinformatics/btg112
- Cao, Y., Li, S., Chen, G., Wang, Y., Bhat, J. A., Karikari, B., et al. (2019). Deciphering the genetic architecture of plant height in soybean using two RIL populations sharing a common M8206 parent. *Plants* 8 (10), 373. doi:10.3390/plants8100373
- Chebrolu, K. K., Fritsch, F. B., Ye, S., Krishnan, H. B., Smith, J. R., and Gillman, J. D. (2016). Impact of heat stress during seed development on soybean seed metabolome. *Metabolomics* 12, 1–14. doi:10.1007/s11306-015-0941-1

Acknowledgements

The authors thank Philip Handly, Hans Hinrichsen, Matt Kersh, Corey Wilkins, and Angelie Davis, who were instrumental in the success of the Stoneville trials (2017, 2018, and 2019). Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2025.1695593/full#supplementary-material>

- Chen, Q. S., Zhang, Z. C., Liu, C. Y., Xin, D. W., Qiu, H. M., Shan, D. P., et al. (2007). QTL analysis of major agronomic traits in soybean. *Agric. Sci. China* 6 (4), 399–405. doi:10.1016/s1671-2927(07)60062-5
- Clerkx, E. J. M., El-Lithy, M. E., Vierling, E., Ruys, G. J., Vries, B.-De. H., Groot, S. P. C., et al. (2004). Analysis of natural allelic variation of arabidopsis seed germination and seed longevity traits between the accessions Landsberg erecta and Shadkara, using a new recombinant inbred line population. *Plant Physiol.* 135, 432–444. doi:10.1104/pp.103.036814
- Cober, E. R., Ablett, G. R., Buzzell, R. I., Luzzi, B. M., Poysa, V., Sahota, A. S., et al. (1998). Imperfect yellow hilum color in soybean is conditioned by II rr TT. *Crop Sci.* 38, 940–941. doi:10.2135/cropsci1998.0011183X003800040010x
- Copeland, L. O., and McDonald, M. B. (2001). Seed germination. *Princ. Seed Sci. Tech.*, 72–123.
- Dargahi, H., Tanya, P., and Srinives, P. (2014). Mapping of the genomic regions controlling seed storability in soybean (*Glycine max* L.). *Jour. Genet.* 93, 365–370. doi:10.1007/s12041-014-0381-0
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6 (5), e19379. doi:10.1371/journal.pone.0019379
- Fang, C., Ma, Y., Wu, S., Liu, Z., Wang, Z., Yang, R., et al. (2017). Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. *Genome Biol.* 18, 1–14. doi:10.1186/s13059-017-1289-9
- Fehr, W. R., Caviness, C. E., and Vorst, J. J. (1977). Response of indeterminate and determinate soybean cultivars to defoliation and half-plant cut-off 1. *Crop Sci.* 17 (6), 913–917. doi:10.2135/cropsci1977.0011183X001700060024x
- Fehr, W. R., Walker, A. K., Schmitthenner, A. F., and Voss, B. K. (1987). Registration of 'Pella 86' soybean. *Crop Sci.* 27, 1313. doi:10.2135/cropsci1987.0011183X002700060057x
- Gao, W., Ma, R., Li, X., Liu, J., Jiang, A., Tan, P., et al. (2024). Construction of genetic map and QTL mapping for seed size and quality traits in soybean (*Glycine max* L.). *Intl. Jour. Mol. Sci.* 25 (5), 2857. doi:10.3390/ijms25052857
- Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv Preprint arXiv:1207.3907*. doi:10.48550/arXiv.1207.3907
- Gillman, J. D., Chebrolov, K., and Smith, J. R. (2020). Quantitative trait locus mapping for resistance to heat-induced seed degradation and low seed phytic acid in soybean. *Crop Sci.* 61, 2023–2035. doi:10.1002/csc2.20419
- Han, Z., Ku, L., Zhang, Z., Zhang, J., Guo, S., Liu, H., et al. (2014). QTLs for seed vigor-related traits identified in maize seeds germinated under artificial aging conditions. *PLoS ONE* 9 (3), e92535. doi:10.1371/journal.pone.0092535
- Han, Z., Bin, W., Zhang, J., Guo, S., Zhang, H., Xu, L., et al. (2018). Mapping of QTLs associated with seed vigor to artificial aging using two RIL populations in maize (*Zea mays* L.). *Agric. Sci.* 9 (04), 397.
