

SABRAO Journal of Breeding and Genetics
 55 (5) 1451-1462, 2023
<http://doi.org/10.54910/sabrao2023.55.5.1>
<http://sabraojournal.org/>
 pISSN 1029-7073; eISSN 2224-8978



GENETIC ENHANCEMENT OF SOYBEAN (*GLYCINE MAX L.*) GERMPLASM FOR ADAPTABILITY AND PRODUCTIVITY

S. RAUF¹, M. SHEHZAD¹, S. FATIMA¹, M.L. WARBURTON^{2*}, and D.P. MALINOWSKI³

¹Department of Plant Breeding and Genetics, College of Agriculture, University of Sargodha, Pakistan

²USDA ARS Plant Germplasm Introduction and Testing Research Unit, Pullman, WA, USA

³Texas A&M AgriLife Research, Texas A&M University, Vernon, TX, USA

*Corresponding author's email: marilyn.warburton@usda.gov

Email addresses of co-authors: saeedbreeder@hotmail.com, smileace133@gmail.com, seeratsayyal@gmail.com, d-malinowski@tamu.edu

SUMMARY

Soybean is an essential protein and oilseed crop, but environmental factors, such as photoperiod and altitude, highly influence its growth. Poor adaptability of soybean cultivars may result in a significant decrease in seed yield due to disruption of flowering or maturity. It is, therefore, vital to select cultivars adapted to specific regions of the world. Global climate change, causing increasing temperature and associated water deficit stress, may further challenge sustainable soybean production worldwide. Therefore, it is necessary to understand genetic responses in soybeans induced by environmental variables. This review highlights previous research showing how several quantitative trait loci (QTLs) activated by photo-thermal responses affected soybean flowering. We discuss cytogenetic stocks, including chromosome segment substitution lines, and their role in introgressing novel genes from wild soybeans. A review of genes showing responses to various environmental variables that affect soybean adaptability is also ensued, with putative functions suggested. These include gene response regulation to temperature and heat stress, identified via genetic mapping, and emerged as tools for developing adapted soybean cultivars.

Keywords: Chromosome segment substitution lines, environmental adaptation, global climate change, soybean (*Glycine max*)

Key Findings: Soybean is highly adaptable and may be an option to meet the demands of a growing global population in changing climate. Genetic and genomic resources exist to allow quick identification and exploitation of valuable allelic diversity. This review presents these resources and suggests efficient ways of applying them to increase abiotic stress resistance in soybeans.

Communicating Editor: Dr. Sajjad Hussain Qureshi

Manuscript received: April 12, 2023; Accepted: August 6, 2023.

© Society for the Advancement of Breeding Research in Asia and Oceania (SABRAO) 2023

Citation: Rauf S, Shehzad M, Fatima S, Warburton ML, Malinowski DP (2023). Genetic enhancement of soybean (*Glycine max* L.) germplasm for adaptability and productivity. *SABRAO J. Breed. Genet.* 55(5): 1451-1462. <http://doi.org/10.54910/sabrao2023.55.5.1>.

INTRODUCTION

Soybean is the number one oilseed crop of the masses, making it mostly grown because of its high protein content, even known as the protein king. Soybeans broadly cover the human and animal diet and many food products, i.e., tofu, soy milk, nuts, and sprouts (Lopes da Silva *et al.*, 2017). Soy milk and meal are the chief contents in formula, feed sausages, and ice creams. Its seed contains about 19%–24% oil and 35%–39% proteins (FAO, 2020). The oil is a rich source of polyunsaturated fatty acid, tocopherols, and sterols that help alleviate diseases due to antioxidant properties. The seed also contains many minerals and vitamins that are good for human health.

Soybean production was 355.60 million MT in 2021 and could reach 391.17 million MT during 2022 (<http://www.worldagriculturalproduction.com/crops/soybean.aspx>). It is the most important source of edible oil, with a worldwide production of 60.02 million MT in 2020 (FAO, 2020). Soybean cultivation occurs in about 97 countries at a wide range of latitudes (FAO, 2020). Despite its origin and center of diversity in Asian countries, soybean cultivation generally concentrates in the North and South American continent, with Brazil (114.26 million tons), USA (96.79 million tons), and Argentina (55.26 million tons) being the key producers in 2019 (USDA, 2020). These countries account for about 80% of the total soybean production. Asia produced 31.72 million tons of soybean seed in 2019, only 10% of the world's soybean production. These statistics show a dangerous situation, as a decrease in production in only one of the three major soy-producing countries could negatively affect global food security. It requires concrete efforts to expand the adaptability of soybeans to other continents, including Africa and Asia, where the popularity of this crop for consumption is increasing (Bhartiya *et al.*, 2012; Alam *et al.*, 2019; Hanafiah *et al.*, 2021).

Soybean demand will continue to rise due to its general positive health benefits, high-quality proteins, and premium-value edible oil (Messina, 2022). Domestic soybeans

may also increase a country's food security. Soybean production improvement needs continuity by developing cultivars with variable breeding objectives, i.e., high yield potential, adaptability to varying climatic conditions, and resistance to diseases and environmental stresses. Meeting these breeding objectives will help develop highly productive and adaptable cultivars in developing countries, where locally produced soybean cultivars often express poor pod formation and seed set (Li *et al.*, 2019). There is also a need to improve the soybean quality, e.g., specific amino acids in proteins and modification of fatty acid and tocopherol or sterol concentration in oil for better health benefits and cooking potential.

