


Review

Progress in soybean functional genomics over the past decade

Min Zhang¹, Shulin Liu¹, Zhao Wang^{1,4}, Yaqin Yuan^{1,4}, Zhifang Zhang^{1,4}, Qianjin Liang^{1,4}, Xia Yang^{1,4}, Zongbiao Duan^{1,4}, Yucheng Liu¹, Fanjiang Kong², Baohui Liu², Bo Ren^{3,4} and Zhixi Tian^{1,4,*} 

¹State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Innovative Academy for Seed Design, Chinese Academy of Sciences, Beijing, China

²Innovative Center of Molecular Genetics and Evolution, School of Life Sciences, Guangzhou University, Guangzhou, China

³State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, Innovative Academy for Seed Design, Chinese Academy of Sciences, Beijing, China

⁴University of Chinese Academy of Sciences, Beijing, China

Received 9 May 2021;

revised 4 August 2021;

accepted 9 August 2021.

*Correspondence (Tel 86 10 64803672;

fax 86 10 64801202; email

zxtian@genetics.ac.cn)

Keywords: soybean, functional genomics, omics, yield components, seed composition, stress resistance, nodulation, domestication, transgenic technology.

Summary

Soybean is one of the most important oilseed and fodder crops. Benefiting from the efforts of soybean breeders and the development of breeding technology, large number of germplasm has been generated over the last 100 years. Nevertheless, soybean breeding needs to be accelerated to meet the needs of a growing world population, to promote sustainable agriculture and to address future environmental changes. The acceleration is highly reliant on the discoveries in gene functional studies. The release of the reference soybean genome in 2010 has significantly facilitated the advance in soybean functional genomics. Here, we review the research progress in soybean omics (genomics, transcriptomics, epigenomics and proteomics), germplasm development (germplasm resources and databases), gene discovery (genes that are responsible for important soybean traits including yield, flowering and maturity, seed quality, stress resistance, nodulation and domestication) and transformation technology during the past decade. At the end, we also briefly discuss current challenges and future directions.

Introduction

Cultivated soybean (*Glycine max* [L.] Merr.) was domesticated from wild soybean (*G. soja* Sieb. & Zucc.) in China approximately 5000 years ago, after which was spread worldwide (Carter *et al.*, 2004; Wilson, 2008). Currently, soybean has become one of the most economically important oilseed and biodiesel crops and also serves as a main source of protein and oil for both human food and animal feed (Hartman *et al.*, 2011). Early soybean breeding mainly relied on repeated selection of preferred seeds by farmers from cultivated population. Starting from the early 1900s, artificial hybridization was applied. The first modern soybean cultivar developed by hybridization was released in North American breeding programmes in 1940s (Rincker *et al.*, 2014; Wolfgang and Charles, 2017). Afterwards, artificial hybridization started to be widely incorporated in soybean breeding (Anderson *et al.*, 2019; Li *et al.*, 2001). Artificial hybridization greatly expanded the genetic base of developed lines and significantly improved soybean adaptation and production (Anderson *et al.*, 2019).

With the arising and developing of molecular biology technology, marker-assisted selection (MAS) was used to speed up the breeding process, particularly in development of disease- and insect pest-resistant cultivars (Li *et al.*, 2020b). Using different kinds of genetic markers, linkage and physical maps were constructed (Chan *et al.*, 2012; Cregan *et al.*, 1999; Marek

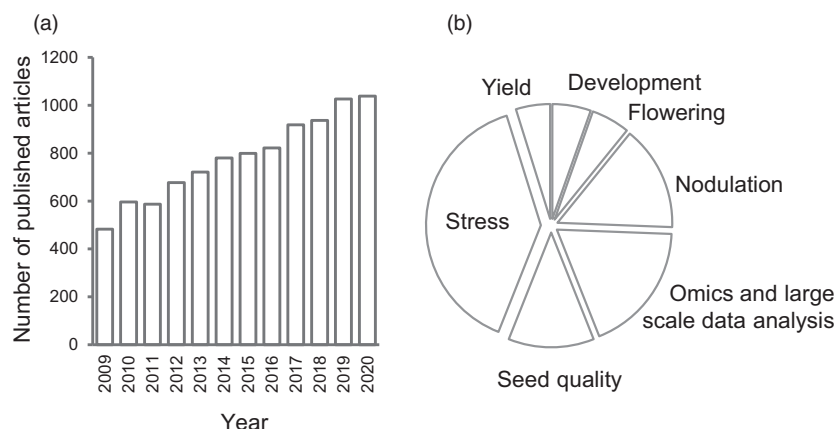
et al., 2001; Song *et al.*, 2004). By integrating available genetic maps and physical maps, the Consensus Map 4.0 was built (Hyten *et al.*, 2010a; Hyten *et al.*, 2010b). Using these markers, large numbers of quantitative trait loci (QTLs) affecting related traits have been identified in soybean. However, the limited number of molecular markers and their uneven distribution limited the efficiency and accuracy of QTL positioning.

The reference genome of a cultivated accession (Williams 82) was released in 2010 (Schmutz *et al.*, 2010), which brought about the era of soybean functional genomics (Chan *et al.*, 2012; Li *et al.*, 2017a; Wang and Tian, 2015; Xia *et al.*, 2013). Benefit of having the reference genome, research publications on soybean have almost doubled compared with ten years ago, with higher ratios on stress, omics and nodulation (Figure 1). Here, we review the advances in soybean functional genomics and transformation technology during the past decade and discuss the challenges and prospects for future soybean functional genomic studies.

Progress in omics

Advances in sequencing technologies, particularly in long-read sequencing, have led to the production of improved assembly genomes (Burton *et al.*, 2013; Yang and Huang, 2018). After the release of the reference genome based on Williams 82, Kim *et al.* (2010) assembled a genome sequence of a wild soybean

Figure 1 Statistics of publications with term 'soybean' from 2009 to 2020. (a) The number of publications with the term 'soybean' in each year. (b) The percentage of publications with the term 'soybean' and the indicated terms in different research fields.



(*G. soja* var. IT182932) using Illumina-GA and GS-FLX. Shen *et al.* *de novo* assembled a high-quality genome for cultivar 'Zhonghuang 13' (Gmax_ZH13) using single-molecule real-time (SMRT) sequencing, optical mapping, chromosome conformation capture sequencing (Hi-C) and next-generation sequencing (HiSeq) (Shen *et al.*, 2019; Shen *et al.*, 2018a). Xie *et al.* (2019) then assembled a high-quality reference genome for wild soybean W05 in 2019. Valliyodan *et al.* (2019) *de novo* assembled references for another two cultivars and one wild soybean using a combination of short- and long-read technologies. In addition to the *de novo* assembly of reference genomes, pan-genome also progressed significantly in soybean. In 2014, Li *et al.* reported the first soybean pan-genome built by assembly of seven wild soybeans decoded using second-generation sequencing technology (Li *et al.*, 2014b). Recently, Liu *et al.* (2020f) constructed a soybean pan-genome by *de novo* assembly of 26 representative wild and cultivated soybeans using long-read sequencing. This assembly produced not only golden-grade genomes for each accession, but also for the first time reported a graph-based genome in plants, which provides a promising platform for future in-depth soybean functional genomic studies (Liu and Tian, 2020; Lyu, 2020; Tao *et al.*, 2020; Willson, 2020).

The reference genome sequences revealed that soybean had undergone at least two whole-genome duplication (WGD) events within the last 60 million years (Schmutz *et al.*, 2010). It was found that approximately 50% of the paralogs arising from the WGD had undergone subfunctionalization at the expression level in soybean, suggesting that the main consequence of WGD in soybean may be at the regulatory level (Roulin *et al.*, 2013). The diversification of WGD genes was confirmed by the evolutionary and functional analyses several gene families (Li *et al.*, 2016b; Wang *et al.*, 2015d; Yang *et al.*, 2016; Zhang *et al.*, 2017). It was also found that, despite the WGD, the genes in networks with a high level of connectivity were more conserved than those with low connectivity (Kim *et al.*, 2015; Severin *et al.*, 2011). Moreover, WGD genes exhibited a decreased frequency of alternative splicing compared with non-WGD genes, which was associated with their reduced intron length, exon number and transcriptional level (Shen *et al.*, 2014). WGD also affected the evolution of transposable elements (Du *et al.*, 2010b; Du *et al.*, 2012; Tian *et al.*, 2012).

Soybean population genetics has also greatly advanced in the past ten years due to the availability of a soybean reference

genome and the development of next-generation sequencing technology. The resequencing of dozens of wild and cultivated soybean accessions revealed the consequences of the artificial selection accompanying domestication and showed that genetic diversity was significantly decreased after domestication (Lam *et al.*, 2010; Li *et al.*, 2013b). Integration of selection sweep and genome-wide association studies (GWAS) on domestication traits using the data from a resequencing of 302 wild soybeans, landraces and improved cultivars disclosed the genetic loci responsible for traits related to soybean domestication, improvement and local breeding (Zhou *et al.*, 2015). Moreover, it was found that naturally occurring introgression was widespread and counteracted genetic bottlenecks during soybean domestication (Wang *et al.*, 2019e). Hundreds of small RNAs had been identified in soybean (Arikat *et al.*, 2014; Golicz *et al.*, 2018; Lin *et al.*, 2020a; Wong *et al.*, 2011; Zhou *et al.*, 2013), some of which showed tissue-specific or time-specific transcriptional patterns, indicating their biological relevance (Arikat *et al.*, 2014; Lin *et al.*, 2020a). Population genetic analyses suggested a co-evolution of *MIRNA* and *miRNA* targets during soybean domestication (Liu *et al.*, 2016a; Zhao *et al.*, 2015a).

So far, over 6000 soybean accessions have been genotyped to investigate the footprints of breeding (Chung *et al.*, 2014; Fang *et al.*, 2017; Han *et al.*, 2016; Kajiya-Kanegae *et al.*, 2021; Liu *et al.*, 2020f; Qi *et al.*, 2021; Qiu *et al.*, 2014; dos Santos *et al.*, 2016; Shimomura *et al.*, 2015; Torkamaneh *et al.*, 2018; Valliyodan *et al.*, 2016; Zhou *et al.*, 2015) (Table 1). The large amount of resequencing data has generated a great number of nucleotide polymorphism (SNP) markers, which greatly facilitate the haplotype map construction (Torkamaneh *et al.*, 2021) and SNP Arrays development (Lee *et al.*, 2015; Song *et al.*, 2020; Wang *et al.*, 2016a), and also increased the efficiency and accuracy of gene/QTL mapping (Baird *et al.*, 2008; Bandillo *et al.*, 2017; Lee *et al.*, 2015; Patil *et al.*, 2016).

In addition to the progresses in genomics, studies on soybean transcriptomics, epigenetics and proteomics were also explored. Transcriptome of different tissues from different developmental stages illuminated gene expression profiling at a whole-genome level (Severin *et al.*, 2010; Wang *et al.*, 2014b). Integration of the gene co-expression network from RNA-seq data of 1978 samples with previously reported QTLs identified a candidate gene that may control flowering time in soybean (Shen *et al.*, 2018a). Meanwhile, large numbers of differentially expression genes (DEGs) that may be related to stresses were identified using the

Table 1 Whole-genome sequencing in soybean.

Accession information	Method	Accession number	Reference
Williams 82 (cultivar)	De novo sequencing and assembly	GCA_000004515.3	Schmutz <i>et al.</i> (2010)
14 cultivars; 17 wilds	Re-sequencing	SRP045129	Lam <i>et al.</i> (2010)
T182932 (wild)	De novo sequencing and assembly	SRP009252	Kim <i>et al.</i> (2010)
10 cultivars; 5 wilds	Re-sequencing	ERP002622	Chung <i>et al.</i> (2014)
7 wilds	De novo sequencing and assembly	PRJNA195632	Li <i>et al.</i> (2014b)
9 semi-wilds; Maliaodou (semi-wild); Lanxi1(wild)	Re-sequencing; De novo sequencing and assembly	PRJNA227063	Qiu <i>et al.</i> (2014)
W05 (wild)	De novo sequencing and assembly	GCA_000722935.2	Qi <i>et al.</i> (2014)
240 cultivars; 62 wilds	Re-sequencing	SRP045129	Zhou <i>et al.</i> (2015)
Enrei (cultivar)	Reference-based assembly	GCA_001269945.2	Shimomura <i>et al.</i> (2015)
28 Brazilian soybean	Re-sequencing	PRJNA294227	dos Santos <i>et al.</i> (2016)
404 fully domesticated; 36 semi-domesticated; 72 non-domesticated	Re-sequencing		Han <i>et al.</i> (2016)
7 wilds; 43 landraces; 56 cultivars	Re-sequencing	SRP062245	Valliyodan <i>et al.</i> (2016)
291 landraces; 278 cultivars	Re-sequencing	PRJCA000205	Fang <i>et al.</i> (2017)
102 cultivars	Re-sequencing	SRP094720	Torkamaneh <i>et al.</i> (2018)
Zhonghuang 13 (cultivar)	De novo sequencing and assembly	CRA001007	Shen <i>et al.</i> (2018a)
Zhonghuang 13 (cultivar)	De novo sequencing and assembly	CRA001810	Shen <i>et al.</i> (2019)
W05 (wild)	De novo sequencing and assembly	SRP158454	Xie <i>et al.</i> (2019)
1 wild; 2 cultivars	De novo sequencing and assembly	GCA_002907465.1; GCA_002905335.1 & PRJNA48389	Valliyodan <i>et al.</i> (2019)
3 wilds; 9 landraces; 14 cultivars	De novo sequencing and assembly	PRJCA002030	Liu <i>et al.</i> (2020f)
177 landraces; 21 breeding lines	Re-sequencing	PRJDB7281	Kajiya-Kanegae <i>et al.</i> (2021)
134 cultivars	Re-sequencing	SRP062560	Qi <i>et al.</i> (2021)

transcriptome data from plants subjected to stress treatments, such as drought, flooding, salt or heat (Chen *et al.*, 2016; Liu *et al.*, 2019; Wang *et al.*, 2018a; Xu *et al.*, 2018). These candidate genes provide clues for further functional study. Recently, 1298 publicly available soybean transcriptome datasets were combined to generate a comprehensive atlas of expression, which can be accessed through the website http://venanciogroup.uenf.br/cgi-bin/gmax_atlas/index.cgi (Machado *et al.*, 2020).

DNA methylation profiling analyses revealed that hypomethylation could affect the expression of neighbouring genes (Song *et al.*, 2013b). Kim *et al.* (2015) found that CG body-methylated genes were abundant in duplicated genes that exhibited higher expression level than single copy genes. Moreover, methylation changes proximal to the transposable elements (TEs) were associated with the divergence of expression profiles of duplicated genes (El Baidouri *et al.*, 2018; Hossain *et al.*, 2017). Shen *et al.* (2018b) investigated the variation of DNA methylation during soybean domestication and improvement and found that differentially methylated regions that are not associated with any genetic variation were enriched in carbohydrate metabolism pathways. It was found that DNA demethylation/methylation also plays critical roles in stress responses, such as continuous cropping (Liang *et al.*, 2019), salinity stress (Song *et al.*, 2012) and cyst nematode infection (Rambani *et al.*, 2020).

Using proteomics approaches, numbers of genes response to various stresses were identified (Khan and Komatsu, 2016; Khatoon *et al.*, 2012; Komatsu *et al.*, 2010; Komatsu *et al.*, 2017; Komatsu *et al.*, 2011; Sobhanian *et al.*, 2010; Wang and Komatsu, 2018). Similarly, the protein changes through different developmental stages were also investigated (Agrawal *et al.*,

2008; Arai *et al.*, 2008). Along with the research progresses in proteomics, several proteomics databases were provided. The Soybean Proteome Database (SPD) stores soybean proteomics data obtained from both gel-based and gel-free techniques (Komatsu *et al.*, 2017; Ohyanagi *et al.*, 2012). A user-intuitive database (<http://oilseedproteomics.missouri.edu>) stores the expression profile data for proteomics research on soybean and other oilseeds plant (Agrawal *et al.*, 2008).

Germplasm resources and databases

Soybean is rich in germplasm resources that carry large amounts of variations. Over the long history of soybean cultivation, more than 60 000 accessions have been developed (Carter *et al.*, 2004; Wilson, 2008). In China, more than 40 000 soybean accessions were collected from all over the nation and stored in the Chinese Crop Germplasm Centre. In the United States, over 20 000 accessions were collected from around the world and stored at the U.S. Department of Agriculture (Song *et al.*, 2015). Investigations of the population structure and genetic diversity of the core collections suggested that the accessions exhibit geographical patterns (Haupt and Schmid, 2020; Li *et al.*, 2008).

Mutant populations are important in facilitating gene cloning and functional analysis. Through fast neutron, gamma radiation and ethylmethane sulphonate (EMS) mutagenesis, several soybean mutant populations had been constructed (Bolon *et al.*, 2014; Espina *et al.*, 2018; Li *et al.*, 2017e; Tsuda *et al.*, 2015). These populations enabled the identification of some genes responsible for important traits (Anderson *et al.*, 2019). For example, using TILLING, Liu *et al.* (2012) identified a soybean cyst nematode-

Table 2 Soybean databases

Database	URL	Description	Reference
Soybean gene expression atlas	http://www.soybase.org/soyseq	a database of soybean 14 tissues specific gene expression	Severin <i>et al.</i> (2010)
Soybean cDNA sequenced	http://digbio.missouri.edu/soybean_atlas	a cDNA database of soybean developmental tissues specifically in root hair and meristem	Libault <i>et al.</i> (2010)
SoyNet	http://www.inetbio.org/soynet	a database for network-based functional predictions	Kim <i>et al.</i> (2017)
Soybean transcriptome data	http://venanciogroup.uenf.br/cgi-bin/gmax_atlas/index.cgi	a database of 1,298 publicly available soybean transcriptome	Machado <i>et al.</i> (2020)
Proteomics of oilseeds	http://oilseedproteomics.missouri.edu	expression profile data for proteomics research on soybean and other oilseeds plants	Agrawal <i>et al.</i> (2008)
Soybean Proteome Database SPD	http://proteome.dc.affrc.go.jp/Soybean/	a database of soybean proteomics	Ohyanagi <i>et al.</i> (2012); Komatsu <i>et al.</i> (2017)
SoyBase	http://www.soybase.org	a database of soybean genetics and genomics	Grant <i>et al.</i> , 2010; Brown <i>et al.</i> (2020)
SoyTEdb	http://www.soytedb.org	a database of soybean transposable elements	Du <i>et al.</i> (2010a)
SoyProDB	http://bioinformatics.towson.edu/Soybean_Seeds_Proteins_2D_Gel_DB/Home.aspx	a database for soybean seed proteins	Tavakolan <i>et al.</i> (2013)
SoyProLow	http://bioinformatics.towson.edu/Soybean_low_abundance_proteins_2D_Gel_DB/Gel1.aspx	a database for soybean low abundant proteins	Tavakolan <i>et al.</i> (2014)
SoyKB	http://soykb.org	a database of soybean translational genomics and for soybean molecular breeding	Joshi <i>et al.</i> (2014)

related gene, *GmSHMT*, at the *Rhg4* locus. Dobbels *et al.* (2017) identified an oil biosynthesis-related gene (*KAS1*) through forward screening of a soybean fast neutron mutant population. Reinprecht *et al.* (2009) found that mutations in two *Fad3* genes lead to a low level of linolenic acid in the EMS mutant *RG10*.

Several soybean databases related to soybean genomics, transcriptomics, proteomics and germplasm analyses have been developed. These freely available databases include SoyBase, a database for genetics and genomics from USDA-ARS (Brown *et al.*, 2020; Grant *et al.*, 2010); SoyTEdb, a database of transposable elements (Du *et al.*, 2010a); SoyNet, a database for co-functional networks (Kim *et al.*, 2017); SoyProDB: a database for seed proteins (Tavakolan *et al.*, 2013); SoyProLow: a database for low-abundance proteins (Tavakolan *et al.*, 2014); and SoyKB, a database for translational genomics and molecular breeding (Joshi *et al.*, 2014; Table 2). These databases provide valuable platforms for soybean research.

Genes responsible for important agronomic traits

In the past decade, great efforts have been made to identify the genetic loci and the genes responsible for different agronomic traits, including yield, seed quality, stress, development and domestication. The efforts led to the identification of hundreds of QTLs (Hacisalihoglu *et al.*, 2018), part of which have been integrated into the Soybase database (<https://www.soybase.org/>). Despite numerous QTLs were identified, only small number of responsible genes for these QTLs have been cloned, leaving most

of them need to be cloned and functionally characterized. Here, we focus on a review of the genes that have been functionally validated (Table S1).

Yield components

Yield is one of the breeding priorities for soybean. Soybean yield is determined mainly by plant architecture, seed weight and size, and seed number per pod. In the past decade, several key genes controlling yield and related traits were cloned.