- Heatherly, L. G. (1999). "Early soybean production system (ESPS)," in *Soybean production system in the midsouth*. Editors L. G. Heatherly and H. F. Hodges (Boca Raton, FL: CRC Press), 103–115.
- Heatherly, L. G., and Spurlock, S. R. (1999). Yield and economics of traditional and early soybean production system (ESPS) seedlings in the mid-southern United States. *Field Crops Res.* 63 (1), 35–45. doi:10.1016/S0378-4290(99)00025-8
- Holland, J. B., Nyquist, W. E., and Cervantes, C. T. (2003). Estimating and interpreting heritability for plant breeding. *Plant Breed. Rev.* 22, 11–112.
- Huang, C. T., Klos, K. E., and Huang, Y. F. (2020). Genome-wide association study reveals the genetic architecture of seed vigor in oats. *G3 Genes Genomes Genet.* 10 (12), 4489–4503. doi:10.1534/g3.120.401602
- Jiamtae, P., Hashiguchi, M., Lelapiyarnit, I., Harada, K., Gondo, T., Tanaka, H., et al. (2022). QTL analysis for soybean (*Glycine max* L. Merr.) seed storability in high-temperature storage conditions. *Euphytica* 218 (12), 169. doi:10.21203/rs.3.rs-4754944/v1
- Jing-bao, L., Zhi-yuan, F., Hui-ling, X., Yan-min, H., Zong-hua, L., Liu-jing, D., et al. (2011). Identification of QTLs for maize seed vigor at three stages of seed maturity using a RIL population. *Euphytica* 178, 127–135. doi:10.1007/s10681-010-0282-0
- Kantolic, A. G., and Slafer, G. A. (2005). Reproductive development and yield components in indeterminate soybean as affected by post-flowering photoperiod. *Field Crop Res.* 93, 212–222. doi:10.1016/j.fcr.2004.10.001
- Kim, J. M., Seo, J. S., Lee, J. W., Lyu, J. I., Ryu, J., Eom, S. H., et al. (2023). QTL mapping reveals key factors related to the isoflavone contents and agronomic traits of soybean (*Glycine max*). *BMC Plant Biol.* 23 (1), 517. doi:10.1186/s12870-023-04519-x
- Kong, L., Lu, S., Wang, Y., Fang, C., Wang, F., Nan, H., et al. (2018). Quantitative trait locus mapping of flowering time and maturity in soybean using next-generation sequencing-based analysis. *Front. Plant Sci.* 9, 995. doi:10.3389/fpls.2018.00995
- Kosambi, D. D. (1944). The estimation of map distances from recombination values. *Ann. Eugen.* 12, 172–175. doi:10.1111/j.1469-1809.1943.tb02321.x
- Lai, Y., Cheng, J., He, Y., Yang, B., Wang, Z., and Zhang, H. (2016). Identification of QTLs with additive, epistatic, and QTL× seed maturity interaction effects for seed vigor in rice. *Plant Mol. Biol. Report.* 34, 160–171. doi:10.1007/s11105-015-0913-7
- Lee, S. H., Bailey, M. A., Mian, M. A. R., Carter Jr, T. E., Ashley, D. A., Hussey, R. S., et al. (1996). Molecular markers associated with soybean plant height, lodging, and maturity across locations. *Crop Sci.* 36 (3), 728–735. doi:10.2135/cropsci1996.0011183X003600030035x
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25 (14), 1754–1760. doi:10.1093/bioinformatics/btp324
- Liu, B., Liu, X. B., Wang, C., Jin, J., Herbert, S. J., and Hashemi, M. (2010). Responses of soybean yield and yield components to light enrichment and planting density. *Intl. J. Plant Prod.* 4 (1), 1–10. doi:10.22069/ijpp.2012.676
- Liu, B., Watanabe, S., Uchiyama, T., Kong, F., Kanazawa, A., Xia, Z., et al. (2010). The soybean stem growth habit gene *Dt1* is an ortholog of arabidopsis *TERMINAL FLOWER1*. *Plant Physiol.* 153 (1), 198–210. doi:10.1104/pp.109.150607
- Liu, L., Lai, Y., Cheng, J., Wang, L., Du, W., Wang, Z., et al. (2014). Dynamic quantitative trait locus analysis of seed vigor at three maturity stages in rice. *PLoS ONE* 9 (12), e115732. doi:10.1371/journal.pone.0115732