Global climate change

Sustainable soybean production will face challenges with changing climate in the key production areas. Global climate change is often characterized as altered weather patterns resulting in excessive heat and drought periods, and diseases and insect outbreaks. Climate change may also manifest itself in an increased likelihood of floods and severe weather events (Donat *et al.*, 2016), seawater encroachment into aquifers, changes in soil composition, and disturbed atmospheric circulation patterns, including ozone levels. High temperature, water-deficiency stress, and increased pest infestation are the principal factors reducing seed yield in soybean production due to climate change. Soybean is very susceptible to soilmoisture deficit, and below a level of -0.67 MPa inhibits germination and reduces seedling growth rate and leaf area, affecting flower, pod formation, and grain yield (St-Marseille *et al.*, 2019). Although soybeans can tolerate a broad range of temperatures (11 °C to 41°C) depending upon specific cultivars, high or broadly fluctuating temperatures cause abscission of flowers, failure of pollination and grain-filling, resulting in significant yield losses. An increase in temperature of 1-2 °C may result in average yield losses of 2.5% (Havlík *et al.*, 2015).

Positive impacts of global climate change may include increased CO₂ concentration that may accelerate the

photosynthetic rate and earlier crop maturity with accelerated degree days of accumulation due to higher CO₂ fixation with rising temperatures. Soybean yield responses to elevated CO₂ levels varied between 5%–90%, depending on cultivars, as determined by the pod number and number of seeds per pod. Despite the potential benefits from increased CO₂ levels, a prediction looms that climate change will decrease overall soybean yield by 5%–6% by 2030 and by 30%–32% by 2080 (Havlík *et al.*, 2015).

New soybean cultivar development is imperative to sustain soybean production under a climate change scenario and resulting altered environmental and growth conditions. Conventional approaches to developing a new cultivar start with characterizing a wide range of soybean accessions to determine the expression of heritable traits related to stress tolerance. For adaptation to climate change, characterizing soybean accessions may be for multiple traits, such as reduced leaf area, hairiness, cuticular waxes, gametophytic fertility, pod formation under adverse conditions, or physiological qualities, such as high canopy temperature discrimination, water use efficiency, production of heat shock proteins, and other molecules related to stress tolerance (Vandana *et al.*, 2017). Genotyping core germplasm sets could use the Single Nucleotide Polymorphism (SNP) markers to link specific SNPs to helpful traits of interest, which will speed line selection with desired features.

New breeding lines that may show high levels of the desired trait(s) require evaluation in targeted cultivation zones before release as new cultivars. Crossing breeding lines with beneficial features can develop segregating populations to combine stress tolerance and high yield through conventional breeding techniques. Producing cultivars that better fit into specific environmental conditions could also help create varieties adaptable to changing climates. In the 1960s, Donald proposed the term "ideotype" to describe different types of wheat, each targeted for diverse (i.e., irrigated or rainfed) conditions (Donald, 1963). Many traits can go into a determination of an ideotype, including correlations between traits, and can include morphological, physiological, metabolic, molecular, and physiological features as correlated with production potential. Ideotypes' creation can serve as a tool to breed for climate adaptation and other broad trait classes (i.e., industrial uses, biotic, and abiotic stresses) (Table 1).

Breeding within and across environments

Genotype × environment (GE) interaction is an integral component of phenotypic variance. Different performances of specific genotypes under contrasting environmental conditions indicate a narrow adaptability of genotypes. Their cultivation may not be recommendable under broader ranges of agroecological

Table 1. List of ideotypes or preferred selection targets of soybeans specific for various environment and industrial uses.

Trait	Use	Methods	Reference
Cultivar with maturity group 7.8 and plant density of 50 m ⁻²	Climatic proof ideotype	Supplementary irrigation was beneficial to improve seed yield in various climatic scenarios	Battisti <i>et al.</i> (2018)
High root depth, root density, and narrow root angle	Drought resistance	Positive relationship between root length, shoot length, and above ground biomass	Dayoub <i>et al.</i> (2021)
High oleic acid ω-9 fatty acid or stearic acid	Improved cooking quality and deep frying, biodiesel production	144 RIL lines were selected though BLUP analysis	Woyam <i>et al.</i> (2019)

* best linear unbiased prediction (BLUP)

conditions. The adaptability of genotypes' assessment can be achieved under varied environmental conditions, helping to estimate genotype \times environment interactions under mixed model arrangements (Li *et al.*, 2017). Estimating GE interactions requires elite soybean breeding material assessment in various environmental or climatic scenarios through mixed model approaches to determine the lines characterized by a low estimate of genotypic variance and/or high GE (Table 1). Although widely adapted cultivars will be tougher to breed from lines with high GE, they may allow the selection of cultivars for specific or niche environments. The stability of the cultivar is a relative term to denote genotypes with minimum differences in the cultivar's performance across the range of environments studied. Stability analyses can identify cultivars with the lowest standard deviation estimated from environmental variances with the best-fit regression coefficient of environments. These are the beneficial lines needed to breed new varieties that can grow across diverse environments.

Conventional approaches have contributed to developing cultivars with high adaptation to targeted environments but are time-consuming. These approaches also do not help determine genes and their role in observed phenological responses to various environmental variables and stresses. Traits, such as plant height, number and placement of nodes on the main stem, number of branches and pods, and 100-grain mass, were the main contributors to high yield under stress-free environments. The number of pods or grain plant⁻¹ directly affects grain yield in soybeans (Li *et al.*, 2020). Improving the buffering capacity of soybean accessions against yield reduction in stress prone environments may help achieve stable outputs. Endorsing grain yield per plant for selecting high-yielding soybean accessions with wider adaptability resulted (Li *et al.*, 2020). Water use efficiency and root penetration traits contribute positively to high yield in water-stressed environments and sharp pan penetration. Genetic studies may help identify genes regulating the environmental responses of the soybean

cultivars, resulting in a rapid development of cultivars with improved productivity and adaptation. Once identifying genes that encode for traits of interest, new options for crop improvement are available, including marker-assisted selection, introgression, and direct-gene editing (described in detail below).