Plant architecture

Plant architecture is a key trait that significantly affects the yield of soybean. *Dt1* and *Dt2* are two key genes for plant height and growth habits (Liu *et al.*, 2010; Ping *et al.*, 2014; Tian *et al.*, 2010). Analyses of population genetics suggested that *Dt1*, which is a homolog of *Arabidopsis Terminal Flower 1 (TFL1)*, underwent artificial selection to create a determinate growth habit at early stages of landrace dissemination (Tian *et al.*, 2010). *Dt2* encodes a MADS-domain factor (Ping *et al.*, 2014) that could bind the promoter of *Dt1* and repress its transcription (Liu *et al.*, 2016b). Further functional investigation showed that *Dt2* was a direct activator or repressor of the precursors of eight miRNAs whose target genes were associated with meristem maintenance, flowering time, stomatal density, water-use efficiency (WUE) and/or stress responses (Zhang *et al.*, 2019a). Other genes regulating plant architecture have also been characterized, such as *SQUAMOSA-Promoter Binding Protein-Like (SPL)* and *Late Elongated Hypocotyl (LHY)*. The *spl9a/spl9b-1/spl9c/spl9d* homozygous quadruple mutant plants had more branches than

WT (Bao et al., 2019). *SPL* was regulated by miR156 to affect plant architecture in soybean (Sun et al. 2019; Wang and Wang 2015). When miR156b was overexpressed, the transgenic lines increased yield per plant by 46%–63%, which was resulted from improved long branches number, nodes and pods number, 100-seed weight (Sun et al., 2019). In soybean, the quadruple mutant of *GmLHY* displayed reduced plant height and shortened internodes (Cheng et al., 2019). Recently, *GmMYB14* were characterized as an important factor to regulate plant architecture, high-density yield and drought tolerance through the brassinosteroid (BR) pathway in soybean (Chen et al., 2021).

Seed size

In *Arabidopsis*, *P450/CYP78A* gene family had been found to control seed size (Fang et al., 2012; Wang et al., 2008). The orthologs of *P450/CYP78A* in soybean, including *GmCYP78A10*, *GmCYP78A72*, *GmCYP78A57* and *GmCYP78A70*, showed conserved function to control seed size (Du et al., 2017; Wang et al., 2015b; Zhao et al., 2016). The genes *BIG SEEDS1 (BS1)* (Ge et al., 2016), *GmKIX8-1* (Nguyen et al., 2020), *GmCIF1* (Tang et al., 2017), *GmPSK1* (Yu et al., 2019), *WRKY15a* (Gu et al., 2017) and *phosphatase 2C-1 (PP2C-1)* (Lu et al., 2017b) were also involved in seed development in soybean. Furthermore, *SoyWRKY15a* and *PP2C-1* were found to have undergone artificial selection during soybean domestication (Gu et al., 2017; Lu et al., 2017b).

Flowering

Soybean is a short-day and photoperiod sensitive plant (Miranda et al., 2020). Genetic analyses of natural variants have identified 12 maturity loci—*E1* to *E11* and *J*—that control flowering time and maturity. In the past decade, significant progress has been made in soybean flowering gene identification and functional characterization (Lin et al., 2020b).

E1 encodes a soybean-specific transcription factor that contains a plant-specific B3 domain (Xia et al., 2012). *E1* could repress flowering under long-day (LD) conditions, while its leaky allele (*e1-as*) and null alleles (such as *e1-fs* and *e1-nl*) caused earlier flowering (Xia et al., 2012; Xu et al., 2013a; Yasutaka et al., 2013). *E2* encodes a homolog of *Arabidopsis GIGANTEA (GI)*, which is a component of the circadian clock (Watanabe et al., 2011). The functionally dominant *E2* delayed flowering time, while the recessive *e2* induced flowering (Watanabe et al., 2011). *E3* and *E4* were found to be functionally redundant and encode *GmPHYA3* (Watanabe et al., 2009) and *GmPHYA2* (Liu et al., 2008; Yasutaka et al., 2013), respectively, which are homologues of the photoreceptor phytochrome A (PHYA) (Franklin and Quail, 2010). *E3* and *E4* were involved in the control of flowering under LDs with high red:far-red (R:FR) quantum ratios and low R:FR ratios, respectively (Cober et al., 1996; Cober and Voldeng, 2001). *J* is an important locus for the adaptation of soybean to lower latitudes. *J* was found to encode an EARLY FLOWERING 3 (ELF3) homolog (Fang et al., 2020; Lu et al., 2017a; Yue et al., 2017). A pair of pseudo-response-regulators (PRRs; named *Tof11* and *Tof12*) were reported to contribute to changes in flowering and early maturity in soybean (Day et al., 2010; Gong, 2020; Li et al., 2020a; Lu et al., 2020b; Wang et al., 2020c).

Other genes have also been reported to function in regulating flowering in soybean, particularly genes homologous to flowering-related genes from other species. *FLOWERING LOCUS T (FT)* is required for flowering and widely conserved among plant

species. In soybean, 10 *FT/TSF* homologs were identified. These *FT/TSF* homologs had divergence functions, and their natural variation might be associated with soybean adaptation to different environments (Jiang et al., 2019; Kong et al., 2010; Wu et al., 2017). Overexpression of *GmFT2a/E9* or *GmFT5a* promoted flowering in soybean, whereas overexpression of *GmFT1a* or *GmFT4* suppressed flowering (Cai et al., 2018; Cai et al., 2020b; Chen et al., 2020b; Kong et al., 2010; Liu et al., 2018; Nan et al., 2014; Takeshima et al., 2016; Zhai et al., 2014a). *FT4* is a candidate gene for the *E10* locus (Samanfar et al., 2017). Although *FT2a* and *FT5a* showed similar functions in inducing flowering, *FT5a* was more effective than *FT2a* in the post-flowering termination of stem growth (Takeshima et al., 2019). Further investigation showed that several floral genes, such as *GmAP1*, *GmSOC1* and *GmLFY*, were significantly induced by *GmFT2a/E9* and *GmFT5a* (Nan et al., 2014). Overexpression of *GmAP1a*, one of the homologs of *Arabidopsis APETALA1 (AP1)* which functions as a class A gene in the ABCDE model of flowering, resulted in early flowering and reduced plant height compared with the wild type under short-day (SD) conditions (Chen et al., 2020c). *GmFDL19*, which is a bZIP transcription factor, interacted with *GmFT2a/E9* and *GmFT5a*. Overexpression of *GmFDL19* in soybean resulted in early flowering through the regulation of the expression of *GmAP1a* (Nan et al., 2014). *GmGBP1* functioned as a positive regulator upstream of *GmFT2a/E9* and *GmFT5a* to activate the expression of *GmFULs* to promote flowering under LDs (Zhao et al., 2018).

In the photoperiod-controlled flowering pathway, the circadian clock-regulated gene *CONSTANS (CO)* is crucial for the induction of the *FT* gene (Turck et al., 2008). The soybean genes *GmCOL1a*/*GmCOL1b* showed high sequence homology to *AtCO*. The late-flowering phenotype of the *Arabidopsis co-1* mutant was fully complemented by overexpression of *GmCOL1a* or *1b*, suggesting that they function similarly to *AtCO* at the protein level (Wu et al., 2014). Further analysis showed that *GmCOL1a* controlled flowering time by suppressing the expression of *GmFT2a/E9* and *GmFT5a* under LDs. In addition, *GmCOL1a/1b* was up-regulated by *E1*, *E2*, *E3* and *E4* (Cao et al., 2015a). It was found that overexpression of *GmAGL1*, which encodes a MADS-box protein, promoted flowering, maturity and led to a smaller floral organ (Chi et al., 2017; Zeng et al., 2018). Moreover, several miRNA families were also demonstrated to play important roles in controlling flowering in soybean. Overexpression of *MIR156b*, which is involved in the PHYA-mediated photoperiod response pathway, delayed flowering under LDs (Cao et al., 2015b). In *Arabidopsis*, miR156 regulated the expression of miR172 via *SPL9* (Wu et al., 2009). In soybean, miR156b down-regulated *MIR172* and *SPL9* and up-regulated the miR172 target gene *GmTOE4a* in soybean (Cao et al., 2015b). In addition, *E2* could promote the maturation of *MIR172a* via increasing the expression of DICER-LIKE 1 and SERRATE homologs (Wang et al., 2016b).

Functional investigations have demonstrated that the above genes are involved in a complicated PHYA-mediated photoperiod response pathway that regulates flowering time in soybean (Kong et al., 2010; Li et al., 2020a; Lu et al., 2020b; Lu et al., 2017a; Xia et al., 2012; Xu et al., 2015). In Figure 2, we summarized the current understanding of this flowering time regulation under LDs. Briefly, *E3* and *E4* were up-regulated *Tof11* and *Tof12* expression. Then, *Tof11* and *Tof12* proteins physically could bind to the promoters of the *GmLHY* genes to suppress their expression, which suppressed their subsequent induction of *J*. At low levels, *J* is not able to suppress the transcription of *E1*. The

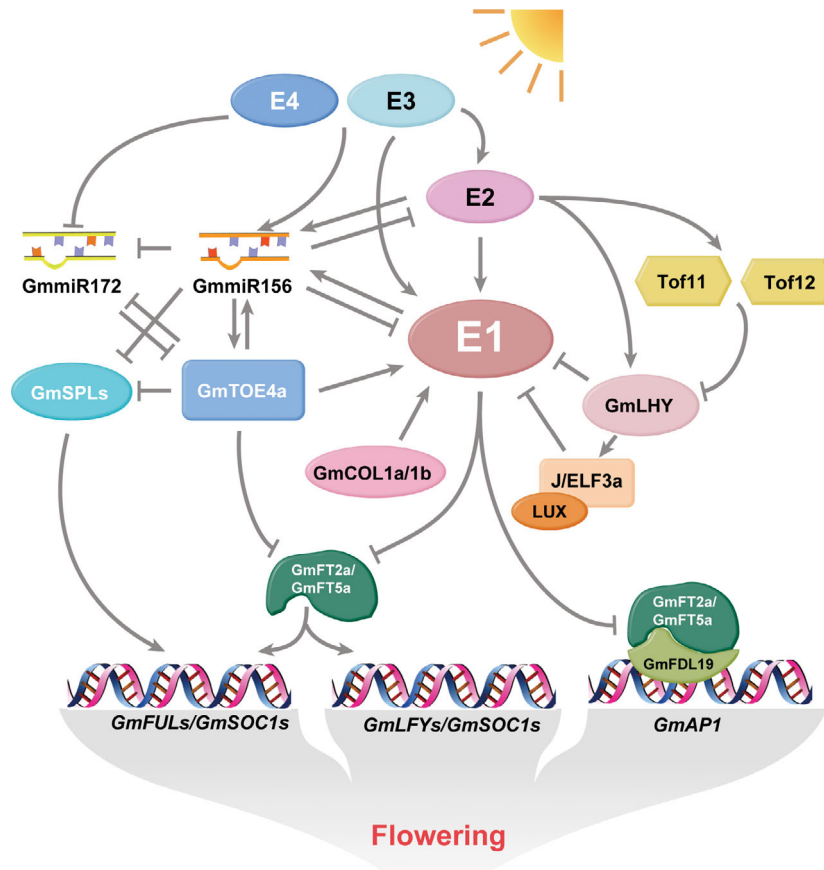


Figure 2 Proposed molecular regulation network of flowering in soybean. Soybean is a short-day flowering plant. E3 and E4 mediate flowering by responding to the ratios of red (R) and far-red (FR) light. Under long days (LDs), E3 and E4 induce the expression of *E1*. GmTof11 and GmTof12 inhibit the expression of *GmLHYs* by binding to their promoters. GmLHY proteins directly bind to the promoter region of *J* to induce its expression and bind to the promoter of *E1* to suppress its transcription. LUX can physically interact with GmELF3a/b to repress *E1* expression. *E1* inhibits the expression of the flowering-inducing factors *GmFT2a* and *GmFT5a*. This suppresses the expression of floral identity genes (*GmAP1*, *GmSOC1s*, *GmLFYs*, *GmFULs*). As a result, flowering is delayed under LDs. Under SDs, the induction of *E1* is decreased, which releases the transcriptional suppression of *GmFT2a* and *GmFT5a*, promoting flowering. The E2 and miRNA-dependent flowering regulation modules are also influenced by photoperiod under PHYA mediation (E3 and E4). GmCOL1a/1b and GmTOE4a inhibit flowering via down-regulation of *GmFT2a* and *GmFT5a* expression. *GmTOE4a* expression is possibly controlled by miR172, which is under the control of miR156b. miR156b may delay flowering by negatively regulating the *GmSPLs*.

abundance of *E1* decreases the expression of *GmFT2a/E9* and *GmFT5a*, ultimately resulting in delayed flowering and later maturity (Figure 2). Recently, it was found that GmLUX2, an ortholog of the *Arabidopsis* Evening Complex (EC) component LUX ARRHYTHMO (LUX), could physically interact with GmELF3a/b to regulate the expression of several circadian clock-associated genes and directly repress *E1* expression (Fang *et al.*, 2021a) (Figure 2).

Others

Other genes that control yield-related developmental traits were also identified. *Ln*, which encodes a C2H2 zinc finger transcription factor, was found to be a key regulator of seed number per pod (Fang *et al.*, 2013; Jeong *et al.*, 2011; Jeong *et al.*, 2012). *GmlPA1*, which encodes an APC8-like protein, was found to control petiole angle (Gao *et al.*, 2017). *Chicken Toes-Like Leaf and Petalody Flower (CTP)* was found to be a novel regulator that controlled leaf and flower development in soybean. Mutation of *CTP* resulted in leaf defects and additional flower parts

developing into petals (Zhao *et al.*, 2017). It was found that *GmCRY1s* could modulate gibberellin metabolism during the response to reduced blue light, affecting yield in a shaded environment (Lyu *et al.*, 2021).

Several genes controlling senescence are also identified. *D1* and *D2*, homologs of *STAY-GREEN (SGR)*, and *GmCHL1*, a subunit of magnesium (Mg)-chelataase, regulated chlorophyll degradation in soybean (Fang *et al.*, 2014; Li *et al.*, 2016a; Slattery *et al.*, 2016). The maternally inherited stay-green gene *cytG* was found to encode the chloroplast protein PsbM (Kohzuma *et al.*, 2017). Under blue light, cryptochrome2 (*CRY2a*) interacts with the soybean basic helix-loop-helix transcription activator CIB1 to regulate leaf senescence in soybean (Meng *et al.*, 2013).

Seed composition

Protein, oil and isoflavone content are the main soybean seed quality traits and are determined via interaction between quantitative loci and the environment. Seed storage proteins usually showed negative correlation with oil content (Chaudhary *et al.*,

2015) and seed yield (Bandillo *et al.*, 2015; Chung *et al.*, 2003; Zhang *et al.*, 2018). Besides, the contents of soluble sugar, mineral nutrition and amino acid are also important for seed quality. Hundreds of QTLs that affected soybean seed quality have been identified using both pedigree-based and association mapping approaches. Some of this information has been integrated into Soybase (<https://www.soybase.org/>).

Oil content and fatty acid

Triacylglycerols (TAGs) are the main components of soybean seed oil. Lines carrying knockdowns of four TAG lipase-encoding genes, namely *GmSDP1-1*, *GmSDP1-2*, *GmSDP1-3* and *GmSDP1-4*, have significantly increased seed oil content (Kanai *et al.*, 2019). Overexpression of the phospholipid diacylglycerol acyltransferase *GmPDAT* increased seed oil and also altered size-related traits (Liu *et al.*, 2020a; Liu *et al.*, 2020b). A comprehensive assessment of the gene co-expression networks showed that *GA20 oxidase* (*GA20OX*) and *nuclear transcription factor Y subunit alpha* (*NFYA*) are two key drivers of seed traits. Overexpression of *GA20OX* and *NFYA* enhanced seed size and weight and oil content, respectively (Lu *et al.*, 2016). It was found that oleosin-encoding *GmOLEO1* was predominantly expressed during seed maturation, and overexpression of *GmOLEO1* significantly increased seed oil content by 10.6% (Zhang *et al.*, 2019b). Overexpression of *GmWRI1a*, an APETALA2/ethylene responsive element-binding protein (AP2/EREBP) encoding gene, under the control of a seed-specific promoter significantly increased total oil and fatty acid content and also changed fatty acid composition (Chen *et al.*, 2018b). Genetic modification of the fatty acid desaturase *FAD* significantly changes the composition of polyunsaturated and polyunsaturated fatty acids (Al Amin *et al.*, 2019; Andreu *et al.*, 2010; Combs and Bilyeu, 2019; Haun *et al.*, 2014; Lakhssassi *et al.*, 2017; Lim *et al.*, 2014; Wagner *et al.*, 2011). Genetic modification of *DGAT*, which encodes a type 1 diacylglycerol acyltransferase, increased the oil content and decreases the soluble carbohydrate content in soybean seeds (Roesler *et al.*, 2016).

Several transcription factors were also found to be associated with seed quality. Overexpression of transcription factors *GmbZIP123*, *GmMYB73*, *GmZF351*, *GmZF392* and *GmWRI1b* enhanced lipid content (Guo *et al.*, 2020; Li *et al.*, 2017b; Liu *et al.*, 2014b; Song *et al.*, 2013a). Further functional investigation showed that the enhancement of oil content by *GmMYB73* was through interaction with *GL3* and *EGL3* and the suppression of *GL2*, a negative regulator of oil accumulation (Liu *et al.*, 2014b). *GmZF351* was more highly expressed in cultivated soybeans compared with wild soybeans, and this was due to a mutation of its promoter, which had undergone artificial selection during soybean domestication (Li *et al.*, 2017b). *GmZF392* physically interacted with *GmZF351*, and both *GmZF392* and *GmZF351* could be further induced by *GmNFYA* (Lu *et al.*, 2021).

Isoflavone

Isoflavone content is highly influenced by the genes involved in chalcone synthesis, including chalcone synthase, chalcone reductase and chalcone isomerase (Dastmalchi *et al.*, 2016). Among 11 chalcone reductase (CHR) proteins of soybean, *CHR5* was found to interact with a 2-hydroxyisoflavanone synthase (IFS) isozyme and most correlated with the distribution patterns of 5-deoxyisoflavonoids, indicating that it may play an important role in the isoflavone pathway (Mameda *et al.*, 2018). In soybean,

some prenyltransferase-encoding genes exhibited isoflavonoid-specific patterns (Sukumaran *et al.*, 2018). When the gene encoding carotenoid cleavage dioxygenase 4 (*GmCCD4*) was knocked out, the loss-of-function lines showed increased carotenoid content and yellow flowers (Gao *et al.*, 2021). Several independent studies demonstrated that the *MYB* gene family played important roles in affecting isoflavone synthesis, which may involve a pathway that consists of 14-3-3s and CHS proteins (Chu *et al.*, 2017; Li *et al.*, 2012; Pandey *et al.*, 2014; Yi *et al.*, 2010). Recently, it was reported that selection of a class B heat-shock factor, *HSFB2b*, in soybean domestication promoted flavonoid biosynthesis and enhanced salt tolerance (Bian *et al.*, 2020). An analysis combining GWAS and QTLs identified *GmMPK1*, which encodes a mitogen-activated protein kinase, as one of the candidate genes responsible for isoflavone content (Wu *et al.*, 2020).

Protein content, amino acid and others

Compared with genes for oil content, fewer genes controlling protein content or amino acid have been functionally identified (Krishnan and Jez, 2018). It was found that MGL, a putative methionine γ -lyase, may be responsible for the accumulation of S-methylmethionine in soybean seed (Teshima *et al.*, 2020). Overexpression of the cytosolic isoform of O-acetylserine sulfhydrylase (OASS) and the plastid ATP sulfurylase isoform 1 improved the cysteine-rich proteins and sulphur amino acid content in transgenic soybean, independently (Kim *et al.*, 2012b; Kim *et al.*, 2020). *Rab5a*, a small GTPase-encoding gene, was reported to be involved in post-Golgi trafficking of storage proteins in developing soybean cotyledons (Wei *et al.*, 2020b). Recently, it was found that SWEET sugar transporters played important roles in soybean seed quality through effects on the contents of both oil and protein. *GmSWEET15* mediated sucrose export from endosperm to early embryo (Wang *et al.*, 2019d). *GmSWEET10a* and *GmSWEET10b* determined oil and protein contents and seed size simultaneously in soybean through affecting sugar allocation from seed coat to embryo (Miao *et al.*, 2020; Wang *et al.*, 2020e; Zhang *et al.*, 2020c).

Stress resistance

Abiotic and biotic stresses greatly affect soybean yield and quality (Anderson *et al.*, 2019). Drought and salinity can reduce soybean yield by 40% through negative impacts on growth, nodulation, flowering, seed quality and seed quantity (Anderson *et al.*, 2019; Papiernik *et al.*, 2005; Specht *et al.*, 1999). Stress (including drought, salt, temperature stress, flooding and disease) has been intensely studied in soybean (Kunert *et al.*, 2016; Li *et al.*, 2021; Li *et al.*, 2014a; Phang *et al.*, 2008; Ramesh *et al.*, 2019; Robison *et al.*, 2019; Shu *et al.*, 2020; Whitham *et al.*, 2016; Widyasari *et al.*, 2020).

Drought

Several transcription factors were reported to play important roles in soybean drought tolerance, such as *WRKY* and *NAC* gene families. *GmWRKY54* increased drought tolerance by activating genes in the abscisic acid and Ca^{2+} signalling pathways (Wei *et al.*, 2019). Overexpression of *WRKY20* from a wild soybean significantly enhanced drought tolerance in *Arabidopsis* (Luo *et al.*, 2013). *GmNAC8* overexpression and knockout transgenic lines exhibit significantly higher and lower drought tolerance, respectively, suggesting that *GmNAC8* is a positive regulator of drought tolerance (Yang *et al.*, 2020).

U-box (PUB) proteins function as E3 ligases in plants. When *GmPUB6* was overexpressed in *Arabidopsis*, the transgenic lines exhibited decreased plant survival under drought stress condition (Wang *et al.*, 2020d). Several reports suggested that circadian clock genes are also involved in drought tolerance, such as the *GmLHYs* and *GmLCL*, could affect drought tolerance (Wang *et al.*, 2021a; Yuan *et al.*, 2021). It is reported that many clock genes respond to both flooding and drought, with the expression patterns of some genes shifting in amplitude and phase (Marcolino-Gomes *et al.*, 2014; Syed *et al.*, 2015).