- Liu, Z., Li, H., Fan, X., Huang, W., Yang, J., Li, C., et al. (2016). Phenotypic characterization and genetic dissection of growth period traits in soybean (*Glycine max*) using association mapping. *PLoS ONE* 11 (7), e0158602. doi:10.1371/journal.pone.0158602
- Liu, J., Jiang, A., Ma, R., Gao, W., Tan, P., Li, X., et al. (2023). QTL mapping for seed quality traits under multiple environments in soybean (*Glycine max* L.). *Agronomy* 13 (9), 2382. doi:10.3390/agronomy13092382
- Mackay, I. J., Cockram, J., Howell, P., and Powell, W. (2021). Understanding the classics: the unifying concepts of transgressive segregation, inbreeding depression and heterosis and their central relevance for crop breeding. *Plant Biotech. J.* 19 (1), 26–34. doi:10.1111/pbi.13481
- Mansur, L. M., Orf, J. H., Chase, K., Jarvik, T., Cregan, P. B., and Lark, K. G. (1996). Genetic mapping of agronomic traits using recombinant inbred lines of soybean. *Crop Sci.* 36 (5), 1327–1336. doi:10.2135/cropsci1996.0011183X003600050042x
- Money, D., Gardner, K., Migicovsky, Z., Schwaninger, H., Zhong, G. Y., and Myles, S. (2015). LinkImpute: fast and accurate genotype imputation for nonmodel organisms. *G3 Genes Genomes Genet.* 5 (11), 2383–2390. doi:10.1534/g3.115.021667
- Nadarajan, J., Walters, C., Pritchard, H. W., Ballesteros, D., and Colville, L. (2023). Seed longevity—the evolution of knowledge and a conceptual framework. *Plants* 12 (3), 471. doi:10.3390/plants12030471
- Nagel, M., Kodde, J., Pistrick, S., Mascher, M., Börner, A., and Groot, S. P. (2016). Barley seed aging: genetics behind the dry elevated pressure of oxygen aging and moist controlled deterioration. *Front. Plant Sci.* 7, 388. doi:10.3389/fpls.2016.00388
- Narayanan, S., Zoong-Lwe, Z. S., Gandhi, N., Welti, R., Fallen, B., Smith, J. R., et al. (2020). Comparative lipidomic analysis reveals heat stress responses of two soybean genotypes differing in temperature sensitivity. *Plants* 9 (4), 453. doi:10.3390/plants9040457
- Ogiso-Tanaka, E., Shimizu, T., Hajika, M., Kaga, A., and Ishimoto, M. (2019). Highly multiplexed AmpliSeq technology identifies novel variation of flowering time-related genes in soybean (*Glycine max*). *DNA Res.* 26 (3), 243–260. doi:10.1093/dnares/dsz005
- Panthee, D. R., Pantalone, V. R., Saxton, A. M., West, D. R., and Sams, C. E. (2007). Quantitative trait loci for agronomic traits in soybean. *Plant Breed.* 126 (1), 51–57. doi:10.1111/j.1439-0523.2006.01305.x
- Paris, R. L., Mengistu, A., Tyler, J. M., and Smith, J. R. (2006). Registration of soybean germplasm line DT97-4290 with moderate resistance to charcoal rot. *Crop Sci.* 46 (5), 2324–2325. doi:10.2135/cropsci2005.09.0297
- Perrot, T., Pauly, M., and Ramirez, V. (2022). Emerging roles of β-glucanases in plant development and adaptive responses. *Plants* 11, 1119. doi:10.3390/plants11091119
- Ping, J., Liu, Y., Sun, L., Zhao, M., Li, Y., She, M., et al. (2014). Dt2 is a gain-of-function MADS-domain factor gene that specifies semideterminacy in soybean. *Plant Cell* 26 (7), 2831–2842. doi:10.1105/tpc.114.126938
- Pinheiro, D. T., Dias, D. C. F. D. S., Medeiros, A. D. D., Ribeiro, J. P. O., Silva, F. L. D., and Silva, L. J. D. (2021). Weathering deterioration in pre-harvest of soybean seeds: physiological, physical, and morpho-anatomical changes. *Sci. Agric.* 78 (Suppl. 1), e20200166. doi:10.1590/1678-992x-2020-0166
- R Foundation (2024). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available online at: <https://www.R-project.org/>.
- Raquid, R., Kohli, A., Reinke, R., Dionisio-Sese, M., Kwak, J., Chebotarov, D., et al. (2021). Genetic factors enhancing seed longevity in tropical japonica rice. *Curr. Plant Biol.* 26, 100196. doi:10.1016/j.cpb.2021.100196
- Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the paleopolyploid soybean. *Nature* 463 (7278), 178–183. doi:10.1038/nature08670
- Shi, H., Guan, W., Shi, Y., Wang, S., Fan, H., Yang, J., et al. (2020). QTL mapping and candidate gene analysis of seed vigor-related traits during artificial aging in wheat (*Triticum aestivum*). *Sci. Rep.* 10 (1), 22060. doi:10.1038/s41598-020-75778-z
- Shipe, E. R., Mueller, J. D., Lewis, S. A., Williams, P. F. Jr., and Tomkins, J. P. (1997). Registration of 'Dillon' soybean. *Crop Sci.* 1983. doi:10.2135/cropsci1997.0011183X003700060062x

- Sleper, D. A., Nickell, C. D., Noel, G. R., Cary, T. R., Thomas, D. J., Clark, K. M., et al. (1998). Registration of 'Maverick' soybean. *Crop Sci.* 38, 549–550. doi:10.2135/cropsci1998.0011183X003800020072x
- Smith, J. R., Mengistu, A., Nelson, R. L., and Paris, R. L. (2008). Identification of soybean accessions with high germinability in high-temperature environments. *Crop Sci.* 48 (6), 2279–2288. doi:10.2135/cropsci2008.01.0026
- Smith, J. R., Gillman, J. D., Bellaloui, N., Gillen, A., Ray, J. D., and Carter, T. E. (2017). Soybean germplasm line DS25-1 with heat tolerance and competitive yield under heat stress. Available online at: <https://npgsweb.ars-grin.gov/gringlobal/accessiondetail.aspx?id=1954416>.
- Smith, J. R., Gillen, A. M., Li, S., Abbas, H. K., Sulyok, M., Shier, W. T., et al. (2025). Registration of soybean germplasm DS1260-2, with improved tolerance to mature seed damage and phomopsis seed decay. *J. Plant Reg.* 19 (1), e20417. doi:10.1002/plr2.20417
- Thabet, S. G., Moursi, Y. S., Karam, M. A., Graner, A., and Alqudah, A. M. (2018). Genetic basis of drought tolerance during seed germination in barley. *PLoS One* 13 (11), e0206682. doi:10.1371/journal.pone.0206682
- Tian, Z., Wang, X., Lee, R., Li, Y., Specht, J. E., Nelson, R. L., et al. (2010). Artificial selection for determinate growth habit in soybean. *Proc. Natl. Acad. Sci.* 107 (19), 8563–8568. doi:10.1073/pnas.1000088107
- Tian, Y., Yang, L., Lu, H., Zhang, B., Li, Y., Liu, C., et al. (2022). QTL analysis for plant height and fine mapping of two environmentally stable QTLs with major effects in soybean. *J. Integra. Agric.* 21 (4), 933–946. doi:10.1016/s2095-3119(21)63693-6
- Toda, K., Yang, D., Yamanaka, N., Watanabe, S., Harada, K., and Takahashi, R. (2002). A single-base deletion in soybean flavonoid 3'-hydroxylase gene is associated with gray pubescence color. *Plant Mol. Biol.* 50, 187–196. doi:10.1023/a:1016087221334
- Tsegaw, M., Zegeye, W. A., Jiang, B., Sun, S., Yuan, S., Han, T., et al. (2023). Progress and prospects of the molecular basis of soybean cold tolerance. *Plants* 12, 459. doi:10.3390/plants12030459
- USDA-ARS. (2020). *Soybase—soybean parentage information*. Ames, IA: Developed by the USDA-ARS SoyBase and Legume Clade Database group at the Iowa State University. Available online at: https://legacy.soybase.org/search/index.php?searchterm=DT98-9102&dist=uniform_trial_listview.
- Wang, W., He, A., Peng, S., Huang, J., Cui, K., and Nie, L. (2018). The effect of storage condition and duration on the deterioration of primed rice seeds. *Front. Plant Sci.* 9, 172. doi:10.3389/fpls.2018.00172
- Wang, F., Nan, H., Chen, L., Fang, C., Zhang, H., Su, T., et al. (2019). A new dominant locus, *E11*, controls early flowering time and maturity in soybean. *Mol. Breed.* 39, 70. doi:10.1007/s11032-019-0978-3
- Wang, R., Wu, F., Xie, X., and Yang, C. (2021). Quantitative trait locus mapping of seed vigor in soybean under -20 °C storage and accelerated aging conditions via RAD sequencing. *Curr. Issues Mol. Biol.* 43, 1977–1996. doi:10.3390/cimb43030136
- Wu, X., Feng, F., Zhu, Y., Xie, F., Yang, J., Gong, J., et al. (2019). Construction of high-density genetic map and identification of QTLs associated with seed vigor after exposure to artificial aging conditions in sweet corn using SLAF-seq. *Genes* 11 (1), 37. doi:10.3390/genes11010037
- Xia, Z., Watanabe, S., Yamada, T., Tsubokura, Y., Nakashima, H., Zhai, H., et al. (2012). Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. *Proc. Nat. Acad. Sci.* 109, E2155–E2164. doi:10.1073/pnas.1117982109
- Xia, Z., Zhai, H., Lü, S., Wu, H., and Zhang, Y. (2013). Recent achievement in gene cloning and functional genomics in soybean. *World Sci. J.* 2013, 1–7. doi:10.1155/2013/281367
- Yang, J., Yang, G., Yang, M., Su, L., Xia, A., Li, D., et al. (2019). Quantitative trait locus analysis of seed germination and early seedling growth in rice. *Front. Plant Sci.* 10, 1582. doi:10.3389/fpls.2019.01582
- Yang, J., Guo, Z., Luo, L., Gao, Q., Xiao, W., Wang, J., et al. (2021). Identification of QTL and candidate genes involved in early seedling growth in rice via high-density genetic mapping and RNA-seq. *Crop J.* 9 (2), 360–371. doi:10.1016/j.cj.2020.08.010
- Yao, Y., Xiong, E., Qu, X., Li, J., Liu, H., Quan, L., et al. (2023). WGCNA and transcriptome profiling reveal hub genes for key development stage seed size/oil content between wild and cultivated soybean. *BMC Genomics* 24 (1), 494. doi:10.1186/s12864-023-09617-6
- Zhang, J., Song, Q., Cregan, P. B., Nelson, R. L., Wang, X., Wu, J., et al. (2015). Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. *BMC Genomics* 16, 1–11. doi:10.1186/s12864-015-1441-4
- Zhang, A., Liu, C., Chen, G., Hong, K., Gao, Y., Tian, P., et al. (2017). Genetic analysis for rice seedling vigor and fine mapping of a major QTL qSSL1b for seedling shoot length. *Breed. Sci.* 67 (3), 307–315. doi:10.1270/jsbbs.16195
- Zhang, X., Hina, A., Song, S., Kong, J., Bhat, J. A., and Zhao, T. (2019). Whole-genome mapping identified novel "QTL hotspots regions" for seed storability in soybean (*Glycine max* L.). *BMC Genomics* 20, 1–14. doi:10.1186/s12864-019-5897-5
- Zhou, W., Chen, F., Luo, X., Dai, Y., Yang, Y., Zheng, C., et al. (2020). A matter of life and death: molecular, physiological, and environmental regulation of seed longevity. *Plant Cell Environ.* 43, 293–302. doi:10.1111/pce.13666
- Zhou, M., Wang, J., Chen, H., Jia, Q., Hu, S., Xiong, Y., et al. (2024). Genome-wide association study on candidate genes associated with soybean stem pubescence and hilum colors. *Agronomy* 14, 512. doi:10.3390/agronomy14030512
- Zimmer, G., Miller, M. J., Steketee, C. J., Jackson, S. A., de Tunes, L. V. M., and Li, Z. (2021). Genetic control and allele variation among soybean maturity groups 000 through IX. *Plant Genome* 14, e220146. doi:10.1002/tpg2.20146