Chromosome segment substitution lines

Chromosome segment substitution lines (CSSLs) represent the movement of complete genomic segments (portions of a chromosome) of a distantly related genotype into an adaptable (usually elite domesticated) genetic background. Each CSSL contains a single chromosome segment of the distantly related genotype and the rest (up to 99.9%, but sometimes less) of the recurrent elite parent genome. They are usually products of backcrossing an initial hybrid to the recurrent parent using high-density, genome-wide marker-assisted selection to select the lines, retaining one discrete part of the donor line only. Each line contains a different portion of the donor genome, and an entire set of lines covers the whole donor genome. A final step of self-pollination fixes the genotypes in a homozygous state (Figure 1). The CSSL set, with molecular analyses, provides a valuable tool to identify and map specific chromosomal segments or QTLs, contributing agronomically crucial traits from the donor.

CSSL evaluation centers on targeted traits and lines showing significant differences from the parental types that can identify the chromosomal location of QTLs and genes of interest (Balakrishnan *et al.*, 2019). However, they may also be functional breeding lines if the effect of the QTL is immense and beneficial. CSSLs are valuable tools for basic and applied genetics and considerably serve as genetic stocks or pre-breeding materials that may help broaden the genetic base of a population. They are also an eminent tool to introgress genes from wild relatives with a minimum linkage drag (or retention of other undesirable traits from the wild parent). A full review of CSSL lines is available for various crops (Balakrishnan *et al.*, 2019).

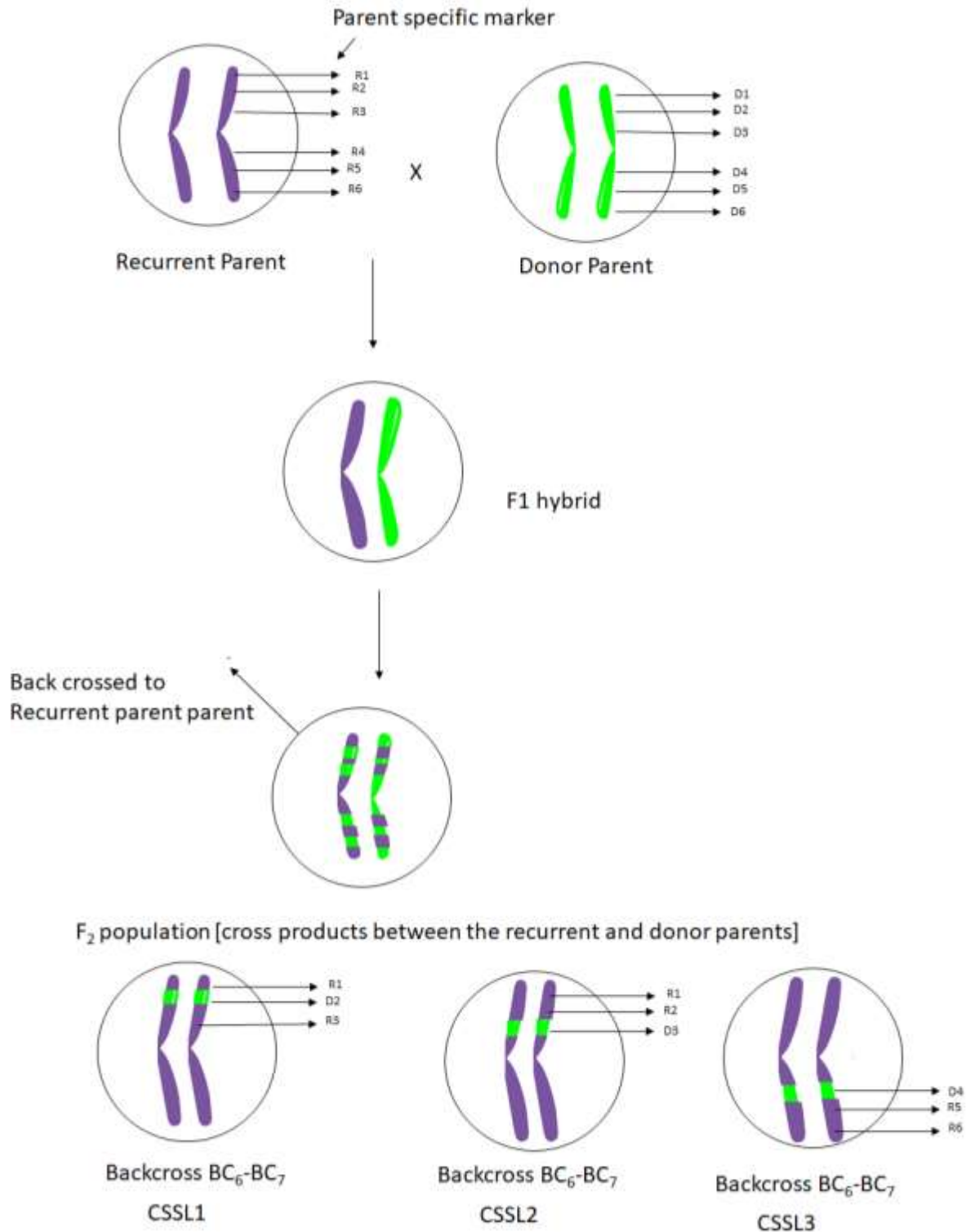


Figure 1. Scheme for the development of chromosome segment substitution lines (CSSLs). D1, 2, 3, etc. represent loci coming from the donor parent, usually a wild or weedy relative of the crop species of interest. R1, 2, 3, etc. represent loci coming from the recurrent parent, usually an adapted and desirable genotype of the crop species. Each CSSL (numbered 1, 2...N) carries a different chromosomal segment from the donor; together, they represent the entire donor genome.

Table 2. Information for chromosome segment substitution lines (CSSLs) that can serve as breeding material in improving soybean adaptability. References are available so that specific data about the CSSLs and traits already phenotyped in each population are accessible.

Parents	CSSLs	Results	Reference
'Khao Dawk Mali 105' ('KDML105')	Lines CSSL11, CSSL12, and CSSL16	DHL212 chromosome 1 was associated with salt tolerance and photosynthesis rate	Chutimanukul <i>et al.</i> (2018)
Jackson × JWS156-1	120	Recessive allele qFT12.1 delayed flowering	Liu <i>et al.</i> (2018)
NN1138-2 (<i>max</i>) × N24852(<i>soja</i>)	177	Six genes regulating days to flowering and two genes regulating seed coat color	Liu <i>et al.</i> (2021)
Enrei × Peking	103	50 reproducible QTLs for flowering, maturity, and seed yield	Watanabe <i>et al.</i> (2018)

Crossing *G. soja* (donor) and *S. max* (recipient) had developed CSSL lines in soybeans. Wang *et al.* (2013) developed 151 lines that covered about 96% of the donor genome; these were subsequently useful in identifying QTLs regulating valuable traits, such as flowering (Liu *et al.*, 2018), nematode resistance (Huang *et al.*, 2021), adaptability (Qingyuan *et al.*, 2021), and seed coat color (Liu *et al.*, 2021) (Table 2). Identifying a major QTL, *qFT12.1*, for flowering time that explained about 40% of the phenotypic variability resulted, and its recessive allele was responsible for the delayed flowering, a beneficial trait for adaptation to certain growing environments (Liu *et al.*, 2019). These CSSL and others presented in Table 2, may undergo phenotyping for other traits that can help improve yield and adaptation to specific environments by soybean breeders later. Since they have also received genotyping, markers for marker-assisted selection of identified genomic segments will be easy to detect.

Genetic and genomic analysis in soybean

Current soybean genetic resources include high-density genetic linkage maps, genome sequences, and high-density mutant libraries. Bi-parental genetic mapping of quantitative traits can identify quantitative trait loci (QTL) for valuable features of interest to breeders. In the case of soybean adaptability, an influential attribute would be the initiation of flowering, which expression depends on environmental factors such as temperature and photoperiod, that can be modified, and the timing that

influences seed yield. Identifying specific genes will aid in developing photoperiod-insensitive cultivars adapted to various agroecological conditions. Genes related to photoperiod sensitivity received an "E" gene designation; examples are presented in Table 3. Knowing the genomic locations of E genes and recognizing the genetic markers linked to them, breeders can select breeding lines containing the advantageous allele of each E gene without having to choose them phenotypically, which can be confounded by the effect of the environment. One QTL mapping study led to the discovery of a genomic region linked with locus E2, with latitude as shown to affect it (Table 3). Mutant plant (e2/e2) for locus E2 initiates early flowering compared with E2/E2 plants. More study of the e2 genotype plants found that they had elevated expression of *GmFt2* (Table 3). *GmFT2* is a photoperiod-responsive gene involved in flower promotion and maintenance, according to transgenic studies in *Arabidopsis thaliana* and soybean. It may be a potential target of soybean breeding to expand its adaptability.

The entire genome sequence assembly of important soybean cultivars, including 'William 82' (USA), 'Enrei' (Japan), and 'Peking' (China) are available online (https://www.ncbi.nlm.nih.gov/assembly/GCF_000004515.6; https://www.naro.go.jp/english/laboratory/nics/genomics/soybeans_genomics/index.html#id). Online and publicly available computing resources aid in genetic and genomic data analyses and storage, which become accessible

Table 3. Traits conferring adaptability to climatic conditions in soybean. Genes or Quantitative Trait Loci (QTL) known to affect these traits are also presented.

Locus	Adaptability	Gene	Reference
<i>PIF4</i>	Early maturity	Reduced plant height, low leaf area, decreased branches, early flowering, robust flowering to pod maturity	Arya <i>et al.</i> (2021)
<i>e3 /e4 , e1 (e1-nl or e1-fs) with e3 or e4</i>	Photoperiod insensitive	Induces photo-insensitivity in accessions, identified through screening of soybean germplasm	Gupta <i>et al.</i> (2017)
<i>E1, E2, E3, and E4</i>	Geographic adaptation	Genes related with flowering time	Kim <i>et al.</i> (2018)
Long-Juvenile traits (LJ)	Low latitude or tropical condition	E1 repressor and J protein suppress E1 inducing early flowering	Lu <i>et al.</i> (2017)
<i>e1</i> and <i>e2</i> alleles in Indian genotypes (<i>e1-as/E2/E3/E4, E1/e2/e3/E4, E1/e2/E3/E4, E1/E2/e3/E4</i>)	Photo insensitive alleles	Reduced days to flowering and photoperiod sensitivity	Tripathi <i>et al.</i> (2021b)
<i>Gmpr37</i> -ZGDD mutant allele	Inducing early flowering under long days	CRISPR/Cas9 induced modification of <i>GmPRR37</i>	Wang <i>et al.</i> (2020)
<i>eI-Lb</i>	Recessive alleles induce photo period insensitivity	Dominant allele retards flowering by suppressing the expression of FT2a and FT5a while recessive allele induces earlier flowering along with <i>e3/E4</i> or <i>E3/E4</i> under long days	Zhu <i>et al.</i> (2019)

through user-friendly genomics databases. These include SoyBase, SoyNet, SoyFN, SoyDB, and SoyKB. The Soybean Knowledge and Experience is a vital database that provides a complete web page for genomics data provided by diverse platforms.

The size of the genome of the Japanese cultivar 'Enrei' is 928 Mb, having been created by sequencing and assembling 221,674 contigs (Shimomura *et al.*, 2015). Other genomes are of similar size, and comparing genomes allows the identification of many insertion/deletion and Single Nucleotide Polymorphism (SNP) differences. Once the sequences of reference soybean genomes were available online, they could help facilitate re-sequencing and high-density genotyping of diverse panels of soybean lines, leading to identifying genes associated with essential traits via the genome-wide association study (GWAS). These genes' selection may be via marker-assisted choice or may one day gain direct manipulation using gene editing.

Adapting soybean to climate change using genomic tools

Soybean evolved as a short-day plant, and its cultivation had confinement to around 22 °N until the 1970s due to precocious flowering (Gupta *et al.*, 2021). Flowering time regulation comes from photo-thermal responsive genes, the environment (daylength), and their interactions. Alleles, such as *E1, E2, E3,* and *E4,* are known to affect maturity in soybeans, and an allelic variation, due to their loss of function within soybeans cultivated in various geographical regions, induces early-maturity and photo-insensitivity to day length (Miladinović *et al.*, 2018). Allelic combinations show complex interaction or additive effects; however, generally, dominant alleles cause late flowering. For example, the genotype *e1-nf* is earlier in flowering than the genotype *E1,* and the average difference between the *e1-asE4* and *E1E4* genotypes in flowering time was about six to 14 days (Zhai *et al.*, 2014).

Genotypes *e1-asE2E3E4* and *E1E2E3E4* differed by 8–28 days (Zhai et al., 2014). Recessive alleles *e3* and *e4* have an inhibitory effect over the allele *E1*. Phenological traits, such as flowering time, tend to show continuous variation due to the interaction of alleles at multiple loci and allele regulation by photoperiod and temperature. Therefore, whole genome screening and identification of markers (simple sequence repeat, SSR; single nucleotide polymorphism, SNP) associated with quantitative trait loci will speed the selection of appropriate plant material for target environmental conditions.

Varieties originating from a specific geographical region carry explicit haplotypes for the alleles of each *E* gene. Particular haplotypes occurred within more than 50% of the cultivars in a geographical area, except for Europe, where more than one haplotype was present (Miladinović et al., 2018). Cultivars originating from Novisad Serbia and carrying the *e1-as/e2/E3/E4* alleles were highly productive, with this combination of alleles recommended for cultivars grown in Central Europe (Miladinović et al., 2018). Similarly, the majority of cultivars grown in Kazakhstan acquired genotyping as *e1-as/e2/E3/E4* alleles, and cultivars carrying either dominant *E3* or *E4* are highly productive there (Abugalieva et al., 2016). Molecular markers (dCAP or CAP)

served in tagging alleles *e1-as* and *e3-fs* in Indian soybean cultivars with early maturity (Gupta et al., 2017).

Genes *FT2a* and *FT5a* participated in the induction of flowering in soybeans, and their controlled expressions came from various allelic combinations of “E” alleles mentioned above in response to photoperiod (Takeshima et al., 2016). Noting a delayed flowering was due to the down-regulation of two genes, *FT2a* and *FT5a*, in the Thai soybean cultivar ‘K3.’ Markers linked to each gene can be a choice for fast-creating combinations of alleles specific to each growing environment, and some are already publicly available (Table 4).

Identification of new markers linked to useful traits

Large-effect QTLs, such as Gm 11, Gm 16, and Gm 20, are also known to influence soybean flowering (Mao et al., 2017). Variation in flowering among cultivars arises due to epistatic × environment and additive × environment interactions of various loci affecting the traits. Genome-wide association studies identified 37 SNPs linked to flowering time and other characteristics associated with maturity in soybeans (Li et al., 2019). Five candidate genes, i.e., *Glyma.05G101800*, *Glyma.11G140100*, *Glyma.11G142900*,

Table 4. Impact of various alleles of genes known to affect flowering, maturity, and productivity of soybean cultivars in different ecological zones.

Allele	Environment	Flowering	Productivity	Reference
<i>e1-n/E2/ E3/E4</i>	Eastern Kazakhstan	670 thermal time length for flowering, Maturity group “0”	Highest yield in Eastern Kazakhstan	Abugalieva et al. (2016)
<i>ft2aft5a</i>	Short day length	Delayed	Increased pod and seed yield	Cai et al. (2020)
Induced modification in <i>E1</i>	Long day length	Early	High harvest index	Han et al. (2019)
<i>J</i> locus <i>Gmpr37</i>	Short day length High altitude	Long Juvenile Early	Yield enhancement Enhanced	Lu et al. (2017) Wang et al. (2020)
Iso-lines developed from <i>e2e2e3e3</i> ‘Enrei’, <i>E2E2e3e3</i> ‘Sachiyutaka’ <i>E2E2E3E3</i> ‘Fukuyutaka’	36°0’26N, 140°1’19E 34°13’47N, 133°46’36E	Early (<i>E1</i>), medium (<i>E3</i>) vs Late flowering (<i>E2</i>)	<i>E2</i> and <i>E3</i> isoline increased adaptability and more productivity in southern Japan	Yamada et al. (2012)

Glyma.19G099700, and *Glyma.19G100900*, were similar to flowering genes identified in *Arabidopsis thaliana* (Li *et al.*, 2019). These genes encode proteins, such as FRI (FRIGIDA), PUB13 (plant U-box 13), MYB59, CONSTANS, and FUS3, and may have a role in controlling the growth period of soybeans (Li *et al.*, 2019). Another association mapping study of flowering revealed loci on various chromosomes, including *Gm04_4497001* (near *GmCOL3a*), *Gm16_30766209* (near *GmFT2a* and *GmFT2b*), *Gm19_47514601* (near *E3* or *GmPhyA3*), and *Gm04_4497001*. These loci interact with other loci or environment variables to induce flowering response in soybeans (Mao *et al.*, 2017). Finally, a GWAS of phenological traits with high-density SNP markers showed complex inheritance of said traits, with as many as 27, 6, 18, and 27 loci correlated with attributes, such as, days to flowering, days to maturity, duration of flowering, and plant height (cm), respectively (Zhang *et al.*, 2021). Ten candidate genes were distinguishable with peak SNPs associated with flowering time.

The discovery of long juvenile (LJ) traits has expanded the latitude boundaries for soybean cultivation. The LJ traits help increase plant biomass under a short day (Tripathi *et al.*, 2021). Genes *E6*, *J*, and *E9* are the chief genes causing the expression of long juvenile traits, such as late flowering and maturity, in varieties grown under a short- day length. The *E6* gene mapping was on GM04 with marker HRM101. The *E6* gene has a suppressive effect over *E1*. Another study linked LJ traits with a SNP variation in FT2a loci (Tripathi *et al.*, 2021). Cultivars with LJ traits contributed to expanding soybean cultivation in Brazil. Short-duration (LJ) long-juvenile varieties combined with photo insensitivity may help to adapt to low rainfall environments. Soybeans' early flowering - is associated with locus e2, which carries three nonsynchronous SNP, identified by sequencing exon e2 in 40 varieties (Kim *et al.*, 2018). SNP variation in the coding sequence of E1Lb causes premature translation termination and induces early flowering under long days. Tagging of genes with molecular markers may help develop soybean cultivars to particularly adjust, leading to yield stability as

cultivars suit well to the best growing environments (Li *et al.*, 2017).

Soybean genome editing

Genome editing helps manipulate the genome by changing one or a few base pairs or inserting or deleting small sections of a specific sequence of DNA. It works with the clustered, regularly interspaced short palindrome (CRISPR) or ZEN (zinc finger nuclease) that targets and modifies desired DNA sequences. It may be a promising avenue for improving crop adaptability to environmental factors (Du *et al.*, 2016). Various genome editing systems generate cleavage sites in targeted genomic regions. The endogenous DNA repairing mechanism causes an addition or a deletion via non-homologous end-joining of the cleavage sites, resulting in a change in the targeted sequence. Additionally, homology-directed DNA repairs can aid in exchanging the target region with a homologous series, allowing longer (not simply single-point) changes. CRISPR RNA (CrRNAs) are extra reliable methods because they target the desired point in the genome more accurately than DNA binding proteins.

Genome editing is considered safer than other genetic modification methods due to the non-introduction of foreign DNA. Employing an editing protocol based on cleavage site analysis of DNA sequences has induced genome modification in soybeans. Utilizing the CRISPR/Cas9 method helped induce mutations in the E1 locus to suppress early flowering of soybeans under long-day lengths. The mutations caused early translation termination and a truncated protein (Han *et al.*, 2019), causing early soybean flowering under long-day length, increasing the harvest index. A decrease in the translation of the E1 protein caused subsequent increases in the expression of loci *GmFT2a* and *GmFT5a*. The E1 allele encodes a bipartite protein comprising a nuclear localizing signal and a region B3 domain, suppressing flowering by reducing the expression of loci *GmFT2a* and *GmFT5a*. Similarly, ethyl methane sulfonate-induced or natural mutations in allele E1 caused a nonsense mutation, resulting in a loss of

localizing specificity of the protein-inducing early flowering (Xia *et al.*, 2012).

A double mutant *ft2aft5a*, induced through CRISPR/Cas9, caused a delay in flowering of 31.3 days under short days, resulting in increased pod and seed yield. The conclusion stated that delayed flowering may be beneficial in tropical environments. In contrast, a modification of flowering inhibiting gene *GmPRR37* through CRISPR/Cas9 induced early flowering in soybeans. A natural mutant allele *Gmpr37* was also notable in soybean through phenotypic screening, and this allele has enabled soybean cultivation at higher altitudes than had been previously possible (Wang *et al.*, 2020); it may point to where gene editing can cause additional valuable variants in soybean.

Abiotic stress tolerance

Soybean is susceptible to various environmental stresses. Global climate change has increased the chance of soybean crop failure at the early establishment or reproductive phases due to heat, causing a significant decline in seed yield. Soybean cultivars of various maturity groups exposed to suboptimal or super-optimal temperatures express a substantial decrease in seedling growth parameters, i.e., seedling vigor index, total biomass, and root elongation (Alsajri *et al.*, 2019). Cultivars belonging to maturity group III showed more responses in seedling growth parameters in response to induced abiotic stresses when compared with cultivars of maturity groups IV and V (Alsajri *et al.*, 2019). Views on differential responses to abiotic stresses are due to various stress signals and activation of molecular pathways that enhance tolerance. Genes, such as ethylene responsive factor *GmERF75*, are triggered by abscisic acid signaling, inducing multi-stress tolerance. Genotypes overexpressing the *GmERF75* transcript had a higher chlorophyll content than wild types (Zhao *et al.*, 2019). *GmWRKY12* is an essential regulator of plant responses to abiotic stresses, with abscisic acid, drought, or salt stress activating it. Transgenic soybeans showing overexpression of *GmHsp90A2* had a high

tolerance to heat stress because of higher chlorophyll and lower malondialdehyde contents. Various transcripts' functions involved soybean's multiple types of stress tolerance.

One of the achievements of soybean genome sequencing projects was identifying the functions of tens of thousands of soybean genes (Bai *et al.*, 2019). The efficient and versatile CRISPR/Cas9 gene-editing technology may allow functional confirmation of the functions of these genes through comparison of modified with wild-type phenotypes. For example, truncated proteins encoded by the gene *GmHsp90A2* caused a loss of heat tolerance in soybean, identifying the possible function of this gene. These may lead directly to suitable gene-edited soybean breeding lines or to others bred by a more conventional marker-assisted selection of the genes following validation.

CONCLUSIONS

Molecular tools help understand gene functions associated with soybean adaptability to varied environmental conditions. The maturity and flowering time of lines containing different alleles at several loci and their interactions with photo-thermal conditions of specific environments are discussed in this review. Genotyping of soybean cultivars has allowed the identification of allelic combinations associated with maturity time and flowering initiation. Molecular genetic markers' development has characterized germplasm carrying various allelic combinations of genes that will aid in breeding cultivars adapted to different parts of the world. More marker development and validation can progress as new soybean mapping studies of abiotic stresses and adaptation reach completion. Global climate change will impact stress tolerance and soybean sustainable production, and breeders will have a continual challenge to keep yields from being reduced. Genetic resources and molecular tools presented here can help.

REFERENCES

- Abugalieva S, Didorenko S, Anuarbek S (2016). Assessment of soybean flowering and seed maturation time in different latitude regions of Kazakhstan. *PLoS one* 11(12): e0166894, <https://doi.org/10.1371/journal.pone.0166894>.
- Alam T, Kurniasih B, Suryanto P, Basunanda P, Supriyanta, Ambarwati E, Widyawan MH, Handayani S, Taryono (2019). Stability analysis for soybean in agroforestry system with kayu putih. *SABRAO J. Breed. Genet.* 51(4): 405-418.
- Alsajri FA, Singh B, Wijewardana C (2019). Evaluating soybean cultivars for low- and high-temperature tolerance during the seedling growth stage. *Agron.* 9(1): 13, <https://doi.org/10.3390/agronomy9010013>.
- Arya H, Singh MB, Bhalla PL (2021). Overexpression of PIF4 affects plant morphology and accelerates reproductive phase transitions in soybean. *Food Energy Secur.* 291, <https://doi.org/10.1002/fes3.291>.
- Bai MY, Yuan JH, Kuang HQ (2019). Generation of a multiplex mutagenesis population via pooled CRISPR-Cas9 in soya bean. *Plant Biotechnol. J.* e13239 <https://doi.org/10.1111/pbi13239>.
- Balakrishnan D, Surapaneni M, Mesapogu S (2019). Development and use of chromosome segment substitution lines as a genetic resource for crop improvement. *Theor. Appl. Genet.* 132(1): 1–25.
- Battisti R, Sentelhas PC, Parker PS (2018). Assessment of crop-management strategies to improve soybean resilience to climate change in Southern Brazil. *Crop Pasture Sci.* 69(2): 154–162.
- Bhartiya A, Aditya JP, Singh G, Gupta A, Agrawal PK, Bhatt JC (2012). Assessment of genetic variability of agronomic characters in indigenous and exotic collection of black soybean (*Glycine max* (L.) Merrill.). *SABRAO J. Breed. Genet.* 43(1) 81-90.
- Cai Y, Wang L, Chen L (2020). Mutagenesis of GmFT2a and GmFT5a mediated by CRISPR/Cas9 contributes for expanding the regional adaptability of soybean. *Plant Biotechnol. J.* 18(1): 298–309.
- Chutimanukul P, Kositsup B, Plaimas K (2018). Photosynthetic responses and identification of salt tolerance genes in a chromosome segment substitution line of 'Khao Dawk Mali 105' rice. *Environ. Exp. Bot.* 155: 497–508.
- Dayoub E, Lamichhane JR, Schoving C (2021). Early-stage phenotyping of root traits provides insights into the drought tolerance level of soybean cultivars. *Agron.* 11(1): 188, <https://doi.org/10.3390/agronomy11010188>.
- Donald CM (1963). Competition among crop and pasture plants. *Adv. Agron.* 15: 1–118.
- Donat MG, Lowry AL, Alexander LV (2016). More extreme precipitation in the world's dry and wet regions. *Nat. Clim. Chang* 6: 508–513. Du HY, Zeng XR, Zhao M (2016). Efficient targeted mutagenesis in soybean by TALENs and CRISPR/ Cas9. *J. Biotechnol.* 217: 90–97.
- FAO (2020). Food and Agriculture Organization of the United Nations. Data retrieved from <https://www.fao.org/faostat/en/#data/QCL>.
- Gupta S, Agrawal N, Tripathi R (2021). Long juvenility trait: A vehicle for commercial utilization of soybean (*Glycine max*) in lower latitudes. *Plant Breed.* <https://doi.org/10.1111/pbr.12926>.
- Gupta S, Bhatia VS, Kumawat G (2017). Genetic analyses for deciphering the status and role of photoperiodic and maturity genes in major Indian soybean cultivars. *J. Genet.* 96(1): 147–154.
- Hanafiah DS, Safni I, Siregar LAM, Damanik RIM, Lestami A (2021). Genetic analysis of soybean mutant lines resistance to stem rot disease (*Athelia rolfsii* Curzi). *SABRAO J. Breed. Genet.* 53(1): 126-138.
- Havlík P, Valin HJP, Gusti M (2015). Climate change impacts and mitigation in the developing world: An integrated assessment of the agriculture and forestry sectors. Policy Research Working Paper Vol 1 Washington DC: World Bank Group.
- Huang M, Qin R, Li C (2021). Transgressive resistance to *Heterodera glycines* in chromosome segment substitution lines derived from susceptible soybean parents. *Plant Genome* e20091, <https://doi.org/10.1002/tpg2.20091>.
- Kim HJ, Cho HS, Pak JH (2018). Confirmation of drought tolerance of ectopically expressed AtABF3 gene in soybean. *Mol. and Cells* 41(5): 413.
- Li M, Liu Y, Tao Y (2019). Identification of genetic loci and candidate genes related to soybean flowering through genome wide association study. *BMC Genom.* 20(1): 1–13, <https://doi.org/10.1186/s12864-019-6324-7>.
- Li M, Liu Y, Wang C, Yang X, Li D, Zhang X, Zhao L (2020). Identification of traits contributing to high and stable yields in different soybean varieties across three Chinese latitudes. *Front. Plant Sci.* 10, 1642.
- Li Y (2017). Genotype by environment interactions in forest tree breeding: Review of methodology

- and perspectives on research and application Tree. *Genet. Genom.* 13(60): doi: 10.1007/s11295-017-1144-x.
- Li Y, Chen Q, Nan H (2017). Overexpression of GmFDL19 enhances tolerance to drought and salt stresses in soybean. *PLoS One* 12(6): e0179554, <https://doi.org/10.1371/journal.pone.0179554>.
- Liu C, Chen X, Wang W (2021). Identifying wild versus cultivated gene-alleles conferring seed coat color and days to flowering in soybean. *Int. J. Mol. Sci.* 22(4): 1559, doi: 10.3390/ijms22041559.
- Liu D, Yan Y, Fujita Y, et al. (2018). A major QTL (qFT12 1) allele from wild soybean delays flowering time. *Mol. Breed.* 38(4): 1–10.
- Lopes da Silva F, Borem A, Sedyama T, et al. (Eds.) (2017). *Soybean Breeding*. Springer, pp. 211.
- Lu S, Zhao X, Hu Y, et al. (2017). Natural variation at the soybean J locus improves adaptation to the tropics and enhances yield. *Nature genetics*, 49(5), 773–779.
- Mao T, Li J, Wen Z (2017). Association mapping of loci controlling genetic and environmental interaction of soybean flowering time under various photo-thermal conditions. *BMC Genom.* 18(1): 1–17.
- Messina M (2022). Perspective: Soybeans can help address the caloric and protein needs of a growing global population. *Front. Nutri.* 9: 909464.
- Miladinović J, Čeran M, Đorđević V (2018). Allelic variation and distribution of the major maturity genes in different soybean collections. *Front Plant Sci.* 9: 1286, <https://doi.org/10.3389/fpls.2018.01286>.
- St-Marseille AFG, Bourgeois G, Brodeur J (2019). Simulating the impacts of climate change on soybean cyst nematode and the distribution of soybean. *Agric. Forest Meteorol.* 264: 178–187.
- Takeshima R, Hayashi T, Zhu J (2016). A soybean quantitative trait locus that promotes flowering under long days is identified as FT5a a FLOWERING LOCUS T ortholog. *J. Exp. Bot.* 67(17): 5247–5258.
- Tripathi R, Agrawal N, Kumawat G (2021). QTL mapping for long juvenile trait in soybean accession AGS 25 identifies association between a functional allele of FT2a and delayed flowering. *Euphytica* 217(3): 1–12.
- US Department of Agriculture (2020). *Oilseeds: World markets and trade*. Available from USDA. Accessed: March 23, 2020.
- Vandana B, Shukla PS, Kamendra S (2017). Morphological characterization and assessment of genetic variability in soybean varieties. *Inte. J. Curr. Microbiol Appl. Sci.* 6(3):361–369.
- Wang L, Sun S, Wu T (2020). Natural variation and CRISPR/Cas9-mediated mutation in GmPRR37 affect photoperiodic flowering and contribute to regional adaptation of soybean. *Plant Biotechnol. J.* 18(9): 1869, <https://doi.org/10.1111/pbi.13346>.
- Wang W, He Q, Yang H (2013). Development of a chromosome segment substitution line population with wild soybean (*Glycine soja* Sieb et Zucc) as donor parent. *Euphytica* 189(2): 293–307.
- Watanabe S, Shimizu T, Machita K (2018). Development of a high-density linkage map and chromosome segment substitution lines for Japanese soybean cultivar Enrei. *DNA Res.* 25(2): 123–136.
- Xia Z, Watanabe S, Yamada T (2012). Positional cloning and characterization reveal the molecular basis for soybean maturity locus E1 that regulates photoperiodic flowering. *Proceed. Nat. Acad. Sci.* 109(32): E2155–E2164, <https://doi.org/10.1073/pnas.1117982109>.
- Zhai H, Lü S, Wang Y (2014). Allelic variations at four major maturity E genes and transcriptional abundance of the E1 gene are associated with flowering time and maturity of soybean cultivars. *PLoS one* 9(5): e97636, <https://doi.org/10.1371/journal.pone.0097636>.
- Zhang S, Singh MB, Bhalla PL (2021). Molecular characterization of a soybean FT homologue GmFT7. *Sci. Rep.* 11(1): 3651, <https://doi.org/10.1038/s41598-021-83305-x>.
- Zhao MJ, Yin LJ, Liu Y (2019). The ABA-induced soybean ERF transcription factor gene GmERF75 plays a role in enhancing osmotic stress tolerance in Arabidopsis and soybean. *BMC Plant Biol.* 19(1): 506, <https://doi.org/10.1186/s12870-019-2066-6>.
- Zhu J, Takeshima R, Harigai K (2019). Loss of function of the E1-Like-b gene associates with early flowering under long-day conditions in soybean. *Front. Plant Sci.* 9: 1867, <https://doi.org/10.3389/fpls.2018.01867>.