Other genes that affect drought tolerance in soybean have also been identified. For instance, the genes responsible for pubescence density, *Pd1*, *P1*, *Ps* and *Pd2*, were also found to be involved in drought tolerance in soybean (Liu *et al.*, 2020c; Pfeiffer *et al.*, 2003). Overexpression of the *LOS5/ABA3* gene, which encodes a molybdenum cofactor sulfuryase, using a constitutive expression promoter in soybean enhanced drought tolerance and increased seed yield by at least 21% under drought stress (Li *et al.*, 2013a). Drought tolerance in soybean through foreign gene transformation was also achieved, such as the overexpression of rice *cystatin oryzacystatin I (OCI)*, *Arabidopsis DREB1A*, and the sunflower transcription factor *HB4* (Quain *et al.*, 2014; Rakocevic *et al.*, 2018; Ribichich *et al.*, 2020).

Salt

Stresses, such as salinity, osmotic stress, imbalance of ions, ion toxicity and excessive reactive oxygen, have significant effects on the growth of plant (Ruiz-Lozano *et al.*, 2012). Under high salinity, plants slow down photosynthesis and ramp up sugar catabolism to provide extra energy for survival (Liu *et al.*, 2019). Balancing the ions and reducing the ion toxicity are the keys to enhance the salt resistant. Through forward genetic approaches, two research groups independently found that *GmSALT3*, which encodes a cation/H⁺ exchanger, could limit the accumulation of sodium ions (Na⁺) in shoots and enhance salt tolerance in soybean (Guan *et al.*, 2014; Qi *et al.*, 2014). Further investigation suggested that *GmSALT3* functioned through exclusion of sodium ions (Na⁺) from the leaf via a root-derived mechanism and exclusion of chloride ions (Cl⁻) via a shoot-derived process (Qu *et al.*, 2021). *GmCDF1*, encoding a cation diffusion facilitator, could negatively regulate salt tolerance by maintaining K⁺-Na⁺ homeostasis in soybean (Zhang *et al.*, 2019c). *GmAKT1*, a K⁺ transporter encoding gene, located in the plasma membrane and was recently reported to play an important role in soybean salt resistance by regulating the K⁺ uptake and Na⁺/K⁺ balance (Wang *et al.*, 2021b).

Other genes that have effects on salt tolerance in soybean were also reported. *GmNAC109* and *SALT INDUCED NAC1 (GmSIN1)* were found to promote root growth and increase abiotic stress tolerance through up-regulation of ABA synthesis-associated and ROS generation genes (Li *et al.*, 2019c; Yang *et al.*, 2019). The salt-inducible gene *GmbZIP110* encodes a protein that could bind to the promoters of genes with an ACGT motif and impact the expression of many stress-related genes and enhance salt tolerance (Xu *et al.*, 2016). Overexpression of *GmNAC20* or *GmNAC11* enhanced salt tolerance in transgenic *Arabidopsis* plants (Hao *et al.*, 2011). Interestingly, miR172 was also involved in the response to salt stress. Under salinity, plants with hairy roots overexpressing the pre-miR172a had healthier leaves and larger roots. Further investigation showed that miR172a promoted salt tolerance mainly through cleaving the AP2/EREBP-type transcription factor *SSAC1* gene, releasing its

inhibition on *THI1*, which encodes a positive regulator of salt tolerance (Pan *et al.*, 2016). Overexpression of soybean *MIR172c* conferred tolerance to both water deficit and salt stress in transgenic *Arabidopsis* (Li *et al.*, 2017d). The plant homeodomain protein *GmPHD6* functioned as a histone code reader and interacted with LHP1 to form a transcriptional activator that regulated genes for salt tolerance. Overexpression of *GmPHD6* improved salt tolerance in soybean (Wei *et al.*, 2017). *HSFB2b*, a B heat-shock factor encoding gene, improved salt tolerance by promoting flavonoid biosynthesis. Interestingly, *HSFB2b* had undergone artificial selection during soybean domestication (Bian *et al.*, 2020). In addition, mitogen-activated protein kinase (MAPK), BURP-domain proteins and NAD(P)H dehydrogenase (NDH) were also found to be involved in salt tolerance in soybean (He *et al.*, 2015; Im *et al.*, 2012; Wang *et al.*, 2012a).

Genes can affect both drought and salt tolerance were also reported. For instance, overexpression of *GmMYB118*, *GmERF135*, *GmCDPK3* improved drought and salt tolerance simultaneously (Du *et al.*, 2018; Wang *et al.*, 2019a). Interestingly, it was also found that even different members from the same WRKY family having differential effects on abiotic stress tolerance in soybean (Wang *et al.*, 2015a; Wei *et al.*, 2019; Zhou *et al.*, 2008).

Minerals

Iron deficiency results in stunting and yield loss and is one of the most common and severe nutritional stresses for soybean (Hacisalihoglu *et al.*, 2018). Soybean plants reprogramme metabolism under iron deficiency (Chu *et al.*, 2019). Moreover, the circadian clock was highly responsive to iron deficiency (Li *et al.*, 2019b). Several genes that may be involved in iron deficiency regulation were identified. Silencing of *GmRPA3*, which encodes subunit 3 of replication protein A, showed reduced iron deficiency chlorosis (IDC) and increased chlorophyll content under iron-deficient conditions (Atwood *et al.*, 2014). It was found that the physical interaction between two soybean bHLH factors, *GmbHLH57* and *GmbHLH300*, was important for Fe homeostasis (Li *et al.*, 2018a). Genetic variation of the genes encoding a Fe deficiency-induced transcription factor and a Fe/Zn-regulated transporter may be responsible for variation of iron use efficiency among different soybean accessions (Liu *et al.*, 2020f; Pfeiffer *et al.*, 2012).

Several reports have focused on the investigation of phosphorus-use efficiency (PUE). *GmACP1* encodes an acid phosphatase and is a candidate for a PUE locus identified via GWAS. Overexpression of *GmACP1* in soybean hairy roots significantly increased PUE (Zhang *et al.*, 2014). Under Pi starvation, the expression of *GmPT1*, *GmPHR25* and *CWPs* was induced, indicating they may be involved in PUE regulation (Song *et al.*, 2014; Wu *et al.*, 2018; Xue *et al.*, 2017). The functions of genes in the *GmALMT* family were also investigated, and *GmALMT5* was found to enhance utilization of soluble P under P-limited conditions (Peng *et al.*, 2018). Moreover, ethylene may also affect PUE. Up-regulating *GmETO1*, which encodes ethylene-overproduction protein 1, could significantly improve phosphorus uptake and use efficiency in soybean (Zhang *et al.*, 2020d). A single point mutation in *GmHMA3*, which encodes a heavy-metal transporter, may be responsible for differential cadmium (Cd) translocation and accumulation in the seeds of different soybean accessions (Wang *et al.*, 2012b). Further, hydrogen sulphide affected aluminium and nitrogen assimilation (Wang *et al.*, 2019b; Zhang *et al.*, 2020e).

Pathogens

Soybean cyst nematode (SCN) is the most devastating soybean pathogen in the United States, with yield losses ranging from 1.9 to 3.5 million tonnes per year (Liu *et al.*, 2015; Wrather and Koenning, 2006). SCN infection induced a large number of genes related to cell wall modification, stress response, defence and signal transduction (Kandoth *et al.*, 2011; Rambani *et al.*, 2015; Tucker *et al.*, 2011). The proteins potentially related to SCN were collected in the database SCNProDB (Natarajan *et al.*, 2014).

Rhg1 and *Rhg4* are two major QTLs/genes conveying SCN resistance (Liu *et al.*, 2012; McHale *et al.*, 2012; Mitchum, 2016). *Rhg4* encodes a serine hydroxymethyltransferase (SHMT), which mediated a novel plant resistance mechanism against a pathogen (Liu *et al.*, 2012). In the *Rhg1* region, the copy number of three genes, *Glyma18g02580*, *Glyma18g02590* and *Glyma18g02610*, was associated with SCN resistance (Cook *et al.*, 2014; Cook *et al.*, 2012; Lee *et al.*, 2016). Further investigation demonstrated that the repeated copies of a gene encoding atypical α -soluble N-ethylmaleimide-sensitive factor (NSF) attachment protein (α -SNAP) were the most likely candidate for conferring resistance to SCN (Liu *et al.*, 2017; Patil *et al.*, 2017). In *Rhg1*(+) germplasm, analysis of an unusual NSF allele [*Rhg1*-associated NSF on chromosome 07; NSF (RAN07)] revealed that NSF (RAN07) exhibited stronger *in vitro* binding with *Rhg1* resistance-type α -SNAPs, suggesting that an atypical co-evolution of the soybean SNARE-recycling machinery balances the acquisition of an otherwise disruptive housekeeping protein, enabling a valuable disease resistance trait (Bayless *et al.*, 2018).

Resistance to SCN involves salicylic acid. Overexpression of a salicylic acid methyltransferase gene conferred resistance to SCN (Lin *et al.*, 2016; Lin *et al.*, 2013). The concentration of 1-aminocyclopropane-1-carboxylic acid (ACC) and expression of ACC synthase were both higher in SCN-colonized root pieces and root tips than in other parts of the root (Tucker *et al.*, 2010). Moreover, *CLE* (encodes an extracellular protein), *GmAFS* (encodes one member of the soybean terpene synthase gene family), *MIR396* and *t-SNAREs* also respond to SCN infection (Dong *et al.*, 2020; Guo *et al.*, 2015; Lin *et al.*, 2017; Noon *et al.*, 2019; Noon *et al.*, 2016). Along with SCN infection, variation of DNA methylation in some genomic regions associated with changes in gene expression (Rambani *et al.*, 2020).

A report on the dissection of the QTL for southern root-knot nematode (RKN) resistance in soybean identified three QTL and mapped the major QTL to a 29.7-kb region on chromosome 10 (Xu *et al.*, 2013b). In addition, there are several reports on soybean resistance to other pests. For instance, defence against *Anticarsia gemmatilis* larvae was modulated by solar UV-B radiation and ethylene (Dillon *et al.*, 2018). Recently, a gene encoding a VQ motif-containing protein, *GmVQ58*, was found to enhance soybean resistance to the common cutworm (Li *et al.*, 2020c).

Asian soybean rust (ASR), caused by the obligate biotrophic fungus *Phakopsora pachyrhizi*, and is one of the most economically important diseases for soybean. Through *in vivo* assessment by Mach-Zehnder double-beam interferometry, Loehrer *et al.* found that *P. pachyrhizi* might be able to forcefully invade a wide range of plants through appressorial turgor pressure (Loehrer *et al.*, 2014). Both a UDP-glucosyl transferase and phenylpropanoid metabolism are essential for the infection of *P. pachyrhizi* (Beyer *et al.*, 2019; Langenbach *et al.*, 2013), while coumarin could be used as a natural fungicide against ASR (Beyer

et al., 2019; Langenbach *et al.*, 2013). Evolutionary analyses indicated that some disease resistance genes have conserved function between soybean and other species, indicating that genes identified from other species could be used in soybean disease resistance (Ashfield *et al.*, 2014; Okutani *et al.*, 2020; Redditt *et al.*, 2019; Wang *et al.*, 2014a; Wei *et al.*, 2020a). When the gene *CcRpp1* (*Cajanus cajan* Resistance against *Phakopsora pachyrhizi* 1) from pigeon pea or *NHR*-linked genes from *Arabidopsis* were transferred to soybean, the transgenic lines exhibited resistance to *P. pachyrhizi* (Kawashima *et al.*, 2016; Langenbach *et al.*, 2016).

Through GWAS and QTL analyses, several loci associated with pattern-triggered immunity (Valdés-López *et al.*, 2011), sudden death syndrome resistance (Zhang *et al.*, 2015) and white mould (Zhao *et al.*, 2015b) were identified. Genetic mapping suggested that *STAY-GREEN* genes may be involved in sudden death syndrome (Chang *et al.*, 2019), and *Rsv4*, which encodes an RNase H family protein with dsRNA-degrading activity, may be responsible for broad-spectrum mosaic virus resistance (Ishibashi *et al.*, 2019). Overexpression of the *HSP40* gene, which encodes a nuclear-localized, type-III DnaJ domain-containing protein, or *GmLMM1*, which encodes a malectin-like receptor kinase, regulated cell death and disease resistance (Liu and Whitham, 2013; Wang *et al.*, 2020a). Overexpression of the plasma membrane protein gene *GmDR1* generated broad-spectrum immunity (Ngaki *et al.*, 2021). Interestingly, genes encoding small peptides, such as *GmSubPep*, encoding a 12-amino acid peptide (Pearce *et al.*, 2010), and *GmPep914*, encoding an 8-amino acid peptide (Yamaguchi *et al.*, 2011), also played important roles in defence against disease. It was recently found that hydroperoxide lyase modulated the defence response and conferred lesion-mimic phenotype in soybean leaves (Wang *et al.*, 2020f).

Phytophthora root and stem rot, caused by *Phytophthora sojae*, is another destructive soybean disease. Overexpression of *GmERF5*, which encodes ethylene response factor 5 (Dong *et al.*, 2015), or *GmMYB29A2*, which encodes a glyceollin transcription factor (Jahan *et al.*, 2020), significantly enhanced resistance to *P. sojae*. Moreover, overexpression of some microRNAs, such as miR393, could promote soybean defence against *P. sojae* (Wong *et al.*, 2014).

Soybean mosaic virus (SMV) is one of the most prevalent viral diseases and could significantly reduce yield losses in soybean. Until now, four dominant SMV resistance loci (*Rsv1*, *Rsv3*, *Rsv4* and *Rsv5*) have been genetically identified (Hayes *et al.*, 2000; Jeong *et al.*, 2002; Klepadlo *et al.*, 2017; Yu *et al.*, 1994). Recently, *GmST1*, which encodes a sulfotransferase, was identified as the responsible gene for conferring the resistance to strains G2 and G3 (Zhao *et al.*, 2021).

R genes specifically activate resistance responses that are effective against diverse pathogens (Zheng *et al.*, 2016). In soybean, Kang *et al.* (2012) predicted that the genome contains a total of 319 nucleotide-binding site/leucine-rich repeat (NBS-LRR) *R* genes. The ULP1-NBS-LRR protein GmRpp1 confers immunity to *P. pachyrhizi* (Pedley *et al.*, 2019). Overexpression of the TIR-NBS-LRR *R* gene *GmKR3* enhanced soybean resistance to several strains of soybean mosaic virus (SMV), which is one of the most prevalent viral diseases and could significantly reduce yield losses in soybean (Xun *et al.*, 2019). Silencing of the soybean *NDR1* homologs (*GmNDR1*) showed that they were required for pathogen resistance (Selote *et al.*, 2014). *GmMPK4*- and *GmMPK6*-silenced plants displayed strong phenotypes, including

induction of PR gene expression and increased SA levels (Liu *et al.*, 2014a; Liu *et al.*, 2011). Fine-tuning the expression of pathogen avirulence (Avr) effector genes using genome editing impacted the compatibility of plant disease, which provided clues to improve crop disease resistance (Ochola *et al.*, 2020).

Other stressors

Salicylic acid (SA) and abscisic acid (ABA) are two important phytohormones for stress resistance. Knockdown of either the phenylalanine ammonia lyase or isochorismate synthase biosynthesis pathway shuts down SA biosynthesis and abrogates pathogen resistance (Shine *et al.*, 2016). In soybean, it was found that *ABA-Sensitive 1* (*GmABAS1*), which encodes a 1R-subtype of MYB, functioned as a transcriptional repressor that enhances ABA sensitivity (Ku *et al.*, 2020). Several transcription factor families were found to be involved in various stress responses. For instance, AP2/ERF-type transcription factor family members, including *GmERF3* and *GsERF7*, played cardinal roles in regulating resistance to diseases as well as salt and drought (Feng *et al.*, 2020; Zhang *et al.*, 2009), whereas the DREB1/CBF-type transcription factors function in heat, drought and cold stresses (Kidokoro *et al.*, 2015).

Other genes that response various stresses were also reported. The root-specific protein kinase-encoding gene *GmWINK1* and the mitogen-activated protein kinase-encoding gene *GmMPK* were found to regulate plant growth and development and, in turn, affect stress responses (Liu *et al.*, 2011; Wang *et al.*, 2010). Malate exudation mediated by *Gm Representative*, which encodes an expansin, was found to be involved in multiple abiotic stresses (Guo *et al.*, 2011). A chaperone binding protein (BiP) functioned as a negative regulator to attenuate stress-induced cell death (Reis *et al.*, 2011). Omega-3 fatty acid desaturase was found to play an important role in cold response in soybean (Román *et al.*, 2012). For competition with weeds, a PIF3-like protein was identified as a candidate that might manipulate the weed stress response in soybean (Horvath *et al.*, 2015).

Nodulation

Symbiotic nitrogen fixation (SNF) through root nodulation is an important feature of legumes and plays important roles in plant growth. Over the past decade, the genes required for symbiotic nitrogen fixation have been thoroughly investigated, and significant progress has been achieved (Roy *et al.*, 2020).

A comprehensive phylogenomic analysis revealed that multiple losses of some key genes, such as *NIN* and *RPG*, were essential for the origin of SNF (Griesmann *et al.*, 2018). A global co-expression analysis suggested that ancient orthologous and duplication events before the origin of legumes had paved the way for nodule formation and nitrogen fixation (Wu *et al.*, 2019). Large-scale transcriptome and metabolome investigations revealed a number of genes and metabolic pathways that are induced or suppressed during nodulation (Agtuca *et al.*, 2020; Hayashi *et al.*, 2012b; Libault *et al.*, 2010). For instance, G protein-encoding genes and a putative beta-carotene hydroxylase gene (*GmBCH1*) exhibited significant transcriptional changes in response to rhizobium infection. RNA interference suppression of the genes encoding G protein and *GmBCHs* severely impaired nitrogen fixation as well as nodule development, suggesting they are positive regulators in nodule formation (Choudhury and Pandey, 2013; Kim *et al.*, 2013). Furthermore, the G protein cycle was regulated by the activity of phosphorylation-dependent G protein

signalling proteins (Choudhury and Pandey, 2015). The *G-box Factor 14-3-3* genes, *SGF14c* and *SGF14l*, had been shown to function as dimers in soybean nodulation (Radwan *et al.*, 2012).

In soybean, the formation of symbiotic root nodules was highly affected by several host genes, referred to as *Rj* or *rj* (Hayashi *et al.*, 2012a). *Rj1* and *Rj5* encode putative Nod factor receptors (NFRs) (Hayashi *et al.*, 2012a; Indrasumunar *et al.*, 2010; Indrasumunar *et al.*, 2011), *Rj2/Rfg1* encodes a Toll-interleukin receptor/nucleotide-binding site/leucine-rich repeat (TIR-NBS-LRR) plant resistance (R) protein (Yang *et al.*, 2010), while *Rj7* encodes a nodule autoregulation receptor kinase (Hayashi *et al.*, 2012a; Nishimura *et al.*, 2002; Searle *et al.*, 2003). One report suggested that a gene encoding thaumatin-like protein (TLP), a pathogenesis-related (PR) protein, might be the candidate for the *Rj4* locus (Hayashi *et al.*, 2014; Tang *et al.*, 2016).

It has been suggested that flavonoids act as chemotactic signals to rhizobia under low-N conditions in legumes (Liu and Murray, 2016; Subramanian *et al.*, 2006). *GmMaT*, a malonyl-CoA:flavonoid acyltransferase-encoding gene, was found to catalyse isoflavone malonylation and affect malonyl isoflavone secretion, which had an effect on soybean nodulation (Ahmad *et al.*, 2021). Flavonoids and/or related compounds such as isoflavones are released from legume roots and attract the Nod factors (NFs), which are secreted by rhizobia to initiate symbiotic nitrogen fixation (Figure 3). Soybean isoflavone synthase (IFS), a key enzyme in the biosynthesis of isoflavones, was induced by *Bradyrhizobium japonicum*. When the expression of *IFS* was knock down in soybean hairy root composite plants, these plants showed severely reduced nodulation (Liu and Murray, 2016; Subramanian *et al.*, 2004; Subramanian *et al.*, 2006). Additionally, glycolysis and lipid biosynthesis may also play essential roles in nodule formation (Chen *et al.*, 2020a; Gillman *et al.*, 2014; Lakhssassi *et al.*, 2020; Zhang *et al.*, 2020b).

The number of nodules formed on the roots of soybean is systemically controlled by autoregulation of nodulation (AON) (Caetano-Anolles and Gresshoff, 1991). *NTS1/GmNARK* was found to be an important sensor that controls nodulation (Searle *et al.*, 2003). Mutation of the *NTS1/GmNARK* locus led to prolific nodulation (Carroll *et al.*, 1985a; Searle *et al.*, 2003). Another two signalling molecules, *Bradyrhizobium-induced and acting systemically 1* (*GmRIC1*) and *GmRIC2*, were involved in the transition of long-distance signals from root to shoot. Further investigation showed that *GmNARK* can be activated by *GmRIC1* and *GmRIC2* (Reid *et al.*, 2012). Overexpression of *GmRIC1* and *GmRIC2* strongly suppressed the nodulation in a *GmNARK*-dependent manner (Reid *et al.*, 2011) (Figure 3). An inoculation- and NARK-dependent gene (*GmUFD1a*) responds in both the bioassay and intact, inoculated plants, indicating that it might be a novel component of the autoregulation pathway (Reid *et al.*, 2012). The transcription factors *GmNF-YA1a* and *b* were also identified as positive regulators in AON (Schaarschmidt *et al.*, 2013). Overall, *GmNIC1* (nitrate-induced and acting locally) and *GmRIC1* played key roles in AON and were relied on the activity of the nodulation autoregulation receptor kinase *GmNARK* (Reid *et al.*, 2013).

Nitrogen and nitrogen compounds such as nitrate have been reported to negatively control nodulation (Carroll *et al.*, 1985a,b; Day *et al.*, 2010; Lim *et al.*, 2014; Reid *et al.*, 2011; Tanaka *et al.*, 1985). The processes of nitrogen regulation of nodulation are distinct from AON and act immediately. Interestingly, a leucine-rich receptor-like kinase, *GmNARK*, was shared between AON and nitrate-dependent regulation of nodulation (Ferguson *et al.*,

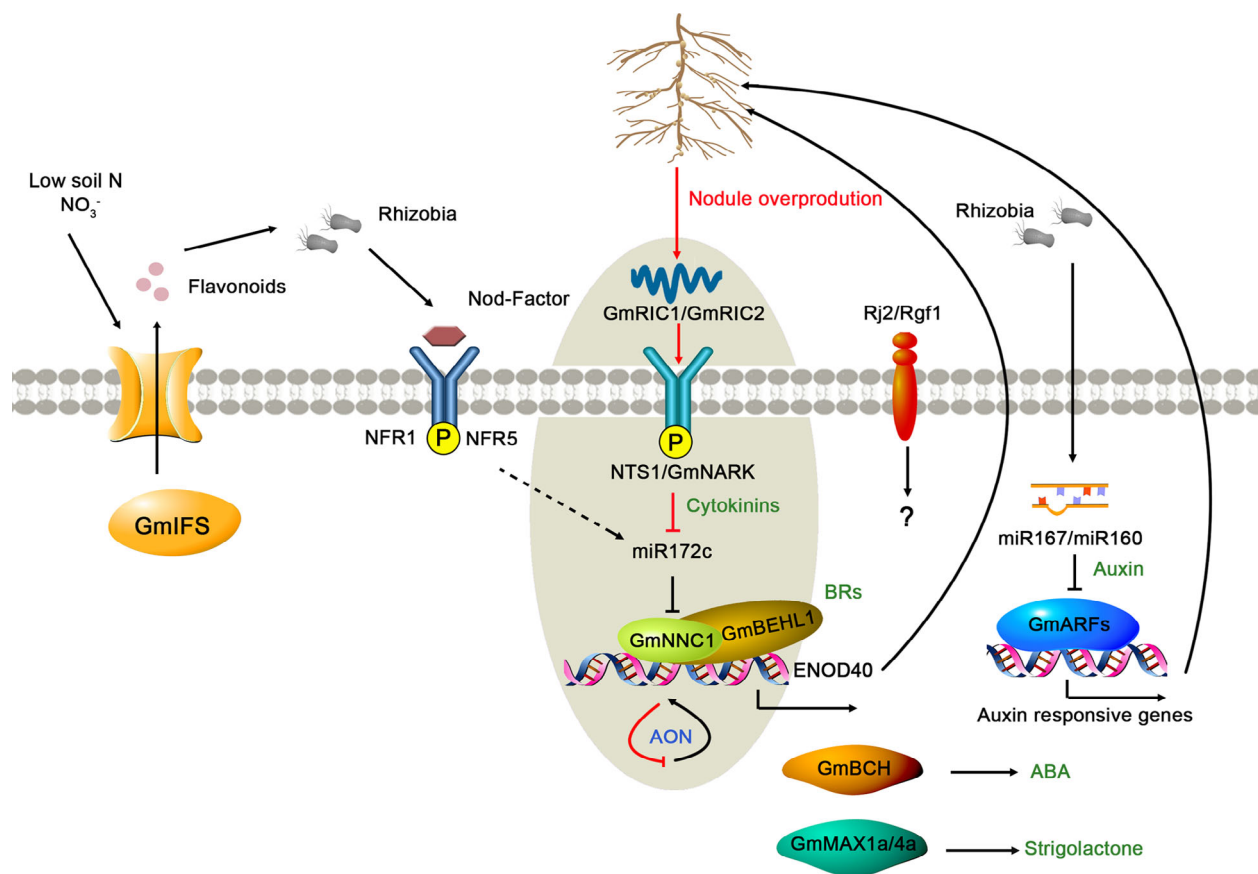


Figure 3 Proposed molecular network of soybean nodulation. Under low soil N, soybean plants produce flavonoids that trigger the production of bacterial Nod factors (NFR1 and NFR5), which, together with other signals, are perceived by receptors at the plasma membrane of plant epidermal cells. *Rj2/Rgf1* encodes a member of the Toll-interleukin receptor/nucleotide-binding site/leucine-rich repeat (TIR-NBS-LRR) class of plant resistance (R) proteins that functions in symbiotic root nodules. Nodule overproduction is prevented by AON signalling, in which *GmRIC1* and *GmRIC2* activate *GmNARK*. *GmNARK* induces shoot-derived cytokinins that, in turn, repress the transcriptional activity of *miR172c*. *miR172c* is a key positive regulator of nodulation and promotes the cleavage of the mRNA encoding its target gene *Nodule Number Control 1* (*NNC1*), which directly suppresses the transcription of the early nodulin gene *ENOD40*. Auxin, brassinosteroids (BRs), strigolactones, cytokinins and abscisic acid (ABA) are reported to function in nodulation in soybean. *GmBEHL1*, related to the BR signal, is an *NNC1*-interacting protein. *GmBEHL1* is suggested to be a co-repressor of *NNC1* and to negatively regulate soybean nodulation.

2019). Under high nitrate, loss-of-function *GmNARK* mutants, *nts* (*nitrate tolerant symbiosis*), still exhibited super nodulation (Carroll et al., 1985b). *GmNARK* perceived the nodulation-suppressing CLE peptides, such as *GmRIC1* and *GmRIC2*, in the shoot through the AON pathway and sensed the nodulation-suppressing CLE peptides, such as *GmNIC1*, in the root through nitrogen regulation of nodulation (Ferguson et al., 2019; Lim et al., 2014; Reid et al., 2011).

The development of nodules is significantly affected by phosphate (Pi). Pi starvation severely inhibited both nodulation and biological N_2 fixation (Hernandez et al., 2009), with decreased nodule numbers, nodule size and nitrogenase activity of soybean. Two phosphate transporters, *GmPT5* and *GmPT7*, regulated phosphate transport and in turn affected nodulation in soybean (Chen et al., 2019; Qin et al., 2012). *GmPT5* controlled Pi entry from roots to nodules, was critical for maintaining Pi homeostasis in nodules and subsequently regulated nodulation and growth performance (Qin et al., 2012). Overexpression of *GmPT7* promoted nodulation and increased plant biomass, shoot nitrogen and phosphorus contents, improving soybean yield by up to 36% (Chen et al., 2019).

The proteins phosphate-transporter 1 (PHT1) and its regulator phosphate-starvation-response 1 (PHR1) worked as a PHR1-PHT1 module to maintain Pi homeostasis and affected nodule development (Lu et al., 2020a).

Hormones have long been known to control nodule organogenesis (Grunewald et al., 2009). The transcription factors BR1-EMS suppressor 1 (*BES1*)/brassinazole-resistant 1 (*BZR1*) played key roles in the brassinosteroids (BRs) signalling pathway (Yan et al., 2018). *GmBEHL1* was identified as an *Arabidopsis BES1/BZR1* homolog and can interact with NODULE NUMBER CONTROL 1 (*NNC1*), a transcriptional repressor that mediates soybean nodulation (Wang et al., 2014c; Yan et al., 2018). Knockdown and overexpression of *GmBEHL1* increased and decreased the number of nodules, respectively (Yan et al., 2018). The strigolactone (SL) biosynthesis enzymes *GmMAX1a* and *GmMAX4a* were apparently regulated by rhizobia infection. *GmMAX1a* and *GmMAX4a* knockdown lines exhibited decreased nodule number (Rehman et al., 2018) (Figure 3). *GmMAX2* interacted with D14 and KAI to influence the SL and karrikins (KARs) signalling pathways to affect soybean root nodulation (Ahmad et al., 2020). Among the *YUCCA* (*YUC*) gene family, *GmYUC2a* functioned as

an important regulator of auxin biosynthesis to modulate nodulation (Wang *et al.*, 2019f). The β -carotene hydroxylase GmBCH catalysed the conversion of β -carotene to β -zeaxanthin, which was related to the ABA synthesis pathway. RNAi-mediated silencing of *GmBCH1/2* impaired nodule development and symbiotic nitrogen fixation (Kim *et al.*, 2013).

It was also found that quite a few miRNA families showed transcript-level responses to nodulation (Jin *et al.*, 2018; Yan *et al.*, 2015; Yan *et al.*, 2016). For example, *miR393j-3p* was significantly up-regulated during nodule formation, and ectopic expression of *miR393j-3p* significantly reduced nodule formation (Yan *et al.*, 2015). The function of miR393 may be through regulation of GmTIR1/GmAFB3-based auxin perception (Cai *et al.*, 2017). miR172 is another important miRNA that was essential for nodule development. Overexpression of *MIR172* or the miRNA-encoded peptide miPEP172c both resulted in an increase in nodule numbers in transgenic soybean roots (Couzigou *et al.*, 2016; Yan *et al.*, 2013). Further investigation showed that the function of miR172c was through the GmNIN-miR172c-NNC1 regulatory module (Wang *et al.*, 2019c; Wang *et al.*, 2014c) (Figure 3). *MIR160* promoted auxin activity by suppressing the levels of the ARF10/16/17 transcription factors to direct proper nodule formation and maturation in soybean (Nizampatnam *et al.*, 2015; Turner *et al.*, 2013). The expression levels of miR167 and its target were up- and down-regulated by auxin in soybean. Moreover, miR167 could positively regulate nodule numbers by repressing the target genes *GmARF8a* and *GmARF8b*, which were homologous genes of the *Arabidopsis* auxin response factor *AtARF8* (Wang *et al.*, 2015c) (Figure 3). In addition, mis-expression of miR482, miR1512 and miR1515 increased nodulation (Li *et al.*, 2010). Recently, Ren *et al.* showed that small RNA fragments (tRFs) derived from rhizobial transfer RNA (tRNA) serve as signal molecules that regulate host nodulation. Three families of rhizobial tRFs (Bj-tRF001, Bj-tRF002 and Bj-tRF003) were confirmed to regulate host genes associated with nodule initiation and development (Ren *et al.*, 2019), which represented a new evidence of a root-shoot-root signalling mechanism during nodulation (Shine *et al.*, 2019; Zhang *et al.*, 2020a).

The soybean gene *early nodulin 40* (*ENOD40*) played a pivotal role in nodule organogenesis (Charon *et al.*, 1999; Kumagai *et al.*, 2006; Wan *et al.*, 2007). NNC1 regulated the expression of *ENOD40* by binding to the AP2 cis-elements of *ENOD40* promoter, which consequently represses *ENOD40* expression and negatively regulated nodulation (Wang *et al.*, 2014c). Several other genes that could affect soybean nodulation were also identified, such as *FW2.2-like 1* (*GmFWL1*) (Libault *et al.*, 2010), LysM-type receptor kinase (*GmNFR1alpha*) (Indrasumunar *et al.*, 2011), ecto-apyrase (*GS52*) (Tanaka *et al.*, 2011), ureide transporter (*UPS1*) (Collier and Tegeder, 2012), *symbiotic ammonium transporter 1* (*SAT1*) (Chiasson *et al.*, 2014), *GmEXPB2* (Li *et al.*, 2015a) and *VAMP721a* and *VAMP721d* (Gavrin *et al.*, 2016). The transcription of the gene *target of rapamycin* (*GmTOR*) and its key downstream effector, *GmS6K1*, were activated during nodulation. When *GmS6K1* was knocked down, nodule development was severely impaired, suggesting an important role for the rapamycin pathway in nodule development (Um *et al.*, 2013). *GmVTL1a*, which function as a transporter of ferrous iron from the infected root cell cytosol to the symbiosome, moved iron across the symbiosome membrane to bacteria's and played a crucial role in nitrogen fixation (Brear *et al.*, 2020; Liu *et al.*, 2020d). In addition, nodulation was

affected under acidic and drought conditions (Gil-Quintana *et al.*, 2013; Lin *et al.*, 2012). A very recent study reported that *Nodule Number Locus 1* (*GmNNL1*), which encodes a novel R protein, may trigger immunity and inhibit nodulation (Zhang *et al.*, 2021).

Domestication

Plant domestication is one of the most important aspects contributing to the development of agriculture (Diamond, 2002). In addition to investigations at the population level through the resequencing of germplasm (Lam *et al.*, 2010; Li *et al.*, 2013b; Sedivy *et al.*, 2017; Zhou *et al.*, 2015), some genes responsible for traits linked to soybean domestication were identified.

One important agronomic trait that was targeted by human selection during crop domestication is decreased pod shattering and seed dispersal (Sedivy *et al.*, 2017). *SHAT1-5*, a gene encoding a NAC (NAM, ATAF1/2 and CUC2) transcription factor, was found to be a prime domestication gene, and the allele in cultivated soybean improves the thickening of the fibre cap cells and suppresses pod shattering (Dong *et al.*, 2014). *Pdh1* is another gene that affects the pod shattering phenotype and showed artificial selection in landraces of Japan, Korea and other South-East Asian countries. The cultivated alleles promoted torsion of dried pods under low humidity, causing higher pod dehiscence (Funatsuki *et al.*, 2014).

Loss of dormancy is another important domestication trait and relates to both physiology and physical, structural changes (Finch-Savage and Leubner-Metzger, 2006). In soybean, *Hs1-1* has long been identified as an important locus corresponding to loss of dormancy through a physical change (Liu *et al.*, 2007). Sun *et al.* performed a genetic analysis and suggested that *Hs1-1* encodes a calcineurin-like metallophosphoesterase transmembrane protein (Sun *et al.*, 2015a), while Jang *et al.* reported that a single nucleotide polymorphism in another gene, a endo-1,4- β -glucanase, in this region may be responsible for *Hs1-1* (Jang *et al.*, 2015). Another key locus responsible for seed dormancy is *G*, which was found to encode a CAAX amino-terminal protease protein. Interestingly, *G* had undergone parallel selection across different plant families (Wang *et al.*, 2018b), which may make it a candidate gene for the acceleration of crop improvement (Lyu, 2018; Rendon-Anaya and Herrera-Estrella, 2018; Wei and Huang, 2018).

As a short-day flowering plant, the spread of soybean cultivation latitudinally requires adaptation to new photoperiods. Several flowering-related genes in soybean have been selected during human cultivation (Cober *et al.*, 2010; Kim *et al.*, 2012a). *GmCRY1a* and *GmCOL7a*, soybean homologs of *Arabidopsis* *CRYPTOCHROME 2* (*CRY2*) and *CONSTANS* (*CO*), respectively, were reported to exhibit strong selection signatures (Li *et al.*, 2013b; Wang *et al.*, 2016c; Wu *et al.*, 2014; Zhang *et al.*, 2008). The major maturity loci *E1*, *E2*, *E3* and *E4*, where several flowering-associated genes are located, have contributed to local adaptation (Kanazawa *et al.*, 2009; Liu *et al.*, 2008; Wang *et al.*, 2016c; Watanabe *et al.*, 2009; Watanabe *et al.*, 2011; Xia *et al.*, 2012; Zhai *et al.*, 2014b). A recent study found that homologous *pseudo-response-regulator* (*PRR*), *Tof11* and *Tof12*, had undergone strong selection (Lu *et al.*, 2020b). Several yield-related genes, such as *GA20OX*, *NFYA* (Lu *et al.*, 2016), *Dt1* (Tian *et al.*, 2010), *SoyWRKY15a* (Gu *et al.*, 2017) and *PP2C-1* (Lu *et al.*, 2017b), had also undergone selection during soybean domestication.

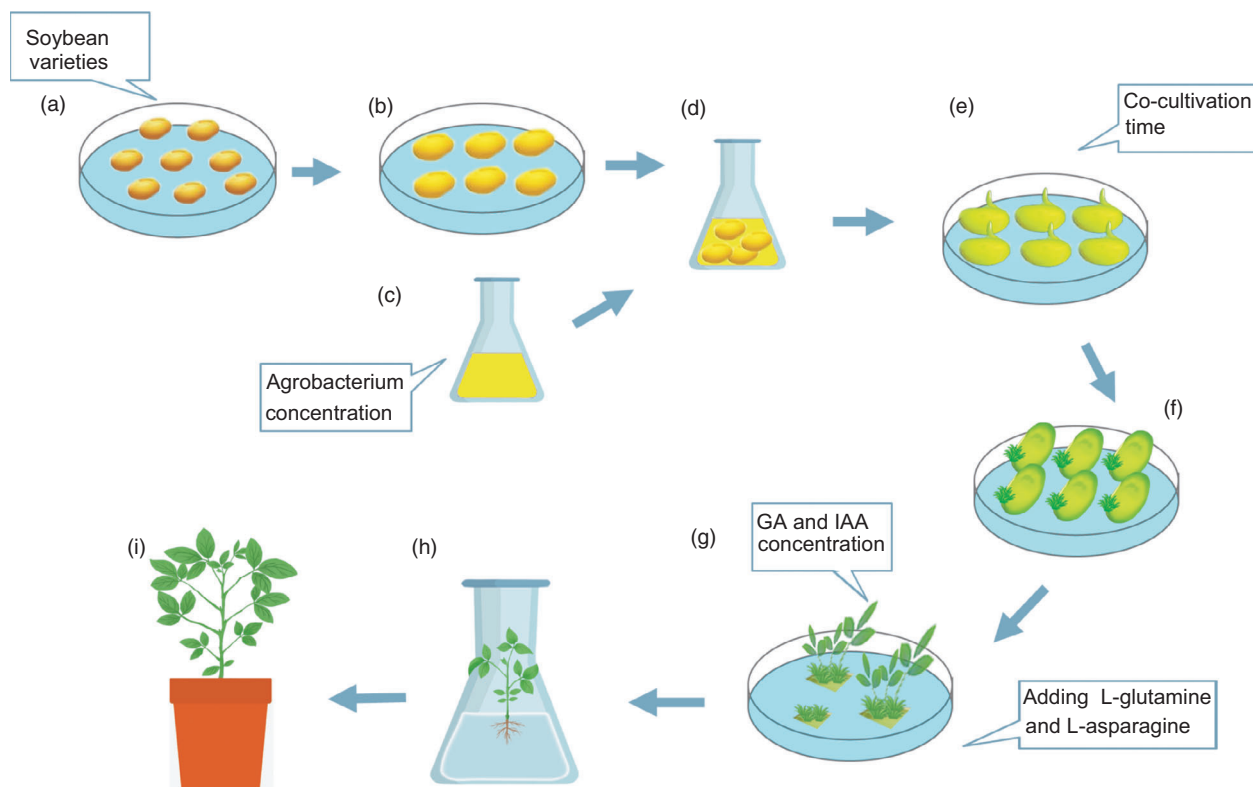


Figure 4 General procedure of *Agrobacterium*-mediated cotyledonary node soybean transformation. (a) Seed sterilization. Selection of soybean varieties dominates the efficiency of transformation affecting on *Agrobacterium* infection and regeneration capacity (Song *et al.*, 2013c). (b) Seed imbibition. (c) Preparation of *Agrobacterium*. (d) *Agrobacterium* infection. It is one of the most important steps, *Agrobacterium* concentrations, soybean explants, *Agrobacterium* suspension medium, and co-cultivation time will affect the infection efficiency. (e) Co-cultivation. Suitable concentration of plant hormone is necessary to (f) Shoot induction and (g) shoot elongation that will improve the efficiency (Li *et al.*, 2017c). Adding L-glutamine and L-asparagine together into culture media will increase the transformation frequency of soybean by suppressed the expression level of *GmPRs* (Chen *et al.*, 2018a). (h) Rooting. (i) Transplanting. Words in the blue boxes indicate the methods of improving infection efficiency.

Transformation technology

Transgenic technology

Soybean transgenic technology is a necessary tool for soybean gene function study. Genetic transformation of soybean has been studied for over two decades; however, the progress remains slow and inefficient, which is why the functional validation in some studies was performed in *Arabidopsis* instead of soybean. Several transformation systems were developed, including using shoot meristems (McCabe *et al.*, 1988; Rech *et al.*, 2008), hypocotyls (Aragão *et al.*, 2000; Dan and Reichert, 1999; Wang and Xu, 2008), embryo (Finer and McMullen, 1991; Trick and Finer, 1998), immature cotyledons, half-seed explants (Liu *et al.*, 2004; Paz *et al.*, 2006) and cotyledonary nodes (Li *et al.*, 2017c; Liu *et al.*, 2004; Sato *et al.*, 1993). Considering the operation, reproducibility, copy number of foreign DNA and experimental cost, *Agrobacterium*-mediated cotyledonary node (CN) soybean transformation was commonly used nowadays (Hinchee *et al.*, 1988; Paz *et al.*, 2006; Somers *et al.*, 2003). The overall average transformation efficiency was 3.8%–8.7% (Li *et al.*, 2017c; Paz *et al.*, 2006). Recently, Pareddy *et al.* (2020) enhanced the average transformation efficiency to 18.7%. But it is still lower than that in rice of 23% (Ge *et al.*, 2006; Lin and Zhang, 2005) and maize of over 30% (Ishida *et al.*, 1996; Yang *et al.*, 2006).

The general transformation process includes seed sterilization and germination; *Agrobacterium* infection; co-cultivate soybean explants and *Agrobacterium*; shoot induction; shoot elongation; rooting; and at last transferring the plants to pots containing soil (Figure 4). In these processes, many factors can affect the efficiency. The first effector is soybean genotypes. Song *et al.* compared transformation efficiency and regeneration rate of twenty soybean varieties and found that transformation efficiency between different varieties differed significantly (0.31%–4.59%) (Song *et al.*, 2013c). Secondly, in the process of *Agrobacterium* infection which is one of the most important steps, all of *Agrobacterium* concentrations, soybean explant status, *Agrobacterium* suspension medium and co-cultivation time will affect the infection efficiency. Another key process in determining the transformation efficiency is explant regeneration. It was reported that plant hormone plays critical role in inducing the regeneration of explant and its suitable concentration could improve the efficiency (Li *et al.*, 2017c). Chen *et al.* (2018a) reported that adding L-glutamine and L-asparagine together into culture media will increase the transformation efficiency by suppressing the expression of *GmPRs*.

Genome editing

Genome editing can introduce precise modifications into the genome to obtain predictable and desired traits, which has been

proved to be a powerful approach for functional study and molecular design breeding than the traditional genetics approaches, such as mutagenesis, transgenic RNAi or overexpression (Gao, 2021; Rodriguez-Leal *et al.*, 2017). Of the different genome editing systems, CRISPR (clustered regularly interspaced short palindromic repeat)/Cas (CRISPR-associated) system shows high efficiency and has been extensively applied in different species (Feng *et al.*, 2013; Jiang *et al.*, 2013; Miao *et al.*, 2013; Shan *et al.*, 2013; Svitashv *et al.*, 2015; Upadhyay *et al.*, 2013).

The first knock out and DNA homology-directed recombination (HDR) soybean plant created by CRISPR/Cas9 technology was successfully obtained in 2015 (Li *et al.*, 2015b). In 2016, Du *et al.* (2016) found that changing the *AtU6-26* promoter into *GmU6-16g-1* promoter of the CRISPR/Cas9 system could significantly improve the efficiency of targeted mutagenesis in soybean. In soybean, nearly 75% of the genes present in multiple copies and knockout of a single gene usually does not exhibit mutant phenotype. It is important to develop a dedicated CRISPR/Cas9 system that can edit multiple homologous genes simultaneously. By optimizing the steps of vector construction, sgRNA assessment, pooled transformation and sgRNA identification, a CRISPR/Cas9 system that can generate multiplex mutagenesis with higher efficiency was developed (Bai *et al.*, 2020). In nature, beside the alleles caused by loss-of-function mutations, large part of the phenotypic variations in agronomic traits are resulted from single nucleotide polymorphism (SNP) variations. Damage of the function of the whole gene using gene-editing system usually leads to severe phenotype, which may not be optimizable for agronomic trait improvement in production. Therefore, generation of point mutations at specific sites affecting important agronomic traits is of great value in molecular breeding (Mishra *et al.*, 2020). Recently, 'base editing' has been developed from CRISPR/Cas9 system, which converts single base into another without requiring DNA double-strand breaks or a donor template (Komor *et al.*, 2016). Cai *et al.* successfully applied the system to create point mutant of *GmFT2a* and *GmFT4* (Cai *et al.*, 2020a).

Nowadays, CRISPR/Cas9 was widely applied in soybean functional studies (Cai *et al.*, 2018; Du *et al.*, 2016; Li *et al.*, 2019a; Michno *et al.*, 2015; Sun *et al.*, 2015b; Xu *et al.*, 2020). For instance in identification of the genes controlling flowering time, frameshift mutations generated by CRISPR/Cas9 demonstrated that *GmFT2a* mainly function under short day (SD), whereas *GmFT5a* had more significant effects under long day (LD) (Cai *et al.*, 2018; Cai *et al.*, 2020b). Similarly, knockout of *GmPRR37* by the CRISPR system suggested that it can repress flowering under LD (Wang *et al.*, 2020c). Two CRISPR/Cas9 gene-editing mutants of *Glyma.13G114200* exhibited male sterility phenotype, confirming that it was the causal gene for *GmMS1* for male sterility (Fang *et al.*, 2021b; Jiang *et al.*, 2021; Nadeem *et al.*, 2021). CRISPR/Cas9 was also applied in yield and seed quality-related trait modifications, such as to alter plant architecture by editing *GmLHYs* (Cheng *et al.*, 2019) or *SPL9* (Bao *et al.*, 2019; Cai *et al.*, 2018; Cheng *et al.*, 2019), to increase seed number per pod by editing *Ln* (Cai *et al.*, 2021), to reduce beany flavour by knocking out *LOXs* (Wang *et al.*, 2020b), to increase isoflavone content by editing *GmF3H1*, *GmF3H2* and *GmFNSII-1* simultaneously (Zhang *et al.*, 2020f) and to alter the fatty acid profiling by editing *FAD2-2* (Al Amin *et al.*, 2019). In the future, more application of 'base editing' for single gene or for multiple genes simultaneously will greatly promote the functional study and molecular design breeding in soybean.

Challenges and future perspectives

The Green Revolution is one of the most remarkable events in agriculture and greatly increased the production of major crops (Hedden, 2003). However, few improvements in yield have been made for soybean over the past six decades. In order to meet the needs of a growing world population, soybean yield must increase at a faster rate than it is at present (Ainsworth *et al.*, 2012; Ray *et al.*, 2013). There is an urgent need for a soybean 'Green Revolution' to breed super varieties with the ideal plant architecture that are adapted to high-density planting environment (Liu *et al.*, 2020e). At this point, most soybean production occurs in South America, North America and Asia. In the future, Africa might become another dominant soybean production area. Therefore, there is a need to start studies on genetic improvement and production technologies for an Africa-adapted soybean. Moreover, there is a need to develop new varieties to meet future environmental changes and to create a more sustainable agricultural system (Bishop *et al.*, 2015; Kumagai *et al.*, 2015; Mourtzinis *et al.*, 2015; Ruiz-Vera *et al.*, 2013; Tian *et al.*, 2020).

To create a super variety, breeders normally need to stack multiple desirable traits into a single line. However, most important traits are quantitatively controlled and exhibit inherited correlations. Bringing about a 'Green Revolution' in soybean may prove to be an exacting task because of the unique plant architecture and the complicated components that determine the final yield in soybean (Liu *et al.*, 2020e). Understanding the inherited mechanism of each trait and the regulatory network among different traits will help us to design the desired crops (Tian *et al.*, 2020). A dissection of the genetic networks underlying 84 agronomical traits has provided insights into the molecular design of soybean (Fang *et al.*, 2017).

De novo domestication of new crops aims to rapidly domesticate wild or semi-wild plants into agricultural crops with favoured agronomic traits by utilization of combines modern technologies, including genomics, gene editing and synthetic biology (Fernie and Yan, 2019; Khan *et al.*, 2019; Li *et al.*, 2018b; Yu *et al.*, 2021; Zsogon *et al.*, 2018). Wild soybean showed higher protein content, lower oil content and higher stress tolerance, whereas most of the current cultivated soybeans exhibited relatively lower protein content and higher oil content. So far, not many genes related to soybean domestication have been identified, leaving the domestication traits are still poorly understood. For instance, switching plant architecture from twinning (sprawl) habit in wild soybeans to upright habit in cultivated soybeans is one of the most prerequisite domestication trait for soybean; however, a genetics dissection of this important trait is unclear yet. With the rapid development of functional genomics in soybean, an ever-increasing number of genes related to agronomic traits are being cloned, which may enable us to redomesticate the wild soybean into a new crop keeping the characters of higher protein content and higher stress tolerance.

At this time, soybean functional studies and breeding still face some technical challenges. The lack of a stable and high-efficiency transgenic system is one of the biggest challenges, which makes functional study more time-consuming than in species with well-developed transformation systems, such as rice and *Arabidopsis*. Another challenge is phenotyping. Because soybean is quite sensitive to photoperiod, the phenotypes of a soybean line usually exhibit significant variations in different environments, which make precise phenotyping and functional

study more difficult. The recent development of new technologies, such as target base editing (Cai *et al.*, 2020a) and a transient expression system (Xiong *et al.*, 2019), will facilitate soybean functional studies. Moreover, the recently developed graph-based soybean pan-genome will both rejuvenate previous omics data and revolutionize functional and evolutionary genomic studies in soybean (Liu *et al.*, 2020f; Liu and Tian, 2020).

Conflict interest

We declare that we have no conflict of interest.

Author contributions

M.Z., S.L., Z.W., Y.Y., Z.Z., Q.L., X.Y., Z.D., F.K., B.L., B.R. and Z.T. drafted the manuscript. Z.T. conceived the article and revised the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (31801380), the National Key Research & Development Program of China (2017YFD0101305), and the State Key Laboratory of Plant Cell and Chromosome Engineering (PCCE-KF-2019-05).

References

- Agrawal, G.K., Hajdich, M., Graham, K. and Thelen, J.J. (2008) In-depth investigation of the soybean seed-filling proteome and comparison with a parallel study of rapeseed. *Plant Physiol.* **148**, 504–518.
- Agtuca, B.J., Stopka, S.A., Evans, S., Samarah, L., Liu, Y., Xu, D., Stacey, M.G. *et al.* (2020) Metabolomic profiling of wild-type and mutant soybean root nodules using laser-ablation electrospray ionization mass spectrometry reveals altered metabolism. *Plant J.* **103**, 1937–1958.
- Ahmad, M.Z., Rehman, N.U., Yu, S., Zhou, Y., Haq, B.U., Wang, J., Li, P. *et al.* (2020) GmMAX2-D14 and -KAI interaction-mediated SL and KAR signaling play essential roles in soybean root nodulation. *Plant J.* **101**, 334–351.
- Ahmad, M.Z., Zhang, Y., Zeng, X., Li, P., Wang, X., Benedito, V.A. and Zhao, J. (2021) Isoflavone malonyl-CoA acyltransferase GmMaT2 is involved in nodulation of soybean (*Glycine max*) by modifying synthesis and secretion of isoflavones. *J. Exp. Bot.* **72**, 1349–1369.
- Ainsworth, E.A., Yendrek, C.R., Skoneczka, J.A. and Long, S.P. (2012) Accelerating yield potential in soybean: potential targets for biotechnological improvement. *Plant Cell Environ.* **35**, 38–52.
- Al Amin, A.N., Ahmad, N., Wu, N., Pu, X., Ma, T., Du, Y., Bo, X. *et al.* (2019) CRISPR-Cas9 mediated targeted disruption of FAD2-2 microsomal omega-6 desaturase in soybean (*Glycine max*L.). *BMC Biotechnol.* **19**, 9.
- Anderson, E.J., Ali, M.L., Beavis, W.D., Chen, P. and Tilmon, K.J. (2019) Soybean [*Glycine max* (L.) Merr.] Breeding: History, Improvement, Production and Future Opportunities, pp. 431–516. Cham: Springer.
- Andreu, V., Lagunas, B., Collados, R., Picorel, R. and Alfonso, M. (2010) The GmFAD7 gene family from soybean: identification of novel genes and tissue-specific conformations of the FAD7 enzyme involved in desaturase activity. *J. Exp. Bot.* **61**, 3371–3384.
- Aragão, F.J.L., Sarokin, L., Vianna, G.R. and Rech, E.L. (2000) Selection of transgenic meristematic cells utilizing a herbicidal molecule results in the recovery of fertile transgenic soybean [*Glycine max* (L.) Merrill] plants at a high frequency. *Theor. Appl. Genet.* **101**, 1–6.
- Arai, Y., Hayashi, M. and Nishimura, M. (2008) Proteomic analysis of highly purified peroxisomes from etiolated soybean cotyledons. *Plant Cell Physiol.* **49**, 526–539.
- Arikit, S., Xia, R., Kakrana, A., Huang, K., Zhai, J., Yan, Z., Valdés-López, O. *et al.* (2014) An atlas of soybean small RNAs identifies phased siRNAs from hundreds of coding genes. *Plant Cell*, **26**, 4584–4601.
- Ashfield, T., Redditt, T., Russell, A., Kessens, R., Rodibaugh, N., Galloway, L., Kang, Q. *et al.* (2014) Evolutionary relationship of disease resistance genes in soybean and Arabidopsis specific for the *Pseudomonas syringae* effectors AvrB and AvrRpm1. *Plant Physiol.* **166**, 235–251.
- Atwood, S.E., O'Rourke, J.A., Peiffer, G.A., Yin, T., Majumder, M., Zhang, C., Cianzio, S.R. *et al.* (2014) Replication protein A subunit 3 and the iron efficiency response in soybean. *Plant Cell Environ.* **37**, 213–234.
- Bai, M., Yuan, J., Kuang, H., Gong, P., Li, S., Zhang, Z., Liu, B. *et al.* (2020) Generation of a multiplex mutagenesis population via pooled CRISPR-Cas9 in soya bean. *Plant Biotechnol. J.* **18**, 721–731.
- Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U. *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, **3**, e3376.
- Bandillo, N., Jarquin, D., Song, Q.J., Nelson, R., Cregan, P., Specht, J. and Lorenz, A. (2015) A population structure and genome-wide association analysis on the USDA soybean germplasm collection. *Plant Genome*, **8**, 3.
- Bandillo, N.B., Lorenz, A.J., Graef, G.L., Jarquin, D., Hyten, D.L., Nelson, R.L. and Specht, J.E. (2017) Genome-wide association mapping of qualitatively inherited traits in a germplasm collection. *Plant Genome*, **10**, 2.
- Bao, A., Chen, H., Chen, L., Chen, S., Hao, Q., Guo, W., Qiu, D. *et al.* (2019) CRISPR/Cas9-mediated targeted mutagenesis of GmSPL9 genes alters plant architecture in soybean. *BMC Plant Biol.* **19**, 131.
- Bayless, A.M., Zapotocny, R.W., Grunwald, D.J., Amundson, K.K., Diers, B.W. and Bent, A.F. (2018) An atypical N-ethylmaleimide sensitive factor enables the viability of nematode-resistant *Rhg1* soybeans. *Proc. Natl Acad. Sci. USA*, **115**, E4512–E4521.
- Beyer, S.F., Beesley, A., Rohmann, P.F.W., Schultheiss, H., Conrath, U. and Langenbach, C.J.G. (2019) The Arabidopsis non-host defence-associated coumarin scopoletin protects soybean from Asian soybean rust. *Plant J.* **99**, 397–413.
- Bian, X.H., Li, W., Niu, C.F., Wei, W., Hu, Y., Han, J.Q., Lu, X. *et al.* (2020) A class B heat shock factor selected for during soybean domestication contributes to salt tolerance by promoting flavonoid biosynthesis. *New Phytol.* **225**, 268–283.
- Bishop, K.A., Betzelberger, A.M., Long, S.P. and Ainsworth, E.A. (2015) Is there potential to adapt soybean (*Glycine max* Merr.) to future [CO₂]? An analysis of the yield response of 18 genotypes in free-air CO₂ enrichment. *Plant Cell Environ.* **38**, 1765–1774.
- Bolon, Y.T., Stec, A.O., Michno, J.M., Roessler, J., Bhaskar, P.B., Ries, L., Dobbels, A.A. *et al.* (2014) Genome resilience and prevalence of segmental duplications following fast neutron irradiation of soybean. *Genetics*, **198**, 967–981.
- Breuer, E.M., Bedon, F., Gavrin, A., Kryvoruchko, I.S., Torres-Jerez, I., Udvardi, M.K., Day, D.A. and *et al.* (2020) GmVTL1a is an iron transporter on the symbiosome membrane of soybean with an important role in nitrogen fixation. *New Phytol.* **228**, 667–681.
- Brown, A.V., Connors, S.I., Huang, W., Wilkey, A.P., Grant, D., Weeks, N.T., Cannon, S.B. *et al.* (2020) A new decade and new data at SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res.* **49**, D1496–D1501.
- Burton, J.N., Adey, A., Patwardhan, R.P., Qiu, R., Kitzman, J.O. and Shendure, J. (2013) Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nat. Biotechnol.* **31**, 1119–1125.
- Caetano-Anolles, G. and Gresshoff, P.M. (1991) Plant genetic control of nodulation. *Annu. Rev. Microbiol.* **45**, 345–382.
- Cai, Y., Chen, L., Liu, X., Guo, C., Sun, S., Wu, C., Jiang, B. *et al.* (2018) CRISPR/Cas9-mediated targeted mutagenesis of *GmFT2a* delays flowering time in soya bean. *Plant Biotechnol. J.* **16**, 176–185.
- Cai, Y., Chen, L., Zhang, Y., Yuan, S., Su, Q., Sun, S., Wu, C. *et al.* (2020a) Target base editing in soybean using a modified CRISPR/Cas9 system. *Plant Biotechnol. J.* **18**, 1996–1998.
- Cai, Y., Wang, L., Chen, L., Wu, T., Liu, L., Sun, S., Wu, C. *et al.* (2020b) Mutagenesis of *GmFT2a* and *GmFT5a* mediated by CRISPR/Cas9 contributes for expanding the regional adaptability of soybean. *Plant Biotechnol. J.* **18**, 298–309.
- Cai, Z., Wang, Y., Zhu, L., Tian, Y., Chen, L., Sun, Z., Ullah, I. *et al.* (2017) GmTIR1/GmAFB3-based auxin perception regulated by miR393 modulates soybean nodulation. *New Phytol.* **215**, 672–686.

- Cai, Z., Xian, P., Cheng, Y., Ma, Q., Lian, T., Nian, H. and Ge, L. (2021) CRISPR/Cas9-mediated gene editing of *GmJAGGED1* increased yield in the low-latitude soybean variety Huachun 6. *Plant Biotechnol. J.* **21**, 1–3.
- Cao, D., Li, Y., Lu, S.J., Wang, J.L., Nan, H.Y., Li, X.M., Shi, D.N. et al. (2015a) *GmCOL1a* and *GmCOL1b* function as flowering repressors in soybean under long-day conditions. *Plant Cell Physiol.* **56**, 2409–2422.
- Cao, D., Li, Y., Wang, J., Nan, H., Wang, Y., Lu, S., Jiang, Q. et al. (2015b) *GmmiR156b* overexpression delays flowering time in soybean. *Plant Mol. Biol.* **89**, 353–363.
- Carroll, B.J., Mcneil, D.L. and Gresshoff, P.M. (1985a) Isolation and properties of soybean [*Glycine-max* (L) Merr] mutants that nodulate in the presence of high nitrate concentrations. *Proc. Natl Acad. Sci. USA*, **82**, 4162–4166.
- Carroll, B.J., McNeil, D.L. and Gresshoff, P.M. (1985b) A Supernodulation and nitrate-tolerant symbiotic (*nts*) soybean mutant. *Plant Physiol.* **78**, 34–40.
- Carter, T.E., Nelson, R., Sneller, C.H. and Cui, Z. (2004) *Soybeans: Improvement, Production and Uses*. Third. Madison: American Society of Agronomy.
- Chan, C., Qi, X., Li, M.W., Wong, F.L. and Lam, H.M. (2012) Recent developments of genomic research in soybean. *J. Genet. Genomics*, **39**, 317–324.
- Chang, H.X., Tan, R., Hartman, G.L., Wen, Z., Sang, H., Domier, L.L., Whitham, S.A. et al. (2019) Characterization of soybean *STAY-GREEN* genes in susceptibility to foliar chlorosis of sudden death syndrome. *Plant Physiol.* **180**, 711–717.
- Charon, C., Sousa, C., Crespi, M. and Kondorosi, A. (1999) Alteration of *enod40* expression modifies medicago truncatula root nodule development induced by sinorhizobium meliloti. *Plant Cell*, **11**, 1953–1966.
- Chaudhary, J., Patil, G.B., Sonah, H., Deshmukh, R.K., Vuong, T.D., Valliyodan, B. and Nguyen, H.T. (2015) Expanding omics resources for improvement of soybean seed composition traits. *Front. Plant Sci.* **6**, 1021.
- Chen, W., Yao, Q., Patil, G.B., Agarwal, G., Deshmukh, R.K., Lin, L., Wang, B. et al. (2016) Identification and comparative analysis of differential gene expression in soybean leaf tissue under drought and flooding stress revealed by RNA-Seq. *Front. Plant Sci.* **7**, 1044.
- Chen, L., Cai, Y., Liu, X., Yao, W., Guo, C., Sun, S., Wu, C. et al. (2018a) Improvement of soybean *Agrobacterium*-mediated transformation efficiency by adding glutamine and asparagine into the culture media. *Int. J. Mol. Sci.* **19**, 3039.
- Chen, L., Zheng, Y., Dong, Z., Meng, F., Sun, X., Fan, X., Zhang, Y. et al. (2018b) Soybean (*Glycine max*) WRINKLED1 transcription factor, *GmWRI1a*, positively regulates seed oil accumulation. *Mol. Genet. Genomics*, **293**, 401–415.
- Chen, L., Qin, L., Zhou, L., Li, X., Chen, Z., Sun, L., Wang, W. et al. (2019) A nodule-localized phosphate transporter *GmPT7* plays an important role in enhancing symbiotic N₂ fixation and yield in soybean. *New Phytol.* **221**, 2013–2025.
- Chen, B., Zhang, G., Li, P., Yang, J., Guo, L., Benning, C., Wang, X. et al. (2020a) Multiple *GmWRI1s* are redundantly involved in seed filling and nodulation by regulating plastidic glycolysis, lipid biosynthesis and hormone signalling in soybean (*Glycine max*). *Plant Biotechnol. J.* **18**, 155–171.
- Chen, L., Cai, Y., Qu, M., Wang, L., Sun, H., Jiang, B., Wu, T. et al. (2020b) Soybean adaption to high-latitude regions is associated with natural variations of *GmFT2b*, an ortholog of *FLOWERING LOCUS T*. *Plant Cell Environ.* **43**, 934–944.
- Chen, L., Nan, H., Kong, L., Yue, L., Yang, H., Zhao, Q., Fang, C. et al. (2020c) Soybean AP1 homologs control flowering time and plant height. *J. Integr. Plant Biol.* **62**, 1868–1879.
- Chen, L., Yang, H., Fang, Y., Guo, W., Chen, H., Zhang, X., Dai, W. et al. (2021) Overexpression of *GmMYB14* improves high-density yield and drought tolerance of soybean through regulating plant architecture mediated by the brassinosteroid pathway. *Plant Biotechnol. J.* **19**, 702–716.
- Cheng, Q., Dong, L., Su, T., Li, T., Gan, Z., Nan, H., Lu, S. et al. (2019) CRISPR/Cas9-mediated targeted mutagenesis of *GmLHY* genes alters plant height and internode length in soybean. *BMC Plant Biol.* **19**, 562.
- Chi, Y., Wang, T., Xu, G., Yang, H., Zeng, X., Shen, Y., Yu, D. et al. (2017) *GmAGL1*, a MADS-Box gene from soybean, is involved in floral organ identity and fruit dehiscence. *Front. Plant Sci.* **8**, 175.
- Chiasson, D.M., Loughlin, P.C., Mazurkiewicz, D., Mohammadidehcheshmeh, M., Fedorova, E.E., Okamoto, M., McLean, E. et al. (2014) Soybean *SAT1* (*Symbiotic Ammonium Transporter 1*) encodes a bHLH transcription factor involved in nodule growth and NH₄⁺ transport. *Proc. Natl Acad. Sci. USA*, **111**, 4814–4819.
- Choudhury, S.R. and Pandey, S. (2013) Specific subunits of heterotrimeric G proteins play important roles during nodulation in soybean. *Plant Physiol.* **162**, 522–533.
- Choudhury, S.R. and Pandey, S. (2015) Phosphorylation-dependent regulation of G-protein cycle during nodule formation in soybean. *Plant Cell*, **27**, 3260–3276.
- Chu, Q., Sha, Z., Maruyama, H., Yang, L., Pan, G., Xue, L. and Watanabe, T. (2019) Metabolic reprogramming in nodules, roots, and leaves of symbiotic soybean in response to iron deficiency. *Plant Cell Environ.* **42**, 3027–3043.
- Chu, S., Wang, J., Zhu, Y., Liu, S., Zhou, X., Zhang, H., Wang, C.E. et al. (2017) An R2R3-type MYB transcription factor, *GmMYB29*, regulates isoflavone biosynthesis in soybean. *PLoS Genet.* **13**, e1006770.
- Chung, J., Babka, H.L., Graef, G.L., Staswick, P.E., Lee, D.J., Cregan, P.B., Shoemaker, R.C. et al. (2003) The seed protein, oil, and yield QTL on soybean linkage group I. *Crop Sci.* **43**, 1053–1067.
- Chung, W.H., Jeong, N., Kim, J., Lee, W.K., Lee, Y.G., Lee, S.H., Yoon, W. et al. (2014) Population structure and domestication revealed by high-depth resequencing of Korean cultivated and wild soybean genomes. *DNA Res.* **21**, 153–167.
- Cober, E.R., Molnar, S.J., Charette, M. and Voldeng, H.D. (2010) A new locus for early maturity in soybean. *Crop Sci.* **50**, 524–527.
- Cober, E.R., Tanner, J.W. and Voldeng, H.D. (1996) Soybean photoperiod-sensitivity loci respond differentially to light quality. *Crop Sci.* **36**, 606–610.
- Cober, E.R. and Voldeng, H.D. (2001) A new soybean maturity and photoperiod-sensitivity locus linked to *E1* and *T*. *Crop Sci.* **41**, 698–701.
- Collier, R. and Tegeder, M. (2012) Soybean ureide transporters play a critical role in nodule development, function and nitrogen export. *Plant J.* **72**, 355–367.
- Combs, R. and Bilyeu, K. (2019) Novel alleles of *FAD2-1A* induce high levels of oleic acid in soybean oil. *Mol. Breed.* **39**, 79.
- Cook, D.E., Bayless, A.M., Wang, K., Guo, X., Song, Q., Jiang, J. and Bent, A.F. (2014) Distinct copy number, coding sequence, and locus methylation patterns underlie *Rhg1*-mediated soybean resistance to soybean cyst nematode. *Plant Physiol.* **165**, 630–647.
- Cook, D.E., Lee, T.G., Guo, X.L., Melito, S., Wang, K., Bayless, A.M., Wang, J.P. et al. (2012) Copy number variation of multiple genes at *Rhg1*-mediates nematode resistance in soybean. *Science*, **338**, 1206–1209.
- Couzigou, J.M., André, O., Guillotin, B., Alexandre, M. and Combiere, J.P. (2016) Use of microRNA-encoded peptide miPEP172c to stimulate nodulation in soybean. *New Phytol.* **211**, 379–381.
- Cregan, P.B., Jarvik, T., Bush, A.L., Shoemaker, R.C., Lark, K.G., Kahler, A.L., Kaya, N. et al. (1999) An integrated genetic linkage map of the soybean genome. *Crop Sci.* **39**, 1464–1490.
- Dan, Y. & Reichert, N.A. (1999) *Soybean transformation and regeneration methods*. United States Patent, PAT: US5968830.
- Dastmalchi, M., Bernards, M.A. and Dhaubhadel, S. (2016) Twin anchors of the soybean isoflavonoid metabolon: evidence for tethering of the complex to the endoplasmic reticulum by IFS and C4H. *Plant J.* **85**, 689–706.
- Day, D.A., Carroll, B.J., Delves, A.C. and Gresshoff, P.M. (2010) Relationship between autoregulation and nitrate inhibition of nodulation in soybeans. *Physiol. Plant*, **75**, 37–42.
- Diamond, J. (2002) Evolution, consequences and future of plant and animal domestication. *Nature*, **418**, 700–707.
- Dillon, F.M., Tejedor, M.D., Ilina, N., Chludil, H.D., Mithöfer, A., Pagano, E.A. and Zavala, J.A. (2018) Solar UV-B radiation and ethylene play a key role in modulating effective defenses against *Anticarsia gemmatilis* larvae in field-grown soybean. *Plant Cell Environ.* **41**, 383–394.
- Dobbels, A.A., Michno, J.M., Campbell, B.W., Viridi, K.S., Stec, A.O., Muehlbauer, G.J., Naeve, S.L. et al. (2017) An induced chromosomal translocation in soybean disrupts a KASI ortholog and is associated with a high-sucrose and low-oil seed phenotype. *G3: Genes - Genomes - Genetics*, **7**, 1215–1223.

- Dong, J., Zielinski, R.E. and Hudson, M.E. (2020) t-SNAREs bind the *Rhg1* α -SNAP and mediate soybean cyst nematode resistance. *Plant J.* **104**, 318–331.
- Dong, L., Cheng, Y., Wu, J., Cheng, Q., Li, W., Fan, S., Jiang, L. et al. (2015) Overexpression of *GmERF5*, a new member of the soybean EAR motif-containing ERF transcription factor, enhances resistance to Phytophthora sojae in soybean. *J. Exp. Bot.* **66**, 2635–2647.
- Dong, Y., Yang, X., Liu, J., Wang, B.H., Liu, B.L. and Wang, Y.Z. (2014) Pod shattering resistance associated with domestication is mediated by a NAC gene in soybean. *Nat. Commun.* **5**, 3352.
- Du, H., Zeng, X., Zhao, M., Cui, X., Wang, Q., Yang, H., Cheng, H. et al. (2016) Efficient targeted mutagenesis in soybean by TALENs and CRISPR/Cas9. *J. Biotechnol.* **217**, 90–97.
- Du, J., Grant, D., Tian, Z., Nelson, R.T., Zhu, L., Shoemaker, R.C. and Ma, J. (2010a) SoyTEdb: a comprehensive database of transposable elements in the soybean genome. *BMC Genom.* **11**, 113.
- Du, J., Tian, Z., Bowen, N.J., Schmutz, J., Shoemaker, R.C. and Ma, J. (2010b) Bifurcation and enhancement of autonomous-nonautonomous retrotransposon partnership through LTR Swapping in soybean. *Plant Cell*, **22**, 48–61.
- Du, J., Tian, Z., Sui, Y., Zhao, M., Song, Q., Cannon, S.B., Cregan, P. et al. (2012) Pericentromeric effects shape the patterns of divergence, retention, and expression of duplicated genes in the paleopolyploid soybean. *Plant Cell*, **24**, 21–32.
- Du, J., Wang, S., He, C., Zhou, B., Ruan, Y.L. and Shou, H. (2017) Identification of regulatory networks and hub genes controlling soybean seed set and size using RNA sequencing analysis. *J. Exp. Bot.* **68**, 1955–1972.
- Du, Y.T., Zhao, M.J., Wang, C.T., Gao, Y., Wang, Y.X., Liu, Y.W., Chen, M. et al. (2018) Identification and characterization of *GmMYB118* responses to drought and salt stress. *BMC Plant Biol.* **18**, 320.
- El Baidouri, M., Kim, K.D., Abernathy, B., Li, Y.H., Qiu, L.J. and Jackson, S.A. (2018) Genic C-methylation in soybean is associated with gene paralogs relocated to transposable element-rich pericentromeres. *Mol. Plant*, **11**, 485–495.
- Espina, M.J., Ahmed, C.M.S., Bernardini, A., Adeleke, E., Yadegari, Z., Arelli, P., Pantalone, V. et al. (2018) Development and phenotypic screening of an ethyl methane sulfonate mutant population in soybean. *Front. Plant Sci.* **9**, 394.
- Fang, C., Li, C., Li, W., Wang, Z., Zhou, Z., Shen, Y., Wu, M. et al. (2014) Concerted evolution of *D1* and *D2* to regulate chlorophyll degradation in soybean. *Plant J.* **77**, 700–712.
- Fang, C., Li, W., Li, G., Wang, Z., Zhou, Z., Ma, Y., Shen, Y. et al. (2013) Cloning of *Ln* gene through combined approach of map-based cloning and association study in soybean. *J. Genet. Genomics*, **40**, 93–96.
- Fang, C., Liu, J., Zhang, T., Su, T., Li, S., Cheng, Q., Kong, L. et al. (2020) A recent retrotransposon insertion of *J* caused *E6* locus facilitating soybean adaptation into low latitude. *J. Integr. Plant Biol.* **63**, 995–1003.
- Fang, C., Ma, Y., Wu, S., Liu, Z., Wang, Z., Yang, R., Hu, G. et al. (2017) Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. *Genome Biol.* **18**, 161.
- Fang, W., Wang, Z., Cui, R., Li, J. and Li, Y. (2012) Maternal control of seed size by *EOD3/CYP78A6* in *Arabidopsis thaliana*. *Plant J.* **70**, 929–939.
- Fang, X., Han, Y., Liu, M., Jiang, J., Li, X., Lian, Q., Xie, X. et al. (2021a) Modulation of evening complex activity enables north-to-south adaptation of soybean. *Sci. China Life Sci.* **64**, 179–195.
- Fang, X.L., Sun, X.Y., Yang, X.D., Li, Q., Lin, C.J., Xu, J., Gong, W.J. et al. (2021b) *MS1* is essential for male fertility by regulating the microsporocyte cell plate expansion in soybean. *Sci. China Life Sci.* <https://doi.org/10.1007/s11427-021-1973-0>
- Feng, X., Feng, P., Yu, H., Yu, X., Sun, Q., Liu, S., Minh, T.N. et al. (2020) GsSnRK1 interplays with transcription factor GsERF7 from wild soybean to regulate soybean stress resistance. *Plant Cell Environ.* **43**, 1192–1211.
- Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.L., Wei, P., Cao, F. et al. (2013) Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res.* **23**, 1229–1232.
- Ferguson, B.J., Mens, C., Hastwell, A.H., Zhang, M., Su, H., Jones, C.H., Chu, X. et al. (2019) Legume nodulation: the host controls the party. *Plant Cell Environ.* **42**, 41–51.
- Fernie, A.R. and Yan, J. (2019) De Novo domestication: an alternative route toward new crops for the future. *Mol. Plant*, **12**, 615–631.
- Finch-Savage, W.E. and Leubner-Metzger, G. (2006) Seed dormancy and the control of germination. *New Phytol.* **171**, 501–523.
- Finer, J.J. and McMullen, M. (1991) Transformation of soybean via particle bombardment of embryogenic suspension culture tissue. *Vitro Cell. Dev. Biol. Plant*, **27**, 175–182.
- Franklin, K.A. and Quail, P.H. (2010) Phytochrome functions in *Arabidopsis* development. *J. Exp. Bot.* **61**, 11–24.
- Funatsuki, H., Suzuki, M., Hirose, A., Inaba, H., Yamada, T., Hajika, M., Komatsu, K. et al. (2014) Molecular basis of a shattering resistance boosting global dissemination of soybean. *Proc. Natl Acad. Sci. USA*, **111**, 17797–17802.
- Gao, C. (2021) Genome engineering for crop improvement and future agriculture. *Cell*, **184**, 1621–1635.
- Gao, J., Yang, S., Cheng, W., Fu, Y., Leng, J., Yuan, X., Jiang, N. et al. (2017) *GmLPA1*, encoding an APC8-like protein, controls leaf petiole angle in soybean. *Plant Physiol.* **174**, 1167–1176.
- Gao, J., Yang, S., Tang, K., Li, G., Gao, X., Liu, B., Wang, S. et al. (2021) *GmCCD4* controls carotenoid content in soybeans. *Plant Biotechnol. J.* **19**, 801–813.
- Gavrin, A., Chiasson, D., Ovchinnikova, E., Kaiser, B.N., Bisseling, T. and Fedorova, E.E. (2016) VAMP721a and VAMP721d are important for pectin dynamics and release of bacteria in soybean nodules. *New Phytol.* **210**, 1011–1021.
- Ge, L., Yu, J., Wang, H., Luth, D., Bai, G., Wang, K. and Chen, R. (2016) Increasing seed size and quality by manipulating *BIG SEEDS1* in legume species. *Proc. Natl Acad. Sci. USA*, **113**, 12414–12419.
- Ge, X., Chu, Z., Lin, Y. and Wang, S. (2006) A tissue culture system for different germplasms of indica rice. *Plant Cell Rep.* **25**, 392–402.
- Gillman, J.D., Stacey, M.G., Cui, Y., Berg, H.R. and Stacey, G. (2014) Deletions of the *SACP-D* locus elevate seed stearic acid levels but also result in fatty acid and morphological alterations in nitrogen fixing nodules. *BMC Plant Biol.* **14**, 1–19.
- Gil-Quintana, E., Larrainzar, E., Seminario, A., Díaz-Leal, J.L., Alamillo, J.M., Pineda, M., Arrese-Igor, C. et al. (2013) Local inhibition of nitrogen fixation and nodule metabolism in drought-stressed soybean. *J. Exp. Bot.* **64**, 2171–2182.
- Golicz, A.A., Singh, M.B. and Bhalla, P.L. (2018) The long intergenic noncoding RNA (lincRNA) landscape of the soybean genome. *Plant Physiol.* **176**, 2133–2147.
- Gong, Z. (2020) Flowering phenology as a core domestication trait in soybean. *J. Integr. Plant Biol.* **62**, 546–549.
- Grant, D., Nelson, R.T., Cannon, S.B. and Shoemaker, R.C. (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res.* **38**, D843–D846.
- Griesmann, M., Chang, Y., Liu, X., Song, Y., Haberer, G., Crook, M.B., Billault-Penneteau, B. et al. (2018) Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. *Science*, **361**, eaat1743.
- Grunewald, W., van Noorden, G., Van Isterdael, G., Beeckman, T., Gheysen, G. and Mathesius, U. (2009) Manipulation of auxin transport in plant roots during Rhizobium symbiosis and nematode parasitism. *Plant Cell*, **21**, 2553–2562.
- Gu, Y., Li, W., Jiang, H., Wang, Y., Gao, H., Liu, M., Chen, Q. et al. (2017) Differential expression of a *WRKY* gene between wild and cultivated soybeans correlates to seed size. *J. Exp. Bot.* **68**, 2717–2729.
- Guan, R., Qu, Y., Guo, Y., Yu, L., Liu, Y., Jiang, J., Chen, J. et al. (2014) Salinity tolerance in soybean is modulated by natural variation in *GmSALT3*. *Plant J.* **80**, 937–950.
- Guo, W., Chen, L., Chen, H., Yang, H., You, Q., Bao, A., Chen, S. et al. (2020) Overexpression of *GmWR1b* in soybean stably improves plant architecture and associated yield parameters, and increases total seed oil production under field conditions. *Plant Biotechnol. J.* **18**, 1639–1641.
- Guo, W., Zhao, J., Li, X., Qin, L., Yan, X. and Liao, H. (2011) A soybean β -expansin gene *GmEXPB2* intrinsically involved in root system architecture responses to abiotic stresses. *Plant J.* **66**, 541–552.
- Guo, X., Chronis, D., De La Torre, C.M., Smeda, J., Wang, X. and Mitchum, M.G. (2015) Enhanced resistance to soybean cyst nematode *Heterodera glycines* in transgenic soybean by silencing putative CLE receptors. *Plant Biotechnol. J.* **13**, 801–810.

- Hacisalihoglu, G., Burton, A.L., Gustin, J.L., Eker, S., Asikli, S., Heybet, E.H., Ozturk, L. *et al.* (2018) Quantitative trait loci associated with soybean seed weight and composition under different phosphorus levels. *J. Integr. Plant Biol.* **60**, 232–241.
- Han, Y., Zhao, X., Liu, D., Li, Y., Lightfoot, D.A., Yang, Z., Zhao, L. *et al.* (2016) Domestication footprints anchor genomic regions of agronomic importance in soybeans. *New Phytol.* **209**, 871–884.
- Hao, Y.J., Wei, W., Song, Q.X., Chen, H.W., Zhang, Y.Q., Wang, F., Zou, H.F. *et al.* (2011) Soybean NAC transcription factors promote abiotic stress tolerance and lateral root formation in transgenic plants. *Plant J.* **68**, 302–313.
- Hartman, G.L., West, E.D. and Herman, T.K. (2011) Crops that feed the World 2. Soybean-worldwide production, use, and constraints caused by pathogens and pests. *Food Secur.* **3**, 5–17.
- Haun, W., Coffman, A., Clasen, B.M., Demorest, Z.L., Lowy, A., Ray, E., Retterath, A. *et al.* (2014) Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. *Plant Biotechnol. J.* **12**, 934–940.
- Haupt, M. and Schmid, K. (2020) Combining focused identification of germplasm and core collection strategies to identify genebank accessions for central European soybean breeding. *Plant Cell Environ.* **43**, 1421–1436.
- Hayashi, M., Saeki, Y., Haga, M., Harada, K., Kouchi, H. and Umehara, Y. (2012a) *Rj (rj)* genes involved in nitrogen-fixing root nodule formation in soybean. *Breed. Sci.* **61**, 544–553.
- Hayashi, M., Shiro, S., Kanamori, H., Mori-Hosokawa, S., Sasaki-Yamagata, H., Sayama, T., Nishioka, M. *et al.* (2014) A thaumatin-like protein, Rj4, controls nodule symbiotic specificity in soybean. *Plant Cell Physiol.* **55**, 1679–1689.
- Hayashi, S., Reid, D.E., Lorenc, M.T., Stiller, J., Edwards, D., Gresshoff, P.M. and Ferguson, B.J. (2012b) Transient Nod factor-dependent gene expression in the nodulation-competent zone of soybean (*Glycine max* [L.] Merr.) roots. *Plant Biotechnol. J.* **10**, 995–1010.
- Hayes, A.J., Ma, G.R., Buss, G.R. and Maroof, M.A.S. (2000) Molecular marker mapping of RSV4, a gene conferring resistance to all known strains of soybean mosaic virus. *Crop Sci.* **40**, 1434–1437.
- He, Y., Fu, J., Yu, C., Wang, X., Jiang, Q., Hong, J., Lu, K. *et al.* (2015) Increasing cyclic electron flow is related to Na⁺ sequestration into vacuoles for salt tolerance in soybean. *J. Exp. Bot.* **66**, 6877–6889.
- Hedden, P. (2003) The genes of the Green Revolution. *Trends Genet.* **19**, 5–9.
- Hernandez, G., Valdes-Lopez, O., Ramirez, M., Goffard, N., Weiller, G., Aparicio-Fabre, R., Fuentes, S.I. *et al.* (2009) Global changes in the transcript and metabolic profiles during symbiotic nitrogen fixation in phosphorus-stressed common bean plants. *Plant Physiol.* **151**, 1221–1238.
- Hinchey, M., Connor-Ward, D.V., Newell, C.A., McDonnell, R.E., Sato, S.J., Gasser, C.S., Fischhoff, D.A. *et al.* (1988) Production of transgenic soybean plants using *Agrobacterium*-mediated DNA transfer. *Biotechnology*, **6**, 915–922.
- Horvath, D.P., Hansen, S.A., Moriles-Miller, J.P., Pierik, R., Yan, C., Clay, D.E., Scheffler, B. *et al.* (2015) RNAseq reveals weed-induced PIF3-like as a candidate target to manipulate weed stress response in soybean. *New Phytol.* **207**, 196–210.
- Hossain, M.S., Kawakatsu, T., Kim, K.D., Zhang, N., Nguyen, C.T., Khan, S.M., Batek, J.M. *et al.* (2017) Divergent cytosine DNA methylation patterns in single-cell, soybean root hairs. *New Phytol.* **214**, 808–819.
- Hyten, D.L., Cannon, S.B., Song, Q.J., Weeks, N., Fickus, E.W., Shoemaker, R.C., Specht, J.E. *et al.* (2010a) High-throughput SNP discovery through deep resequencing of a reduced representation library to anchor and orient scaffolds in the soybean whole genome sequence. *BMC Genom.* **11**, 38.
- Hyten, D.L., Choi, I.Y., Song, Q.J., Specht, J.E., Carter, T.E., Shoemaker, R.C., Hwang, E.Y. *et al.* (2010b) A High density integrated genetic linkage map of soybean and the development of a 1536 universal soy linkage panel for quantitative trait locus mapping. *Crop Sci.* **50**, 960–968.
- Im, J.H., Lee, H., Kim, J., Kim, H.B. and An, C.S. (2012) Soybean MAPK, GMK1 is dually regulated by phosphatidic acid and hydrogen peroxide and translocated to nucleus during salt stress. *Mol. Cells*, **34**, 271–278.
- Indrasumunar, A., Kereszt, A., Searle, I., Miyagi, M., Li, D., Nguyen, C.D., Men, A. *et al.* (2010) Inactivation of duplicated nod factor receptor 5 (NFR5) genes in recessive loss-of-function non-nodulation mutants of allotetraploid soybean (*Glycine max* L. Merr.). *Plant Cell Physiol.* **51**, 201–214.
- Indrasumunar, A., Searle, I., Lin, M.H., Kereszt, A., Men, A., Carroll, B.J. and Gresshoff, P.M. (2011) Nodulation factor receptor kinase 1alpha controls nodule organ number in soybean (*Glycine max* L. Merr.). *Plant J.* **65**, 39–50.
- Ishibashi, K., Saruta, M., Shimizu, T., Shu, M., Anai, T., Komatsu, K., Yamada, N. *et al.* (2019) Soybean antiviral immunity conferred by dsRNase targets the viral replication complex. *Nat. Commun.* **10**, 4033.
- Ishida, Y., Saito, H., Ohta, S., Hiei, Y., Komari, T. and Kumashiro, T. (1996) High efficiency transformation of maize (*Zea mays* L.) mediated by *Agrobacterium tumefaciens*. *Nat. Biotechnol.* **14**, 745–750.
- Jahan, M.A., Harris, B., Lowery, M., Infante, A.M., Percifield, R.J. and Kovich, N. (2020) Glyceollin transcription factor GmMYB29A2 regulates soybean resistance to *Phytophthora sojae*. *Plant Physiol.* **183**, 530–546.
- Jang, S.J., Sato, M., Sato, K., Jitsuyama, Y., Fujino, K., Mori, H., Takahashi, R. *et al.* (2015) A single-nucleotide polymorphism in an endo-1,4-beta-glucanase gene controls seed coat permeability in soybean. *PLoS One*, **10**, e0128527.
- Jeong, N., Moon, J.K., Kim, H.S., Kim, C.G. and Jeong, S.C. (2011) Fine genetic mapping of the genomic region controlling leaflet shape and number of seeds per pod in the soybean. *Theor. Appl. Genet.* **122**, 865–874.
- Jeong, N., Suh, S.J., Kim, M.H., Lee, S., Moon, J.K., Kim, H.S. and Jeong, S.C. (2012) *Ln* is a key regulator of leaflet shape and number of seeds per pod in soybean. *Plant Cell*, **24**, 4807–4818.
- Jeong, S.C., Kristipati, S., Hayes, A.J., Maughan, P.J., Noffsinger, S.L., Gunduz, I., Buss, G.R. *et al.* (2002) Genetic and sequence analysis of markers tightly linked to the soybean mosaic virus resistance gene, Rsv3. *Crop Sci.* **42**, 265–270.
- Jiang, B., Chen, L., Yang, C., Wu, T., Yuan, S., Wu, C., Zhang, M. *et al.* (2021) The cloning and CRISPR/Cas9-mediated mutagenesis of a male sterility gene *MS1* of soybean. *Plant Biotechnol. J.* **19**, 1098–1110.
- Jiang, B., Zhang, S., Song, W., Khan, M.A.A., Sun, S., Zhang, C., Wu, T. *et al.* (2019) Natural variations of *FT* family genes in soybean varieties covering a wide range of maturity groups. *BMC Genom.* **20**, 230.
- Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B. and Weeks, D.P. (2013) Demonstration of CRISPR/Cas9/sgrRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res.* **41**, e188.
- Jin, D., Meng, X., Wang, Y., Wang, J., Zhao, Y. and Chen, M. (2018) Computational investigation of small RNAs in the establishment of root nodules and arbuscular mycorrhiza in leguminous plants. *Sci. China Life Sci.* **61**, 706–717.
- Joshi, T., Fitzpatrick, M.R., Chen, S., Liu, Y., Zhang, H., Endacott, R.Z., Gaudiello, E.C. *et al.* (2014) Soybean knowledge base (SoyKB): a web resource for integration of soybean translational genomics and molecular breeding. *Nucleic Acids Res.* **42**, D1245–1252.
- Kajiya-Kanegae, H., Nagasaki, H., Kaga, A., Hirano, K., Ogiso-Tanaka, E., Matsuoka, M., Ishimori, M. *et al.* (2021) Whole-genome sequence diversity and association analysis of 198 soybean accessions in mini-core collections. *DNA Res.* **28**, dsaa032.
- Kanai, M., Yamada, T., Hayashi, M., Mano, S. and Nishimura, M. (2019) Soybean (*Glycine max* L.) triacylglycerol lipase GmSDP1 regulates the quality and quantity of seed oil. *Sci. Rep.* **9**, 8924.
- Kanazawa, A., Liu, B., Kong, F., Arase, S. and Abe, J. (2009) Adaptive evolution involving gene duplication and insertion of a novel Ty1/copia-like retrotransposon in soybean. *J. Mol. Evol.* **69**, 164–175.
- Kandath, P.K., Ithal, N., Recknor, J., Maier, T., Nettleton, D., Baum, T.J. and Mitchum, M.G. (2011) The Soybean *Rhgt1* locus for resistance to the soybean cyst nematode *Heterodera glycines* regulates the expression of a large number of stress- and defense-related genes in degenerating feeding cells. *Plant Physiol.* **155**, 1960–1975.
- Kang, Y.J., Kim, K.H., Shim, S., Yoon, M.Y., Sun, S., Kim, M.Y., Van, K. *et al.* (2012) Genome-wide mapping of *NBS-LRR* genes and their association with disease resistance in soybean. *BMC Plant Biol.* **12**, 139.
- Kawashima, C.G., Guimarães, G.A., Nogueira, S.R., MacLean, D., Cook, D.R., Steuernagel, B., Baek, J. *et al.* (2016) A pigeonpea gene confers resistance to Asian soybean rust in soybean. *Nat. Biotechnol.* **34**, 661–665.
- Khan, M.N. and Komatsu, S. (2016) Proteomic analysis of soybean root including hypocotyl during recovery from drought stress. *J. Proteomics*, **144**, 39–50.

- Khan, M.Z., Zaidi, S.S., Amin, I. and Mansoor, S. (2019) A CRISPR way for fast-forward crop domestication. *Trends Plant Sci.* **24**, 293–296.
- Khatoun, A., Rehman, S., Hiraga, S., Makino, T. and Komatsu, S. (2012) Organ-specific proteomics analysis for identification of response mechanism in soybean seedlings under flooding stress. *J. Proteomics*. **75**, 5706–5723.
- Kidokoro, S., Watanabe, K., Ohori, T., Moriwaki, T., Maruyama, K., Mizoi, J., Myint Phyu Sin Htwe, N. et al. (2015) Soybean DREB1/CBF-type transcription factors function in heat and drought as well as cold stress-responsive gene expression. *Plant J.* **81**, 505–518.
- Kim, E., Hwang, S. and Lee, I. (2017) SoyNet: a database of co-functional networks for soybean *Glycine max*. *Nucleic Acids Res.* **45**, D1082–D1089.
- Kim, K.D., El Baidouri, M., Abernathy, B., Iwata-Otsubo, A., Chavarro, C., Gonzales, M., Libault, M. et al. (2015) A comparative epigenomic analysis of polyploidy-derived genes in soybean and common bean. *Plant Physiol.* **168**, 1433–1447.
- Kim, M.Y., Lee, S., Van, K., Kim, T.H., Jeong, S.C., Choi, I.Y., Kim, D.S. et al. (2010) Whole-genome sequencing and intensive analysis of the undomesticated soybean (*Glycine soja* Sieb. and Zucc.) genome. *Proc. Natl Acad. Sci. USA*, **107**, 22032–22037.
- Kim, M.Y., Shin, J.H., Kang, Y.J., Shim, S.R. and Lee, S.H. (2012a) Divergence of flowering genes in soybean. *J. Biosci.* **37**, 857–870.
- Kim, W.S., Chronis, D., Juergens, M., Schroeder, A.C., Hyun, S.W., Jez, J.M. and Krishnan, H.B. (2012b) Transgenic soybean plants overexpressing O-acetylserine sulfhydrylase accumulate enhanced levels of cysteine and Bowman-Birk protease inhibitor in seeds. *Planta*, **235**, 13–23.
- Kim, W.S., Sun-Hyung, J., Oehrle, N.W., Jez, J.M. and Krishnan, H.B. (2020) Overexpression of ATP sulfurylase improves the sulfur amino acid content, enhances the accumulation of Bowman-Birk protease inhibitor and suppresses the accumulation of the beta-subunit of beta-conglycinin in soybean seeds. *Sci. Rep.* **10**, 14989.
- Kim, Y.K., Kim, S., Um, J.H., Kim, K., Choi, S.K., Um, B.H., Kang, S.W. et al. (2013) Functional implication of beta-carotene hydroxylases in soybean nodulation. *Plant Physiol.* **162**, 1420–1433.
- Klepado, M., Chen, P.Y., Shi, A.N., Mason, R.E., Korth, K.L., Srivastava, V. and Wu, C.J. (2017) Two tightly linked genes for soybean mosaic virus resistance in soybean. *Crop Sci.* **57**, 1844–1853.
- Kohzuma, K., Sato, Y., Ito, H., Okuzaki, A., Watanabe, M., Kobayashi, H., Nakano, M. et al. (2017) The non-mendelian green cotyledon gene in soybean encodes a small subunit of photosystem II. *Plant Physiol.* **173**, 2138–2147.
- Komatsu, S., Kobayashi, Y., Nishizawa, K., Nanjo, Y. and Furukawa, K. (2010) Comparative proteomics analysis of differentially expressed proteins in soybean cell wall during flooding stress. *Amino Acids*, **39**, 1435–1449.
- Komatsu, S., Wang, X., Yin, X., Nanjo, Y., Ohyanagi, H. and Sakata, K. (2017) Integration of gel-based and gel-free proteomic data for functional analysis of proteins through Soybean Proteome Database. *J. Proteomics*. **163**, 52–66.
- Komatsu, S., Yamamoto, A., Nakamura, T., Nouri, M.Z., Nanjo, Y., Nishizawa, K. and Furukawa, K. (2011) Comprehensive analysis of mitochondria in roots and hypocotyls of soybean under flooding stress using proteomics and metabolomics techniques. *J. Proteome Res.* **10**, 3993–4004.
- Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A. and Liu, D.R. (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*, **533**, 420–424.
- Kong, F., Liu, B., Xia, Z., Sato, S., Kim, B.M., Watanabe, S., Yamada, T. et al. (2010) Two coordinately regulated homologs of *FLOWERING LOCUS T* are involved in the control of photoperiodic flowering in soybean. *Plant Physiol.* **154**, 1220–1231.
- Krishnan, H.B. and Jez, J.M. (2018) Review: the promise and limits for enhancing sulfur-containing amino acid content of soybean seed. *Plant Sci.* **272**, 14–21.
- Ku, Y.S., Ni, M., Muñoz, N.B., Xiao, Z., Lo, A.W., Chen, P., Li, M.W. et al. (2020) ABA51 from soybean is a 1R-subtype MYB transcriptional repressor that enhances ABA sensitivity. *J. Exp. Bot.* **71**, 2970–2981.
- Kumagai, E., Aoki, N., Masuya, Y. and Shimono, H. (2015) Phenotypic plasticity conditions the response of soybean seed yield to elevated atmospheric CO₂ concentration. *Plant Physiol.* **169**, 2021–2029.
- Kumagai, H., Kinoshita, E., Ridge, R.W. and Kouchi, H. (2006) RNAi knock-down of ENOD40s leads to significant suppression of nodule formation in *Lotus japonicus*. *Plant Cell Physiol.* **47**, 1102–1111.
- Kunert, K.J., Vorster, B.J., Fenta, B.A., Kibido, T., Dionisio, G. and Foyer, C.H. (2016) Drought stress responses in soybean roots and nodules. *Front. Plant Sci.* **7**, 1015.
- Lakhssassi, N., Zhou, Z., Liu, S., Colantonio, V., AbuGhazaleh, A. and Meksem, K. (2017) Characterization of the *FAD2* gene family in soybean reveals the limitations of gel-based TILLING in genes with high copy number. *Front. Plant Sci.* **8**, 324.
- Lakhssassi, N., Zhou, Z., Liu, S., Piya, S., Cullen, M.A., El Baze, A., Knizia, D. et al. (2020) Soybean TILLING-by-sequencing + reveals the role of novel GmSACPD members in the unsaturated fatty acid biosynthesis while maintaining healthy nodules. *J. Exp. Bot.* **71**, 6969–6987.
- Lam, H.M., Xu, X., Liu, X., Chen, W., Yang, G., Wong, F.L., Li, M.W. et al. (2010) Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat. Genet.* **42**, 1053–1059.
- Langenbach, C., Campe, R., Schaffrath, U., Goellner, K. and Conrath, U. (2013) UDP-glucosyltransferase UGT84A2/BRT1 is required for Arabidopsis nonhost resistance to the Asian soybean rust pathogen *Phakopsora pachyrhizi*. *New Phytol.* **198**, 536–545.
- Langenbach, C., Schultheiss, H., Rosendahl, M., Tresch, N., Conrath, U. and Goellner, K. (2016) Interspecies gene transfer provides soybean resistance to a fungal pathogen. *Plant Biotechnol. J.* **14**, 699–708.
- Lee, T.G., Diers, B.W. and Hudson, M.E. (2016) An efficient method for measuring copy number variation applied to improvement of nematode resistance in soybean. *Plant J.* **88**, 143–153.
- Lee, Y.G., Jeong, N., Kim, J.H., Lee, K., Kim, K.H., Pirani, A., Ha, B.K. et al. (2015) Development, validation and genetic analysis of a large soybean SNP genotyping array. *Plant J.* **81**, 625–636.
- Li, C., Li, Y.H., Li, Y., Lu, H., Hong, H., Tian, Y., Li, H. et al. (2020a) A domestication-associated gene *GmPRR3b* regulates the circadian clock and flowering time in soybean. *Mol. Plant*, **13**, 745–759.
- Li, C., Nguyen, V., Liu, J., Fu, W., Chen, C., Yu, K. and Cui, Y. (2019a) Mutagenesis of seed storage protein genes in Soybean using CRISPR/Cas9. *BMC Res. Notes*, **12**, 176.
- Li, H., Deng, Y., Wu, T., Subramanian, S. and Yu, O. (2010) Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation. *Plant Physiol.* **153**, 1759–1770.
- Li, K.P., Wong, C.H., Cheng, C.C., Cheng, S.S., Li, M.W., Mansveld, S., Bergsma, A. et al. (2021) GmDNJ1, a type-I heat shock protein 40 (HSP40), is responsible for both Growth and heat tolerance in soybean. *Plant Direct*, **5**, e00298.
- Li, L., Gao, W., Peng, Q., Zhou, B., Kong, Q., Ying, Y. and Shou, H. (2018a) Two soybean bHLH factors regulate response to iron deficiency. *J. Integr. Plant Biol.* **60**, 608–622.
- Li, M., Cao, L., Mwimba, M., Zhou, Y., Li, L., Zhou, M., Schnable, P.S. et al. (2019b) Comprehensive mapping of abiotic stress inputs into the soybean circadian clock. *Proc. Natl Acad. Sci. USA*, **116**, 23840–23849.
- Li, M.W., Wang, Z., Jiang, B., Kaga, A., Wong, F.L., Zhang, G., Han, T. et al. (2020b) Impacts of genomic research on soybean improvement in East Asia. *Theor. Appl. Genet.* **133**, 1655–1678.
- Li, M.W., Xin, D., Gao, Y., Li, K.P., Fan, K., Muñoz, N.B., Yung, W.S. et al. (2017a) Using genomic information to improve soybean adaptability to climate change. *J. Exp. Bot.* **68**, 1823–1834.
- Li, P.S., Yu, T.F., He, G.H., Chen, M., Zhou, Y.B., Chai, S.C., Xu, Z.S. et al. (2014a) Genome-wide analysis of the Hsf family in soybean and functional identification of GmHsf-34 involvement in drought and heat stresses. *BMC Genom.* **15**, 1009.
- Li, Q., Fang, C., Duan, Z., Liu, Y., Qin, H., Zhang, J., Sun, P. et al. (2016a) Functional conservation and divergence of *GmCHLI* genes in polyploid soybean. *Plant J.* **88**, 584–596.
- Li, Q.T., Lu, X., Song, Q.X., Chen, H.W., Wei, W., Tao, J.J., Bian, X.H. et al. (2017b) Selection for a zinc-finger protein contributes to seed oil increase during soybean domestication. *Plant Physiol.* **173**, 2208–2224.
- Li, S., Cong, Y., Liu, Y., Wang, T., Shuai, Q., Chen, N., Gai, J. et al. (2017c) Optimization of *Agrobacterium*-mediated transformation in soybean. *Front. Plant Sci.* **8**, 246.
- Li, S., Wang, N., Ji, D., Xue, Z., Yu, Y., Jiang, Y., Liu, J. et al. (2016b) Evolutionary and functional analysis of membrane-bound NAC transcription factor genes in soybean. *Plant Physiol.* **172**, 1804–1820.

- Li, S., Wang, N., Ji, D., Zhang, W., Wang, Y., Yu, Y., Zhao, S. *et al.* (2019c) A GmSIN1/GmNCE3s/GmRbohBs feed-forward loop acts as a signal amplifier that regulates root growth in soybean exposed to salt stress. *Plant Cell*, **31**, 2107–2130.
- Li, T., Yang, X., Yu, Y., Si, X., Zhai, X., Zhang, H., Dong, W. *et al.* (2018b) Domestication of wild tomato is accelerated by genome editing. *Nat. Biotechnol.* **36**, 1160–1163.
- Li, W., Wang, T., Zhang, Y. and Li, Y. (2017d) Overexpression of soybean miR172c confers tolerance to water deficit and salt stress, but increases ABA sensitivity in transgenic *Arabidopsis thaliana*. *J. Exp. Bot.* **68**, 4727–4729.
- Li, X., Chen, L. and Dhaubhadel, S. (2012) 14-3-3 proteins regulate the intracellular localization of the transcriptional activator GmMYB176 and affect isoflavonoid synthesis in soybean. *Plant J.* **71**, 239–250.
- Li, X., Qin, R., Du, Q., Cai, L., Hu, D., Du, H., Yang, H. *et al.* (2020c) Knockdown of GmVQ58 encoding a VQ motif-containing protein enhances soybean resistance to the common cutworm (*Spodoptera litura Fabricius*). *J. Exp. Bot.* **71**, 3198–3210.
- Li, X., Zhao, J., Tan, Z., Zeng, R. and Liao, H. (2015a) GmEXPB2, a cell wall beta-expansin, affects soybean nodulation through modifying root architecture and promoting nodule formation and development. *Plant Physiol.* **169**, 2640–2653.
- Li, Y.H., Guan, R.X., Liu, Z.X., Ma, Y.S., Wang, L.X., Li, L.H., Lin, F.Y. *et al.* (2008) Genetic structure and diversity of cultivated soybean (*Glycine max* (L.) Merr.) landraces in China. *Theor. Appl. Genet.* **117**, 857–871.
- Li, Y., Zhang, J., Zhang, J., Hao, L., Hua, J., Duan, L., Zhang, M. *et al.* (2013a) Expression of an *Arabidopsis* molybdenum cofactor sulphurase gene in soybean enhances drought tolerance and increases yield under field conditions. *Plant Biotechnol. J.* **11**, 747–758.
- Li, Y.H., Zhao, S.C., Ma, J.X., Li, D., Yan, L., Li, J., Qi, X.T. *et al.* (2013b) Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing. *BMC Genom.* **14**, 579.
- Li, Y.H., Zhou, G., Ma, J., Jiang, W., Jin, L.G., Zhang, Z., Guo, Y. *et al.* (2014b) *De novo* assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. *Nat. Biotechnol.* **32**, 1045–1052.
- Li, Z., Jiang, L., Ma, Y., Wei, Z., Hong, H., Liu, Z., Lei, J. *et al.* (2017e) Development and utilization of a new chemically-induced soybean library with a high mutation density. *J. Integr. Plant Biol.* **59**, 60–74.
- Li, Z., Liu, Z.B., Xing, A., Moon, B.P., Koellhoffer, J.P., Huang, L., Ward, R.T. *et al.* (2015b) Cas9-guide RNA directed genome editing in soybean. *Plant Physiol.* **169**, 960–970.
- Li, Z.L., Qiu, L.J., Thompson, J.A., Welsh, M.M. and Nelson, R.L. (2001) Molecular genetic analysis of US and Chinese soybean ancestral lines. *Crop Sci.* **41**, 1330–1336.
- Liang, X., Hou, X., Li, J., Han, Y., Zhang, Y., Feng, N., Du, J. *et al.* (2019) High-resolution DNA methylome reveals that demethylation enhances adaptability to continuous cropping comprehensive stress in soybean. *BMC Plant Biol.* **19**, 79.
- Libault, M., Zhang, X.C., Govindarajulu, M., Qiu, J., Ong, Y.T., Brechenmacher, L., Berg, R.H. *et al.* (2010) A member of the highly conserved FWL (*tomato* FW2.2-like) gene family is essential for soybean nodule organogenesis. *Plant J.* **62**, 852–864.
- Lim, C.W., Lee, Y.W., Lee, S.C. and Hwang, C.H. (2014) Nitrate inhibits soybean nodulation by regulating expression of CLE genes. *Plant Sci.* **229**, 1–9.
- Lin, J., Mazarei, M., Zhao, N., Zhu, J.J., Zhuang, X., Liu, W., Pantalone, V.R. *et al.* (2013) Overexpression of a soybean salicylic acid methyltransferase gene confers resistance to soybean cyst nematode. *Plant Biotechnol. J.* **11**, 1135–1145.
- Lin, J., Mazarei, M., Zhao, N., Hatcher, C.N., Wuddineh, W.A., Rudis, M., Tschaplinski, T.J. *et al.* (2016) Transgenic soybean overexpressing GmSAMT1 exhibits resistance to multiple-HG types of soybean cyst nematode *Heterodera glycines*. *Plant Biotechnol. J.* **14**, 2100–2109.
- Lin, J., Wang, D., Chen, X., Köllner, T.G., Mazarei, M., Guo, H., Pantalone, V.R. *et al.* (2017) An (*E*, *E*)- α -farnesene synthase gene of soybean has a role in defence against nematodes and is involved in synthesizing insect-induced volatiles. *Plant Biotechnol. J.* **15**, 510–519.
- Lin, M.H., Gresshoff, P.M. and Ferguson, B.J. (2012) Systemic regulation of soybean nodulation by acidic growth conditions. *Plant Physiol.* **160**, 2028–2039.
- Lin, X., Lin, W., Ku, Y.S., Wong, F.L., Li, M.W., Lam, H.M., Ngai, S.M. *et al.* (2020a) Analysis of soybean long non-coding RNAs reveals a subset of small peptide-coding transcripts. *Plant Physiol.* **182**, 1359–1374.
- Lin, X., Liu, B., Weller, J.L., Abe, J. and Kong, F. (2020b) Molecular mechanisms for the photoperiodic regulation of flowering in soybean. *J. Integr. Plant Biol.* **63**, 981–994.
- Lin, Y.J. and Zhang, Q. (2005) Optimising the tissue culture conditions for high efficiency transformation of indica rice. *Plant Cell Rep.* **23**, 540–547.
- Liu, A., Xiao, Z., Li, M.W., Wong, F.L., Yung, W.S., Ku, Y.S., Wang, Q. *et al.* (2019) Transcriptomic reprogramming in soybean seedlings under salt stress. *Plant Cell Environ.* **42**, 98–114.
- Liu, B., Fujita, T., Yan, Z.H., Sakamoto, S., Xu, D. and Abe, J. (2007) QTL mapping of domestication-related traits in soybean (*Glycine max*). *Ann. Bot.* **100**, 1027–1038.
- Liu, B., Kanazawa, A., Matsumura, H., Takahashi, R., Harada, K. and Abe, J. (2008) Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene. *Genetics*, **180**, 995–1007.
- Liu, B., Watanabe, S., Uchiyama, T., Kong, F., Kanazawa, A., Xia, Z., Nagamatsu, A. *et al.* (2010) The soybean stem growth habit gene *Dt1* is an ortholog of *Arabidopsis* *TERMINAL FLOWER1*. *Plant Physiol.* **153**, 198–210.
- Liu, C.W. and Murray, J.D. (2016) The role of flavonoids in nodulation host-range specificity: an update. *Plants*, **5**, 33.
- Liu, H.K., Yang, C. and Wei, Z.M. (2004) Efficient *Agrobacterium tumefaciens*-mediated transformation of soybeans using an embryonic tip regeneration system. *Planta*, **219**, 1042–1049.
- Liu, J.Z., Braun, E., Qiu, W.L., Shi, Y.F., Marcelino-Guimaraes, F.C., Navarre, D., Hill, J.H. *et al.* (2014a) Positive and negative roles for soybean MPK6 in regulating defense responses. *Mol. Plant. Microbe Interact.* **27**, 824–834.
- Liu, J.Z., Graham, M.A., Pedley, K.F. and Whitham, S.A. (2015) Gaining insight into soybean defense responses using functional genomics approaches. *Brief. Funct. Genomics*, **14**, 283–290.
- Liu, J.Z., Horstman, H.D., Braun, E., Graham, M.A., Zhang, C., Navarre, D., Qiu, W.L. *et al.* (2011) Soybean homologs of MPK4 negatively regulate defense responses and positively regulate growth and development. *Plant Physiol.* **157**, 1363–1378.
- Liu, J.Y., Li, P., Zhang, Y.W., Zuo, J.F., Li, G., Han, X., Dunwell, J.M. *et al.* (2020a) Three-dimensional genetic networks among seed oil-related traits, metabolites and genes reveal the genetic foundations of oil synthesis in soybean. *Plant J.* **103**, 1103–1124.
- Liu, J.Z. and Whitham, S.A. (2013) Overexpression of a soybean nuclear localized type-III DnaI domain-containing HSP40 reveals its roles in cell death and disease resistance. *Plant J.* **74**, 110–121.
- Liu, J.Y., Zhang, Y.W., Han, X., Zuo, J.F., Zhang, Z., Shang, H., Song, Q. *et al.* (2020b) Evolutionary population structure model reveals pleiotropic effects of GmPDAT for seed oil- and size-related traits in soybean. *J. Exp. Bot.* **71**, 6988–7002.
- Liu, S., Fan, L., Liu, Z., Yang, X., Zhang, Z., Duan, Z., Liang, Q. *et al.* (2020c) A Pd1-Ps-P1 feedback loop controls pubescence density in soybean. *Mol. Plant*, **13**, 1768–1783.
- Liu, S., Kandoth, P.K., Warren, S.D., Yeckel, G., Heinz, R., Alden, J., Yang, C. *et al.* (2012) A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature*, **492**, 256–260.
- Liu, S., Kandoth, P.K., Lakhssassi, N., Kang, J., Colantonio, V., Heinz, R., Yeckel, G. *et al.* (2017) The soybean GmSNAP18 gene underlies two types of resistance to soybean cyst nematode. *Nat. Commun.* **8**, 14822.
- Liu, S., Liao, L.L., Nie, M.M., Peng, W.T., Zhang, M.S., Lei, J.N., Zhong, Y.J. *et al.* (2020d) A VIT-like transporter facilitates iron transport into nodule symbiosomes for nitrogen fixation in soybean. *New Phytol.* **226**, 1413–1428.
- Liu, S., Zhang, M., Feng, F. and Tian, Z. (2020e) Toward a “Green Revolution” for soybean. *Mol. Plant*, **13**, 688–697.
- Liu, T., Fang, C., Ma, Y., Shen, Y., Li, C., Li, Q., Wang, M. *et al.* (2016a) Global investigation of the co-evolution of MIRNA genes and microRNA targets during soybean domestication. *Plant J.* **85**, 396–409.
- Liu, W., Jiang, B., Ma, L., Zhang, S., Zhai, H., Xu, X., Hou, W. *et al.* (2018) Functional diversification of Flowering Locus T homologs in soybean: GmFT1a and GmFT2a/5a have opposite roles in controlling flowering and maturation. *New Phytol.* **217**, 1335–1345.

- Liu, Y., Du, H., Li, P., Shen, Y., Peng, H., Liu, S., Zhou, G.A. et al. (2020f) Pan-genome of wild and cultivated soybeans. *Cell*, **182**, 162–176.e113.
- Liu, Y.F., Li, Q.T., Lu, X., Song, Q.X., Lam, S.M., Zhang, W.K., Ma, B. et al. (2014b) Soybean *GmMYB73* promotes lipid accumulation in transgenic plants. *BMC Plant Biol.* **14**, 73.
- Liu, Y.C. and Tian, Z.X. (2020) From one linear genome to a graph-based pan-genome: a new era for genomics. *Sci. China Life Sci.* **12**, 1938–1941.
- Liu, Y.F., Zhang, D.J., Ping, J.Q., Li, S., Chen, Z.X. and Ma, J.X. (2016b) Innovation of a regulatory mechanism modulating semi-determinate stem growth through artificial selection in soybean. *PLoS Genet.* **12**, e1005818.
- Loehrer, M., Botterweck, J., Jahnke, J., Mahlmann, D.M., Gaetgens, J., Oldiges, M., Horbach, R. et al. (2014) In vivo assessment by Mach-Zehnder double-beam interferometry of the invasive force exerted by the Asian soybean rust fungus (*Phakopsora pachyrhizi*). *New Phytol.* **203**, 620–631.
- Lu, L., Wei, W., Li, Q.T., Bian, X.H., Lu, X., Hu, Y., Cheng, T. et al. (2021) A transcriptional regulatory module controls lipid accumulation in soybean. *New Phytol.* **231**, 661–678.
- Lu, M., Cheng, Z., Zhang, X.M., Huang, P., Fan, C., Yu, G., Chen, F. et al. (2020a) Spatial divergence of PHR-PHT1 modules maintains phosphorus homeostasis in soybean nodules. *Plant Physiol.* **184**, 236–250.
- Lu, S., Dong, L., Fang, C., Liu, S., Kong, L., Cheng, Q., Chen, L. et al. (2020b) Stepwise selection on homeologous *PRR* genes controlling flowering and maturity during soybean domestication. *Nat. Genet.* **52**, 428–436.
- Lu, S.J., Zhao, X.H., Hu, Y.L., Liu, S.L., Nan, H.Y., Li, X.M., Fang, C. et al. (2017a) Natural variation at the soybean *J* locus improves adaptation to the tropics and enhances yield. *Nat. Genet.* **49**, 773–779.
- Lu, X., Li, Q.T., Xiong, Q., Li, W., Bi, Y.D., Lai, Y.C., Liu, X.L. et al. (2016) The transcriptomic signature of developing soybean seeds reveals the genetic basis of seed trait adaptation during domestication. *Plant J.* **86**, 530–544.
- Lu, X., Xiong, Q., Cheng, T., Li, Q.T., Liu, X.L., Bi, Y.D., Li, W. et al. (2017b) A *PP2C-1* allele underlying a quantitative trait locus enhances soybean 100-Seed weight. *Mol. Plant*, **10**, 670–684.
- Luo, X., Bai, X., Sun, X., Zhu, D., Liu, B., Ji, W., Cai, H. et al. (2013) Expression of wild soybean *WRKY20* in *Arabidopsis* enhances drought tolerance and regulates ABA signalling. *J. Exp. Bot.* **64**, 2155–2169.
- Lyu, J. (2018) A skeleton key to domestication. *Nat. Plants*, **4**, 858.
- Lyu, J. (2020) Pan-genome upgrade. *Nat. Plants*, **6**, 732.
- Lyu, X., Cheng, Q., Qin, C., Li, Y., Xu, X., Ji, R., Mu, R. et al. (2021) *GmCRY1s* modulate gibberellin metabolism to regulate soybean shade avoidance in response to reduced blue light. *Mol. Plant*, **14**, 298–314.
- Machado, F.B., Moharana, K.C., Almeida-Silva, F., Gazara, R.K., Pedrosa-Silva, F., Coelho, F.S., Grativol, C. et al. (2020) Systematic analysis of 1298 RNA-Seq samples and construction of a comprehensive soybean (*Glycine max*) expression atlas. *Plant J.* **103**, 1894–1909.
- Mameda, R., Waki, T., Kawai, Y., Takahashi, S. and Nakayama, T. (2018) Involvement of chalcone reductase in the soybean isoflavone metabolism: identification of *GmCHR5*, which interacts with 2-hydroxyisoflavonone synthase. *Plant J.* **96**, 56–74.
- Marcolino-Gomes, J., Rodrigues, F.A., Fuganti-Pagliarini, R., Bendix, C., Nakayama, T.J., Celaya, B., Molinari, H.B. et al. (2014) Diurnal oscillations of soybean circadian clock and drought responsive genes. *PLoS One*, **9**, e86402.
- Marek, L.F., Mudge, J., Darnielle, L., Grant, D., Hanson, N., Paz, M., Huihuang, Y. et al. (2001) Soybean genomic survey: BAC-end sequences near RFLP and SSR markers. *Genome*, **44**, 572–581.
- Mccabe, D.E., Swain, W.F., Martinell, B.J. and Christou, P. (1988) Stable Transformation of soybean (*Glycine Max*) by particle acceleration. *Nat. Biotechnol.* **6**, 923–926.
- McHale, L.K., Haun, W.J., Xu, W.W., Bhaskar, P.B., Anderson, J.E., Hyten, D.L., Gerhardt, D.J. et al. (2012) Structural variants in the soybean genome localize to clusters of biotic stress-response genes. *Plant Physiol.* **159**, 1295–1308.
- Meng, Y., Li, H., Wang, Q., Liu, B. and Lin, C. (2013) Blue light-dependent interaction between cryptochrome2 and CIB1 regulates transcription and leaf senescence in soybean. *Plant Cell*, **25**, 4405–4420.
- Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., Zhang, X., Wan, J. et al. (2013) Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res.* **23**, 1233–1236.
- Miao, L., Yang, S., Zhang, K., He, J., Wu, C., Ren, Y., Gai, J. et al. (2020) Natural variation and selection in *GmSWEET39* affect soybean seed oil content. *New Phytol.* **225**, 1651–1666.
- Michno, J.M., Wang, X.B., Liu, J.Q., Curtin, S.J., Kono, T.J.Y. and Stupar, R.M. (2015) CRISPR/Cas mutagenesis of soybean and *Medicago truncatula* using a new web-tool and a modified Cas9 enzyme. *Gm Crops Food*, **6**, 243–252.
- Miranda, C., Scaboo, A., Cober, E., Denwar, N. and Bilyeu, K. (2020) The effects and interaction of soybean maturity gene alleles controlling flowering time, maturity, and adaptation in tropical environments. *BMC Plant Biol.* **20**, 65.
- Mishra, R., Joshi, R.K. and Zhao, K. (2020) Base editing in crops: current advances, limitations and future implications. *Plant Biotechnol. J.* **18**, 20–31.
- Mitchum, M.G. (2016) Soybean resistance to the soybean cyst nematode heterodera glycines: an update. *Phytopathology*, **106**, 1444–1450.
- Mourtzinis, S., Specht, J.E., Lindsey, L.E., Wiebold, W.J., Ross, J., Nafziger, E.D., Kandel, H.J. et al. (2015) Climate-induced reduction in US-wide soybean yields underpinned by region- and in-season-specific responses. *Nat. Plants*, **1**, 14026.
- Nadeem, M., Chen, A.D., Hong, H.L., Li, D.D., Li, J.J., Zhao, D., Wang, W. et al. (2021) *GmMts1* encodes a kinesin-like protein essential for male fertility in soybean (*Glycine max* L.). *J. Integr. Plant Biol.* **63**, 1054–1064.
- Nan, H.Y., Cao, D., Zhang, D.Y., Li, Y., Lu, S.J., Tang, L.L., Yuan, X.H. et al. (2014) *GmFT2a* and *GmFT5a* redundantly and differentially regulate flowering through interaction with and upregulation of the bZIP transcription factor *GmFDL19* in soybean. *PLoS One*, **9**, e97669.
- Natarajan, S., Tavakolan, M., Alkharouf, N.W. and Matthews, B.F. (2014) SCNProDB: A database for the identification of soybean cyst nematode proteins. *Bioinformatics*, **10**, 387–389.
- Ngaki, M.N., Sahoo, D.K., Wang, B. and Bhattacharyya, M.K. (2021) Overexpression of a plasma membrane protein generated broad-spectrum immunity in soybean. *Plant Biotechnol. J.* **19**, 502–516.
- Nguyen, C.X., Paddock, K.J., Zhang, Z. and Stacey, M.G. (2020) *GmKIX8-1* regulates organ size in soybean and is the causative gene for the major seed weight QTL *qSw17-1*. *New Phytol.* **229**, 920–934.
- Nishimura, R., Hayashi, M., Wu, G.J., Kouchi, H., Imaizumi-Anraku, H., Murakami, Y., Kawasaki, S. et al. (2002) *HAR1* mediates systemic regulation of symbiotic organ development. *Nature*, **420**, 426–429.
- Nizampatnam, N.R., Schreier, S.J., Damodaran, S., Adhikari, S. and Subramanian, S. (2015) microRNA160 dictates stage-specific auxin and cytokinin sensitivities and directs soybean nodule development. *Plant J.* **84**, 140–153.
- Noon, J.B., Hewezi, T. and Baum, T.J. (2019) Homeostasis in the soybean miRNA396-GRF network is essential for productive soybean cyst nematode infections. *J. Exp. Bot.* **70**, 1653–1668.
- Noon, J.B., Qi, M., Sill, D.N., Muppirala, U., Eves-van den Akker, S., Maier, T.R., Dobbs, D. et al. (2016) A Plasmodium-like virulence effector of the soybean cyst nematode suppresses plant innate immunity. *New Phytol.* **212**, 444–460.
- Ochola, S., Huang, J., Ali, H., Shu, H., Shen, D., Qiu, M., Wang, L. et al. (2020) Editing of an effector gene promoter sequence impacts plant-Phytophthora interaction. *J. Integr. Plant Biol.* **62**, 378–392.
- Ohyanagi, H., Sakata, K. and Komatsu, S. (2012) Soybean Proteome Database 2012: update on the comprehensive data repository for soybean proteomics. *Front. Plant Sci.* **3**, 110.
- Okutani, F., Hamamoto, S., Aoki, Y., Nakayasu, M., Nihei, N., Nishimura, T., Yazaki, K. et al. (2020) Rhizosphere modelling reveals spatiotemporal distribution of daidzein shaping soybean rhizosphere bacterial community. *Plant Cell Environ.* **43**, 1036–1046.
- Pan, W.J., Tao, J.J., Cheng, T., Bian, X.H., Wei, W., Zhang, W.K., Ma, B. et al. (2016) Soybean miR172a improves salt tolerance and can function as a long-distance signal. *Mol. Plant*, **9**, 1337–1340.
- Pandey, A., Misra, P., Khan, M.P., Swarnkar, G., Tewari, M.C., Bhambhani, S., Trivedi, R. et al. (2014) Co-expression of *Arabidopsis* transcription factor, *AtMYB12*, and soybean isoflavone synthase, *GmIFS1*, genes in tobacco leads to enhanced biosynthesis of isoflavones and flavonols resulting in osteoprotective activity. *Plant Biotechnol. J.* **12**, 69–80.
- Papiernik, S.K., Grieve, C.M., Lesch, S.M. and Yates, S.R. (2005) Effects of salinity, imazethapyr, and chlorimuron application on soybean growth and yield. *Commun. Soil Sci. Plant Anal.* **36**, 951–967.
- Paredy, D., Chennareddy, S., Anthony, G., Sardesai, N., Mall, T., Minnicks, T., Karpova, O. et al. (2020) Improved soybean transformation for efficient and high throughput transgenic production. *Transgenic Res.* **29**, 267–281.

- Patil, G., Do, T., Vuong, T.D., Valliyodan, B., Lee, J.D., Chaudhary, J., Shannon, J.G. *et al.* (2016) Genomic-assisted haplotype analysis and the development of high-throughput SNP markers for salinity tolerance in soybean. *Sci. Rep.* **6**, 19199.
- Patil, G., Mian, R., Vuong, T., Pantalone, V., Song, Q., Chen, P., Shannon, G.J. *et al.* (2017) Molecular mapping and genomics of soybean seed protein: a review and perspective for the future. *Theor. Appl. Genet.* **130**, 1975–1991.
- Paz, M.M., Martinez, J.C., Kalvig, A.B., Fonger, T.M. and Kan, W. (2006) Improved cotyledonary node method using an alternative explant derived from mature seed for efficient *Agrobacterium*-mediated soybean transformation. *Plant Cell Rep.* **25**, 248.
- Pearce, G., Yamaguchi, Y., Barona, G. and Ryan, C.A. (2010) A subtilisin-like protein from soybean contains an embedded, cryptic signal that activates defense-related genes. *Proc. Natl Acad. Sci. USA*, **107**, 14921–14925.
- Pedley, K.F., Pandey, A.K., Ruck, A., Lincoln, L.M., Whitham, S.A. and Graham, M.A. (2019) *Rpp1* encodes a ULP1-NBS-LRR protein that controls immunity to phakopsora pachyrhizi in soybean. *Mol. Plant. Microbe Interact.* **32**, 120–133.
- Pfeiffer, G.A., King, K.E., Severin, A.J., May, G.D., Cianzio, S.R., Lin, S.F., Lauter, N.C. *et al.* (2012) Identification of candidate genes underlying an iron efficiency quantitative trait locus in soybean. *Plant Physiol.* **158**, 1745–1754.
- Peng, W., Wu, W., Peng, J., Li, J., Lin, Y., Wang, Y., Tian, J. *et al.* (2018) Characterization of the soybean GmALMT family genes and the function of GmALMT5 in response to phosphate starvation. *J. Integr. Plant Biol.* **60**, 216–231.
- Pfeiffer, T.W., Peyyala, R., Ren, Q.X. and Ghabrial, S.A. (2003) Increased soybean pubescence density: Yield and soybean mosaic virus resistance effects. *Crop Sci.* **43**, 2071–2076.
- Phang, T.H., Shao, G. and Lam, H.M. (2008) Salt tolerance in soybean. *J. Integr. Plant Biol.* **50**, 1196–1212.
- Ping, J.Q., Liu, Y.F., Sun, L.J., Zhao, M.X., Li, Y.H., She, M.Y., Sui, