

# Chapter 1

## Genomic Designing for Abiotic Stress Tolerant Soybean



**Milind B. Ratnaparkhe, Gyanesh K. Satpute, Giriraj Kumawat, Subhash Chandra, Viraj G. Kamble, Rucha Kavishwar, Vijayata Singh, Jogendra Singh, Ajay K. Singh, S. V. Ramesh, Virender Kumar, Sreeja Sudhakaran, Manoj K. Srivastava, Nishtha Shesh, Anjana Jajoo, Sanjay Gupta, Maharaj Singh, Donghe Xu, Madan Bhattacharya, and Henry T. Nguyen**

**Abstract** Soybean is an agronomically important crop which is rich in seed protein (about 40%) and oil (about 20%), enriches the soil by fixing nitrogen through symbiosis with bacteria. It is widely used as food, feed, and for industrial purpose. In addition to human consumption, soybean is a major protein source in animal feeds. Soybean is also becoming a major crop for biodiesel production. In soybean, abiotic stresses including drought, temperature extremes, floods, salinity, acidity, mineral

---

M. B. Ratnaparkhe (✉) · G. K. Satpute · G. Kumawat · S. Chandra · V. G. Kamble · R. Kavishwar · M. K. Srivastava · N. Shesh · S. Gupta · M. Singh  
ICAR—Indian Institute of Soybean Research, Indore, Madhya Pradesh, India  
e-mail: [milind.ratnaparkhe@icar.gov.in](mailto:milind.ratnaparkhe@icar.gov.in)

V. Singh · J. Singh  
ICAR—Central Soil Salinity Research Institute, Karnal, Haryana, India  
e-mail: [Jogendra.Singh@icar.gov.in](mailto:Jogendra.Singh@icar.gov.in)

A. K. Singh  
ICAR—National Institute of Abiotic Stress Management, Baramati, Maharashtra, India

S. V. Ramesh  
ICAR—Central Plantation Crops Research Institute, Kasaragod, Kerala, India

V. Kumar · S. Sudhakaran  
National Agriculture Food Technology Institute, Mohali, Punjab, India

N. Shesh · A. Jajoo  
School of Life Sciences, Devi Ahilya University, Indore, Madhya Pradesh, India

G. Kumawat · D. Xu  
Japan International Research Center for Agricultural Sciences, Tsukuba, Ibaraki, Japan  
e-mail: [xudh@jircas.affrc.go.jp](mailto:xudh@jircas.affrc.go.jp)

M. Bhattacharya  
Iowa State University, Ames, IA, USA  
e-mail: [mbhattac@iastate.edu](mailto:mbhattac@iastate.edu)

H. T. Nguyen  
University of Missouri, Columbia, USA  
e-mail: [nguyenhenry@missouri.edu](mailto:nguyenhenry@missouri.edu)

toxicity and nutrient deficiency have emerged as the major challenges for achieving the increased productivity. Breeding for tolerance to abiotic stresses is cumbersome due to their genetic control by multiple genes and are also very much influenced by the environment. The novel genomic designing approaches have enabled the improvement of soybean at a faster pace than traditional approaches. Genomic-assisted breeding, genomic selection, genome sequencing, marker-assisted selection, genetic engineering approaches, and genomics tools have been used to improve tolerance to biotic and abiotic stresses, yield and seed composition traits. Genomic designing overcomes the limitations of traditional breeding methods and accelerate the development of climate-smart soybean crops. Developing abiotic stress-tolerant soybean varieties have become convenient with the availability of a complete genomic sequence of soybean and functional genomics studies. This chapter discusses the major milestones in soybean genetics, genome mapping and recent developments in comparative and functional genomics and genome editing related to abiotic stresses.

**Keywords** Soybean · Abiotic stress · Drought tolerance · GWAS · Functional genomics

## 1.1 Introduction

### 1.1.1 *Economic Importance of the Crop*

Soybean [*Glycine max* (L.) Merr] is an important oilseed crop in the world and serves as a major source of protein and oil for both humans and animals. Soybean forms a raw material for several human health and industrial applications. Besides the edible oil (18–22%), the seed comprising around 38–45% of protein, and ash, carbohydrate minerals along with antioxidants are major component with potential nutraceutical applications for human health. Hence, soybean has been gaining wide attention in various industries such as food, feed wellness and pharmaceuticals which are attributable to its unique components of minerals, isoflavones, tocopherols etc. Ecologically soybean is involved in biological nitrogen fixation hence improves the soil fertility. Considering its diverse uses the crop is aptly named “miracle bean”. Although the crop is cultivated globally, the United States of America, Brazil, Argentina, China and India are major global producers. Also considering the multiple sectors the crop serves as raw materials, sustainable soybean production is imperative for ensuring global food security.

### 1.1.2 *Reduction in Yield and Quality Due to Abiotic Stresses*

Multiple abiotic stressors such as drought, elevated temperature, freezing conditions, floods, soil salinity, acidity and the consequent mineral toxicity or nutrient deficiency

are some of the challenges soybean production encounters worldwide. Further, the anticipated frequent extreme in weather conditions due to global climate change is another serious concern for sustainable soybean production. It has been observed that millions of acres of soybean crop loss occur every year due to multiple abiotic factors. Crop loss due to various abiotic stresses demands developing strategies to increase soybean yield or maintain yield stability under multitude of abiotic stresses. Therefore, genomic design of soybean for climate resilience and sustainable production with higher yield potential and nutritional value is mandatory.

### ***1.1.3 Growing Importance in the Face of Climate Change and Increasing Population***

The multitude of biochemical characteristics and good quality oil makes soybean a desired oil seed crop and rising its demand worldwide. Nevertheless, the requirement of doubled food production by the end of the year 2050 owing to population explosion will severely squeeze the sufficient production of oil seed crops even more so in the context of changing climatic conditions (Deshmukh et al. 2014). Climate change and extreme weather conditions negatively impact the crop yield while temperature, precipitation, and solar radiation are the main drivers of crop growth and development. Therefore, the breeders are entrusted to provide emphasis on the development of not only high yielding and nutritionally superior soybean genotypes but also the genotypes which are expected to tolerate extreme weather conditions.

### ***1.1.4 Limitations of Traditional Breeding and Rational of Genome Designing***

The conventional plant breeding strategies such as single pod descent, back-cross breeding, pedigree breeding and bulk population breeding have undoubtedly contributed to the improved soybean yield and tolerance to various abiotic stresses. Nevertheless, these strategies are time consuming and warrant screening of huge plant population that consumes land, labour and water resources. Moreover the breeding for complex traits that are governed by multiple genes are severely influenced by the environment. Further, the complexity of multiple abiotic stresses affecting the standing crop due to climate change have instilled a sense of new urgency into accelerating the rates of genetic gain in molecular breeding programs. Hence, regardless of the conventional breeding efforts, it is imperative to integrate the genome designing based breeding approaches to enhance the production potential of the soybean. To facilitate the advances in soybean breeding, it is indispensable to exploit the molecular breeding techniques such as marker-assisted breeding, recombinant DNA technology, genome editing and multiple “omics” to improve the soybean quality and

yield. Hence, these limitations of traditional breeding strategies warrant the large-scale application of genomics science in the improvement of soybean for abiotic stresses.

## **1.2 Abiotic Stresses and Related Traits in Soybean**

### ***1.2.1 Root Characters***

Considering the water-deficit stress or flooding stress due to climate change, characteristic features of soybean roots are important to tide over the abiotic stresses. The observed root architecture traits of soybean have revealed that narrow root angle to the soil surface is preferred as it enhances development of lateral roots in the upper root regions where penetration of sunlight is ample. Other root traits such as number of forks, crossings are imperative for good soil penetration, coupled with appreciable root length density (RLD) due to enhanced root surface and root volume. Root characteristic features are important for absorption of soil moisture during stress conditions. Nonetheless deeper soybean roots have not yielded desired results when the soil is shallow or no water at depth or during the conditions of mild water stress (Vadez et al. 2015).

### ***1.2.2 Drought Tolerance***

An estimate states that around 40% reduction in soybean production worldwide is due to decrease in water supply and it is also anticipated that such losses would further aggravate due to frequent droughts and water shortages under the scenario of future climate change. Enhancing the irrigation potential is not a viable approach considering the poor resource conditions of the many of the developing countries. This scenario warrants the development of drought-tolerant varieties as an important research urgency. Drought in soybean reduces the economic yield levels by 40% (Specht et al. 1999), however, depending upon the intensity of water-deficit stress and the stage of occurrence, yield losses could be as high as 80%. Phenotyping for drought resistance assumed significance in this context, wherein physiological and biochemical aspects of dehydration avoidance and dehydration tolerance are measured. Breeding for drought tolerance depends on phenotyping methods which are reliable, relatively fast and economical. Generally, the measure of dehydration avoidance involves investigating plant water status, in terms of visual symptoms of leaf senescence, relative water content and analyzing other constitutive traits such as root architecture attributes.

### ***1.2.3 Flooding and Submergence Tolerance***

Water logging/flooding is a most deleterious stress next only to drought. Flooding affects the plant health and yield of soybeans in 16% of the soybean production area causes severe economic losses. In US alone flooding stress in soybean causes an annual loss of approximately \$1.5 billion (Boyer 1982, 1983; Oosterhuis et al. 1990; Rosenzweig et al. 2002; Bailey-Serres et al. 2012; Ahmed et al. 2013). Flooding stress could be due to submergence or water logging though the former in soybean is a rare occurrence (Oosterhui et al. 1990; VanToai et al. 1994; Linkermer et al. 1998). Water logging or flooding results in reduction in root and shoot growth, decline in atmospheric nitrogen fixation, photosynthetic potential, stomatal conductance and nutrient uptake consequently severely affects the yield of soybean and it may cause death of plant in severe conditions (Sullivan et al. 2001; Shannon et al. 2005; VanToai et al. 2012; Rhine et al. 2012; Wu et al. 2017a).

### ***1.2.4 Heat Tolerance***

Yield reduction in soybean due to extreme temperature conditions has been estimated to be around 40% (Specht et al. 1999). Heat stress during vegetative stage affects the growth of soybean. Soybean is highly sensitive to elevated temperature conditions (>35 °C) during reproductive stages as heat stress cause flower and pod abortion during early stages, however the prolonged heat stress during pod filling stages leads to severe reduction in seed size and seed vigour (Boyer 1982; Chebrolu et al. 2016). Therefore, improving heat tolerance of soybean varieties is very crucial to improve the yield levels.

### ***1.2.5 Cold Tolerance***

In order to expand the soybean cultivation area from its traditional stronghold it is essential to impart cold tolerance trait so that cultivars could adapt to growing under low temperature conditions. The multiple effects of low temperature on soybean include poor germination, less seedling vigour, flower abortion and poor grain filling at reproductive stages (Yamamoto and Narikawa 1966). Northern hemisphere is characterized with short growing seasons and hence efforts are required to develop soybean varieties having traits such as good emergence and early seedling vigor. Seedling emergence test and early seedling weight are the traits evaluated in soybean germplasm. Genetic dissection of these traits and introgression in cultivated varieties through marker assisted breeding programme is a viable approach to enable the growth of soybean in northern regions.

### 1.2.6 Salinity Tolerance

Salinity stress severely affects the yields of soybean. High salinity level poses serious damage to the life cycle of soybean whereas low salt levels could cause significant reduction in soybean yield levels (Abel and Mackenzie 1964; Pitman and Läuchli 2002). Various agronomic features of the crop that are affected due to high salinity are significant reduction in plant height, leaf size, biomass, number of pods.plant<sup>-1</sup>, number of internodes.plant<sup>-1</sup>, number of branches.plant<sup>-1</sup>, weight.plant<sup>-1</sup> and 100 seed weight (Shao et al. 1986; Shao et al. 1993; Parida and Das 2005; Blanco et al. 2007; Bustingorri and Lavado 2011; El-Sabagh et al. 2015). Salt stress observed during the nodulation stage greatly reduces the efficiency of biological nitrogen fixation as severe reduction in number and biomass of root nodules documented (Singleton and Bohlool 1984; Rabie and Kumazawa 1988; Yang and Blanchar 1993; Delgado et al. 1994; Elsheikh and Wood 1995). Soybean germplasm display a spectrum of salt tolerance capability (Yang et al. 1993; Pitman and Läuchli 2002; Lenis et al. 2011).

## 1.3 Genetic Resources of Resistance/Tolerance Genes

The diverse morphological, cytological and genetical features of wild species of soybean and also the cultivated soybean display wide array of genetic sources of resistance to multiple biotic and abiotic stresses. Thus, wild species form an important component of gene pool for the exploration of useful genes and alleles conferring abiotic stress tolerance. The annual and perennial soybean species are significantly distantly related. Wild perennial *Glycine* species offer immense potential for soybean improvement. The genus *Glycine* has two subgenera, *Glycine* Willd. (perennial) and *Soja* (Moench) F.J. Herm (annual). The subgenus *Soja* includes two species: the cultivated soybean [(*G. max* (L.) Merr.)] and its wild annual progenitor *G. soja* Sieb. & Zucc. (Ratnaparkhe et al. 2010). The subgenus *Glycine* comprises 30 wild perennial species thus, the genetic resources of soybean may be categorized into four plausible gene pools (GP).

### 1.3.1 Soybean GP-1

Soybean gene pool -1 (GP-1) comprise all of the biological species which could be crossed among them to yield vigorous hybrids characterized with normal meiotic chromosome pairing and possess total seed fertility. All soybean (*G. max*) germplasm and the wild soybean, *G. soja*, constitute GP-1 with the qualification that seed sterility can be associated with chromosomal structural changes such as inversions and translocations. Gene segregation is normal and gene exchange is generally easy.

### **1.3.2 Soybean GP-2**

GP-2 include those species which can hybridize with GP-1 with relative ease and the resultant F<sub>1</sub> plants exhibit at least some seed fertility (Harlan and de Wet 1971). *Glycine max* is devoid of GP-2 because no known species exhibit such a relationship with soybean. Though it is plausible for existence of such species in Southeast Asia where the *Glycine* genus may have originated, extensive germplasm exploration is indispensable to validate this suggestion.

### **1.3.3 Soybean GP-3**

GP-3 is a potential genetic resource of soybean even though the hybrids between GP-1 and GP-3 are lethal. Gene transfers between GP-1 and GP-3 are not possible without resorting to in vitro culture based techniques such as embryo rescue etc. (Harlan and de Wet 1971). GP-3 includes the 26 wild perennial species of the subgenus *Glycine* which are indigenous to Australia and remain geographically isolated from *G. max* and *G. soja*. The three species *G. argyrea*, *G. canescens*, and *G. tomentella* have been successfully hybridized with soybean embryo culture based rescue techniques ensured the survival of F<sub>1</sub> hybrids. However, much progress has not been made beyond the amphidiploid stage, with the exception of Singh et al. (1998a, b) suggesting that only these three species belong to GP-3.

### **1.3.4 Soybean GP-4**

GP-4 is considered an extremely outer limit of genetic resources of soybean. Several pre- and post-hybridization barriers arrests the process of embryo development resulting in premature abortion of embryo. Bridge crosses with genus *Glycine* could circumvent the problems of seedling lethality, seed inviability and inviable F<sub>1</sub> plants (Singh et al. 2007). Hence restricting the utility of GP-4 only few wild perennial *Glycine* species have been hybridized with soybean. Thus, majority of species belong to soybean GP-4 as they have not been hybridized with GP-1 when hybridized did not produce viable F<sub>1</sub> plants (Singh et al. 1987). Although the wild perennial species exhibit resistance to several biotic and abiotic stresses, the transfer of useful genes into soybean has not been accomplished. Thus, the breeders/geneticists have access to the primary gene pool for expanding the germplasm base.

## 1.4 Classical Genetics and Traditional Breeding of Abiotic Stress Tolerance in Soybean

### 1.4.1 Classical Mapping Efforts

Soybean has been under continuous investigation for its genetic improvement by plant breeders. The crop encounters various biotic and abiotic stresses and hence improving their tolerance to stresses along with seed composition traits is pertinent. Improving agronomic performance of the crop would ensure higher productivity, and improved consumption of soybean and soy products leading to realization of greater economic benefits. Plant breeders have been traditionally using the practices of hybridization and meticulous selection methods to ensure better performing genotypes resulting in development of many soybean varieties. Classical genetics and traditional breeding approaches have been used to develop tolerance for drought, waterlogging stress, salt tolerance and for other abiotic stresses. Table 1.1 lists the soybean lines and resources used for the genetic improvement of abiotic stress tolerance.

### 1.4.2 Breeding Objectives

Designing highly productivity genotypes under water-limited conditions is an important breeding objective. It warrants introgression of physiological traits that define plant water relations and hydraulic processes, into a common genetic background (Satpute et al. 2021). Water deficit condition is an outcome of complex interplay of several factors including low soil moisture and extreme temperatures and other edaphic factors. Breeding promising soybean genotypes through transfer of gene(s) conferring drought tolerance is an effective approach to alleviating the ill-effects of drought. Under drought stress, soybean plants suffer from oxidative injury, membrane system damage, cellular ion leakage and protein denaturation, declined photosynthetic rates, and CO<sub>2</sub> uptake rates consequently causing reduction in biomass accumulation and yield levels. Hence, under drought, among the various physiological processes, photosynthesis is severely down-regulated with wider ramifications for the economic yield levels of the crops.

Breeding approaches to develop drought tolerant soybean involve diverse strategies namely recurrent selection and evaluation of segregating population under imposed drought-stress environment, and investigating the secondary traits for efficient selection, molecular breeding for drought tolerance, genomics-based and transgenic technologies to improve the drought tolerance trait. Advance phenotyping-based breeding approaches are pre-requisite and are being adopted systematically by developing early generation biparental, backcross or multi-parent intercross populations using identified candidate drought tolerant parental lines and wider-adaptable high yielding varieties. The populations are advanced through F<sub>2</sub> generation by mass selection where bulks are subjected to chemical desiccation process using potassium



**Table 1.1** Potential genetic resources for abiotic stress tolerance

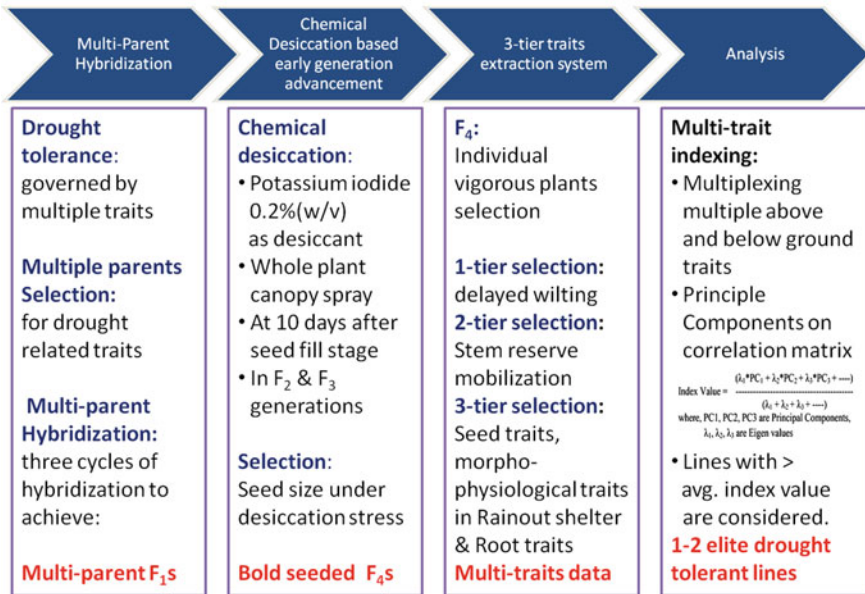
Tolerant genetic resource/genotype	Source	References
<i>Drought tolerance</i>		
PI 416,937	USDA, ARS	Sloane et al. (1990), Patterson and Hudak (1996), Sinclair et al. (2007), King et al. (2009)
Young	USDA, ARS	Mian et al. (1996)
Jackson	–	Purcell et al. (1997)
PI 407,155	–	Chen et al. (2006)
R01-416F and R01-581F	USA	Chen et al. (2007)
93,705-36 and PI471938	USDA	King et al. (2009)
PI 468,917	–	Seversike et al. (2014)
C12 and W05	CUHK, China	Hossain et al. (2014)
PI 567,690 and PI 567,731	China	Pathan et al. (2014)
EC 538,828, JS 97–52 and EC 602,288	India	Bhatia et al. (2014), Bhatia and Jumrani (2016)
NTPR94-5157, N09-13,890, NC-Raleigh, and SC07-1518RR	USA	Fried et al. (2019)
PK 1180 and SL 46	India	Sreenivasa et al. (2020)
TGX 709-7E and EC 389,174	India	Satpute et al. (2020a)
<i>Water logging tolerance</i>		
Edison, GR 8836, CX 415	USA	VanToai et al. (1994)
Archer	USA	VanToai et al. (2001)
PI 408105A	Korea	Shannon et al. (2005)
Misuzudaizu	Japan	Githiri et al. (2006)
91,210–350, 91,210–316	USA	Henshaw et al. (2007)
Kefeng No. 1	China	Wang et al. (2008b)
Peking	China	Sayama et al. (2009)
Nam Vang, VND2, ATF15-1	Cambodia	VanToai et al. (2010)
AGS 313	Bangladesh	Ara et al. (2015)
JS 97–52 and JS 20–38	India	Anonymous (2015)
Iyodaizu	Japan	Nguyen et al. (2017)
Kokubu 7, Maetsuezairai 90B, Yahagi	Japan	Suematsu et al. (2017)

(continued)

**Table 1.1** (continued)

Tolerant genetic resource/genotype	Source	References
UA 5615C, R10-4892, R13-12552, R07-6669, Walters, R04-342, S11-25,108, S12-1362, S11-25,615	USA	Wu et al. (2017a)
JS 20–38, C 2797, Hardee, IS 128 and JS 71–05	India	Chandra et al. (2020)
PI 561,271, PI 567,651, PI 567,343, PI 407,184, PI 603910C, PI 567394B, PI 467,162, PI 479,751, PI 407,229, PI 597459C, PI 424,082, PI 378699A, PI 424107A, PI 366,124	USA	Valliyodan et al. (2017)
<i>Salt tolerance</i>		
S-100, Centennial, Cook, D49-2491, Dillon, Forrest, Gordon, Haskell, Hill, Hutton, Johnston, Lee, Manokin, Wright	USA	Lee et al. (2004)
Nannong 1138–2	China	Chen et al. (2008)
Jindou no. 6 (PI574484)	China	Hamwihet et al. (2011)
FT-Abyara (PI628838)	Brazil	Hamwihet et al. (2011)
JWS156-1	Japan	Hamwihet and Xu (2008)
PI483463	China	Lee et al. (2009)
Tiefeng 8	China	Guan et al. (2014a)
W05	China	Qi et al. (2014)
Fiskeby III	Sweden	Do et al. (2018)
Osage	USA	Zeng et al. (2017)
Jidou 12	China	Shi et al. (2018)
SL 1226, SL 1258	India	Singh et al. (2020)
PI597458B, PI342434, PI548198, PI561389B, PI407202, PI407220, PI424107A, PI479752, PI407083, PI468908, PI080837, PI417500, PI424116, PI483460B, PI562551	USDA Soybean Germplasm Collection	Do et al. (2019)

### Drought Tolerant Soybean Pipeline: A Complete Workflow



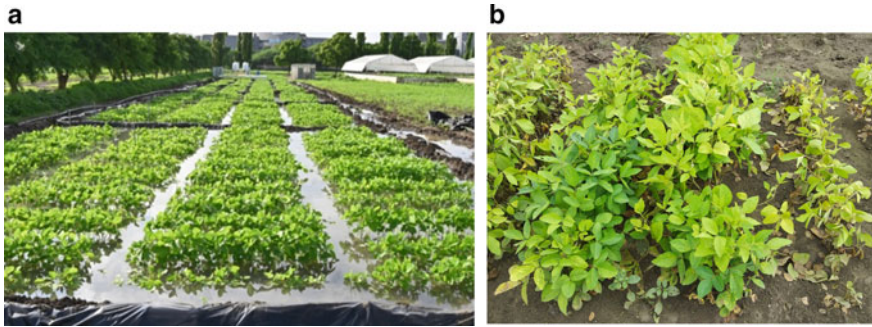
**Fig. 1.1** Workflow for developing drought tolerant parental lines and varieties specially for target population of environments (TPE) of drought conditions occurring at seed filling growth stage

iodide 0.2% (Blum et al. 1991; Bhatia et al. 2014; Satpute et al. 2019) followed by selections. Mass selection for the trait large seeds following chemical desiccation has significantly improved the seed weight and grain yield under chemical desiccation stress compared to control set up wherein seed selection was performed without chemical desiccation (Blum 2011). After two cycles of selection, intensive investigation of candidate genotypes for multiple drought tolerance-related traits is practiced in advance generations using three-tier selection scheme followed by multi-traits indexing. Figure 1.1 shows the scheme involved in developing drought tolerant soybean (Satpute et al. 2021). Development of soybean varieties with enhanced tolerance to drought, heat, salinity and cold has become a high research priority in major breeding programs worldwide (Fig. 1.2).

## 1.5 Genetic Diversity Analysis

### 1.5.1 Phenotype and Genotype Based Diversity Analysis

During the past three decades, genetic diversity studies in soybean has been dominated by phenotyping-based diversity analysis, cytogenetics and molecular studies,



**Fig. 1.2** a and b Screening of soybean germplasm under field condition for tolerance to water logging

including variation in isozymes and, seed proteins, use of restriction fragment length polymorphism (RFLP), random amplification of polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) markers. The geographic differentiation in Chinese cultivated soybean and genetic diversity was studied using the coefficient of parentage (Cui et al. 2000a), morphological traits (Dong et al. 2004), SSR markers (Li et al. 2008a; Wang et al. 2008a; Li et al. 2010; Wang et al. 2015) and SNP markers (Kajiyama-Kanegae et al. 2021; Saleem et al. 2021). The diversity analysis of Asian soybean landraces and North American cultivars revealed a low level of diversity in the American pools than in the Asian pools, based on phenotypic characterization (Cui et al. 2000a, 2001) or the coefficient of parentage analysis (Cui et al. 2000b). Low diversity was further confirmed in DNA sequence-based analyses showing successive genetic bottlenecks between wild and cultivated soybeans and between Asian landraces and North American cultivars (Hyten et al. 2006). Genetic diversity studies in soybean have been discussed in detail by Carter et al. (2004). Comparison of Chinese and American Soybean Accessions using High-Density SNPs in population structure analysis, and cluster analysis revealed that the genetic basis of Chinese soybeans is entirely distinct from that of the USA (Liu et al. 2017).

### ***1.5.2 Relationship with Other Cultivated Species and Wild Relatives***

Comprehensive biosystematics based relationship analysis of all species in the genus *Glycine* reveal that annual (subgenus *Soja*) and perennial (subgenus *Glycine*) soybean species are significantly distantly related (Doyle et al. 2003), having diverged from a common ancestor around 5 MYA (Innes et al. 2008). As stated above in genetic resources section, several attempts to hybridize species between the subgenus *Soja* and subgenus *Glycine* were unsuccessful. The pods resulted from interspecific

hybridization eventually aborted and got abscised although pod development was found to be initiated (Ladizinsky et al. 1979a; b). Later, the intersubgeneric F<sub>1</sub> hybrids of *G. max* × *G. clandestina*, *G. max* × *G. tomentella* and *G. max* × *G. canescens* were successfully obtained either following embryo rescue technique (Newell and Hymowitz 1982; Singh and Hymowitz 1985; Singh et al. 1987) or using transplanted endosperm as a nurse layer (Broué et al. 1982). In general, the cultivated soybean could only hybridize with members of the subgenus *Glycine* imperfectly. The progeny of such inter-subgeneric hybrids were completely sterile and obtained with a great difficulty. Studies have proven that cultivated soybean does not hybridize with any of the wild relatives in other genera of the tribe (Hymowitz et al. 1995; Hymowitz 2004). The wild soybean (*G. soja*) has accumulated rich genetic diversity in the process of evolution and adaptation (Kofsky et al. 2018). This adaptive evolutionary process has resulted in wide diversification in the traits of wild soybean. The diversity for multiple morphological features includes flower, pubescence, seed and hilum color, disease and insect resistance traits, physiological and biochemical traits (protein, oil and carbohydrates and their constituents content) (Boerma and Specht 2004).

## 1.6 Association Mapping Studies

### 1.6.1 Extent of Linkage Disequilibrium

Linkage disequilibrium (LD) describes changes in the genetic variation within a population over time. Variation in LD either at genome scale or at a particular-genomic region is influenced by various factors such as mutation, domestication, level of inbreeding and selection, confounding effects, population admixture, and population substructure (Rafalski and Morgante 2004). A strong correlation is anticipated between inter-locus distance and LD if the recombination rates do not vary across the genome particularly in a constant population size. Soybean, owing to its ineffective recombination and homozygous genetic background, exhibit less decay of LD (longer region is in LD).

### 1.6.2 Genome-Wide LD Studies

SNPs are choice markers due to its abundant DNA polymorphism and hence are useful in genetic diversity studies and in determination of genetic relatedness among the individuals. To investigate the genetic frequency of SNP in soybean genome, ~28.7 kbp of coding region, 37.9 kbp of non-coding perigenic region, and 9.7 kbp of random non-coding genomic regions were evaluated in 25 diverse soybean genotypes (Zhu et al. 2003). This study divulged that the nucleotide diversity ( $\theta$ ) observed in coding and in non-coding perigenic DNA was 0.00053 and 0.00111, respectively.

The combined nucleotide diversity of whole sequence was 0.00097. Squared allele frequency correlations ( $r^2$ ) among haplotypes at 54 loci with two or more SNPs indicated low genome-wide LD. A haplotype map of soybean (GmHapMap) was constructed using whole-genome sequence data from 1007 *Glycine max* accessions yielding 14.9 million variants as well as 4.3 M tag SNPs (Torkamanhe et al. 2021). A lower level of genome wide genetic diversity was observed in soybean compared to other major crops. Genome-wide LD investigations in soybean have facilitated identification of molecular markers and key genes associated with various abiotic stresses.

### 1.6.3 Genome Wide LD Studies for Drought Tolerance

Quantitative trait locus (QTL) mapping using bi-parental populations is limited by restricted allelic diversity of parental genotypes and mapping resolution. The allelic diversity among mapping populations can be increased to some extent by using multi-parental crosses (Deshmukh et al. 2014). The genome wide association study (GWAS) approach provides opportunities to explore the tremendous allelic diversity present in soybean germplasm. Since millions of crossing events and natural mutations have been fixed in the germplasm during evolution, the mapping resolution of GWAS is comparatively higher. The recent advances in sequencing have played an important role in performing the genome-wide association studies (Abdurakhmonov and Abdukarimov 2008). GWAS is now routinely being used in soybean and other plant species, however fewer studies have been reported with regards to different abiotic stresses. GWAS for quantitative traits like drought tolerance are predicted to be affected by population structure. GWAS models like mixed linear model (MLM) and compressed mixed linear model (CMLM) have been developed which takes into account the population structure, kinship and spurious allelic associations (Deshmukh et al. 2014). Recent development in statistical tools involving larger set of genotypes and high throughput genotyping approaches will definitely improve GWAS power.

Dhanpal et al. (2015) analyzed a population of 373 genotypes in four environments for carbon isotope ratio ( $\delta^{13}\text{C}$ ), an important physiological trait linked with water use efficiency (WUE). An association of 39 SNPs, linked to 21 different loci involved in conferring drought tolerance trait has been found. Similarly, Kaler et al. (2017) reported 54 SNPs associated with  $\delta^{13}\text{C}$  & 47 SNPs associated with  $\delta^{18}\text{O}$ . These SNPs were tagged with 46 putative loci and 21 putative genetic loci for  $\delta^{13}\text{C}$  and  $\delta^{18}\text{O}$ , respectively. Several markers and loci have been reported for various drought related traits viz., chlorophyll fluorescence (Hao et al. 2012; Herritt et al. 2018), canopy temperature (Kaler et al. 2017), delayed canopy wilting (Steketee et al. (2020) and drought susceptibility index (Chen et al. 2020) (Table 1.2). GWAS analysis in soybean for drought tolerance was reported using 259 Chinese cultivars for drought related traits. This investigation was based on a total of 4,616 SNPs, and 15 SNP-trait

**Table 1.2** Details of genome wide association studies (GWAS) carried out for abiotic stress tolerance

Trait	GWAS loci	Markers	Genotypes	Method	References
<i>Drought tolerance</i>					
Chlorophyll and chlorophyll fluorescence parameters	51	1536 SNP	168	MLM	Hao et al. (2012)
Ureide concentration	53	33,957 SNP	374	PROC GLIMMIX	Roy et al. (2014)
Carbon isotope ratio ( $\delta^{13}C$ )	21	12,347 SNP	373	GLM & MLM	Dhanpal et al. (2015)
Chlorophyll contents	27	31,253 SNP	332	MLM	Dhanpal et al. (2016)
Carbon isotope ratio ( $\delta^{13}C$ ) Oxygen isotope ratio ( $\delta^{13}C$ )	46 21	31,260 SNP	373	Farm-CPU	Kaler et al. (2017)
Canopy temperature	34	31,260 SNP	345	Farm-CPU	Kaler et al. (2017)
Chlorophyll fluorescence	53	32,453 SNP	189	CMLM	Herritt et al. (2018)
Delayed canopy wilting	44	34,379 SNP	162	MLM	Steketee et al. (2020)
Germination under drought	15	4,616 SNP	259	MLM	Liu et al. (2020a)
Drought susceptibility index and yield traits	302	105,970 SNP	136	MLM	Chen et al. (2020)
<i>Water logging tolerance</i>					
Seed-flooding tolerance in terms of germination rate, normal seedling rate and electric conductivity	25	60,109 SNPs	347	MLM and mrMLM	Yu et al. (2019)
Foliar damage score	14	31,125 SNPs	384	GLM, MLM, CMLM., ECMLM	Wu et al. (2020)
<i>Salinity tolerance</i>					
Seed germination under salt stress	8	1142	191	MLM	Kan et al. (2015)

(continued)

**Table 1.2** (continued)

Trait	GWAS loci	Markers	Genotypes	Method	References
Leaf chloride concentrations and leaf chlorophyll concentrations	45	33,009	283	GLM, MLM	Zeng et al. (2017)
Leaf scorch score	62	37,281	192	GLM, MLM	Huang et al. (2018)
leaf scorching score (LSS), chlorophyll content ratio(CCR), leaf sodium content(LSC) and leaf chloride content (LCC)	8	42,000	305	EMMAX, MLM	Do et al. (2019)
LSS, CCR, LSC and LCC	29	3.7 M	305	EMMAX, MLM	Do et al. (2019)
Salt tolerance at germination stage	18	207,608	211	CMLM	Zhang et al. (2019)

associations were identified by GWAS, among which three SNPs were suggestively associated with two of the drought-tolerance indices (Liu et al. 2020a).

#### ***1.6.4 Genome-Wide Association Mapping for Flooding Tolerance***

Genome-wide association mapping has advantages over bi-parental QTL mapping as the former exploits the historical and evolutionary recombination (Zhu et al. 2008). Yu et al. (2020) conducted GWAS in a panel of 347 soybean genotypes to identify SNPs associated with seed-flooding tolerance related traits, viz., germination rate (GR), normal seedling rate (NSR) and electrical conductivity (EC). Use of 60,109 SNPs identified three major QTNs, viz., QTN13, qNSR-10 and qEC-7-2 linked to the traits. Further, QTN13 was consistently identified in all three traits investigated and in multiple environments. Wu et al. (2020a) applied GWAS in a panel of 384 soybean lines, using 42,291 SNP markers and models viz. regression linear model (GLM), mixed linear model (MLM), compressed mixed linear model (CMLM), and enriched compressed mixed linear model (ECMLM) for dissecting flooding tolerance. It has resulted in identification of 14 SNPs associated with flooding tolerance across all the environments and models (Table 1.2).



## ***1.6.5 Genome-Wide Association Mapping for Salt Tolerance***

Seed germination under salt stress was used for an association mapping study by Kan et al. (2015). Under salt stress, three loci significantly associated with three traits namely the ratio of imbibition rate, the ratio of germination index, and the ratio of germination rate, were identified and mapped to chromosomes Gm08, Gm09 and Gm18. Using 283 diverse lines of worldwide soybean accessions, In another GWAS study, Zeng et al. (2017a) identified eight genetic loci (mapped on to chromosomes Gm02, Gm7, Gm08, Gm10, Gm13, Gm14, Gm16, and Gm20) associated with leaf chloride and leaf chlorophyll concentrations by using sing 283 diverse lines of soybean accessions. Huang et al. (2018) used a diverse set of 192 soybean germplasm and identified six genomic regions (Gm02, Gm03, Gm05, Gm06, Gm08 and Gm18) associated with salt tolerance based on visual leaf scorch score. The study by Do et al. (2019), using two GWAS populations for association mapping of salt tolerance, confirmed the major locus on chromosome Gm03 and identified three novel loci on Gm01, Gm08 and Gm18. Several SNPs have been identified to be significantly associated with traits, leaf scorching score (LSS), chlorophyll content ratio (CCR), leaf sodium content (LSC) and leaf chloride content (LCC) (Do et al. 2019). Zhang et al. (2019) identified genomic regions associated with salt tolerance at germination stage and showed 18 significant SNPs were located on chromosome Gm08 and Gm18. Seventeen of the 18 significant SNPs were located in a major QTL $qST-8$ , which was identified by linkage mapping in recombinant inbred lines (RILs) (Zhang et al. 2019). Though GWAS studies for salinity stress are relatively few in soybean, besides confirming major genetic determined by linkage mapping, GWAS has provided information of tolerant loci from new germplasm sources, which are quite useful in QTL pyramiding (Table 1.2).

## **1.7 Molecular Mapping of Tolerance Genes and QTLs**

### ***1.7.1 Brief History of Molecular Mapping in Soybean***

The first report of utilization of molecular markers in soybean is use of restriction fragment length polymorphism (RFLP) for the assessment of genetic diversity of the soybean nuclear genome (Apuya et al. 1988). Subsequently, RFLP markers were used extensively for genetic diversity analysis (Kiem et al. 1989; Skorupska et al. 1993; Lorenzen et al. 1995) and linkage mapping (Kiem et al. 1990; Diers et al. 1992; Lark et al. 1993; Akkaya et al. 1995; Shoemaker and Specht 1995; Mansur et al. 1996; Kiem et al. 1997; Cregan et al. 1999; Ferreira et al. 2000; Yamanaka et al. 2001; Lightfoot et al. 2005) until SSR and SNP markers become popular (Hyten et al. 2010a), Lee et al. (2015), Sun et al. (2019a), Ratnaparkhe et al. (2020), Kumawat et al. 2021).

### ***1.7.2 Evolution of Marker Types and Genetic Diversity Studies***

Various marker-based technologies such as RFLPs, RAPDs, AFLPs, SSRs and SNPs were used for genetic mapping and diversity studies in soybean. Apuya et al. (1988) analyzed randomly chosen 300 RFLP probes in genomic DNA of the genetically distant cultivars Minosy and Noir 1. RAPDs were also used extensively by soybean geneticists, mainly for germplasm classification (Thompson et al. 1998; Brown–Guedira et al. 2000; Li and Nelson 2002). Construction of soybean linkage maps was done using SSR and AFLP markers (Morgante et al. 1994; Keim et al. 1997; Matthews et al. 2001). Interestingly, the first report of SSR allelic variation and their use as marker system in plant species was in soybean (Akkaya et al. 1992; Morgante and Oliveri 1993). SSR polymorphism showed high level of allelic variation in cultivated and wild soybean genotypes (Morgante et al. 1994; Maughan et al. 1995; Rongwen et al. 1995; Li et al. 2010). Akkaya et al. (1995) developed SSRs in soybean and integrated them into the linkage map. Subsequently, Cregan et al. (1999) mapped 606 SSR loci to develop an integrated soybean linkage map which was subsequently improved by addition of 420 SSRs (Song et al. 2004; Cregan et al. 1999). Hisano et al. (2007) used EST sequences to map a total 668 EST-derived SSR marker loci on soybean linkage map. Further, the availability of BAC-end sequence facilitated development of additional SSRs leading to integration of physical map and genetic map (Shultz et al. 2007; Shoemaker et al. 2006). Utilizing the whole genome sequence of soybean, a SSR database (BARCSOYSSR\_1.0) was developed by Song et al. (2010). This genome-wide SSR database provides informative SSRs at any genomic position required for fine mapping as well as for MAS. Choi et al. (2007) identified SNPs via the resequencing of sequence-tagged sites (STSs) developed from EST sequences. In the total 2.44 Mbp of aligned sequence, a total of 5,551 SNPs were discovered, including 4712 single-base changes and 839 InDels resulting in an average nucleotide diversity of  $\theta = 0.000997$ . Exploiting these SNPs, a total of 1,141 genes were placed on the genetic map by virtue of a SNP segregating among one or more RIL mapping populations, thus constructed a transcript map in soybean. Recent advances in whole genome sequencing and high throughput genotyping helped in the large scale genetic diversity studies of soybean germplasm collections.

### ***1.7.3 Mapping Populations***

Various mapping populations in soybean have been developed independently based upon the interests and needs of individual researchers, i.e., the degree of polymorphism required and specific agronomic traits for analysis.  $F_2$  populations or recombinant inbred lines (RILs) have been employed for the construction of linkage maps in soybean. While interspecific mapping populations have contributed enormously

to the saturation of the soybean linkage map, intraspecific linkage maps have also been developed. Recently, Nested association mapping (NAM) have been used for genetic mapping in soybean (Diers et al. 2018; Beche et al. 2020).

### 1.7.4 QTL Mapping Studies

Molecular markers especially DNA-based markers have been used extensively to identify the genomic locations of major QTLs governing different traits in soybean. RILs which are developed following several generations of selfing (typically up to F<sub>6</sub> or F<sub>7</sub>) are used in mapping QTLs. RILs are helpful in dissecting the QTLs and the estimate of influence of single or few QTL is possible depending on the population size. More than thousand QTLs governing over 100 agronomically and physiologically important traits have been characterized or mapped in soybean (Grant et al. 2010). Information pertaining to the QTLs mapped in soybean is available on database *SoyBase* (<http://soybase.org>). Recently, the advent of SNP-based genetic markers has facilitated the QTL analysis of many agronomic traits of soybean (<https://soybase.org>, <http://soykb.org>). The developments in the field of whole genome sequencing and the popularity of high throughput technologies have facilitated the genetic mapping in soybean in a great way yielding millions of SNP markers (Schmutz et al. 2010).

QTL mapping and molecular marker development have advanced in dissecting several agronomic traits and in studying the genetic basis of resistance against drought and water logging along with improved yield. In the pursuit to develop genotyping tools for investigating mapping population, Hyten et al. (2008) has developed a multiplex assay designated as soybean oligo pool all-1 (SoyOPA-1). This custom-made 384-SNP GoldenGate assay was developed utilizing SNPs discovered through resequencing of five diverse soybean accessions. Later, Hyten et al. (2010a) sequenced a total of 3,268 SNP-containing robust STS in six diverse genotypes, resulting in identification of 13,042 SNPs with an average of 3.5 SNP per polymorphic STS. These SNPs along with 5,551 SNPs discovered by Choi et al. (2007) were used to design two Illumina custom 1536 SNP GoldenGate assays designated as SoyOPA-2 and SoyOPA-3. A set of 1,536 SNPs (from the 3456 SNPs in three SoyOPAs) designated as Universal Soy Linkage Panel 1.0 (USLP1.0), ensured sufficient polymorphic markers at genome scale for use in QTL mapping applications. Hyten et al. (2010b) sequenced a reduced representation library of soybean to identify SNPs using high throughput sequencing methods. A total of 1,536 SNPs were selected to create an Illumina GoldenGate assay (SoyOPA-4). The SoyOPA-4 produced 1,254 successful GoldenGate assays suggesting an assay conversion rate of 81.6% for the predicted SNPs. Chaisan et al. (2010) used ESTs derived from 18 genotypes for EST clustering and SNP identification resulting in a total of 3,219 EST contigs and a total of 26,735 SNPs. The confirmation of in silico identified SNPs by Sanger sequencing yielded 15.7% accuracy rate between two cultivars Williams 82 and Harosoy. SNP markers in soybean which could be utilized for mapping of complex traits as well as

molecular breeding applications have been developed in recent investigations (Song et al. 2012; Li et al. 2019; Song et al. 2020).

### ***1.7.5 QTL Mapping Software***

QTL mapping in soybean has progressed swiftly in last three decades or so nonetheless, a large fraction of QTLs remains unutilized in breeding programs because of issues such as low accuracy and false-positives. However, the QTL accuracy could be improved by adopting various QTL mapping methods and effective statistical models such as single marker analysis (SMA), simple interval mapping (SIM), composite interval mapping (CIM), multiple interval mapping (MIM), and Bayesian interval mapping (BIM). Various QTL mapping softwares and QTL network have been developed to perform the task. “Meta-QTL analysis” compile QTL data from multiple reports onto a same map to ensure precise identification of QTL regions (Deshmukh et al. 2012; Sosnowski et al. 2012). Meta-QTL was effectively utilized by Hwang et al. (2015) to identify QTLs linked to Canopy wilting using 1 five different populations (RILs). Among the QTLs identified, one QTL on chromosome 8 in the 93,705 KS4895 × Jackson population co-segregated with already known QTL linked with wilting identified in a Kefeng1 × Nannong 1138–2 population. The advances in statistical approaches and software resulted in exponential increase in soybean genetic mapping studies to understand plants response to extreme climatic conditions for abiotic stress such as drought, water logging and high temperature stress.

## **1.8 Marker-Assisted Breeding for Resistance/Tolerance Traits**

Marker-assisted selection (MAS) is an indirect selection method where the linked molecular marker is utilized to transfer important agronomic traits from one genotype to another genotype. Marker-assisted backcrossing (MABC) is an important approach employed in soybean for transferring trait of interest. The high-throughput genotyping technologies have greatly assisted in the process of molecular marker identification and QTL mapping for different traits in soybean. The molecular breeding approaches such as Marker-assisted backcrossing and marker-assisted recurrent selection have aided in the introgression of the trait of economic or agronomic interest in soybean. In the past decades, several studies have focused on the genetic and molecular mechanisms of drought tolerance, flooding tolerance, salt tolerance where several QTLs have been identified to be associated with various abiotic stresses.

### ***1.8.1 QTL Mapping for Drought Tolerance***

Drought tolerance is a complex trait influenced by multiple genetic locations or governed by polygenes/QTLs, introgression of minor QTLs from donor to recipient cultivar is not an easy task. QTL mapping identified a total of 10 genomic regions associated with canopy wilting under drought stress (Table 1.3). Majority of these QTLs (9/10) have donor alleles conferring slow wilting traits from PI 416,937, Jackson, or both (Charlson et al. 2009; Abdel-Haleem et al. 2012; Hwang et al. 2015). Molecular markers associated with these QTLs could be explored for use in MAS to introgress the slow canopy wilting phenotypes from the donor into the elite backgrounds. However, transferring these QTLs is challenging task owing to the complex and, quantitative nature of the trait along with its sensitivity to prevailing environmental factors. Most minor QTLs were found to be unstable across the environments and populations. For instance, even major QTLs on chromosome 12 ( $R^2 = 0.27$ ) identified in all five environments from Benning  $\times$  PI 416,937 (Abdel-Haleem et al. 2012) was not detected in any populations or site-years (Hwang et al. 2015). Accordingly, it is mandatory for QTL confirmation in more advanced generations to validate each individual QTL. It also suggests that molecular stacking of all confirmed QTLs in the genetic background of an elite cultivar is imperative to develop drought tolerance in soybean (Valliyodan et al. 2016). Ren et al. (2020) identified 23 QTL linked to drought tolerance of which seven QTLs were identified on chromosomes 2, 6, 7, 17, and 19 while five QTL were found on chromosomes 2, 6, 13, 17, and 19 respectively.

### ***1.8.2 QTL Mapping for Root System Architecture and Canopy Characteristics***

Mapping of genomic regions controlling root system architecture (RSA) and canopy characteristics is critical to develop soybean that is cultivable in water-limited environment (Song et al. 2016a). In an interspecific RIL population derived from cross *G. max* (V71-370)  $\times$  *G. soja* (PI407162), four significant QTLs associated with different root architectural traits were identified on chromosome Gm06 and Gm 07 (Prince et al. 2015a). In another study, Manavalan et al. (2015) identified a major QTL on chromosome Gm08 controlling tap root length, lateral root number and shoot length. Six transcription factors and two key cell wall expansion-related genes were identified as candidate genes in the confidence interval of this QTL. Recently, Dhanpal et al. (2021) conducted first genome-wide association study reporting genetic loci for RSA traits for field-grown soybean and identified key candidate genes.

**Table 1.3** Overview of QTLs identified for abiotic stress (Drought, water logging) tolerance in soybean depicting parents, mapping population, associated trait, chromosome and markers and phenotypic variance explained

Parents involved	Population/Size	Traits related to drought tolerance	Linked marker(s)	Type of marker	LG/Ch No	QTLs/genes	PVE(%)	References
<i>Drought tolerance</i>								
Young × P1416937	F <sub>4</sub> (120)	Water use efficiency, leaf ash	cr497-I	RFLP	J	3 QTLs	13.2	Mian et al. (1996)
S100' × "Tokyo,"	F <sub>2</sub> (116)	Water use efficiency	A489H	RFLP	L	2 QTLs	14	Mian et al. (1998)
Minsoy' × Noir 1'	RIL (236)	Transpiration efficiency, yield and carbon isotope discrimination	Satt205, satt489	SSR	C2	3 QTLs	7–12.8	Specht et al. (2001)
Kefeng 1 × Nannong 1138-2	RIL	Dry root weight, total root length and root volume, canopy-wilting trait	–	–	N6-C2 N8-D1b + W N11-E, N18-K	13 QTLs	–	Liu et al. (2005)
Jackson × KS4895	RIL (81)	Leaf wilting	Sat_044	SSR	K	1 QTL	17	Bhatnagar et al. (2005)
Hutcheson × P1471938	F <sub>4</sub> (140)	Yield and wilting	Satt226, Sat_375, Sat_074	SSR	D2, F1, F2	6 QTLs	–	Monteros et al. (2006)
KS 4895 × Jackson	RIL (92)	Canopy wilting	Satt177, Satt362	SSR	A2, B2, D2, F	4 QTLs	47	Charlson et al. (2009)
Kefeng 1 × Nannong 1138-2	RIL (184)	Seed yield per plant, drought susceptibility index	–	–	C2, H	10 QTLs	–	Du et al. (2009a)
Kefeng 1 × Nannong 1138-2	RIL (184)	Leaf water and seed yield	–	–	C2, D1b, H and A2	17 QTLs	–	Du et al. (2009b)

(continued)

Table 1.3 (continued)

Parents involved	Population/Size	Traits related to drought tolerance	Linked marker(s)	Type of marker	LG/Ch No	QTLs/genes	PVE(%)	References
PI 416,937 × “Benning”	RIL (128)	Root traits	Satt383, Satt339, Sct_191, Satt429, Sat_299	SSR	1, 3, 4, 8, 20	5 QTLs	51	Abdel-Haleem et al. (2011)
Hongfeng 11 × Clark	BC	RWC, water holding capacity, root traits canopy wilting	Sat_136, Satt167, Satt398, Satt694, GMSL514	SSR	A1, A2, B1, B2, D1a, C2, E, F, K, G, I, L, M, N, O	40 QTLs	–	Li et al. (2011)
Hongfeng 11 × Clark	BC (46)	Germination rate and Seedling stage drought	Satt449, Satt499, Satt440, Sat_180	SSR	A1, K, I, H	31 QTLs	–	Qiu et al. (2011)
PI 416,937 × Benning	RIL (150)	Canopy wilting	Satt 302	SSR	2, 4, 5, 12, 14, 17, 19	7 QTLs	75	Abdel-Haleem et al. (2012)
Hongfeng 11 × Harosoy	BIL	–	Satt253, Satt513, Satt693, Satt240, Satt323, and Satt255	SSR	–	18 QTLs	–	Zhang et al. (2012)
SNWS 0048 × Jinda73	BIL	Physiological trait	–	–	D2, G, M, N	9 QTLs	–	Yang et al. (2014)
Jingdou23 × ZDD2315	RIL (447)	Root and shoot traits o seedlings	Satt333, Satt327, Satt519, Satt597	SSR	–	24 QTLs	7.05–38.91	Liang et al. (2014)

(continued)

Table 1.3 (continued)

Parents involved	Population/Size	Traits related to drought tolerance	Linked marker(s)	Type of marker	LG/Ch No	QTLs/genes	PVE(%)	References
93,705 KS4895 × Jackson, 08,705 KS4895 × Jackson, KS4895 × PI 424,140, A5959 × PI 416,937, Benning × PI 416,937	–	Canopy-wilting trait	–	–	2, 5, 11, 14, 17, 19	13 QTLs	–	Hwang et al. (2015)
Dunbar × PI 326582A	BIL (251)	Root and shoot traits	Sat315, Sat253, Sat142	SSR	8, 12	4 QTLs	7.7–20.8	Manavalan et al. (2015)
CPI 26,671 × G 2120; Valder × G 2120	RIL	Epidermal conductance, relative water content, plant survival after stress	soPt-856602	DART	6	106 QTLs; 34 QTLs	–	Vu et al. (2015)
PK 1180 × UPSL 298	F <sub>2</sub> /BSA	Seedling survivability	Satt277	SSR	6	One gene	–	Sreenivasa et al. (2020)
PI 416,997 × PI 567201D	RIL (196)	Carbon isotope discrimination	–	SNP	6, 7, 10, 11, 15, 17, 20	16 QTLs	2.5–29.9	Bazzer et al. (2020)
<i>Water logging tolerance</i>								
Archer × Minsoy and Archer × Minsoy	RILs (122), RILs (86)	Flooding tolerance in terms of yield and plant height	Sat_064	SSR	18	1 QTL	–	VanToai et al. (2001)

(continued)



Table 1.3 (continued)

Parents involved	Population/Size	Traits related to drought tolerance	Linked marker(s)	Type of marker	LG/Ch No	QTLs/genes	PVE(%)	References
A5403 × Archer and P9641 × Archer	RILs (103), RILs (67)	Water logging injury score	Satt269, Satt385	SSR	3, 5, 13	3 QTLs	10–16	Cornelious et al. (2005)
Misuzudaizu × Moshidou Gong 503	RILs (60)	Flooding tolerance index based on seed yield	Satt100, Sat_238, A520	SSR, RFLP, AFLP	2, 6, 7, 10, 11, 14, 19	7 QTLs	6.8–49.2	Githiri et al. (2006)
Kefeng No. 1 × Nannong	RILs (184)	Flooding tolerance in terms of plant survival rate	Satt531, Satt648, Satt038	SSR	1, 5, 18	3 QTLs	4.4–7.6	Wang et al. (2008b)
Peking × —	—	Flooding tolerance	Sat_175, Satt187, Satt338, Sat_279	SSR	2, 4, 8, 12	4 QTLs	—	Sayama et al. (2009)
Peking × Tamahomare	RILs (96)	Flooding tolerance in terms of Alcohol dehydrogenase content	Satt184, Sat_134, Sat_309	SSR	1, 8, 9, 13, 19	5 QTLs	14–35	Rizal and Karki (2011)
PI 408105A × S99-2281	RILs (188)	Flooding tolerance in terms of plant injury and yield index	BARC-016279-02, 316, BARC-024569-04, 982	SSR	11, 13	2 QTLs	7.7–19.7	Nguyen et al. (2012)
Tachinagaha × —	—	Flooding tolerance in terms of Root traits	Satt342, Satt469, Sat_177, Sat_342	SSR	4, 9, 11, 12, 13, 14	17 QTLs	1–23	Nguyen et al. (2017)

(continued)

**Table 1.3** (continued)

Parents involved	Population/Size	Traits related to drought tolerance	Linked marker(s)	Type of marker	LG/Chr No	QTLs/genes	PVE(%)	References
S99-2281 × PI 561,271	RILs (182)	Flooding injury score	Gm03_3087237_A/G, Gm10_43840376_T/C	SNP	3, 10	2 QTLs	8.5–33.1	Ye et al. (2018)
5002 T × 91,210–350 and RA-452 × Osage	RILs (111), RILs (79)	Flooding damage score, Plant survival and chlorophyll content	Gm11_27896148_A_G, Gm13_28041039_G_A	SNP	1, 3, 6, 7, 11, 13, 14, 17, 19, 20	29 QTLs	9–29	Hummer (2018)
Linhe × M8206 and Zhengyang × M8206	RILs (104), RILs (126)	Seed flooding tolerance in terms of seedling length	LDB_1_46012724, LDB_13_29032373_29059156, LDB_17_8879391_9078832	SNPLDB/SNP	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 19, 20	33 QTLs	0.5–4	Ali et al. (2020)
Paldalkong × NTS1116	RILs	Foliar damage at V <sub>1</sub> /V <sub>2</sub> stage	–	SNP	10, 12, 13 and other 6 chr	20 QTLs	5.8 – 33.3%	Dhungana et al. (2020)

### 1.8.3 QTL Mapping for Flooding Tolerance

Several studies have focused on understanding the genetic and molecular mechanisms of flooding tolerance in soybean identifying underlying major QTLs (<http://www.soybase.org>). VanToai et al. (2001) identified one QTL linked to molecular marker Sat\_064 located on chromosome 18 associated with flooding tolerance. However, Reyna et al. (2003) could not find this QTL (Sat\_064) associated to water logging tolerance in near-isogenic line (NIL) populations due to different genetic background or location/soil types of studies. Cornelious et al. (2005) reported five QTLs associated with flooding tolerance. The marker Satt485 on chromosome 3, marker Satt599 on chromosome 5, and three markers Satt160, Satt269, and Satt252 on chromosomes 13 were identified to be linked with the QTL. They are associated with flooding tolerance in two RIL populations (Table 1.3). Githiri et al. (2006) identified seven QTLs associated with yield under flooding stress resulting in a proposed QTL near Satt100. Wang et al. (2008b) mapped three QTLs, Satt531-A941V (chromosome 1), Satt648-K418\_2V (chromosome 5), and Satt038-Satt275 (chromosome 18) associated with soybean flooding tolerance. Sayama et al. (2009) detected four putative QTLs viz. Sft1, Sft2, Sft3, and Sft4 associated with flooding tolerance and were mapped on to the chromosomes 2, 4, 8, and 12, respectively. Two new QTLs associated with both flooding injury score and flooding yield index were mapped on chromosomes 11 and 13 (Nguyen et al. 2012). However, these QTLs were discovered using bi-parental population characterized with a restricted mapping resolution due to limited recombination events. Later several novel QTLs associated with root system architecture, water-logging tolerance and yield in soybean have been identified (Ye et al. 2018; Wu et al. 2017b; Wu et al. 2020; Sharmin et al. 2020).

### 1.8.4 QTL Mapping for Salt Tolerance

Dissecting the genetic mechanism of salt tolerance in various stages of crop growth critical for the breeding of salt-tolerant soybeans (Munns and Tester 2008; Deinlein et al. 2014). Genetic architecture of salt tolerance in soybean has been dissected in several studies through bi-parental mapping and genome-wide association studies. An overview of the salt tolerant QTLs identified in soybean through bi-parental mapping is given in Table 1.4. In an F<sub>2:5</sub> population derived from a cross of the salt-tolerant cultivar S-100 and salt-sensitive cultivar Tokyo, Lee et al. (2004) mapped a major locus on Gm03, explaining 29% and 35% of phenotypic variation in green house and field conditions, respectively. Chen et al. (2008) identified four QTLs for salt tolerance at the seedling stage on Gm03, Gm07, Gm09, and Gm18. Subsequently, several studies have confirmed the major locus on Gm03, in different genetic backgrounds using bi-parental mapping populations, including interspecific cross mapping population of *G. max* × *G. soja* (Hamwieh and Xu 2008; Hamwieh et al. 2011; Ha et al. 2013; Qi et al. 2014; Guan et al. 2014a; Zeng et al. 2017; Do et al.

**Table 1.4** Overview of QTLs identified for salt tolerance in soybean

Parents	Population type	Trait <sup>a</sup>	Chromosome	Nearest marker/ marker interval of QTL	LOD score	Phenotypic variance (%)	Positive allele parent	References
Kefeng No.1 × Nannong 1138-2*	RIL	ST-GR, ST-GP	Gm01	Satt468, Satt268	3.7, 4.12	13.12, 22.95	-	Zhang et al. (2019)
Kefeng No.1 × Nannong 1138-2*	RIL	ST-GP, ST-IR	Gm02	Satt611, Satt266	3.21, 3.23	9.23, 7.03	-	Zhang et al. (2019)
S-100* × Tokyo	F <sub>2:5</sub>	LSS	Gm03	Satt237-Sat_091	7.2-12.9	29-45	S-100	Lee et al. (2004)
Kefeng 1 × Nannong 1138-2*	RIL	STR	Gm03	Satt237-Satt312	-	7.8	Nannong 1138-2	Chen et al. (2008)
Jackson × JWS156-1*	F <sub>2</sub>	STR, SPAD	Gm03	Satt237-Satt255	43.4, 33.4	68.7, 49.6	JWS156-1	Hamwiah and Xu (2008)
FT-Abyara* × C01	RIL	STR, SPAD	Gm03	Sat_091	12.1, 10	44, 38.2	FT-Abyara	Hamwiah et al. (2011)
Jindou No. 6* × 0197	RIL	STR, SPAD	Gm03	Sat_091	11.2, 8.6	47.1, 40.9	Jindou No. 6	Hamwiah et al. (2011)
PI483463* × Hutcheson	RIL	LSS	Gm03	Satt255-BARC-038333-10,036	18.8	56.5	PI 483,463	Ha et al. (2013)

(continued)

Table 1.4 (continued)

Parents	Population type	Trait <sup>a</sup>	Chromosome	Nearest marker/ marker interval of QTL	LOD score	Phenotypic variance (%)	Positive allele parent	References
85-140 × Tiefeng 8*	F <sub>2:3</sub> and RIL	STR	Gm03	QS08064-BARCSOYSSR_03_1301	-	-	Tiefeng 8	Guan et al. (2014a)
W05* × C08	RIL	Salt damage index	Gm03	40,204,091-41,182,426	4.3	54.6	W05	Qi et al. (2014)
Williams 82 × Fiskeby III*	F2	LSS, CCR, LSC, LCC	Gm03	Salt-20-Gm03_41135466	19, 11, 7.6, 25.5	48.2, 31.3, 20.6, 58.9	Williams 82, Fiskeby III, Williams 82, Fiskeby III	Do et al. (2018)
RA-452 × Osage*	RIL	LCC	Gm03	Gm03_39491355-Gm03_41605831	13	36-54	Osage	Zeng et al. (2017)
Jidou 12* × Ji NF 58	RIL	STR, SPAD	Gm03	GMABAB-Barcsoyssr_03_1421	19.8-20.1, 10.6-11	44.3-44.7, 27-27.6	Jidou 6	Shi et al. (2018)
Kefeng 1 × Nannong 1138-2*	RIL	STR	Gm07	Satt655-Satt210	-	19.7	Nannong 1138-2	Chen et al. (2008)
Kefeng No.1 × Nannong 1138-2*	RIL	ST-IR, ST-GP, ST-GR	Gm07	Satt245, Satt245, Satt590	3.01, 3.04, 3.15	10.69, 9.08, 16.5	-	Zhang et al. (2019)

(continued)

Table 1.4 (continued)

Parents	Population type	Trait <sup>a</sup>	Chromosome	Nearest marker/ marker interval of QTL	LOD score	Phenotypic variance (%)	Positive allele parent	References
Kefeng No.1 × Nannong 1138-2*	RIL	ST-IR, ST-GI, ST-GP, ST-GR	Gm08	Satt162, Satt162, Satt162 and Satt310, Satt162 and BES20148	3.14-17.06, 3.34-3.72, 2.84-7.63, 8.41-8.85	6.73-46.8, 7.11-8.46, 11.2-19.8, 18-19.8	-	Zhang et al. (2019)
Kefeng 1 × Nannong 1138-2*	RIL	STR	Gm09	Sat_363-S att710	-	19.2	Nannong 1138-2	Chen et al. (2008)
Kefeng No.1 × Nannong 1138-2*	RIL	ST-GP, ST-GR	Gm10	Satt331, Sat_274	4.44, 4.8	33.4, 30.6	-	Zhang et al. (2019)
Williams 82 × Fiskeby III*	F <sub>2</sub>	LSC	Gm13	Gm13_27665585 -Gm13_28206014	4.56	11.5	Fiskeby III	Do et al. (2018)
RA-452 × Osage*	RIL	LCC	Gm13	Gm03_39491355 - Gm03_41605831	3.2	6	RA-452	Zeng et al. (2017)
RA-452 × Osage*	RIL	LCC	Gm15	Gm15_6646246-Gm15_7147226	2.4	9	Osage	Zeng et al. (2017)
Kefeng No.1 × Nannong 1138-2*	RIL	ST-GI	Gm15	Satt606	2.64	6.02	-	Zhang et al. (2019)
Kefeng No.1 × Nannong 1138-2*	RIL	ST-GI, ST-GP, ST-GR	Gm17	Satt669, Satt669 and Sat_222, Satt669	2.56, 2.85-3.04, 3.93	10.94, 6.08-16.9, 13.73	-	Zhang et al. (2019)

(continued)

Table 1.4 (continued)

Parents	Population type	Trait <sup>a</sup>	Chromosome	Nearest marker/ QTL	marker interval of	LOD score	Phenotypic variance (%)	Positive allele parent	References
Jackson × JWS156-1*	RIL	STR SPAD	Gm17	Satt447		14.5 11.8	50.2 43.0	JWS156-1	Tuyen et al. (2010)
Kefeng 1 × Nannong 1138-2*	RIL	STR	Gm18	Sat_164-Sat_358		–	17.9	Nannong 1138-2	Chen et al. (2008)
Kefeng No.1 × Nannong 1138-2*	RIL	ST-GR	Gm18	Sat_358		2.5	6.74	–	Zhang et al. (2019)

\* Salt tolerant parent. <sup>a</sup> STR–Salt tolerance rating, PPS–Percentage plant survival, SPAD–Leaf chlorophyll content (SPAD value), LSS–Leaf scorching score, CCR–Chlorophyll content ratio, LSC–Leaf sodium content, LCC–Leaf chloride content, ST-IR: the ratio of the imbibition rate under salt conditions to the imbibition rate under no-salt conditions, ST-GI: the ratio of the germination index under salt conditions to the germination index under no-salt conditions, ST-GP: the ratio of the germination potential under salt conditions to the germination potential under no-salt conditions, ST-GR: the ratio of the germination rate under salt conditions to the germination rate under no-salt conditions

2018; Shi et al. 2018). Zeng et al. (2017b) identified two new QTLs for leaf chloride content on Gm13 and Gm15, using KCl and NaCl treatments. Do et al. (2018) identified a QTL for salt tolerance on Gm13, linked with leaf sodium content.

To identify salt tolerance QTLs at the germination stage, Zhang et al. (2019) used a RIL population and mapped 25 QTLs associated with four different salt tolerance indices during the soybean germination stage. Out of 25 QTLs identified for four salt tolerance indices at seedling stage, nine QTLs were located in an overlapping region on Gm08 (named *qST-8*, Zhang et al. 2019). A wild soybean (*Glycine soja*) accession JWS156-1 with high saline and alkaline salt tolerance was identified, and a significant QTL for alkaline salt tolerance was detected on Gm17 (Tuyen et al. 2010). The QTL for alkaline salt tolerance was different from the QTL for saline tolerance found on Gm03, previously in this genotype. This study demonstrated that saline and sodic stress tolerances are controlled by different genes in soybean. DNA markers associated with these QTLs can be used for marker-assisted pyramiding of tolerance genes in soybean for both saline and sodic stresses. Bi-parental linkage mapping has successfully mapped two major locus and several minor loci for salt tolerance, however bi-parental linkage mapping can detect alleles from parents only (Korte and Farlow 2013). Nevertheless, salt-tolerant loci identified by linkage mapping are highly useful for marker-assisted selection and gene cloning.

DNA markers tightly linked with the salt tolerance QTLs and the genes characterized can be used in the selection of salt-tolerant lines. The major QTLs identified on Gm03 and Gm08, are stable QTLs identified in several studies, therefore, highly useful for MAS. Marker-assisted pyramiding of the identified major and minor QTLs may provide higher salt tolerance than single QTL. Marker-assisted development of NILs for major QTL on Gm03, and their evaluation showed higher salt tolerance (Guan et al. 2014b; Do et al. 2016), and higher grain yield in saline field conditions (Do et al. 2016; Liu et al. 2016). The salinity tolerance of tolerant NILs, NIL-T, was associated with the maintenance of seed size under salt stress and could be attributed to the ability to regulate  $\text{Na}^+$  and  $\text{Cl}^-$  in both vegetative and reproductive tissues (Liu et al. 2016). Haplotype-based markers for the identified salt-tolerant QTLs were successfully developed and utilized for new tolerant germplasm identification (Patil et al. 2016; Kumawat et al. 2020).

## 1.9 Map-Based Cloning of Tolerance Genes

### 1.9.1 Strategies: Landing and Walking

Availability of genomic clone libraries with large DNA inserts is one of the essential requirements for plant genome analysis, primarily for physical mapping, gene isolation, and gene structure and function analysis. The Bacterial Artificial chromosome (BAC) vectors have been used widely for generating genomic DNA libraries in economically important crop plants including soybean. Development of BAC



libraries is considered as critical step towards physical mapping and positional cloning of important genes.

### **1.9.2 Libraries: BAC/YAC Libraries**

Several BAC libraries have been developed from different soybean genotypes and wild species. These soybean BAC libraries have been developed with different objectives including general genomic research as well as specifically for cloning of biotic and abiotic stress tolerance loci. These libraries have provided a good resource for positional cloning of agronomical and biologically important QTL genes that the representative genotype possesses. BAC libraries have also been constructed for several wild species including *G. soja*, *G. syndetika*, *G. canescens*, *G. stenophita*, *G. cyrtoloba*, *G. tomentella*, *G. falcata*, and the polyploid, *G. dolichocarpa*. All BAC libraries are publicly available to soybean researchers. The physical map generation of soybean was initiated with the development of early genetic maps characterized by the even distribution on the whole genome of the crop. Yeast artificial chromosomes (YACs) were initially developed with a view to utilize the resource for chromosome walking and in situ hybridization (Zhuet al. 1996). BAC libraries covering the whole soybean genome were generated by early genomic researchers (Marek and Shoemaker 1997; Danesh et al. 1998; Tomkins et al. 1999; Salimath and Bhattacharyya 1999; Meksem et al. 2000). BAC libraries encompassing variety of genotypes have led to the development of early physical contigs (Marek and Shoemaker 1997). Efforts were made to develop physical map of soybean genome using BAC and BIBAC based libraries (Wu et al. 2004). A physical map of soybean cultivar Williams 82 was in place that was generated from 67,968 BAC clones from a *Bst*YI library and 40,320 clones from a *Hind*III library (<http://soybeanphysicalmap.org/>). Furthermore, SSR markers derived from BAC ends sequence (BES) were mapped and integrated into the physical map to improve its quality (Shoemaker et al. 2008). Six-dimensional BAC clones pools were employed to demonstrate the anchoring of genetic markers to the soybean BAC clones (Wu et al. 2008). On the parallel lines soybean unigene sets from NCBI were computationally anchored to Williams 82 BES resulting in anchoring of additional 305 contigs thereby complementing 1,184 anchored contigs achieved through 6-D pool screening efforts (Wu et al. 2008). Thus, the physical framework was accomplished by associating the contigs to the molecular markers which in turn was achieved by finger printing of the BAC clones through overgo hybridization, RFLP hybridization and SSR amplification (Song et al. 2004). The soybean physical map was updated and available at Soybean Breeders Toolbox (SBT) in soybase website (<http://www.soybase.org>) for the greater benefit of research community. Later, physical maps of soybean and related wild species were used for comparative and functional genomics studies (Innes et al. 2008; Ha et al. 2012; Ashfield et al. 2012).

## 1.10 Genomics-Aided Breeding for Tolerance Traits

### 1.10.1 *Details of Genome Sequencing*

Soybean genome sequencing project was accomplished by US Department of Energy-Joint Genome Initiative (DOE-JGI)-Community Sequencing Program (CSP) (Schmutz et al. 2010). Peptides from other flowering plants, TIGR legume EST data base were used and aligned with soybean genome data to obtain the gene rich regions. The resultant regions were fed in to the gene prediction algorithms to find putative genic regions. The homologous regions were integrated with EST sequences using PASA program (Haas et al. 2003). The genome sequence data and gene annotation of soybean is housed in Phytosome database (Goodstein et al. 2012) (<http://www.phytosome.net/>). It provides access to genes and gene families either by keyword-based search or sequence similarity-based programs like BLAST and BLAT (BLAST like Alignment Tool). The sequence analysis via shared functional domain or consensus sequence similarity enables the study on the evolutionary history of each gene family and identification of the closely linked gene families. Gbrowser in the database facilitates EST alignments, utility of VISTA tracks that helps in assessing the extent of nucleotide conservation in related plant genera. The Biomart- open-source data retrieval software allows the research community to download complete data from phytosome.

### 1.10.2 *Application of Structural and Functional Genomics in Genomics-Assisted Breeding*

New sequencing technologies have the potential to rapidly change the molecular research landscape in soybean (Lam et al. 2010; Libault et al. 2010; Li et al. 2013; Chung et al. 2014). Several research projects include genome re-sequencing, gene expression, and whole transcript profiling have provided large scale datasets for comparative and functional genomics studies (Valliyodan et al. 2016, 2019; Kim et al. 2019; Kajiya-Kanegae et al. 2021). Structural variations play important roles in driving genome evolution and gene structure variation which in turn contribute to agronomic trait variations. Liu et al. (2020) selected 26 accessions and performed de novo genome assembly for soybean accession. Through a comparative genome analysis, a total of 14,604,953 SNPs and 12,716,823 Indels, 27,531 copy number variations and 723,862 present and absent variations, were identified.

In addition to structural variations, gene expression studies are imperative constituent of any crop improvement program. Expression studies on single and global gene expression pattern analysis is an integral part of any crop improvement program. The gene expression patterns are investigated using the global expression analysis techniques like high-density expression arrays, Serial Analysis of Gene Expression and other functional genomics approaches. Usage of microarray on

soybean gene expression studies were conducted for functional studies of key genes (Maguire et al. 2002; Thibaud-Nissen et al. 2003; Vodkin et al. 2004). Functional genomics studies were also conducted to identify the role of microRNAs. MicroRNAs (miRNAs) are key regulators of gene expression and play important roles in many aspects of plant biology. Turner et al. (2012), identified number of novel miRNAs and previously unknown family members for conserved miRNAs in the recently released soybean genome sequence. They classified all known soybean miRNAs based on their phylogenetic conservation (conserved, legume- and soybean-specific miRNAs) and examined their genome organization, family characteristics and target diversity. Comparative and functional genomics have been applied extensively in soybean for identification of genes associated with key agronomic and physiological traits and for understanding the genome structure (Ma et al. 2010; Livingstone et al. 2010; Kim et al. 2010; Deshmukh et al. 2014; Ratnaparkhe et al. 2013; Valliyodan et al. 2016; Li et al. 2017; Zhou et al. 2019; Kim et al. 2019; Lin et al. 2019; Ferreira-Neto et al. 2019; Schmutz et al. 2019; Chaudhary et al. 2019; Paganon et al. 2020; Liu et al. 2020a; Valliyodan et al. 2021).

### ***1.10.3 Transcriptomic Approaches to Develop Drought Tolerance***

Characterization of genetic elements defining the root traits and related transcriptional responses to drought tolerance has gained greater interests in soybean (Thao et al. 2013). Initial exploration of genetic tool box for drought tolerance in soybean showed strong upregulation of around 3000 root-derived genes and metabolite coumestrol (Tripathi et al. 2016). In another study, a complex response of root tissues subjected to drought tolerance was identified along with the involvement of multiple biochemical pathways (Stolf-Moreira et al. 2010). In addition, early transcriptional responses of soybean roots to drought stress have been investigated in detail by Neto et al. (2013). Further, molecular basis of canopy wilting tolerance was studied through whole transcriptome sequences of leaf tissues of contrasting soybean genotypes (Prince et al. 2015b). Among the various differentially expressed genes, gene encoding UDP glucuronosyl transferase was specific to the drought tolerant line PI 567690. Comparison of root transcriptome profiles of genotypes DT2008 and William 82 indicated that the drought tolerant ability of DT2008 roots could be ascribed to the expression of high number of genes of root origin during early dehydration than during the prolonged dehydration. Also, differential expression of genes involved in osmo-protectant biosynthesis, transcription factors among others conferred drought tolerance (Ha et al. 2015). Root-specific transcriptome changes were observed in soybean lines subjected to drought stress. It identified several transcription factors that were differentially regulated during drought stress paving way for development of transcription factor-cis element network (Song et al. 2016b).

To gain further molecular insights about the aquaporin family proteins (AQPs), the plant specific AQPs, 23 soybean tonoplast intrinsic proteins (TIPs) genes were analyzed (Song et al. 2016b). Analysis identified 81 SNPs and many InDels in coding regions of TIP genes and their functional validation have provided key information regarding the roles of AQPs in soybean under various abiotic stresses (Song et al. 2016b). Similarly, exploration of AQPs in *Glycine soja* yielded 62 *GsAQP* genes. Comparative expression and protein–protein interaction analysis of AQPs in cultivated and wild soybean have helped in identifying *GmTIP2;1* as a novel candidate gene, conferring salt and water stress tolerance (Zhang et al. 2017). The comprehensive list of investigations exploring the drought tolerance mechanism in soybean utilizing transcriptomic approaches are presented in Table 1.5.

### ***1.10.4 Applications of Structural and Functional Genomics***

Plants have evolved an integrated strategy including signal perception and transduction, regulation of gene expression and biochemical and physiological responses adapting to drought stress. An effective and direct strategy to endure drought stress is to reduce water loss through closing stomata. The stomatal aperture is modulated by multiple factors including environmental signals, biotic/abiotic stress, CO<sub>2</sub> concentration, light and plant hormones. Several hormones are involved in stomatal regulation, among which the stress hormone abscisic acid (ABA) plays the main role. During the signal transduction and adaptive response, the expressional changes of a large number of drought responsive genes occur. Chen et al. (2020) identified soybean drought-tolerant genotypes and new candidate genes for breeding. Total 422 SNPs and 302 genes were correlated with drought associated traits through GWAS studies. In addition, thirteen genes were identified which were associated with the node number of main stem trait. By qRT-PCR, the expression level of Glyma.03G018000 and Glyma.03G018900 in drought-tolerant varieties was significantly increased. This study provides important drought-tolerant genotypes, traits, SNPs and potential genes, possibly useful for soybean genetic breeding.

#### **1.10.4.1 Reverse Genetics Approaches**

Recent advances in gene isolation, plant transformation, and genetic engineering are being used extensively to alter metabolic pathways in plants by tailor made modifications to single or multiple genes. Many of these modifications are directed toward increasing the nutritional value of plant-derived foods and feeds. These methodologies are based on quickly growing information based on molecular findings, understanding, and predictions of metabolic fluxes and network pathways. The application of recombinant DNA and related techniques to plants opened up the potential to improve agronomic characters, drought tolerance, heat tolerance and salt stress resistance.

**Table 1.5** Differentially expressed genes related to abiotic stress tolerance in soybean

Abiotic stress	Genotype(s) investigated	Tissue	Major finding(s)	References
<i>Drought</i>				
Drought related trait	-	Root tips	Genes related to membrane transport, defence signaling and metabolism associated to roots	Haerizadeh et al. (2011)
Drought-late vegetative (V6) to reproductive stage (R2)	Williams 82	Leaf	Both conserved and non-conserved pathways might be involved in drought tolerance in soybean Expression of <i>GmNAC</i> and hormone-related genes were modulated during drought	Le et al. (2012)
Canopy wilting tolerance	Drought-susceptible (DS) Pana and drought-tolerant (DT) PI 567,690	Leaf	Down-regulated genes (protein binding, hydrolase activity, carbohydrate/lipid metabolism, xyloglucan endo-transglycosylases associated with cell-wall, apoplast, and chlorophyll <i>a/b</i> binding proteins). Upregulated genes (ion binding and transport, the oxido-reductive process), SNPs in aquaporins genes	Prince et al. (2015b)
Long term and short term Dehydration and drought	Williams 82 and DT2008	Root	Early dehydration and short-term dehydration caused expression of more number of genes in DT2008 than in model Williams 82	Ha et al. (2015)

(continued)

Table 1.5 (continued)

Abiotic stress	Genotype(s) investigated	Tissue	Major finding(s)	References
Water-deficit stress (Different intensities)	Williams 82	Root	Root-specific transcriptome changes. Genes associated with hormones, carbohydrate metabolism, cell wall synthesis. Transcription factors and Cis-elements network	Song et al. (2016)
Combined drought and heat stress	Drought-resistant cultivar 'Heinong44'	Leaf	Genes involved in defence, photosynthesis, metabolic pathways differentially regulated. Combined stresses did not cause additive genetic response	Wang et al. (2018b)
Drought stress	Cultivar Prima 2000	Crown nodules	Expression of C1 cysteine protease ( <i>Glyma.14G085800</i> ) and two vacuole processing enzymes (VPEs), [ <i>Glyma.17G230700</i> and <i>Glyma.05G055700</i> ], were highly up-regulated. C1 cysteine protease ( <i>Glyma.14G085800</i> )-indicator of drought-induced premature senescence of root nodules	Cilliers et al. (2018)
Drought	Williams 82	Leaf	Increase in proline (Pro) content and decreased malondialdehyde content	Shi et al. (2018)
Drought	Williams 82	Leaf	Physiological and transcriptomic responses of soybean leaves during Reproductive stage (R2) to drought	Xu et al. (2018)

(continued)

Table 1.5 (continued)

Abiotic stress	Genotype(s) investigated	Tissue	Major finding(s)	References
Dehydration stress at V2 developmental stage	Embrapa 48' (drought-tolerant) and 'BR 16' (drought-sensitive)	Leaf	Utilized both the High-Throughput Serial Analysis of Gene Expression and the RNA-Seq approaches to decipher the role of inositol and inositol phosphates and raffinose family oligosaccharides pathways in dehydration stress	Ferreira-Neto et al. (2019)
Drought	<i>Glycine soja</i> drought-tolerant (PI342618B/DTP and A214/DTL) and drought-sensitive (NN86-4/DSP and A195/DSL)	Root	547 differentially expressed genes showed consistently contrasting expression between the genotypes. Twenty-eight DEGs were found located on chromosomal region rich in drought stress associated QTLs and "Hot spot regions"	Aleem et al. (2020)
<i>Salinity</i>				
Alkalinity (NaHCO <sub>3</sub> -50 mmol/L)	<i>Glycine soja</i>	Roots	Affymetrix® Soybean GeneChip®-based expression profiling revealed upregulation of 3307 probes and 5720 down-regulation. Series of molecular events starting from signal transduction through changes in metabolism, defence, intracellular traffic leading to modulations in protein synthesis	Ge et al. (2010)

(continued)

**Table 1.5** (continued)

Abiotic stress	Genotype(s) investigated	Tissue	Major finding(s)	References
Salinity (75 mM NaCl (low stress), and 150 mM NaCl (high stress))	Lindou 10, Qihuang 34, Weidou 8	Roots	Identified key genes (Glyma06G01990, Glyma08G22730, Glyma019G05140, and Glyma06G20160) involved in salt stress tolerance in soybean	Li et al. (2020c)
Soybean under salt (120 mM NaCl), saline-alkali (70 mM NaCl and 50 mM NaHCO <sub>3</sub> ) and drought (2% PEG 8000)	Inbred line HJ-1	Leaves and roots	Calcium signaling and nucleic acid pathway genes are upregulated in all the three stresses indicating molecular cross-talk	Fan et al. (2013)
NaHCO <sub>3</sub> stress (90 mM) at two time points (12 and 24 h after treatment)	Wild soybean variety N24852 and Lee 68 (control)	Roots	Identified DEG related to TFs, and ion transport, KEGG pathway analysis showed upregulation of phenylpropanoid biosynthesis” and “phenylalanine metabolism” pathways in soybean in response to alkalinity	Zhang et al. (2016)
200 mM NaCl solution for 7 days	Williams 82	Leaf	Increase in proline (Pro) content and decreased malondialdehyde content	Shi et al. (2018)



## RNAi Technology

In functional genomics, RNA interference (RNAi) is a propitious gene regulatory approach that plays a substantial role in crop improvement by permitting down-regulation of gene expression by small molecules of interfering RNA without affecting the expression of other genes. The discovery and study of the RNA interference phenomenon, in which double-stranded RNAs (*dsRNA*) elicits degradation of a target mRNA containing homologous sequence, led to development of more effective *dsRNA*-mediated gene silencing methods. RNAi is a less complicated, quick and efficient method of silencing gene expression in a range of organism including prokaryotes and eukaryotes. The silencing of a gene is a result of degradation of RNA into short RNA fragments that binds to specific nuclease which activates ribonucleases to target homologous mRNA. Specific gene silencing has been shown to be related to two ancient processes, co-suppression in plants and quelling in fungi, and has also been associated with regulatory processes such as transposon silencing, antiviral defense mechanisms, gene regulation, and chromosomal modification (Agrawal et al. 2003). The insertion of a functional intron region in the nuclear genome as a spacer fragment additionally increases the efficiency of the gene silencing induction, due to generation of an intron spliced hairpin RNA (*ihpRNA*) (Wesley et al. 2003). In plants, biotic stress is caused by living organisms, especially, viruses, bacteria, fungi, insects, arachnids, nematodes, and weeds. These organisms account for about a 40% loss in the overall yield of six major food and cash crops. RNAi technology has opened up new prospects for crop protection against biotic stresses.

Plants in their natural field conditions constantly get exposed to various abiotic factors such as high salinity, variation in temperature, flood, drought, and heavy metals, which hinders proper growth and development in plants. These factors are also one of the major causes behind huge crop losses globally. The changing climatic conditions and rapidly expanding population demand creates an urgent need to develop more stress-tolerant cultivars. Hence, RNAi technology can be utilized to develop transgenic cultivars that can cope with different abiotic stresses. Functional genomics studies revealed that novel genetic determinants are involved in stress adaptation in plants, which can be used to attain stress tolerance.

Receptor for activated C-kinase 1 (RACK-1) is a highly conserved scaffold protein that plays a significant role in plant growth and development. Rice plants generated through transgenic method (RNAi technology- a reverse genetic approach) where RNAi mediated downregulation of RACK-1 gene carried out, has shown more tolerance to drought stress as compared to the non-transgenic rice plants (Li et al. 2009). Likewise, disruption of rice farnesyltransferase/squalene synthase (SQS) by maize squalene synthase via RNAi, resulted in enhanced drought tolerance at vegetative and reproductive stages (Manavalan et al. 2012).

Stress tolerance and development in plants are regulated by miRNA and negatively affect the expression of the post-transcriptional gene. Wang et al. (2011a) examined that miRNA are involved in the very early stage during seed germination and identified that miRNA-mediated regulation of gene expression is present in maize imbibed seed. Wang et al. (2011b), reported 32 known members of 10 miRNA families and 8

new miRNAs/new members of known miRNA families that were found to be responsive to drought stress by high-throughput sequencing of small RNAs from *Medicago truncatula*. These findings suggest the importance of miRNAs in the response of plants to abiotic stress in general and drought stress in particular.

*OsTZF1* gene is a member of the CCCH-type zinc finger gene family in rice (*Oryza sativa*). Conditions like drought, high-salt stress, and hydrogen peroxide can induce the expression of *OsTZF1*. Expression of *OsTZF1* gene was also induced by abscisic acid, salicylic acid, and methyl jasmonate. *OsTZF1* gene overexpressed transgenic plants showed enhanced tolerance to high salt and drought stresses; whereas transgenic rice plants in *OsTZF1* gene were silenced using RNAi technology has shown less tolerance. This suggests the role played by *OsTZF1* gene in abiotic stress tolerance (Jan et al. 2013). Dehydrin proteins play a significant role in protecting plants from osmotic damage. Various research results suggest that overexpression of dehydrin gene *WZY2* provides more tolerance to plant against osmotic stress. A study conducted by Yu et al. (2019) suggests that RNAi mediated silencing of *WZY2* gene in *Arabidopsis thaliana* makes plant intolerant to osmotic stress.

Several researchers have focused on functional genomics studies of drought responsive genes (Le et al. 2012; Barbosa et al. 2013; Hua et al. 2018; Wang et al. 2018a; Wei et al. 2019). Drought responsive genes consist of regulatory genes encoding plenty of transcription factors (TFs), effector genes encoding chaperones, enzymes and ion/water channels etc. Several groups of TFs, such as ABA-responsive element-binding (AREB), dehydration responsive element binding (DREB), MYB, bZIP, NAC, and WRKY, respond to drought stress and act in an ABA-dependent or ABA-independent manner. Transcription factors are being used to develop genetically modified plants more tolerant to abiotic stresses. DREB and AREB TFs were introduced in soybean showing improved drought tolerance, under controlled conditions. Soybean, transgenic lines containing *AtDREB1A*, showed higher survival rate after a severe water deficit and important physiological responses to water deprivation, such as higher stomatal conductance and the maintenance of photosynthesis and photosynthetic efficiency (Polizel et al. 2011; de Pavia Rolla et al. 2014). Higher survival rates of DREB plants are because of lower water use due to lower transpiration rates under well-watered conditions. In addition to physiological studies, molecular analysis revealed that drought-response genes were highly expressed in *DREB1A* plants subjected to severe water deficit (Polizel et al. 2011). Mizoi et al. (2012) identified *GmDREB2A*, and showed that its heterologous expression in *Arabidopsis* induced stress-inducible genes such as *RD29A*, *RD29B*, *HsfA3*, and *HSP70* and improved stress tolerance. These findings indicate that plants overexpressing *AtDREB2A* and *DREB2Alike* proteins have increased tolerance to abiotic stresses. In soybean, the overexpression of *AREB1* gene indicated drought tolerance and exhibiting no leaf damage. It showed better growth and physiological performance under water-deficit as compared to the wild type (Barbosa et al. 2013).

Other transcription factor, WRKY, plays important roles in response to various abiotic stresses (Zhou et al. 2008). Previous studies have proved that soybean

GmWRKY54 can improve stress tolerance in transgenic *Arabidopsis*. Soybean transgenic plants were generated and further investigated for biological mechanisms of GmWRKY54 in response to drought stress (Wei et al. 2019). This study demonstrated that expression of *GmWRKY54*, driven by either a constitutive promoter (*pCm*) or a drought-induced promoter (*RD29a*), confers drought tolerance. Recently, genes as candidate biomarkers have also been identified to screen for drought-tolerant genotypes (Hua et al. 2018). Using a GeneChip Soybean Genome Array, Hua et al. (2018) identified 697 differentially expressed genes. These genes are mainly involved in the metabolic and hormone signaling pathways. Ten DEGs were validated in a sample of 20 soybean cultivars varying in the level of drought tolerance. This research provided a new set of transcriptomic data and biomarkers for early diagnosis of drought damage and molecular breeding of drought tolerance in soybean.

Major advancement has also been made in the structural and functional genomics studies for salt tolerance (Roy et al. 2014; Wang et al. 2018b; Zhang et al. 2019; Li et al. 2020a, b). Several loci for salt tolerance have been mapped in soybean and among them candidate genes for two major loci have been cloned (Guan et al. 2014b; Qi et al. 2014; Do et al. 2016; Zhang et al. 2019). A major and consistent salt tolerance locus on Gm03, was fine mapped and candidate gene was cloned and characterized as a sodium transporter (Guan et al. 2014b; Qi et al. 2014; Do et al. 2016; Patil et al. 2016). Qi et al. (2014) fine mapped and identified the gene underlying this QTL in a salt tolerant wild soybean accession W05. The candidate gene named *GmCHX1*, is a counterpart of *Glyma03g32900* in Williams 82 and homolog of the  $\text{Na}^+/\text{H}^+$  antiporter gene family. Genomic sequence analysis of *GmCHX1* for W05 and Williams 82 revealed that Williams 82 had a ~3.8 Kb Ty1/copia retrotransposon inserted into exon 3, but not in its counterpart *Glysoja01g005509* in W05 (Qi et al. 2014). In another study, Guan et al. (2014b) resolved this QTL into a salt tolerant variety Tiefeng 8, identifying the same gene *Glyma03g32900* (named as *GmSALT3*) having similar insertion of a 3.78-kb copia retrotransposon in exon 3 of salt sensitive parent. Subsequently, Do et al. (2016) characterized this locus in salt tolerant cultivar FT-Abhayra and identified *Glyma03g32900* (named *Ncl*) as causal gene. Insertion of a ~3.8-kb Ty1/copia type retrotransposon was responsible for the loss of gene function and salt sensitivity. Association of *Glyma03g32900* functional alleles and salt tolerance was confirmed in near isogenic lines (Guan et al. 2014b; Do et al. 2016). genetic Overexpression of *Glyma03g32900* by genetic transformation in the sensitive genotype Kariyutaka showed improved salt tolerance (Do et al. 2016). Fine mapping of major locus for salt tolerance *qST-8* was conducted and a candidate gene *Glyma.08g102000* (named *GmCDF1*), belonging to the cation diffusion facilitator (CDF) family, was identified (Zhang et al. 2019). RNA interference mediated down-regulation of *GmCDF1* in soybean hairy roots resulted in tolerance to salt stress (Zhang et al. 2019).

## 1.11 Recent Concepts and Strategies Developed

Genomic-assisted breeding, genomic selection (GS), genome sequencing, marker-assisted selection (MAS), genetic engineering approaches, and genomics tools have been used to improve soybean yield and quality. Genomic selection is a simple, reliable, and powerful approach that enables the rapid selection of superior genotypes, bringing bigger benefits to the breeders. The marker-assisted selection also has an advantage in screening difficult traits and identification of recessive alleles. Recent advancement in genomic tools and next-generation sequencing techniques makes it easier to develop new varieties with the superior trait. Genomic approaches, along with bioinformatics tools, allow a gigantic leap forward in plant breeding. Genomic designing overcomes the limitations of traditional breeding methods and accelerated the development of climate-smart soybean crops. Developing abiotic stress-tolerant soybean varieties have become convenient with the availability of a complete genomic sequence of soybean. Recently, gene editing tools such as modified meganucleases, hybrid DNA/RNA oligonucleotides, zinc finger nucleases, TAL effector nucleases and modified CRISPR/Cas9 are used for developing abiotic stress tolerance (Bao et al. 2021). Each of these tools has the ability to precisely target one specific DNA sequence within a genome and to create a double-stranded DNA break. DNA repair to such breaks sometimes leads to gene knockouts or gene replacement by homologous recombination. Genome rearrangements are also possible to engineer. Creation and use of such genome rearrangements, gene knockouts and gene replacements by the soybean researchers is gaining significant momentum (Carrizo et al. 2021).

### 1.11.1 Genome Editing—A Magic Bullet

Genome editing is at the dawn of its golden age. It is described as the ability to modify and manipulate DNA sequences with higher precision in living cells (Segal and Meckler 2003). The ability to remove, insert or even edit DNA sequences easily and accurately has attracted the interest of the scientific community in a wide range of biotechnology areas, such as medicine, environmental studies and even agriculture. Targetable nucleases enable scientists to target and modify theoretically any gene in any organism. In the past few years, rapid development of molecular understanding with the aid of advanced computational technology and instrumentation with multiplexing and higher precision has led to the development of sequence specific DNA nucleases has progressed rapidly and such nucleases like zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems have been used in plant species such as Arabidopsis (Zhang et al. 2010; Li et al. 2013), tobacco (Nekrasov et al. 2013; Zhang et al. 2013), rice (Li et al. 2012;

Shan et al. 2013, 2014), barley (Wendt et al. 2013), soybean (Sun et al. 2015; Curtin et al. 2011), *Brachypodium* (Shan et al. 2013) and maize (Shukla et al. 2009).

All these nucleases involved in the genome editing technology are consist of DNA binding domains together with non-specific nuclease domains that generate double- strand breaks (DSBs). The DSBs are mainly repaired by non-homologous end-joining (NHEJ) or homologous recombination (HR) pathway (Chen and Gao 2013). NHEJ simply re-joins the broken DNA ends in an error-prone fashion and often results in small deletions or insertions. In the HR pathway, DSBs are correctly repaired using a homologous donor DNA as template. So far most genome editing has utilised the NHEJ pathway to knockout genes and only a few illustrations of gene insertion by HDR have been reported (Hyun 2020). The reasons may be that the mass of tissues to which DNA is delivered are often composed of determinate cells in which HDR is not the preferred repair mechanism.

#### 1.11.1.1 ZFNs (Zinc Finger Nucleases)

Zinc finger proteins were considered as the very first of the “genome editing” nucleases to hit the scene in the end of the twentieth century. The Zinc fingers, class of protein which is found the most commonly as a DNA binding protein domain in eukaryotes. Zinc finger nuclease (ZFN) is made up of two domains: DNA binding domain with repeated zinc fingers and FokI restriction enzyme-derived nuclease domain which is considered one of the most abundant DNA binding motifs in eukaryotic genome having the ability to recognize any sequence (Bitinaite et al. 1998). It is generally comprised of ~30 amino acid modules that interact with nucleotide triplets ie codons. ZFNs have been designed in such way that that it can recognize all of the 64 possible trinucleotide combinations, and by stringing different zinc finger moieties, one can create ZFNs that specifically recognize any specific sequence of DNA triplets (Segal et al. 2003). Each ZFN typically recognizes 3–6 nucleotide triplets, binds to the nuclease functions only as dimer, are required to target any specific locus. The first half part that recognizes the sequence upstream and the later one recognize the sequence downstream of the site to be modified (Szczepek et al. 2007).

#### 1.11.1.2 TALENs (Transcription Activator-Like Effector Nucleases)

Transcription activator-like effector nucleases (TALENs) have made a huge impact on the genomic engineering (Bedell et al. 2012). TALENs, like ZFNs contain the FokI nuclease fused to the DNA binding protein domain which can be exploited for targeted cleavage. This DNA binding domain known as Transcription activator-like effectors (TALE) derived from plant pathogenic *Xanthomonas* bacterium contains 33–35 amino acid repeat domains that recognizes a single base pair of the DNA (Joung and Sander 2013). Two hyper variable amino acids which are known as the repeat-variable di-residues (RVD) determine the TALE specificity found at positions

12 and 13. The TALE repeats use four RVD domains NN, NI, HD and NG which recognize guanine, adenine, cytosine and thymidine, respectively (Deng et al. 2012).

Although TALENs are effective tools for genome editing, there are some limitations regarding the potential target sites, such as the need for T at position 1 (Doyle et al. 2012) and the fact that some TALENs fail to cause mutations at the desired location despite of engineering nuclease and DNA binding domain. The latest developed genome editing technology- CRISPR/Cas system seems to provide a complementary approach to ZFNs and TALENs, as it only requires the PAM (NGG) motif preceding the recognition sequence.

### 1.11.1.3 CRISPR/Cas System

The research into the defence mechanisms of bacteria brought CRISPR to the scientific community. First discovered in 1987, the CRISPR-Cas system is an adaptive immunity prokaryotic defence system. As a result, it has been the focus of aggressive research that provided compelling insights into its function, as well as the promise of new molecular techniques (Ishino et al. 1987). CRISPR immunity has been categorized into three stages: adaptation, expression and interference. During the adaptation stage new spacer sequences are incorporated into the CRISPR locus. During the expression stage the CRISPR locus is transcribed to generate, or mature, the CRISPR RNA (crRNA). Finally, in the interference stage the invading nucleic acid is destroyed using the processed crRNA in some form of effector complex containing Cas proteins.

The most commonly used RGN in genome editing is the Cas9 nuclease from the type II CRISPR/Cas9 system of *Streptococcus pyogenes* (Jinek et al. 2012). With this system, there are two components that enable targeted DNA cleavage: a Cas9 protein and an RNA complex consisting of a CRISPR RNA (crRNA; contains 20 nucleotides of RNA that are homologous to the target site) and a transactivating CRISPR RNA (tracrRNA). For genome engineering purposes, the system can be reduced in complexity by fusing the crRNA and tracrRNA to generate a single-guide RNA (sgRNA) (Jinek et al. 2012). Also protospacer adjacent motif (PAM) sequences (5'- NGG-3'), an essential targeting component is situated upstream of the crRNA which is recognized by the cas9. The CRISPR/Cas systems can therefore cleave 23 bps target DNA sequence.

In contrast to ZFNs and TALENs, which require recoding of proteins using large DNA segments (500–1500 bp) for each new target site, CRISPR-Cas9 can be easily altered to target any genomic sequence by changing the 20-bp protospacer of the guide RNA, which can be accomplished by subcloning this nucleotide sequence into the guide RNA plasmid backbone. The Cas9 protein component remains unchanged. This ease of use for CRISPR-Cas9 is a significant advantage over ZFNs and TALENs, especially in generating a large set of vectors to target numerous sites (Mali et al. 2013). Another potential advantage of CRISPR-Cas9 is the ability to multiplex, i.e., to use multiple guide RNAs in parallel to target multiple sites simultaneously in the same cell (Cong et al. 2013 and Mali et al. 2013). With respect to site selection,

CRISPR-Cas9 compares favourably with ZFNs and TALENs. With the most flexible version of the *S. pyogenes* CRISPR-Cas system, site selection is limited to 23-bp sequences on either strand that end in an NGG motif (the PAM for *S. pyogenes* Cas9), which occurs on average once every 8 bp (Cong et al. 2013).

The targeted plant genome editing using sequence specific nucleases has a great potential for crop improvement to meet the increasing global food demands and to provide sustainable productive agriculture system. Immediately after its early use to edit the genomes of bacteria and animals (Hwang et al. 2013; Mali et al. 2013), its efficacy was validated in the model plant systems of *Arabidopsis*, rice and tobacco (Feng et al. 2013; Nekrasov et al. 2013; Xie and Yang, 2013).

## 1.12 Genetic Engineering for Tolerance Traits

Genetic modification of soybean utilizing various genes has resulted in the improved salt and drought tolerance traits (Table 1.6). Confirmation of drought tolerance in soybean was performed by ectopic expression of AtABF3 Gene (Kim et al. 2018). Several genes and TFs have been ectopically expressed in other model plants to study their functional significance. For example, over-expression of soybean-derived calmodulin gene (GmCaM4) in *Arabidopsis* enhanced tolerance to salinity owing to upregulation of AtMYB2-regulated genes, namely P5CS1 ( $\Delta^1$ -pyrroline-5-carboxylate synthetase-1) (Yoo et al. 2005). Similarly, soybean-derived S-phase kinase-associated protein 1 (SKP1) gene GmSK1 was over expressed in *Nicotiana tabacum* cv. Samsun showing improved tolerance to salinity and drought stress (Chen et al. 2018). Pitman and Läuchli (2002) suggested that genetic modification for enhanced salt tolerance is an important approach. In dry regions, irrigation of moderately salt tolerant crops with brackish water is feasible and will be helpful for increasing the crop production. Identification of orthologs and their functional analysis will provide opportunity to improve salt tolerance in soybean through genetic engineering. Based on the knowledge of monovalent cation/proton antiporter (CPA) family in *Arabidopsis*, several genes have been identified and functionally characterized for their involvement in salt tolerance in soybean. Jia et al. (2017) demonstrated that *GsCHX19.3*, a member of cation/H<sup>+</sup> exchanger super family from wild soybean provide tolerance to high salinity and carbonate alkaline stress. *GsCHX19.3* mediates K<sup>+</sup> uptake and Na<sup>+</sup> excretion under carbonate alkaline stress when over-expressed in *Arabidopsis*. Sun et al. (2019a) found that a Na<sup>+</sup>/H<sup>+</sup> exchanger, *GmNHX1*, was upregulated under salt stress in soybean genotype Jidou 7. Overexpression of *GmNHX1* in *Arabidopsis*, enhances salt tolerance by maintaining K<sup>+</sup>/Na<sup>+</sup> ratio in root (Sun et al. 2019b). Similarly, overexpression of transcription factor *GmNAC15*, a member of the NAC transcription factor family in soybean, enhances salt tolerance in soybean hairy roots (Ming et al. 2018).

Jia et al. (2020) characterized *GmCHX20a*, a paralog of salt tolerant gene *GmCHX1*, and found that the ectopic expression of *GmCHX20a* in soybean hairy roots and *Arabidopsis* led to an increase in salt sensitivity and osmotic tolerance. It

**Table 1.6** Genetic engineering studies related to abiotic stress tolerance

Sl. No.	Gene and organism	Recipient genome	Features	Reference
1	<i>Pumax ginseng</i> ( <i>PgTIP1</i> )	Hybrid soybean line 4076 (F <sub>5</sub> ) ( <i>G. max</i> N23674 × <i>G. soja</i> BB52)	Transgenic lines under salt-stress showed better leaf stomatal movement and water–gas exchange capacity; less Na <sup>+</sup> , Cl <sup>-</sup> ; a lower Na <sup>+</sup> /K <sup>+</sup> ratio; and enhanced enzymatic antioxidant activities. Under PEG induced stress, transgenics demonstrated greater leaf water-retention capacity and reduced cell membrane damage	An et al. (2018)
2	Soybean activated C kinase 1 ( <i>GmRACK1</i> )	Zhonghuang 13	Over expression and RNAi lines of <i>GmRACK1</i> were developed Silencing of <i>GmRACK1</i> gene improved drought tolerance during vegetative stage	Li et al. (2018)
3	<i>Arabidopsis thaliana</i> : ABRE-binding factors 3 <i>AtABF3</i>	Korean cultivar 'Kwangankong'	Enhanced drought tolerance due to prevention of cell membrane damage and maintenance of chlorophyll	Kim et al. (2018)
4	Glycine max <i>GmFDL19</i> ( <i>bZIP transcription factor</i> )	–	Drought and salt tolerance at seedling stage. Reduced accumulation of Na <sup>+</sup> ion and up-regulated the expression of ABA/stress-responsive genes	Li et al. (2017)
5	Soybean <i>BRASSINOSTEROID-INSENSITIVE 2</i> ( <i>GmBIN2</i> ) (member of the GSK3 protein kinase family)	Transgenic Soybean hairy root	Higher root growth rate when treated with salt and mannitol	Wang et al. (2018)
6	<i>Glycine soja</i> ( <i>GsWRKY20</i> )	Soybean	Reduced MDA, low relative permeability, and higher antioxidant enzymes	Ning et al. (2017)

(continued)



Table 1.6 (continued)

Sl. No.	Gene and organism	Recipient genome	Features	Reference
7	<i>Soybean NAC transcription factor GmNAC15</i>	Soybean hairy roots	Overexpression enhances salt tolerance in soybean hairy roots	Ming et al. (2018)
8	<i>Soybean Cation/H<sup>+</sup> + exchanger gene GmCHX20a</i>	Soybean hairy roots	Overexpression leads to salt sensitivity and osmotic tolerance	Jia et al. (2020)
9	<i>Soybean bZIP transcription factor GmbZIP15</i>	C03-3 ( <i>G. max</i> )	Overexpression resulted in hypersensitivity to abiotic stresses (Salt and drought)	Zhang et al. (2020)
10	<i>Soybean cation diffusion facilitator; GmCDF1</i>	Soybean hairy roots	Overexpression and RNA interference of <i>GmCDF1</i> in soybean hairy roots resulted in increased sensitivity and tolerance to salt stress, respectively	Zhang et al. (2019)
11	<i>Soybean NAC transcription factor GmNAC06</i>	Soybean hairy roots	Overexpression reduces ROS production and enhance salt tolerance	Li et al. (2020)
12	<i>Soybean Peroxidase Gene GsPRX9</i>	Soybean Hairy roots	Overexpression of the <i>GsPRX9</i> gene enhanced the salt tolerance and antioxidant response	Jin et al. (2019)

was suggested that *GmCHX20a* and *GmCHX1* together addresses both osmotic stress and ionic stress at different times of salinity stress exposure (Jia et al. 2020). Higher expression of *GmCHX20a* led to an increase in salt sensitivity and osmotic tolerance in early stage of salinity stress, whereas higher expression of *GmCHX1* protected plants via  $\text{Na}^+$  exclusion under salt stress in later stage. Jin et al. (2019) characterized *GsPRX9*, a class III peroxidase which upregulated significantly under salt stress. Overexpression of the *GsPRX9* in soybean hairy roots resulted in higher root fresh weight, primary root length, activities of peroxidase and superoxide dismutase, and glutathione level, but had shown lower  $\text{H}_2\text{O}_2$  content than those in control roots under salt stress. This suggests that the overexpression of the *GsPRX9* gene results in enhanced salt tolerance and activation of antioxidant response in soybean. These examples provide insight into the mechanism of salt tolerance in soybean and various genes playing important role in maintaining ion ratio and antioxidant properties in plant, which can be utilized for genetic engineering of salt tolerance in soybean. To improve salt tolerance through genetic engineering, the negative regulators of salt tolerance could be down-regulated by gene editing and positive regulators could be overexpressed through genetic transformation.

The availability of large number of salinity tolerant genotypes makes it possible to develop salt tolerant soybean cultivars. Further, genetic characterization for trait inheritance and QTL identification made it feasible to introgress single or multiple salinity stress tolerant QTLs in desirable genetic background through DNA marker-assisted backcrossing and marker assisted recurrent selection (Lee et al. 2009). Identification of progeny lines which have shown higher tolerance than tolerant parental genotypes in some of the studies indicated that when positive alleles from tolerant and susceptible parents come together, higher tolerance is achievable (Hamwiesh et al. 2011; Do et al. 2018). Therefore, identification of positive alleles from both types of parents is desirable for QTL pyramiding for higher salt tolerance. It is also possible to identify different positive loci from two different tolerant genotypes to increase the threshold of stress tolerance, and in such cases QTL mapping may be performed in populations derived from tolerant  $\times$  tolerant parents. Functional characterization of positive regulators of salinity stress tolerance like *GmCHX1*, *GmCHX19.3*, *GmNAC15* and *GmNHX1*, made it feasible to genetically engineer target soybean cultivars in a short period of time. However, identification of negative regulators of salinity tolerance indicates that target genetic background should be carefully characterized to overcome the negative interaction of these negative loci, when introgression or modification of positive genes and alleles is planned.

### 1.13 Prospectus and Limitations of Genomic Designing for Soybean

Genomic designing approaches have enabled the improvement of soybean at a faster pace than traditional approaches. Introgression of genes and QTLs become much easier with the genomics advances. Marker-based QTL mapping is a powerful method

to recognize regions of the genome that co-segregate with a given trait and mapping of QTL for abiotic stress tolerance can be utilized for the elevation of tolerance against drought (Carpentieri-Pipolo et al. 2012; Zhang et al. 2012), salt (Hamwiah et al. 2011; Ha et al. 2013; Tuyenet al. 2013), flood (Guzman et al. 2007; Li et al. 2008b), and heavy metal stress (Sharma et al. 2011) in soybean. QTL mapping is more efficient compared to traditional mapping approaches since it does not require large numbers of progenies and generations of segregation populations. Genome-wide association study is an excellent approach to explore the allelic diversity present in the natural accessions of soybean. Furthermore, GWAS mapping resolution is higher than QTL mapping resolution due to millions of crossing events accumulated in the germplasm in the course of evolution (Deshmukh et al. 2014). Genome-wide association study has a great advantage in the dissection of the complex genetic architecture (Korte and Farlow 2013). Genome-assisted breeding in soybean helps in selecting superior genotypes which in turn improve the quality and yield of soybean crops on a large scale.

Although genome designing approaches have many benefits and are less time-consuming, more reliable, and easier methods, it has some limitations also. For instance, the resolution of QTL mapping is not very high due to biased mapping of QTL. Also, this method is limited to map allelic diversity that tends to segregate in a biparental population (Borevitz and Nordborg 2003). From a single QTL mapping experiment, it is hard to isolate perfect candidate genes. Moreover, genes that are identified by QTL mapping experiments are limited to those that segregate in the considered cross (Brachi et al. 2010). Genome-wide association study can overcome these limitations of QTL mapping, although it has its limitations such as the risk of many false positives as a result of population structure, unpredictable power to detect QTL, and the background LD can confound the results. The main drawback of MAS is linkage drag which can be minimized by marker-assisted backcrossing (MAB) and GS limitation is high cost and low accuracy (Staub et al. 1996; Deshmukh et al. 2014). Genome editing and other genomic methods undoubtedly set a milestone that solves all new challenges in the stream of science, however, it has some major ethical issues and negative side effects. In the future, advancement in genomic designing tools and methodologies may overcome the above-mentioned limitations (Bao et al. 2021; Carrijo et al. 2021).

## 1.14 Bioinformatic Resources for Soybean Improvement

Bioinformatics plays an inevitable role in the modern genomics era. It is a science of collecting, storing, and developing algorithmic tools to analyze and understand complex biological data. There are several databases and bioinformatics tools available for various purposes.

### ***1.14.1 Gene and Genome Databases***

*Arabidopsis* was the first plant species and the third multicellular organism to be completely sequenced and published (Kaul et al. 2000). Later, with the advancement of next-generation sequencing, several plant genomes were sequenced, and most of them are available in public databases. Biological databases are stores of biological information, and are mainly of two types, primary and secondary database. In the primary database, the sequence information is stored, and the secondary database utilizes this information. The secondary database uses the genome sequence information and performs the downstream analysis like functional annotations. The most important databases where genome and gene sequences can be submitted and retrieved are NCBI, Phytozome, and Ensemble. SoyKB and SoyBase are secondary databases that are specific to soybean. Most of these databases were generated for easy retrieval of specific genomic sequences, annotated genes, and putative functions of the genes possess marker information, QTL, transcriptomic data and can perform other downstream analysis. These databases play an important role in the identification of homologous genes using the information of functionally characterized genes.

### ***1.14.2 Comparative Genome Databases***

Genome sequencing of a large number of plant species and whole-genome resequencing of different cultivars of a crop generates new scopes of comparative genomics. Several studies have been published for comprehensive gene family analysis and duplication among the plant species. These types of studies are very important for the evolutionary fingerprinting of plant species. On the other hand, whole genome resequencing helps to explore genomic variants within a species. The comparative genomic variants would help in the dissection of biochemical pathways. The variant information of around 20,000 soybean accessions is available at SoyBase (Grant et al. 2010; Brown et al. 2020; <https://www.soybase.org/>) and SoyKB (Joshi et al. 2017; <http://soykb.org/>) database generated by SoySNP50K chip (Song et al. 2013). These single nucleotide variant data can be downloaded from SoyBase and SoyKB databases using Plant Introduction (PI) ID, genomic coordinate, and SNP ID. Further, variant information can be utilized for various studies like genome-wide association study, genomic selection, and superior haplotype identification. The comparative genomic analysis also provides evolutionary information, polyploidization, copy number variation, and presence-absent variations (PAV). Ha et al. (2019) developed a database Soybean-VCF2Genomes to identify the closest accession in soybean germplasm collection.

### ***1.14.3 Gene Expression Databases***

The transcriptional data provides the information about different gene interaction in diverse biological conditions, their role in biochemical pathways, and their function. The microarray and expressed sequence tags (EST) data information was dominant over a decade. Later, the advancement of NGS techniques replaced these conventional techniques. Next-generation sequencing is based on whole tissue mRNA sequencing and generate large amounts of sequencing data related to gene expression in various environmental conditions that can play important role in predicting gene function. There are several methods for gene expression analysis such as microarrays, Gene Chips, EST, serial analysis of gene expression (SAGE), massive parallel signature sequencing (MPSS), and RNAseq (Chaudhary et al. 2015). RNA-seq data related to various environment and stress conditions are available at different public sites like NCBI (<https://www.ncbi.nlm.nih.gov/sra/>), EMBL-ENA (<https://www.ebi.ac.uk/ena/browser/home>) and DDBJ (<https://www.ddbj.nig.ac.jp/index-e.html>). These databases provide the RNA-seq data of sequence read archive (SRA) raw files which can be analyzed using various publicly available RNA-seq pipelines. However, some databases like BAR (<http://bar.utoronto.ca/>), SoyKB, and SoyBase provides the publicly available analyzed data in the form of gene expression profile in different tissues and conditions. Several studies have been performed using publicly available RNAseq data and identified various key genes related to specific conditions (Machado et al. 2019). The biotic and abiotic stress related RNAseq data is also available in future, meta-transcriptomics analysis would result in the understanding of precise gene function, gene-environment interaction, and complex biological pathways.

### ***1.14.4 Protein or Metabolome Databases***

Proteins are the most important biomolecules as they directly control biological pathways and act as a functional unit. There are several hundred different proteins present in soybean seed but the major is glycinin (11S legumin type) and conglycinin (7S vicilin type), both comprise 65–80% of total protein content and 25–35% of seed content (Hammond et al. 2003). Soybean also has anti-nutrient content like kunitz trypsin inhibitors, lectin, P34 allergen, urease, and some other transporter protein, oil storage protein oleosins, sucrose binding and many others. Many studies have been conducted in soybean and different crops for the identification of protein expression in different tissue at various time intervals under stress conditions. The different techniques like 2D gel electrophoresis, HPLC, UPLC, LCMS, and GCMS have been used for the identification of proteins/metabolome in different environmental conditions. Several metabolites are available in Kyoto Encyclopedia of Genes and Genomes (KEGG: <https://www.genome.jp/kegg/>), Arabidopsis acyl-lipid metabolism (<http://aralip.plantbiology.msu.edu/pathways/pathways>), BRENDA (<https://www.brenda-enzymes.org/index.php>),

MassBank (<http://www.massbank.jp/>). Medicine plant (<http://medicinalplantgenomics.msu.edu/>), MetabolomeXchange (<http://www.metabolomexchange.org/site/>), Plant Metabolic Network (PMN: <https://plantcyc.org/>), Plant/Eukaryotic and Microbial Systems Resource (PMR: <http://metnetweb.gdcb.iastate.edu/PMR/>), PRIME ([http://prime.psc.riken.jp/?action=metabolites\\_index](http://prime.psc.riken.jp/?action=metabolites_index)), MetaboLights (<https://www.ebi.ac.uk/metabolights/index>). SoyMetDB is a metabolomic database for soybean and provide a one-stop web resource for integrating, mining and visualizing soybean metabolomic data, including identification and expression of various metabolites across different experiments and time courses (Joshi et al. 2017). These databases give the idea about metabolite biochemical and physiological properties.

### ***1.14.5 Integration of Data from Multiple Sources***

The advancement of different modern techniques in genomics, proteomics, ionomics, metabolomics, and phenomics develops a large amount of data that can be integrated to find precise identification of the target. There are several studies that successfully identified target by integrating two or more techniques. The genome-wide association studies (GWAS) along with transcriptomics data have been successfully explored for the identification of candidate genes governing particular traits. A computation approach, “camoco” has been developed which is the integration of GWAS and gene co-expression network (Schaefer et al. 2018). The integrated use of GWAS and RNAseq data identified 7 promising candidate genes for drought tolerance in maize, from the 62 loci identified in GWAS (Guo et al. 2020). Similar studies are also available in *Brassica* for yield (Lu et al. 2017) and in linseed for seed fatty acid metabolism (Xie et al. 2019). In recent study, integration of GWAS, digital phenotyping and transcriptomics was done for the identification of drought resistance genes in cotton (Li et al. 2020). Further, the integration of WGRS, transcriptome, and metabolite at different seed development stages have been utilised for the dissection of seed component related traits (Chaudhary et al. 2015). SoyBase provides the data of genetics, genomics, and USDA germplasm information. The loci information of nearly 100 traits for QTLs mapping and GWAS studies are available on SoyBase (Grant et al. 2010). The SoyKB is a web-based database that provides data of genomics, transcriptomics, metabolomics, and molecular breeding (Joshi et al. 2017). A recently developed SoyTD integrated database ([http://artemis.cyverse.org/soykb\\_dev/SoyTD/](http://artemis.cyverse.org/soykb_dev/SoyTD/)) of WGRS and transcriptomics gives the information of natural variations and expression of soybean transporter genes (Deshmukh et al. 2020). Lai et al. (2020) developed a comprehensive framework consisting of of bioinformatics big data mining, meta-analysis, and a gene prioritization algorithm. A total of 36,705 test genes set collected from multidimensional data platforms were analysed and candidate genes for flooding tolerance were identified. In the future, integration of more databases would help to accurately understand the complex biochemical pathways and identification of candidate genes for a specific trait.

## 1.15 Future Perspectives

In just the past few years we have witnessed tremendous progress in soybean comparative and functional genomics and an explosive expansion of new resources. We have seen large scale whole genome sequencing, development of high-density genetic maps using high through put approaches, construction of physical and transcript maps, development of high-density cDNA and oligo arrays, and advancement in functional genomics studies. These resources and the research outcome have shed much light on the structure, organization and evolution of the soybean genome and key genes associated with biotic, abiotic stresses and other traits. With the availability of the whole-genome sequence of the soybean genome, emerging functional genomic data and large-scale re-sequencing data, genome-wide comparisons are being achieved. These approaches will allow researchers to decipher the evolutionary history and genomic complexity of soybean. We will be able to further explore genomic approaches to the elucidation of key genes or functional components that control complex agronomical and physiological traits.

## References

- Abdurakhmonov IY, Abdugarimov A (2008) Application of association mapping to understanding the genetic diversity of plant germplasm resources. *Intl J Plant Genom* 2008: 574927
- Abdel-Haleem H, Lee GJ, Boerma RH (2011) Identification of QTL for increased fibrous roots in soybean. *Theor Appl Genet* 122(5):935–946
- Abdel-Haleem H, Carter TE Jr, Purcell LC, King CA, Ries LL, Chen P, Schapaugh W Jr, Sinclair TR, Boerma HR (2012) Mapping of quantitative trait loci for canopy-wilting trait in soybean (*Glycine max* L. Merr). *Theor Appl Genet* 125(5): 837–846
- Abel GH, Mackenzie AJ (1964) Salt tolerance of soybean varieties (*Glycine max* L. Merrill) during germination and later growth. *Crop Sci* 14:157–161
- Agrawal N, Dasaradhi PVN, Mohammed A, Malhotra P et al (2003) RNA interference: biology, mechanism, and applications. *Microbiol Mol Biol Rev* 67(4):657–685
- Ahmed F, Rafii MY, Ismail MR, Juraimi AS, Rahim HA, Asfaliza R, Latif MA (2013) Waterlogging tolerance of crops: breeding, mechanism of tolerance, molecular approaches, and future prospects. *Biomed Res Int* 2013:1–10
- Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132:1131–1139
- Akkaya MS, Shoemaker RC, Specht JE, Bhagwat AA, Cregan PB (1995) Integration of simple sequence repeat (SSR) DNA markers into a soybean linkage map. *Crop Sci* 35:1439–1445
- Aleem M, Raza MM, Haider MS, Atif RM, Ali Z, Bhat JA, Zhao T (2020) Comprehensive RNA-seq analysis revealed molecular pathways and genes associated with drought tolerance in Wild Soybean (*Glycine soja* Sieb. & Zucc.) *Physiol Planta* <https://doi.org/10.1111/ppl.13219>
- Ali MJ, Xing G, He J, Zhao T, Gai J (2020) Detecting the QTL-allele system controlling seed-flooding tolerance in a nested association mapping population of soybean. *Crop J* 8(5):781–792
- An J, Cheng C, Hu Z, Chen H, Cai W, Yu B (2018) The Panax ginseng PgTIP1 gene confers enhanced salt and drought tolerance to transgenic soybean plants by maintaining homeostasis of water, salt ions and ROS. *Environ Exp Bot* 155:45–55
- Anonymous (2015) Annual Report 2014–2015. Directorate of Soybean Research, Indore

- Apuya NR, Frazier BL, Keim P, Roth EJ, Lark KG (1988) Restriction fragment length polymorphisms as genetic markers in soybean, *Glycine max* (L.) Merrill. *Theor Appl Genet* 75:889–901
- Ara R, Mannan MA, Khaliq QA, Miah MU (2015) Waterlogging tolerance of soybean. *Bangladesh Agrono J* 18(2):105–109
- Ashfield T, Egan AN, Pfeil BE, Chen NW, Podicheti R, Ratnaparkhe MB, Ameline Torregrosa C, Denny R, Cannon S, Doyle JJ et al (2012) Evolution of a complex disease resistance gene cluster in diploid phaseolus and tetraploid glycine. *Plant Physiol* 159:336–354
- Bailey-Serres J, Voisenek LACJ (2008) Flooding stress: acclimations and genetic diversity. *Annu Rev Plant Biol* 59:313–339
- Bailey-Serres J, Lee SC, Brinton E (2012) Waterproofing crops: effective flooding survival strategies. *Plant Physiol* 160:1698–1709
- Bao A, Zhang C, Huang Y, Chen H, Zhou X, Cao D (2021) Genome editing technology and application in soybean improvement. *Oil Crop Sci* 5(1):31–40
- Barbosa EGG, Leita JP, Marin SRR et al (2013) Overexpression of the ABA-dependent AREB1 transcription factor from *Arabidopsis thaliana* improves soybean tolerance to water deficit. *Plant Mol Biol Rep* 31:719–730
- Bazzer SK, Kaler AS, Ray JD, Smith JR, Fritsch FB, Purcell LC (2020) Identification of quantitative trait loci for carbon isotope ratio ( $\delta^{13}C$ ) in a recombinant inbred population of soybean. *Theor Appl Genet* 15:1–5
- Beche E, Gillman JD, Song Q et al (2020) Nested association mapping of important agronomic traits in three interspecific soybean populations. *Theor Appl Genet* 133:1039–1054
- Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG et al (2012) In vivo genome editing using a high-efficiency TALEN system. *Nature* 491(7422):114–118
- Bhatia VS, Jumrani K, Pandey GP (2014) Evaluation of the usefulness of senescing agent potassium iodide as a screening tool for tolerance to terminal drought in soybean. *Plant Knowl J* 3(1):23–30
- Bhatia VS, Jumrani K (2016) A maximin-minimax approach for classifying soybean genotypes for drought tolerance based on yield potential and loss. *Plant Breed* 136:691–700
- Bhatnagar S, King CA, Purcell L, Ray JD (6–10 Nov 2005) Identification and mapping of quantitative trait loci associated with crop responses to water-deficit stress in soybean [*Glycine max* (L.) Merr.]. The ASACSSA-SSSA international annual meeting poster abstract. Salt lake city, UT, USA
- Bitinaite J, Wah DA, Aggarwal AK, Schildkraut I (1998) FokI dimerization is required for DNA cleavage. *Proc Natl Acad Sci USA* 95:10570–10575
- Blanco FF, Folegatti MV, Gheyi HR, Fernandes PD (2007) Emergence and growth of corn and soybean under saline stress. *Sci Agr* 64(5):451–459
- Blum A, Shpiler L, Golan G, Mayer J, Sinmena B (1991) Mass selection of wheat for grain filling without transient photosynthesis. *Euphytica* 54:111–116
- Blum A (2011) *Plant breeding for water-limited environments*. Springer. ISBN 978-1-4419-7490-7. <https://doi.org/10.1007/978-1-4419-7491-4>
- Boerma HR, Specht JE (2004) *Soybeans: improvement, production, and uses*, 3rd edn, Agron. Monogr. No. 16. American society of agronomy, Inc./Crop science society of America, Inc./Soil science society of America, Inc., Madison, Wisconsin
- Borevitz JO, Nordborg M (2003) The impact of genomics on the study of natural variation in *Arabidopsis*. *Plant Physiol* 32(2):718–725
- Boyer JS (1982) Plant productivity and environment. *Science* 218:443–448
- Boyer J (1983) Environmental stress and crop yields. In: Raper C, Kramer P (eds) *Crop reactions to water and temperature stress in humid, temperate climates*. West view Press, Boulder, CO, pp 3–7
- Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, et al (2010) Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genet* 6: e1000940
- Broue P, Douglass J, Grace JP, Marshall DR (1982) Interspecific hybridization of soybean and perennial glycine species indigenous to Australia via embryo culture. *Euphytica* 31:715–724



- Brown-Guedira GL, Thompson JA, Nelson RL, Warburton ML (2000) Evaluation of genetic diversity of soybean introductions and north American ancestors using RAPD and SSR markers. *Crop Sci* 40:815–823
- Brown AV, Conners SI, Huang W, Wilkey AP, Grant D, Weeks NT, Cannon SB, Graham MA, Nelson RT (2021) A new decade and new data at SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res* 8:49
- Bustingorri C, Lavado RS (2011) Soybean growth under stable versus peak salinity. *Sci Agri* 68:102–108
- Carpentieri-Pipolo V, Pipolo AE, Abdel-Haleem H et al (2012) Identification of QTLs associated with limited leaf hydraulic conductance in soybean. *Euphytica* 186:679–686
- Carrijo J, Illa-Berenguer E, LaFayette P et al (2021) Two efficient CRISPR/Cas9 systems for gene editing in soybean. *Transgenic Res* 3:239–249
- Carter TE Jr, Nelson, R L, Sneller CH, and Cui Z (2004) Genetic diversity in soybean In: Boerma HR, Specht JE (eds) *Soybeans: improvement, production, and uses*, 3rd edn, Agron Monogra vol 16. American Society of Agronomy, Inc/Soil Science Society of America, Inc, Madison, Wisconsin, pp 303–416
- Carthew RW, Sontheimer EJ (2009) Origins and mechanisms of miRNAs and siRNAs. *Cell* 136(4):642–655
- Chaisan T, Van K, Kim MY, Kim KD, Choi BS, Lee SH (2010) *In silico* single nucleotide polymorphism discovery and application to marker-assisted selection in soybean. *Mol Breed* 29:221–233
- Chandra S, Satpute GK, Nagar S, Singh M, Kumawat G, Rajesh V et al (2020) Reproductive stage water-logging tolerance: a critical assessment of traits in Soybean. *Soybean Research (In Press)*
- Chaudhary J, Patil GB, Sonah H, Deshmukh RK, Vuong TD, Valliyodan B, Nguyen HT (2015) Expanding omics resources for improvement of soybean seed composition traits. *Front Plant Sci* 6:1021
- Chaudhary J, Shivaraj S, Khatri P, Ye H, Zhou L, Klepadlo M, Dhakate P, Kumawat G, Patil G, Sonah H, Ratnaparkhe MB et al (2019) Approaches, applicability, and challenges for development of climate-smart soybean Genomic designing of climate-smart oilseed crops. Springer, Berlin, pp 1–74
- Charlson DV, Bhatnagar S, King CA, Ray JD, Sneller CH, Carter TE Jr, Purcell L (2009) Polygenic inheritance of canopy wilting in soybean [*Glycine max* (L.) Merr.]. *Theor Appl Genet* 119:587–594
- Chebrolu KK, Fritschi FB, Ye S, Krishnan HB, Smith JR, Gillman JD (2016) Impact of heat stress during seed development on soybean seed metabolome. *Metabolomics* 12:28
- Chen Y, Chen P, de los Reyes BG (2006) Differential responses of the cultivated and wild species of soybean to dehydration stress. *Crop Sci* 46:2041–2046
- Chen HT, Cui SY, Fu SX, Gai JY, Yu DY (2008a) Identification of quantitative trait loci associated with salt tolerance during seedling growth in soybean (*Glycine max* L.). *Aust J Agric Res* 59(12):1086–1091
- Chen L, Fang Y, Li X, Zeng K, Chen H, Zhang H, Yang H, Cao D, Hao Q, Yuan S, Zhang C (2020a) Identification of soybean drought-tolerant genotypes and loci correlated with agronomic traits contributes new candidate genes for breeding. *Plant Mol Biol* 102(1–2):109–122
- Chen P, Sneller CH, Purcell LC, Sinclair TR, King CA, Ishibashi T (2007) Registration of soybean germplasm lines R01–416F and R01–581F for improved yield and nitrogen fixation under drought stress. *J Plant Regist* 1:166–167
- Chen K, Gao C (2013) TALENs: customizable molecular DNA scissors for genome engineering of plants. *J of Genet and Genomi* 40(6):271–279
- Chen Y, Chi Y, Meng Q, Wang X, Yu D (2018) *GmSK1*, an *SKP1* homologue in soybean, is involved in the tolerance to salt and drought. *Plant Physiol Biochem* 127:25–31
- Chen HT, Cui SY, Fu SX, Gai JY, Yu DY (2008b) Identification of quantitative trait loci associated with salt tolerance during seedling growth in soybean (*Glycine max* L.). *Aust J Agr Res* 59:1086–1091

- Chen L, Fang Y, Li X, Zeng K, Chen H, Zhang H, Yang H, Cao D, Hao Q, Yuan S, Zhang C, Guo W, Chen S, Yang Z, Shan Z, Zhang X, Qiu D, Zhan Y, Zhou XA (2020b) Identification of soybean drought-tolerant genotypes and loci correlated with agronomic traits contributes new candidate genes for breeding. *Plant Mol Biol* 102(1–2):109–122
- Choi IY, Hyten DL, Lakshmi KM, Qijian S, Julian MC, Charles VQ, Kevin C, Lark KG, Robert SR, Mun SY, Eun YH, Seung IY, Nevin DY, Randy CS, Curtis PVT, James ES, Cregan PB (2007) A soybean transcript map: gene distribution, haplotype and SNP analysis. *Genetics* 176:685–696
- Choi M, Yun JY, Kim JH et al (2021) The efficacy of CRISPR-mediated cytosine base editing with the RPS5a promoter in *Arabidopsis thaliana*. *Sci Rep* 11:8087
- Chung WH, Jeong N, Kim J, Lee WK, Lee YG, Lee SH, Yoon W, Kim JH, Choi IY, Choi HK, Moon JK, Kim N, Jeong SC (2014) Population structure and domestication revealed by high-depth resequencing of Korean cultivated and wild soybean genomes. *DNA Res* 21(2):153–167
- Cilliers M, van Wyk SG, van Heerden PDR, Kunert KJ, Vorster BJ (2018) Identification and changes of the drought-induced cysteine protease transcriptome in soybean (*Glycine max*) root nodules. *Environ Exp Bot* 148:59–69
- Cornelius B, Chen P, Chen Y, De Leon N, Shannon JG, Wang D (2005) Identification of QTLs underlying water-logging tolerance in soybean. *Mol Breeding* 16(2):103–112
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121):819–823
- Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, Specht JE (1999) An integrated genetic linkage map of the soybean. *Crop Sci* 39:1464–1490
- Cui Z, Carter TE Jr, Burton JW (2000a) Genetic base of 651 Chinese soybean cultivars released during 1923 to 1995. *Crop Sci* 40:1470–1481
- Cui Z, Carter TE Jr, Burton JW (2000b) Genetic diversity patterns in Chinese soybean cultivars based on coefficient of parentage. *Crop Sci* 40:1780–1793
- Cui Z, Carter TE Jr, Burton JW, Wells R (2001) Phenotypic diversity of modern Chinese and North American soybean cultivars. *Crop Sci* 41:1954–1967
- Curtin SJ, Zhang F, Sander JD, Haun WJ, Starker C et al (2011) Targeted mutagenesis of duplicated genes in soybean with zinc-finger nucleases. *Plant Physiol* 156:466–473
- Danesh D, Penula S, Mudge J, Denny RL, Nordstrom H, Martinez JP, Young ND (1998) A bacterial artificial chromosome library for soybean and identification of clones near a major cyst nematode resistance gene. *Theor Appl Genet* 96:196–202
- de Paiva Rolla AA, Carvalho JdFC, Fuganti-Pagliarini R, Engels C, do Rio A, Marin SRR, et al (2014) Phenotyping soybean plants transformed with rd29A: AtDREB1A for drought tolerance in the greenhouse and field. *Transgenic Res* 23: 75–87
- Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK et al (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* 335:720–723
- Deshmukh R, Sonah H, Patil G, Chen W, Prince S, Mutava R et al (2014) Integrating omic approaches for abiotic stress tolerance in soybean. *Front Plant Sci* 5:244
- Deshmukh R, Rana N, Liu Y, Zeng S, Agarwal G, Sonah H, Varshney R, Joshi T, Patil GB, Nguyen HT (2020) Soybean transporter database (SoyTD): a comprehensive database for identification and exploration of natural variants in soybean transporter genes. *Plant Physiol* 171(4):756–770
- Deinlein U, Stephan AB, Horie T, Luo W, Xu G, Schroeder JI (2014) Plant salt-tolerance mechanisms. *Trends Plant Sci* 19(6):371–379
- Delgado MJ, Ligerio F, Lluch C (1994) Effects of salt stress on growth and nitrogen-fixation by pea, faba-bean, common bean and soybean plants. *Soil Biol Biochem* 26:371–376
- Dhanapal AP, Ray JD, Singh SK, Hoyos-Villegas V, Smith JR et al (2015) Genome-wide association study (GWAS) of carbon isotope ratio ( $\delta^{13}\text{C}$ ) in diverse soybean [*Glycine max* (L.) Merr.] genotypes. *Theor Appl Genet* 128:73–91
- Dhanapal AP, Ray JD, Singh SK, Hoyos-Villegas V, Smith JR, Purcell LC, Fritschi FB (2016) Genome-wide association mapping of soybean chlorophyll traits based on canopy spectral reflectance and leaf extracts. *BMC Plant Biol* 16(1):174

- Dhanapal AP, York LM, Hames KA, Fritschi FB (2021) Genome-wide association study of topsoil root system architecture in field-grown soybean [*Glycine max* (L.) Merr.]. *Front Plant Sci.* 10(11):590179
- Dhungana SK, Kim HS, Kang BK, Seo JH, Kim HT, Shin SO, Park CH, Kwak DY (2020) Quantitative trait loci mapping for flooding tolerance at an early growth stage of soybean recombinant inbred line population. *Plant Breed* 139:626–638
- Diers BW, Keim P, Fehr WR, Shoemaker RC (1992) RFLP analysis of soybean seed protein and oil content. *Theor Appl Genet* 83:608–612
- Diers BW, Specht J, Rainey KM, Cregan P, Song Q, Ramasubramanian V, Graef G, Nelson R, Schapaugh W, Wang D, Shannon G, McHale L, Kantartzi SK, Xavier A, Mian R, Stupar RM, Michno JM, An YC, Goettel W, Ward R, Fox C, Lipka AE, Hyten D, Cary T, Beavis WD (2018) Genetic architecture of soybean yield and agronomic traits. *G3: Genes Genomes Genetics* 3, 8(10):3367–3375
- Do TD, Chen H, Hien VT, Hamwieh A, Yamada T, Sato T, Yan Y, Cong H, Shono M, Suenaga K, Xu D (2016) Ncl synchronously regulates Na(+), K(+), and Cl(-) in soybean and greatly increases the grain yield in saline field conditions. *Sci Rep* 6:19147
- Do TD, Vuong TD, Dunn D, Smothers S, Patil G, Yungbluth DC, Chen P, Scaboo A, Xu D, Carter TE et al (2018) Mapping and confirmation of loci for salt tolerance in a novel soybean germplasm. Fiskeby III. *Theor Appl Genet* 131(3):513–524
- Do TD, Vuong TD, Dunn D et al (2019) Identification of new loci for salt tolerance in soybean by high-resolution genome-wide association mapping. *BMC Genomics* 20:318
- Dong YS, Zhao LM, Liu B, Wang ZW, Jin ZQ, Sun H (2004) The genetic diversity of cultivated soybean grown in China. *Theor Appl Genet* 108:931–936
- Donovan S, Mao Y, Orr D J, Carmo-Silva E, McCormick AJ (2020) CRISPR-Cas9-mediated mutagenesis of the rubisco small subunit family in nicotiana tabacum. *Front Genome Editing* 2. <https://doi.org/10.3389/fgeed.2020.605614>
- Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, Bogdanove AJ (2012) TAL effector-nucleotide targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic acids research* 40 (Web Server issue), W117–W122
- Doyle JJ, Doyle JL, Harbison C (2003) Chloroplast-expressed glutamine synthetase in *Glycine* and related Leguminosae: phylogeny, gene duplication, and ancient polyploidy. *Syst Bot* 28:567–577
- Du W, Wang M, Fu S, Yu D (2009a) Mapping QTLs for seed yield and drought susceptibility index in soybean (*Glycine max*) across different environments. *J Genet Genom* 36:721–731
- Du W, Yu D, Fu S (2009b) Detection of quantitative trait loci for yield and drought tolerance traits in soybean using a recombinant inbred line population. *J Integr Plant Biol* 51:868–878
- Elsheikh EAE, Wood M (1995) Nodulation and N-2 fixation by soybean inoculated with salt-tolerant Rhizobia or salt-sensitive *Bradyrhizobia* in saline soil. *Soil Biol Biochem* 27:657–661
- Fan XD, Wang JQ, Yang N, Dong YY, Liu L, Wang FW, Wang N, Chen H, Liu WC, Sun YP, Wu JY (2013) Gene expression profiling of soybean leaves and roots under salt, saline-alkali and drought stress by high-throughput Illumina sequencing. *Gene* 512(2):392–402
- Feng Z, Zhang B, Ding W, Liu X, Yang DL, Wei P, Cao F, Zhu S, Zhang F, Mao Y, Zhu JK (2013) Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res* 10:1229–1232
- Ferreira-Neto JRC, da Silva MD, Benko-Iseppon AM, Pandolfi V, Binneck E, Nepomuceno, AL, Abdelnoor RV and Kido EA (2019) Inositol phosphates and Raffinose family oligosaccharides pathways: Structural genomics and transcriptomics in soybean under root dehydration. *Plant Gene* 20:100202
- Ferreira AR, Foutz KR, Keim P (2000) Soybean genetic map of RAPD markers assigned to an existing scaffold RFLP map. *J Hered* 91:392–396
- Fried HG, Narayanan S, Fallen B (2019) Evaluation of soybean [*Glycine max* (L.) Merr.] genotypes for yield, water use efficiency, and root traits. *PLoS One* 14(2): e0212700
- Githiri S, Watanabe S, Harada K, Takahashi R (2006) QTL analysis of flooding tolerance in soybean at an early vegetative growth stage. *Plant Breed* 125:613–618

- Grant D, Nelson RT, Cannon SB, Shoemaker RC (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res* 38:D843–D846
- Ge Y, Li Y, Zhu YM et al (2010) Global transcriptome profiling of wild soybean (*Glycine soja*) roots under NaHCO<sub>3</sub> treatment. *BMC Plant Biol* 10:153
- Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res* 40(D1):D1178–D1186
- Guan R, Chen J, Jiang J, Qiu L (2014a) Mapping and validation of a dominant salt tolerance gene in the cultivated soybean (*Glycine max*) variety Tiefeng 8. *Crop J* 2:358–365
- Guan RX, Qu Y, Guo Y, Yu LL, Liu Y, Jiang JH, Chen JG, Ren YL, Liu GY, Tian L, Jin LG, Liu ZX, Hong HL, Chang RZ, Gilliam M, Qiu LJ (2014b) Salinity tolerance in soybean is modulated by natural variation in GmSALT3. *Plant J* 80:937–950
- Guo J, Li C, Zhang X, Li Y, Zhang D, Shi Y, Song Y, Li Y, Yang D, Wang T (2020) Transcriptome and GWAS analyses reveal candidate gene for seminal root length of maize seedlings under drought stress. *Plant Sci* 292:110380
- Guzman PS, Diers B, Neece D, St Martin S, Leroy A, Grau C et al (2007) QTL associated with yield in three backcross-derived populations of soybean. *Crop Sci* 47:111–122
- Ha B, Vuong TD, Velusamy V et al (2013) Genetic mapping of quantitative trait loci conditioning salt tolerance in wild soybean (*Glycine soja*) PI 483463. *Euphytica* 193:79–88
- Ha CV, Watanabe Y, Tran UT, Le DT, Tanaka M, Nguyen KH, Seki M, Nguyen DV, Tran LS (2015) Comparative analysis of root transcriptomes from two contrasting drought-responsive williams 82 and DT2008 soybean cultivars under normal and dehydration conditions. *Front Plant Sci* 6:551
- Ha J, Abernathy B, Nelson W, Grant D, Wu X, Nguyen HT, Stacey G, Yu Y, Wing RA, Shoemaker RC, Jackson SA (2012) Integration of the draft sequence and physical map as a framework for genomic research in soybean (*Glycine max* (L) Merr) and wild soybean (*Glycine soja* Sieb and Zucc). *G3: Genes Genomes Genetics* 2(3):321–329
- Ha J, Jeon HH, Woo DU, Lee Y, Park H, Lee J, Kang YJ (2019) Soybean-VCF2Genomes: a database to identify the closest accession in soybean germplasm collection (2019). *BMC Bioinformatics* 24 (20)(Suppl 13):384
- Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RK Jr, Hannick LI, Maiti R, Ronning CM, Rusch DB, Town CD et al (2003) Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res* 31:5654–5666
- Haerizadeh F, Singh MB, Bhalla PL (2011) Transcriptome profiling of soybean root tips. *Funct Plant Biol* 38(6):451–461
- Hao D, Chao M, Yin Z, Yu D (2012) Genome-wide association analysis detecting significant single nucleotide polymorphisms for chlorophyll and chlorophyll fluorescence parameters in soybean (*Glycine max*) landraces. *Euphytica* 186:919–931
- Hammond EG, Murphy PA and Johnson LA, (2003) SOY (SOYA) BEANS. Properties and Analysis 5389–5392. <https://doi.org/10.1016/B0-12-227055-X/01111-1>
- Hamwieh A, Tuyen DD, Cong H, Benitez ER, Takahashi R, Xu DH (2011) Identification and validation of a major QTL for salt tolerance in soybean. *Euphytica* 179:451–459
- Hamwieh A, Xu DH (2008) Conserved salt tolerance quantitative trait locus (QTL) in wild and cultivated soybeans. *Breed Sci* 58:355–359
- Harlan JR, de Wet JMJ (1971) Toward a rational classification of cultivated plants. *Taxon* 20(4):509–517
- Herritt M, Dhanapal AP, Purcell LC, Fritschi FB (2018) Identification of genomic loci associated with 21 chlorophyll fluorescence phenotypes by genome-wide association analysis in soybean. *BMC Plant Biol* 18(1):312
- Henshaw TL, Gilbert RA, Scholberg JMS, Sinclair TR (2007) Soybean (*Glycine max* L Merr) genotype response to early-season flooding: I root and nodule development. *J Agron Crop Sci* 193(3):177–188

- Hisano H, Sato S, Isobe S, Sasamoto S, Wada T, Matsuno A, Fujishiro T, Yamada M, Nakayama S, Nakamura Y, Watanabe S, Harada K, Tabata S (2007) Characterization of the soybean genome using EST-derived microsatellite markers. *DNA Res* 14:271–281
- Hossain MM, Liu X, Qi X, Lam HM, Zhang J (2014) Differences between soybean genotypes in physiological response to sequential soil drying and rewetting. *Crop J* 2(6):366–380
- Hua L, Challa GS, Subramanian S et al (2018) Genome-wide identification of drought response genes in soybean seedlings and development of biomarkers for early diagnoses. *Plant Mol Biol Rep* 36:350–362
- Huang L, Zeng A, Chen P, Wu C, Wang D, Wen Z (2018) Genome-wide association analysis of salt tolerance in soybean [*Glycine max* (L) Merr]. *Plant Breed* 137:714–720
- Hummer WS Linkage mapping for soybean (*Glycine max*) flood tolerance (2018) MSc thesis and dissertations, University of Arkansas, 3074
- Hwang S, King CA, Ray JD, Cregan PB, Chen P, Carter TE, Li Z, Abdel-Haleem H, Matson KW, Schapaugh W (2015) Confirmation of delayed canopy wilting QTLs from multiple soybean mapping populations. *Theor Appl Genet* 128:2047–2065
- Hwang W, Fu Y, Reyon D et al (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31:227–229
- Hymowitz T (2004) Speciation and cytogenetics Soybeans: improvement, production, and uses, vol 16. pp 97–136
- Hymowitz T, Singh RJ, Larkin RP (1990) Long-distance dispersal: the case for the allopolyploid *Glycine tabacina* (Labill) Benth and *G tomentella* Hayata in the West-Central Pacific. *Micronesica* 23:5–13
- Hyten DL, Song Q, Zhu Y, Choi IY, Nelson RL et al (2006) Impacts of genetic bottlenecks on soybean genome diversity. *Proc Natl Acad Sci USA* 103:16666–16671
- Hyten DL, Choi I-Y, Song Q, Specht JE, Carter TE, Shoemaker RC, Hwang EY, Matukumalli LK, Cregan PB (2010a) A high density integrated genetic linkage map of soybean and the development of a 1,536 Universal Soy Linkage Panel for QTL mapping. *Crop Sci* 50:960–968
- Hyten DL, Cannon SB, Song Q, Weeks N, Fickus EW, Shoemaker RC, Specht JE, Farmer AD, May GD, Cregan PB (2010b) High-throughput SNP discovery through deep resequencing of a reduced representation library to anchor and orient scaffolds in the soybean whole genome sequence. *BMC Genomics* 11:38
- Hyten DL, Song Q, Choi IY, Yoon MS, Specht JE, Matukumalli LK, Nelson RL, Shoemaker RC, Young ND, Cregan PB (2008) High-throughput genotyping with the goldengate assay in the complex genome of soybean. *Theor Appl Genet* 116:945–952
- Hyun T (2020) CRISPR/Cas-based genome editing to improve abiotic stress tolerance in plants. *Botanica Serbica* 44(2):121–127
- Innes RW, Ameline-Torregrosa C, Ashfield T, Cannon E, Cannon SB, Chacko B, Chen NW, Couloux A, Dalwani A, Denny R, Deshpande S, Egan AN, Glover N, Hans CS, Howell S, Ilut D, Jackson S, Lai H, Mammadov J, Del Campo SM, Metcalf M, Nguyen A, O’Bleness M, Pfeil BE, Podicheti R, Ratnaparkhe MB, Samain S, Sanders I, Segurens B, Sevignac M, Sherman-Broyles S, Thareau V, Tucker DM, Walling J, Wawrzynski A, Yi J, Doyle JJ, Geffroy V, Roe BA, Maroof MA, Young ND (2008) Differential accumulation of retroelements and diversification of NB-LRR disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean. *Plant Physiol* 148:1740–1759
- Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A (1987) Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* 169:5429–5433
- Jackson AL, Burchard J, Schelter J, Chau BN, Cleary M, Lim L, Linsley PS (2006) Widespread siRNA “off-target” transcript silencing mediated by seed region sequence complementarity. *RNA* (new York) 12(7):1179–1187
- Jia B, Sun M, DuanMu H, Ding X, Liu B, Zhu Y et al (2017) *GsCHX193*, a member of cation/H<sup>+</sup> exchanger superfamily from wild soybean contributes to high salinity and carbonate alkaline tolerance. *Sci Rep* 7:9423

- Jia Q, Li MW, Zheng C, Xu Y, Sun S, Li Z, Wong FL, Song J, Lin WW, Li Q, Zhu Y, Liang K, Lin W, Lam HM (2020) The soybean plasma membrane-localized cation/H<sup>+</sup> exchanger GmCHX20a plays a negative role under salt stress. *Physiol Plant* 171(4):714–727
- Jin T, Sun Y, Zhao R, Shan Z, Gai J, Li Y (2019) Overexpression of peroxidase gene GsPRX9 confers salt tolerance in soybean. *Intl J Mol Sci* 20(15):3745
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096):816–821
- Joshi T, Wang J, Zhang H, Chen S, Zeng S, Xu B, Xu D (2017) The evolution of soybean knowledge base (SoyKB) in plant genomics databases. Humana Press, New York, NY, pp 149–159
- Joung JK, Sander JD (2013) TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol* 14(1):49–55
- Kajija-Kanegae H, Nagasaki H, Kaga A, Hirano K, Ogiso-Tanaka E, Matsuoka M, Ishimori M, Ishimoto M, Hashiguchi M, Tanaka H, Akashi R, Isobe S, Iwata H (2021) Whole-genome sequence diversity and association analysis of 198 soybean accessions in mini-core collections. *DNA Res* 19(1):28
- Kaler AS, Ray JD, Schapaugh WT, King CA, Purcell LC (2017a) Genome-wide association mapping of canopy wilting in diverse soybean genotypes. *Theor Appl Genet* 130:2203–2217
- Kaler AS, Dhanapal AP, Ray JD, King CA, Fritsch FB, Purcell LC (2017b) Genome-wide association mapping of carbon isotope and oxygen isotope ratios in diverse soybean genotypes. *Crop Sci* 57:3085–3100
- Kan G, Zhang W, Yang W, Ma D, Zhang D, Hao D, Hu Z, Yu D (2015) Association mapping of soybean seed germination under salt stress. *Mol Gen Genom* 290(6):2147–2162
- Kaul S, Koo HL, Jenkins J, Rizzo M, Rooney T, Tallon LJ, Feldblyum T, Nierman W, Benito MI, Lin X, Town CD (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408(6814):796–815
- King CA, Purcell LC, Brye KR (2009) Differential wilting among soybean genotypes in response to water deficit. *Crop Sci* 49(1):290–298
- Kofsky J, Zhang H, Song B-H (2018a) The untapped genetic reservoir: the past, current, and future applications of the wild soybean (*Glycine soja*). *Front Plant Sci* 9:949
- Keim P, Shoemaker RC, Palmer RG (1989) Restriction fragment length polymorphism diversity in soybean. *Theo App Gene* 77:786–792
- Keim P, Diers BW, Olson TC, Shoemaker RC (1990) RFLP mapping in soybean: association between marker loci and variation in quantitative traits. *Genetics* 126:735–742
- Keim P, Schupp JM, Travis SE, Clayton K, Zhu T, Shi L, Ferreira A, Webb DM (1997) A high-density soybean genetic map based on AFLP markers. *Crop Sci* 37:537–543
- Keller R, Brearley CA, Trethewey RN, Muller-Rober B (1998) Reduced inositol content and altered morphology in transgenic potato plants inhibited for 1D-myo-inositol 3-phosphate synthase. *Plant J* 16(4):403–410
- Kim MY, Lee S, Van K, Kim TH, Jeong SC, Choi IY, Kim DS, Lee YS, Park D, Ma J, Kim WY, Kim BC, Park S, Lee KA, Kim DH, Kim KH, Shin JH, Jang YE, Kim KD, Liu WX, Chaisan T, Kang YJ, Lee YH, Kim KH, Moon JK, Schmutz J, Jackson SA, Bhak J, Lee SH (2010) Whole-genome sequencing and intensive analysis of the undomesticated soybean (*Glycine soja* Sieb and Zucc) genome. *Proc Natl Acad Sci USA* 107(51):22032–22037
- Kim HJ, Cho HS, Pak JH, Kwon T, Lee JH, Kim DH, Lee DH, Kim CG, Chung YS (2018) Confirmation of drought tolerance of ectopically expressed *AtABF3* gene in soybean. *Mol Cells* 41(5):413–422
- Kim JY, Jeong S, Kim KH, Lim WJ, Lee HY, Jeong N, Moon JK, Kim N (2019) Dissection of soybean populations according to selection signatures based on whole-genome sequences. *Gigascience* 8(12):giz151
- Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9:29

- Kumawat G, Maranna S, Gupta S, Tripathi R, Agarwal N, Singh V, Rajesh V, Chandra S, Kamble V, Nataraj V, Bharti A, Sharma MP, Jadhav PV, Ratnaparkhe MB, Satpute GK, Bhatia VS (2020) Identification of novel genetic sources for agronomic and quality traits in soybean using multi-trait allele specific genic marker assays. *J Plant Biochem Biotechnol* 30:160–171
- Kuwano M, Ohyama A, Tanaka Y et al (2006) Molecular breeding for transgenic rice with low-phytic-acid phenotype through manipulating *myo*-inositol 3-phosphate synthase gene. *Mol Breeding* 18:263–272
- Ladizinsky G, Newell CA, Hymowitz T (1979a) Giemsa staining of soybean chromosomes. *J Hered* 70:415–416
- Ladizinsky G, Newell CA, Hymowitz T (1979b) Wide crosses in soybean: prospects and limitations. *Euphytica* 28:421–423
- Lai MC, Lai ZY, Jhan LH, Lai YS, Kao CF (2021) Prioritization and evaluation of flooding tolerance genes in soybean [*Glycine max* (L.) Merr.]. *Front Genet* 11:612131
- Lam HM, Xu X, Liu X, Chen W, Yang G, Wong FL, Li MW, He W, Qin N, Wang B, Li J, Jian M, Wang J, Shao G, Wang J, Sun SS, Zhang G (2010) Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat Genet* 42(12):1053–1059
- Lark KG, Weisemann JM, Matthews BF, Palmer R, Chase K, Macalma T (1993) A genetic map of soybean (*Glycine max* L) using an intraspecific cross of two cultivars: ‘Minsoy’ and ‘Noir 1.’ *Theor Appl Genet* 86:901–906
- Larson SR, Rutger JN, Young KA, Raboy V (2000) Isolation and genetic mapping of a non-lethal rice (*Oryza sativa* L.) low phytic acid 1 Mutation. *Crop Sci* 40(5):1397–1405
- Le DT, Nishiyama R, Watanabe Y, Tanaka M, Seki M, Ham LH, et al (2012) Differential gene expression in soybean leaf tissues at late developmental stages under drought stress revealed by genome-wide transcriptome analysis. *PLoS One* 7(11):e49522
- Lee JD, Shannon JG, Vuong TD, Nguyen HT (2009) Inheritance of salt tolerance in wild soybean (*Glycine soja* Sieb and Zucc) accession PI483463. *J Hered* 100:798–801
- Lee GJ, Carter TE Jr, Villagarcia MR, Li Z, Zhou X, Gibbs MO, Boerma HR (2004) A major QTL conditioning salt tolerance in S-100 soybean and descendent cultivars. *Theor Appl Genet* 109:1610–1619
- Lee S, Freewalt KR, McHale LK et al (2015) A high-resolution genetic linkage map of soybean based on 357 recombinant inbred lines genotyped with BARCSoySNP6K. *Mol Breeding* 35:58
- Lenis JM, Ellersieck M, Blevins DG, Sleper DA, Dunn D, Lee JD, Shannon JG (2011) Differences in ion accumulation and salt tolerance among *Glycine* Accessions. *J Agron Crop Sci* 197:302–310
- Li B, Chen L, Sun W, Wu D, Wang M, Yu Y, Chen G, Yang W, Lin Z, Zhang X, Duan L (2020a) Phenomics-based GWAS analysis reveals the genetic architecture for drought resistance in cotton. *Plant Biotechnol J* 18(12):2533–2544
- Li CD, Jiang HW, Liu CY, Guo T, Xin WZ, Hong WX et al (2011) QTL identification of drought tolerance to soybean in selection population. *Acta Agron Sin* 37:603–611
- Li D, Pfeiffer T, Cornelius P (2008a) Soybean QTL for yield and yield components associated with *Glycine soja* alleles. *Crop Sci* 48:571–581
- Li DH, Li W, Li HY, Guo JJ, Chen FJ (2018) The soybean GmRACK1 gene plays a role in drought tolerance at vegetative stages. *Russ J Plant Physiol* 65:541–552
- Li JF, Norville J, Aach J et al (2013a) Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat Biotechnol* 31:688–691
- Li M, Chen R, Jiang Q, Sun X, Zhang H, Hu Z GmNAC06 (2020b) A NAC domain transcription factor enhances salt stress tolerance in soybean. *Plant Mol Biol* 5
- Li N, Li Z, Fan S, Pu Y, Gong Y, Tian R, Guo X, Ding H (2020c) Identification of potential key genes affecting soybean growth under salt stress via transcriptome study. *Biol Planta* 64:692–700
- Li R, Jiang H, Zhang Z, Zhao Y, Xie J, Wang Q, Zheng H, Hou L, Xiong X, Xin D, Hu Z, Liu C, Wu X, Chen Q (2019) Combined linkage mapping and BSA to identify QTL and candidate genes for plant height and the number of nodes on the main stem in soybean. *Int J Mol Sci* 21(1):42

- Li T, Liu B, Spalding M et al (2012) High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat Biotechnol* 30:390–392
- Li W, Han Y, Zhang D, Yang M, Teng W, Jiang Z, Qiu L, Sun G (2008b) Genetic diversity in soybean genotypes from north-eastern China and identification of candidate markers associated with maturity rating. *Plant Breed* 127:56–61
- Li Y, Chen Q, Nan H, Li X, Lu S, Zhao X, et al (2017) Overexpression of *GmFDL19* enhances tolerance to drought and salt stresses in soybean. *PLoS One* 12(6):e0179554
- Li YH, Li W, Zhang C et al (2010) Genetic diversity in domesticated soybean (*Glycine max*) and its wild progenitor (*Glycine soja*) for simple sequence repeat and single-nucleotide polymorphism loci. *New Phytol* 188(1):242–253
- Li YH, Zhao SC, Ma JX, Li D, Yan L, Li J, Qi XT, Guo XS, Zhang L, He WM, Chang RZ, Liang QS, Guo Y, Ye C, Wang XB, Tao Y, Guan RX, Wang JY, Liu YL, Jin LG, Zhang XQ, Liu ZX, Zhang LJ, Chen J, Wang KJ, Nielsen R, Li RQ, Chen PY, Li WB, Reif JC, Purugganan M, Wang J, Zhang MC, Wang J, Qiu LJ (2013b) Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing. *BMC Genomics* 28(14):579
- Li Z, Nelson RL (2002) RAPD marker diversity among cultivated and wild soybean accessions from four Chinese provinces. *Crop Sci* 42:1737–1744
- Liang H, Yu Y, Yang H, Xu L, Dong W, Du H, Cui W, Zhang H (2014) Inheritance and QTL mapping of related root traits in soybean at the seedling stage. *Theor App Genet* 10:2127–2137
- Libault M, Farmer A, Joshi T, Takahashi K, Langley RJ, Franklin LD, He J, Xu D, May G, Stacey G (2010) An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants. *Plant J* 63(1):86–99
- Lightfoot DA, Njiti VN, Gibson PT, Kassem MA, Iqbal JM, Meksem K (2005) Registration of the Essex by Forrest recombinant inbred line mapping population. *Crop Sci* 45:1678–1681
- Lin Y, Li W, Zhang Y et al (2019) Identification of genes/proteins related to submergence tolerance by transcriptome and proteome analyses in soybean. *Sci Rep* 9:14688
- Linkemer G, Board JE, Musgrave ME (1998) Waterlogging effects on growth and yield components in late-planted soybean. *Crop Sci* 38:1576–1584
- Liu Y, Gai JY, Lü HN, Wang YJ, Chen SY (2005) Identification of drought tolerant germplasm and inheritance and QTL mapping of related root traits in soybean (*Glycine max* (L.) Merr.). *Yi Chuan Xue Bao* 32(8):855–863
- Liu Y, Du H, Li P, Shen Y, Peng H, Liu S, Zhou GA, Zhang H, Liu Z, Shi M, Huang X, Li Y, Zhang M, Wang Z, Zhu B, Han B, Liang C, Tian Z (2020a) Pan-genome of wild and cultivated soybeans. *Cell J* 9(1):182
- Liu D, Li M, Liu Y, Shi L (2020b) Integration of the metabolome and transcriptome reveals the resistance mechanism to low nitrogen in wild soybean seedling roots. *Environ Exp Bot* 175:104043
- Liu Y, Yu L, Qu Y, Chen J, Liu X, Hong H, Liu Z, Chang R, Gilliam M, Qiu L et al (2016) GmSALT3, which confers improved soybean salt tolerance in the field, increases leaf Cl<sup>-</sup> exclusion prior to Na<sup>+</sup> exclusion but does not improve early vigor under salinity. *Front Plant Sci* 7(1485):1485
- Liu, ZX, Li, HH, Wen, ZX, Fan, XH, Li, YH, Guan, RX, Guo, Y, Wang, SM, Wang, DC, Qiu, LJ (2017) Comparison of genetic diversity between Chinese and American soybean (*Glycine max* (L.) accessions revealed by high-density SNPs. *Front Plant Sci* 8:2014
- Livingstone JM, Cheng KC, Strömviik MV (2010) Bioinformatics as a tool. In: Bilyeu K, Ratnaparkhe MB, Kole C (eds) *Genetics, genomics, and breeding of soybean*. CRC, New Hampshire, UK
- Liu K, Peng L, Zhang C, Lu J, Yang B, Xiao Z, Liang Y, Xu X, Qu C, Zhang K et al (2017) Genome-wide association and transcriptome analyses reveal candidate genes underlying yield-determining traits in *Brassica napus*. *Front Plant Sci* 8:206
- Lorenzen L, Boutin S, Young N, Specht JE, Shoemaker RC (1995) Soybean pedigree analysis using map-based molecular markers: tracking RFLP markers in cultivars. *Crop Sci* 35:1326–1336



- Ma J, Shoemaker R, Jackson S, Cannon S (2010) Comparative genomics. In: Bilyeu K, Ratnaparkhe MB, Kole C (eds) Genetics, genomics & breeding of soybean. CRC, New Hampshire, UK, pp 245–262
- Machado FB, Moharana KC, Almeida-Silva F, Gazara RK, Pedrosa-Silva F, Coelho FS, Grativolo C, Venancio TM (2019) Systematic analysis of 1,298 RNA-Seq samples and construction of a comprehensive soybean (*Glycine max*) expression atlas. *Plant J* 103(5):1894–1909
- Maguire TL, Grimmond S, Forrest A, Iturbe-Ormaetxe I, Meksem K, Gresshoff P (2002) Tissue-specific gene expression in soybean (*Glycine max*) detected by cDNA microarray analysis 159:1361–1364
- Mali P, Esvelt K, Church G (2013) Cas9 as a versatile tool for engineering biology. *Nat Methods* 10:957–963
- Manavalan LP, Prince SJ, Musket TA, Chaky J, Deshmukh R, Vuong TD et al (2015) Identification of novel QTL governing root architectural traits in an interspecific soybean population. *PLoS One* 10(3):e0120490
- Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB, Lark KG (1996) Genetic mapping of agronomic traits using recombinant inbred lines of soybean. *Crop Sci* 36:1327–1336
- Marek LF, Shoemaker RC (1997) BAC contig development by fingerprint analysis in soybean. *Genome* 40:420–427
- Matthews BF, Devine TE, Weisemann JM, Beard HS, Lewers KS, MacDonald MH, Park YB, Maiti R, Lin JJ, Kuo J (2001) Incorporation of sequenced cDNA and genomic markers into the soybean genetic map. *Crop Sci* 41:516–521
- Maughan PJ, Saghai Maroof MA, Buss GR (1995) Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean. *Genome* 38:715–723
- Meksem K, Zobrist K, Ruben E, Hyten D, Quanzhou T, Zhang HB, Lightfoot DA (2000) Two large-insert soybean genomic libraries constructed in a binary vector: applications in chromosome walking and genome wide physical mapping. *Theor Appl Genet* 101:747–755
- Meng F, Zhao H, Zhu B, Zhang T, Yang M, Li Y, Han Y, Jiang J (2021) Genomic editing of intronic enhancers unveils their role in fine-tuning tissue-specific gene expression in *Arabidopsis thaliana*. *Plant Cell* 33(6):1997–2014
- Mian M, Bailey M, Ashley D, Wells R, Carter T, Parrott W, Boerma H (1996) Molecular markers associated with water use efficiency and leaf ash in soybean. *Crop Sci* 36:1252–1257
- Mian M, Ashley D, Boerma H (1998) An additional QTL for water use efficiency in soybean. *Crop Sci* 38:390–393
- Ming LI, Zheng HU, Jiang QY et al (2018) GmNAC15 overexpression in hairy roots enhances salt tolerance in soybean. *J Integr Agri* 17(3):530–538
- Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K (2012) AP2/ERF family transcription factor in plant abiotic stress responses. *Biochim Biophys Acta* 1819:86–96
- Monteros MJ, Lee G, Missaoui AM, Carter TE, Boerma HR (2006) Identification and confirmation of QTL conditioning drought tolerance in Nepalese soybean. In: The 11th biennial conference on the molecular and cellular biology of the soybean, abstract PI471938. August 5–8. Lincoln, NE
- Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. *Plant J* 3:175–182
- Morgante M, Rafalski A, Biddle P, Tngey S, Olivieri AM (1994) Genetic mapping and variability of seven soybean simple sequence repeat loci. *Genome* 37:763–769
- Munns R, Tester M (2008) Mechanisms of salinity tolerance. *Annu Rev Plant Biol* 59:651–681
- Nekrasov V, Staskawicz B, Weigel D, Jones JD, Kamoun S (2013) Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat Biotechnol* 8:691–693
- Newell CA, Hymowitz T (1980) A taxonomic revision in the genus *Glycine* subgenus *Glycine* (Leguminosae). *Brittonia* 32:63–69
- Neto JRFC, Pandolfi V, Guimaraes FCM, Benko-Iseppon AM, Romero C, de Oliveira Silva RL, Rodrigues FA, Abdelnoor RV, Nepomuceno AL, Kido EA (2013) Early transcriptional response of soybean contrasting accessions to root dehydration. *PLoS ONE* 8(12):83466

- Nguyen VT, Vuong TD, VanToai T, Lee JD, Wu X, Mian MR, Dorrance AE, Shannon JG, Nguyen HT (2012) Mapping of quantitative trait loci associated with resistance to *Phytophthora sojae* and flooding tolerance in soybean. *Crop Sci* 52(6):2481–2493
- Nguyen VL, Takahashi R, Githiri SM et al (2017) Mapping quantitative trait loci for root development under hypoxia conditions in soybean (*Glycine max* L Merr). *Theor Appl Genet* 130:743–755
- Ning W, Zhai H, Yu J, Liang S, Yang X, Xing X, Huo J, Pang T, Yang Y, Bai X (2017) Overexpression of *Glycine soja* WRKY20 enhances drought tolerance and improves plant yields under drought stress in transgenic soybean. *Mol Breed* 37(2):19
- Nunes ACS, Vianna GR, Cuneo F et al (2006) RNAi-mediated silencing of the *myo*-inositol-1-phosphate synthase gene (*GmMIP1*) in transgenic soybean inhibited seed development and reduced phytate content. *Planta* 224:125–132
- Oosterhuis DM, Scott HD, Hampton RE, Wullschlegel SD (1990) Physiological response of two soybean [*Glycine max* L Merr] cultivars to short-term flooding. *Environ Exp Botany* 30:85–92
- Pagano MC, Miransari M, Corrêa EJA, Duarte NF, Yelikbayev BK (2020) Genomic research favoring higher soybean production. *Curr Genomics* 21(7):481–490
- Parida AK, Das AB (2005) Salt tolerance and salinity effects on plants: a review ecotoxicol. *Environ Safety* 60(3):324
- Pathan SM, Lee JD, Slepner DA, Fritsch FB, Sharp RE, Carter TE Jr, Nelson RL, King CA, Schapaugh WT, Ellersieck MR, Nguyen HT, Shannon JG (2014) Two soybean plant introductions display slow leaf wilting and reduced yield loss under drought. *J Agron Crop Sci* 200(3):231–236
- Patil G, Do T, Vuong TD, Valliyodan B, Lee JD, Chaudhary J, Shannon JG, Nguyen HT (2016) Genomic-assisted haplotype analysis and the development of high-throughput SNP markers for salinity tolerance in soybean. *Sci Rep* 6:19199
- Patterson PR, Hudak CM (1996) Drought-avoidant soybean germplasm maintains nitrogen fixation capacity under water stress. *Plant Soil* 186(1):39–43
- Phang TH, Shao G, Lam HM (2008) Salt tolerance in soybean. *J Integr Plant Biol* 50:1196–1212
- Pitman MG, Läuchli A (2002) Global impact of salinity and agricultural ecosystems In: Läuchli A, Lüttge U (eds) *Salinity: environment—plants—molecules* Kluwer. Dordrecht, Netherlands, pp 3–20
- Polizel AM, Medri ME, Nakashima K, Yamanaka N, Farias JRB, Oliveira MCN, Marin SRR, Abdelnoor RV, Marcelino-Gui-marães FC, Fuganti R et al (2011) Molecular, anatomical and physiological properties of a genetically modified soybean line transformed with rd29A:AtDREB1A for the improvement of drought tolerance. *Genet Mol Res* 10:3641–3656
- Prince SJ, Song L, Qiu D, dos Santos JVM, Chai C, Joshi T, Patil G, Valliyodan B, Vuong TD, Murphy M (2015a) Genetic variants in root architecture-related genes in a *Glycine soja* accession, a potential resource to improve cultivated soybean. *BMC Genomics* 16:132
- Prince SJ, Joshi T, Mutava RN, Syed N, Vitor MDSJ, Patil G, Song L, Wang J, Lin L, Chen W, Shannon JG (2015b) Comparative analysis of the drought-responsive transcriptome in soybean lines contrasting for canopy wilting. *Plant Sci* 240:65–78
- Purcell LC, DeSilva M, King CA, Kim WH (1997) Biomass accumulation and allocation in soybean associated with genotypic differences in tolerance of nitrogen fixation to water deficits. *Plant Soil* 196(1):101–113
- Qadir M, Quillerou E, Nangia V, Murtaza G, Singh M, Thomas RJ, Drechsel P, Noble AD (2014) Economics of salt-induced land degradation and restoration. *Nat Resour Forum* 38:282–295
- Qi X, Li MW, Xie M, Liu X, Ni M, Shao G, Song C, Kay-Yuen Yim A, Tao Y, Wong FL, Isobe S, Wong CF, Wong KS, Xu C, Li C, Wang Y, Guan R, Sun F, Fan G, Xiao Z, Zhou F, Phang TH, Liu X, Tong SW, Chan TF, Yiu SM, Tabata S, Wang J, Xu X, Lam HM (2014) Identification of a novel salt tolerance gene in wild soybean by whole-genome sequencing. *Nat Commun* 5:4340
- Qiu PC, Zhang WB, Li CD, Jiang HW, Liu CY, Fan DM, Zeng QL, Hu HG, Chen QS (2011) Genetic overlap of drought-tolerance loci between germination stage and seedling stage analyzed using introgression lines in soybean. *Acta Agron Sin* 37(3):477–483

- Rabie RK, Kumazawa K (1988) Effect of salt stress on nitrogen nutrition and yield quality of nodulated soybeans. *Soil Sci Plant Nutr* 34:385–391
- Rafalski A, Morgante M (2004) Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. *Trends Genet* 20:103–111
- Ratnaparkhe MB, RJ Singh, JJ Doyle (2010) *Glycine*. In: Kole C (eds) *Wild crop relatives*. Springer, Genomic and Breeding Resources pp 83–116
- Ratnaparkhe MB, Ramesh SV, Giriraj K, Husain SM, Gupta S (2013) In: Gupta (ed) *Advances in soybean genomics*. Springer, Legumes in the Omic Era pp 41–72
- Ratnaparkhe MB, Marmat N, Kumawat G, Shivakumar M, Kamble VG, Nataraj V, Ramesh SV, Deshmukh MP, Singh AK, Sonah H, Deshmukh RK, Prasad M, Chand S, Gupta S (2020) Whole genome re-sequencing of soybean accession EC241780 providing genomic landscape of candidate genes involved in rust resistance. *Curr Genomics* 21(7):504–511
- Ray JD, Dhanapal AP, Singh SK, Hoyos-Villegas V, Smith JR, Purcell LC, King CA, Boykin D, Cregan PB, Song Q, Fritschi FB (2015) Genome-wide association study of ureide concentration in diverse maturity group IV soybean [*Glycine max* (L.) Merr.] Accessions. G3 (Bethesda). 5(11): 2391–403
- Ren H, Han J, Wang X, Zhang B et al (2020) QTL mapping of drought tolerance traits in soybean with SLAF sequencing. *Crop J* 8(6):977–989
- Reyna N, Cornelious B, Shannon JG, Sneller CH (2003) Evaluation of a QTL for waterlogging tolerance in Southern soybean germplasm. *Crop Sci* 43(6):2077–2082
- Rhine MD, Stevens G, Shannon G, Wrather A, Slep D (2010) Yield and nutritional responses to waterlogging of soybean cultivars. *Irrig Sci* 28:135–142
- Riaz A, Kanwal F, Börner A, Pillen K, Dai F, Alqudah AM (2021) Advances in genomics-based breeding of barley: molecular tools and genomic databases. *Agronomy* 11(5):894
- Rizal G, Karki S (2011) Research article alcohol dehydrogenase (ADH) activity in soybean (*Glycine max* [L] Merr) under flooding stress. *Electr J Plant Breed* 2(1):50–57
- Rosenzweig C, Tubiello FN, Goldberg R, Mills E, Bloomfield J (2002) Increased crop damage in the US from excess precipitation under climate change. *Global Environ Change* 12:197–202
- Rongwen J, Akkaya MS, Bhagwat AA, Lavi U, Cregan PB (1995) The use of microsatellite DNA markers for soybean genotype identification. *Theor Appl Genet* 90:43–48
- Roy SJ, Negrao S, Tester M (2014) Salt resistant crop plants. *Curr Opin Biotechnol* 26:115–124
- Saleem A, Muylle H, Aper J, Ruttink T, Wang J, Yu D, Roldán-Ruiz I (2021) A genome-wide genetic diversity scan reveals multiple signatures of selection in a European soybean collection compared to Chinese collections of wild and cultivated soybean accessions. *Front Plant Sci* 26(12):631767
- Salimath SS, Bhattacharyya MK (1999) Generation of a soybean BAC library, and identification of DNA sequences tightly linked to the Rps1-k disease resistance gene. *Theor Appl Genet* 98:712–720
- Satpute GK, Gupta S, Bhatia VS, Devdas R, Kumawat G, Tiwari VK, Purwar JP, Agrawal A (2019) Multi-tiered selection scheme for breeding drought tolerance in soybean In: *Abstr golden jubilee international salinity conference (GJISC): resilient agriculture in saline environments under changing climate: challenges & opportunities, held during February 7–9, 2019 at ICAR–CSSRI, Karnal, Haryana*, p 127
- Satpute GK, Arya M, Gupta S, Bhatia VS, Ramgopal D, Ratnaparkhe MB, Chandra S, Singh M, Nagar S, Kamble VG, Pandey S, Kumawat G, Shivakumar M, Nataraj V, Rajesh V (2020) Identifying drought tolerant germplasm through multiplexing polygenic traits in soybean (*Glycine max* L Merrill). *J Oilseeds Res* 37(SI):56–57
- Satpute GK, Ratnaparkhe MB, Chandra S, Kamble VG, Kavishwar R, Singh AK, Gupta S, Devdas R, Arya M, Singh M, Sharma MP, Kumawat G, Shivakumar M, Nataraj V, Kuchlan MK, Rajesh V, Srivastava MK, Chitikineni A, Varshney RK, Nguyen HT (2021) Breeding and molecular approaches for evolving drought-tolerant soybeans. In: Giri B, Sharma MP (eds) *Plant stress biology*. Springer, Singapore, pp 83–130

- Sayama T, Nakazaki T, Ishikawa G, Yagasaki K, Yamada N, Hirota N, Hirata K, Yoshikawa T, Saito H, Teraishi M (2009) QTL analysis of seed-flooding tolerance in soybean (*Glycine max* [L] Merr). *Plant Sci* 176:514–521
- Schaefer RJ, Michno JM, Jeffers J, Hoekenga O, Dilkes B, Baxter I, Myers CL (2018) Integrating coexpression networks with GWAS to prioritize causal genes in maize. *Plant Cell* 30(12):2922–2942
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463:178
- Segal DJ, Meckler JF (2013) Genome engineering at the dawn of the golden age. *Annu Rev Genomics Hum Genet* 14(1):135–158
- Seversike TM, Sermons SM, Sinclair TR, Carter TE Jr, Rufty TW (2014) Physiological properties of a drought-resistant wild soybean genotype: transpiration control with soil drying and expression of root morphology. *Plant Soil* 374(1–2):359–370
- Shan Q, Wang Y, Chen K, Liang Z, Li J, Zhang Y, Zhang K, Liu J, Voytas DF, Zheng X, Zhang Y, Gao C (2013) Rapid and efficient gene modification in rice and *Brachypodium* using TALENs. *Mol Plant* 6(4):1365–1368
- Shannon JG, Stevens WE, Wiebold WJ, McGraw RL, Sleper DA, Nguyen HT (2005) Breeding soybeans for improved tolerance to flooding proceedings of the 30th soybean research conference, American Seed Trade Association. Chicago, IL, USA
- Shao GH, Song JZ, Liu HL (1986) Preliminary studies on the evaluation of salt tolerance in soybean varieties. *Acta Agron Sin* 6:30–35
- Shao GH, Wan CW, Chang RZ, Chen YW (1993) Preliminary study on the damage of plasma membrane caused by salt stress. *Crops* 1:39–40
- Sharmin RA, Bhuiyan MR, Lv W, Yu Z, Chang F, Kong J, Bhat JA, Zhao T (2020) RNA-Seq based transcriptomic analysis revealed genes associated with seed-flooding tolerance in wild soybean (*Glycine soja* Sieb & Zucc). *Environ Exp Bot* 171:p103906
- Sharma SK, Goloubinoff P, Christen P (2011) Non-native proteins as newly-identified targets of heavy metals and metalloids. In: Banfalvi G (ed) Cellular effects of heavy metals. Springer, pp 263–274
- Shi X, Yan L, Yang C et al (2018) Identification of a major quantitative trait locus underlying salt tolerance in ‘Jidou 12’ soybean cultivar. *BMC Res Notes* 11:95
- Shoemaker RC, Specht JE (1995) Integration of the soybean molecular and classical genetic linkage groups. *Crop Sci* 35:436–446
- Shoemaker RC, Schlueter J, Doyle JJ (2006) Paleopolyploidy and gene duplication in soybean and other legumes. *Curr Opin Plant Biol* 9:104–109
- Shoemaker RC, Grant D, Olson T, Warren WC, Wing R, Yu Y, Kim H, Cregan P, Joseph B, FutrellGriggs M, Nelson W, Davito J, Walker J, Wallis J, Kremitski C, Scheer D, Clifton SW, Graves T, Nguyen H, Wu X, Luo M, Dvorak J, Nelson R, Cannon S, Tomkins J, Schmutz J, Stacey G, Jackson S (2008) Microsatellite discovery from BAC end sequences and genetic mapping to anchor the soybean physical and genetic maps. *Genome* 51:294–302
- Shukla VK, Doyon Y, Miller JC, DeKolver RC, Moehle EA, Worden SE, Mitchell JC, Arnold NL, Gopalan S, Meng X, Choi VM, Rock JM, Wu YY, Katibah GE, Zhifang G, McCaskill D, Simpson MA, Blakeslee B, Greenwalt SA, Urnov FD (2009) Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature* 459(7245):437–441
- Shultz JL, Kazi S, Bashir R, Afzal JA, Lightfoot DA (2007) The development of BAC-end sequence-based microsatellite markers and placement in the physical and genetic maps of soybean. *Theor Appl Genet* 114:1081–1090
- Sigoillot FD, King RW (2010) Vigilance and validation: keys to success in RNAi screening. *ACS Chem Biol* 6(1):47–60
- Sinclair TR, Purcell LC, King CA, Sneller CH, Chen P, Vadez V (2007) Drought tolerance and yield increase of soybean resulting from improved symbiotic N<sub>2</sub> fixation. *Field Crop Res* 101(1):68–71
- Singh RJ, Hymowitz T (1985) An intersubgeneric hybrid between *Glycine tomentella* Hayata and the soybean, *G max* (L) Merr. *Euphytica* 34:187–192

- Singh RJ, Kollipara KP, Hymowitz T (1987) Intersubgeneric hybridization of soybeans with a wild perennial species, *Glycine clandestina* Wendl. *Theor Appl Genetics* 74:391–396
- Singh RJ, Klein TM, Mauvais CJ, Knowlton S, Hymowitz T, Kostow CM (1998a) Cytological characterization of the transgenic soybean. *Theor Appl Genet* 96:319–324
- Singh RJ, Kollipara KP, Hymowitz T (1998b) The genomes of *Glycine canescens* F.J. Herm and *G. tomentella* Hayata of Western Australia and their phylogenetic relationships in the genus *Glycine* Willd. *Genome* 41:669–679
- Singh RJ, Chung GH, Nelson RL (2007) Landmark research in Legumes. *Genome* 50:525–537
- Singh V, Sanwal SK, Kumawat G, Shivakumar M, Satpute GK, Gill BS, Singh Z, Panwar S, Singh J, Sharma PC (2020) Assessing the effect of salt stress on Soybean [*Glycine max* (L) Merrillis] genotypes using AMMI and GGE biplot analysis. *J Soil Salinity Water Quality* 12(1):95–100
- Singleton PW, Bohloul BB (1984) Effect of salinity on nodule formation by soybean. *Plant Physiol* 74:72–76
- Skorupska HT, Shoemaker RC, Warner A, Shipe ER, Bridges WC (1993) Restriction fragment length polymorphism in soybean germplasm of the southern USA. *Crop Sci* 33:1169–1176
- Sloane RJ, Patterson RP, Carter TE Jr (1990) Field drought tolerance of a soybean plant introduction. *Crop Sci* 30(1):118–123
- Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. *Theor Appl Genet* 109:122–128
- Song L, Nguyen N, Deshmukh RK, Patil GB, Prince SJ, Valliyodan B, Mutava R, Pike SM, Gassmann W, Nguyen HT (2016a) Soybean TIP gene family analysis and characterization of GmTIP1; 5 and GmTIP2; 5 water transport activity. *Front Plant Sci* 7:1564
- Song L, Prince S, Valliyodan B et al (2016b) Genome-wide transcriptome analysis of soybean primary root under varying water-deficit conditions. *BMC Genomics* 17:57
- Song Q, Jia G, Zhu Y, Grant D, Nelson RT, Hwang E, Hyten DL, Cregan PB (2010) Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR\_10) in soybean. *Crop Sci* 50:1950–1960
- Song Q, Hyten DL, Jia G, Quigley CV, Fickus EW, Nelson RL, Cregan PB (2013) Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. *PLoS One* 8(1):e54985
- Song Q, Yan L, Quigley C, Fickus E, Wei H, Chen L, Dong F, Araya S, Liu J, Hyten D, Pantalone V, Nelson RL (2020) Soybean BARCSoySNP6K: an assay for soybean genetics and breeding research. *Plant J* 104(3):800–811
- Sosnowski O, Charcosset A, Joets J (2012) BioMercator v3: an upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms. *Bioinformatics* 28:2082–2083
- Specht J, Hume D, Kumudini SV (1999) Soybean yield potential—a genetic and physiological perspective. *Crop Sci* 39:1560–1570
- Specht JE, Chase K, Macrander M, Graef GL, Chung J, Markwell JP, German M, Orf JH, Lark KG (2001) Soybean response to water: a QTL analysis of drought tolerance. *Crop Sci* 41(2):493–509
- Sreenivasa V, Lal SK, Babu PK, Swamy HKM, Yadav RR, Talukdar A, Rathod DR (2020) Inheritance and mapping of drought tolerance in soybean at seedling stage using bulked segregant analysis. *Plant Genet Resour*:1–8
- Steketeer CJ, Schapaugh WT, Carter TE, Li Z (2020) Genome-wide association analyses reveal genomic regions controlling canopy wilting in soybean. *G3 Genes Genomes Genetics* 10:1413–1425
- Stolf-Moreira R, Medri ME, Marcelino FC, de Oliviera MLN, Farias JRB, Abdelnoor RV, Nepomuceno AL (2010) Cloning and quantitative expression of drought induced genes in soybean. *Genet Mol Res* 9:858–867
- Staub JE, Serquen FC, Gupta M (1996) Genetic markers, map construction, and their application in plant breeding. *Hort Sci* 31(5):729–739
- Suematsu K, Abiko T, Nguyen VL, Mochizuki T (2017) Phenotypic variation in root development of 162 soybean accessions under hypoxia condition at the seedling stage. *Plant Prod Sci* 20(3):323–335

- Sun F, Xu M, Park C, Dwiyantri MS, Nagano AJ, Zhu J, Watanabe S, Kong F, Liu B, Yamada T, Abe J (2019a) Characterization and quantitative trait locus mapping of late-flowering from a Thai soybean cultivar introduced into a photoperiod-insensitive genetic background. *PLoS One* 14(12):e0226116
- Sun TJ, Fan L, Yang J, Cao RZ, Yang CY, Zhang J et al (2019b) A *Glycine max* sodium/hydrogen exchanger enhances salt tolerance through maintaining higher Na<sup>+</sup> efflux rate and K<sup>+</sup>/Na<sup>+</sup> ratio in *Arabidopsis*. *BMC Plant Biol* 19:496
- Sun X, Hu Z, Chen R et al (2015) Targeted mutagenesis in soybean using the CRISPR-Cas9 system. *Sci Rep* 5:10342
- Sullivan M, VanToai T, Fausey N, Beuerlein J, Parkinson R, Soboyejo A (2001) Evaluating on-farm flooding impacts on soybean. *Crop Sci* 41:93–100
- Szczypek M, Brondani V, Buchel J, Serrano L, Segal DJ, Cathomen T (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol* 25:786–793
- Thao NP, Thu NB, Hoang XL, Van Ha C, Tran LS (2013) Differential expression analysis of a subset of drought-responsive GmNAC genes in two soybean cultivars differing in drought tolerance. *Intl J Mol Sci* 14:23828–23841
- Thibaud-Nissen F, Shealy RT, Khanna A, Vodkin LO (2003) Clustering of microarray data reveals transcript patterns associated with somatic embryogenesis in soybean. *Plant Physiol* 132(1):118–136
- Thompson JA, Nelson RL, Vodkin LO (1998) Identification of diverse soybean germplasm using RAPD markers. *Crop Sci* 38:1348–1355
- Tomkins JP, Mahalingham R, Miller-Smith H, Goicoechea JL, Knapp HT, Wing RA (1999) A soybean bacterial artificial chromosome library for PI 437654 and the identification of clones associated with cyst nematode resistance. *Plant Mol Biol* 41:25–32
- Torkamaneh D, Laroche J, Valliyodan B, O'Donoghue L, Cober E, Rajcan I, Vilela Abdelnoor R, Sreedasyam A, Schmutz J, Nguyen HT, Belzile F (2021) Soybean (*Glycine max*) Haplotype Map (GmHapMap): a universal resource for soybean translational and functional genomics. *Plant Biotechnol J* 19(2):324–334
- Tripathi P, Rabara RC, Reese RN, Miller MA, Rohila JS, Subramanian S et al (2016) A toolbox of genes, proteins, metabolites and promoters for improving drought tolerance in soybean includes the metabolite coumestrol and stomatal development genes. *BMC Genomics* 17:102
- Turner M, Yu O, Subramanian S (2012) Genome organization and characteristics of soybean microRNAs. *BMC Genom* 13:169
- Tuyen DD, Lal SK, Xu DH. Identification of a major QTL allele from wild soybean (*Glycine soja* Sieb. & Zucc.) for increasing alkaline salt tolerance in soybean. *Theor Appl Genet* 121:229–236
- Tuyen DD, Zhang HM, Xu DH (2013) Validation and high-resolution mapping of a major quantitative trait locus for alkaline salt tolerance in soybean using residual heterozygous line. *Mol Breed* 31:79–86
- Vadez V, Kholová J, Hummel G, Zhokhavets U, Gupta SK, Hash CT (2015) LeasyScan: a novel concept combining 3D imaging and lysimetry for high-throughput phenotyping of traits controlling plant water budget. *J Exp Bot* 66(18):5581–5593
- Valliyodan B, Dan Qiu, Patil G, Zeng P, Huang J, Dai L, Chen C, Li Y, Joshi T, Song L, Vuong TD, Musket TA, Xu D, Shannon JG, Shifeng C, Liu X, Nguyen HT (2016) Landscape of genomic diversity and trait discovery in soybean. *Sci Rep* 31(6):23598
- Valliyodan B, Ye H, Song L, Murphy M, Shannon JG, Nguyen HT (2017) Genetic diversity and genomic strategies for improving drought and waterlogging tolerance in soybeans. *J Exp Bot* 68(8):1835–1849
- Valliyodan B, Cannon SB, Bayer PE, Shu S, Brown AV, Ren L, Jenkins J, Chung CY, Chan TF, Daum CG, Plott C, Hastie A, Baruch K, Barry KW, Huang W, Patil G, Varshney RK, Hu H, Batley J, Yuan Y, Song Q, Stupar RM, Goodstein DM, Stacey G, Lam HM, Jackson SA, Schmutz J, Grimwood J, Edwards D, Nguyen HT (2019) Construction and comparison of three reference-quality genome assemblies for soybean. *Plant J* 100(5):1066–1082

- Valliyodan B, Brown AV, Wang J et al (2021) Genetic variation among 481 diverse soybean accessions, inferred from genomic re-sequencing. *Sci Data* 8:50
- VanToai TT, Beuerlein JE, Schmitthenner AF, St Martin SK (1994) Genetic variability for flooding tolerance in soybean. *Crop Sci* 34:1112–1115
- VanToai TT, St Martin SK, Chase K, Boru G, Schnipke V, Schmitthenner AF, Lark KG (2001) Identification of a QTL associated with tolerance of soybean to soil waterlogging. *Crop Sci* 41:1247–1252
- VanToai TT, Hoa TC, Hue TN, Nguyen HT (2010) Flooding tolerance of soybean [*Glycine max* (L) Merr] germplasm from Southeast Asia under field and screen-house environments. *The Open Agriculture Journal* 4(1):38–46
- Vodkin LO, Khanna A, Shealy R, Clough SJ, Gonzalez DO, Philip R, Zabala G, Thibaud-Nissen F, Sidarous M, Stromvik MV, et al (2004) Microarray analysis for global expression constructed with a low redundancy set of 27,500 sequenced cDNAs representing an array of developmental stages and physiological conditions of the soybean plant. *BMC Genomics* 5:73
- Vu HTT, Kilian A, James AT, Bielig LM, Lawn RJ (2015) Use of DArT molecular markers for QTL analysis of drought-stress responses in soybean. II. Marker identification and QTL analyses. *Crop Pasture Sci* 66:817–830
- Wang YH, Zhang XJ, Fan SJ (2015) Genetic diversity of wild soybean populations in Dongying, China, by simple sequence repeat analysis. *Genet Mol Res* 14(3):11613–11623
- Wang LS, Chen QS, Xin DW, Qi ZM, Zhang C, Li SN, Jin Y, Li M, Mei HY, Su AY et al (2018a) Overexpression of *GmBIN2*, a soybean glycogen synthase kinase 3 gene, enhances tolerance to salt and drought in transgenic *Arabidopsis* and soybean hairy roots. *J Integr Agri* 17:1959–1971
- Wang L, Liu L, Ma Y, Li S, Dong S, Zu W (2018b) Transcriptome profiling analysis characterized the gene expression patterns responded to combined drought and heat stresses in soybean. *Comput Biol Chem* 77:413–429
- Wang LX, Guan RX, Li YH, Lin FY, Luan WJ, Li W, Ma YS, Liu ZX, Chang RZ, Qiu LJ (2008a) Genetic diversity of Chinese spring soybean germplasm revealed by SSR markers. *Plant Breed* 127:56–61
- Wang F, Zhao TJ, Yu DY, Chen SY, Gai JY (2008b) Inheritance and QTL analysis of submergence tolerance at seedling stage in soybean. *Acta Agron Sin* 34:748–753
- Wei W, Liang DW, Bian XH, Shen M, Xiao JH, Zhang WK, Ma B, Lin Q, Lv J, Chen X, Chen SY, Zhang JS (2019) *GmWRKY54* improves drought tolerance through activating genes in abscisic acid and Ca<sup>2+</sup> signaling pathways in transgenic soybean. *Plant J* 100(2):384–398
- Wendt T, Holm PB, Starker CG et al (2013) TAL effector nucleases induce mutations at a pre-selected location in the genome of primary barley transformants. *Plant Mol Biol* 83:279–285
- Wesley SV, Liu Q, Wielopolska A, Ellacott G, Smith N, Singh S, Helliwell C (2003) Custom knock-outs with hairpin RNA-mediated gene silencing. *Methods Mol Biol* 236:273–286
- Wu C, Sun S, Nimmakayala P, Santos FA, Meksem K, Springman R, Ding K, Lightfoot DA, Zhang HB (2004) A BAC- and BIBAC-based physical map of the soybean genome. *Genome Res* 14:319–326
- Wu XL, Zhong GH, Findley SD, Cregan P, Stacey G, Nguyen HT (2008) Genetic marker anchoring by six-dimensional pools for development of a soybean physical map. *BMC Genomics* 9:28
- Wu C, Zeng A, Chen P, Florez-Palacios L, Hummer W, Moku J, Klepadlo M, Yan L, Ma Q, Cheng Y (2017a) An effective field screening method for flooding tolerance in soybean. *Plant Breed* 136(5):710–719
- Wu C, Chen P, Hummer W, Zeng A, Klepadlo M (2017b) Effect of flood stress on Soybean seed germination in the field. *Am J Plant Sci* 8:53–68
- Wu C, Mozzoni LA, Moseley D, Hummer W, Ye H, Chen P, Shannon G, Nguyen H (2020) Genome-wide association mapping of flooding tolerance in soybean. *Mol Breed* 40(1):1–4
- Xie D, Dai Z, Yang Z, Tang Q, Deng C, Xu Y, Wang J, Chen J, Zhao D, Zhang S, Zhang S (2019) Combined genome-wide association analysis and transcriptome sequencing to identify candidate genes for flax seed fatty acid metabolism. *Plant Sci* 286:98–107

- Xie K, Yang Y (2013) RNA-guided genome editing in plants using a CRISPR-Cas system. *Mol Plant* 6:1975–1983
- Xu C, Xia C, Xia Z et al (2018) Physiological and transcriptomic responses of reproductive stage soybean to drought stress. *Plant Cell Rep* 37:1611–1624
- Yamamoto T, Narikawa T (1966) Survey report on cool weather injury in Hokkaido Part III upland crops agriculture, forestry and fisheries research council, ministry of agriculture and forestry. pp 161–180
- Yamanaka N, Ninomiya S, Hoshi M, Tsubokura Y, Yano M, Nagamura Y, Sasaki T, Harada K (2001) An informative linkage map of soybean reveals QTLs for flowering time, leaflet morphology and regions of segregation distortion. *DNA Res* 8:61–72
- Yang J, Blanchar RW (1993) Differentiating chloride susceptibility in soybean cultivars. *Agron J* 85:880–885
- Yang W, Wang M, Yue A, Wu J, Li S, Li G, Du W (2014) QTLs and epistasis for drought-tolerant physiological index in soybean (*Glycine max* L.) across different environments. *Caryologia* 67(1):72–78
- Ye H, Song L, Chen H, Valliyodan B, Cheng P, Ali L, Vuong T, Wu C, Orlowski J, Buckley B, Chen P, Shannon JG, Nguyen HT (2018) A major natural genetic variation associated with root system architecture and plasticity improves waterlogging tolerance and yield in soybean. *Plant Cell Environ* 41(9):2169–2182
- Yoo JH, Park CY, Do Heo W, Cheong MS, Park HC, Kim MC, Moon BC, Choi MS, Kang YH, et al (2005) Direct interaction of a divergent *CaM* isoform and the transcription factor, MYB2, enhances salt tolerance in *Arabidopsis*. *J Biol Chem* 280:3697–3706
- Yu Z, Chang F, Lv W, Sharmin RA, Wang Z, Kong J, Bhat JA, Zhao T (2019) Identification of QTN and candidate gene for seed-flooding tolerance in soybean [*Glycine max* (L) Merr] using genome-wide association study (GWAS). *Genes* 10(12):957
- Zeng A, Chen P, Korth K, Hancock F, Pereira A, Brye K, Wu C, Shi A (2017a) Genome-wide association study (GWAS) of salt tolerance in worldwide soybean germplasm lines. *Mol Breed* 37(3):30
- Zeng A, Lara L, Chen P, Luan X, Hancock F, Korth K, Brye K, Pereira A, Wu C (2017b) Quantitative trait loci for chloride tolerance in ‘Osage’ soybean. *Crop Sci* 57:2345–2353
- Zhang J, Wang J, Jiang W, Liu J, Yang S, Gai J, Li Y (2016) Identification and analysis of NaHCO<sub>3</sub> stress responsive genes in wild soybean (*Glycine soja*) roots by RNA-seq. *Front Plant Sci* 7:1842
- Zhang M, Liu Y, Cai H, Guo M, Chai M, She Z, Ye L, Cheng Y, Wang B, Qin Y (2020) The bZIP transcription factor GmbZIP15 negatively regulates salt- and drought-stress responses in soybean. *Intl J Mol Sci* 21(20):7778
- Zhang WB, Qiu PC, Jiang HW, Liu CY, Li CD, Hu GH, Chen QS (2012) Dissection of genetic overlap of drought and low-temperature tolerance QTLs at the germination stage using backcross introgression lines in soybean. *Mol Biol Rep* 39(5):6087–6094
- Zhang DY, Kumar M, Xu L et al (2017) Genome-wide identification of major intrinsic proteins in *Glycine soja* and characterization of GmTIP2;1 function under salt and water stress. *Sci Rep* 7:4106
- Zhang W, Liao X, Cui Y, Ma W, Zhang X, Du H, et al (2019) A cation diffusion facilitator, GmCDF1, negatively regulates salt tolerance in soybean. *PLoS Genet* 15(1):e1007798
- Zeng A, Chen P, Korth K, Hancock F, Pereira A, Brye K et al (2017c) Genome-wide association study (GWAS) of salt tolerance in worldwide soybean germplasm lines. *Mol Breed* 37:30
- Zhou QY, Tian AG, Zou HF, Xie ZM, Lei G, Huang J, Wang CM, Wang HW, Zhang JS, Chen SY (2008) Soybean WRKY-type transcription factor genes, GmWRKY13, GmWRKY21, and GmWRKY54, confer differential tolerance to abiotic stresses in transgenic *Arabidopsis* plants. *Plant Biotechnol J* 6:486–503
- Zhou Z, Jiang Y, Wang Z et al (2015) Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat Biotechnol* 33:408–414
- Zhu C, Core M, Buckler E, Yu J (2008) Status and prospects of association mapping in plants. *Plant Genome* 1:5–20



- Zhu T, Shi I, Gresshoff P, Keim P (1996) Characterization and application of soybean YACs to molecular cytogenetics. *Mol Gen Genet* 252:483–488
- Zhu YL, Song QJ, Hyten DL, Van Tassell CP, Matukumalli LK, Grimm DR, Hyatt SM, Fickus EW, Young ND, Cregan PB (2003) Single-nucleotide polymorphisms in soybean. *Genetics* 163:1123–1134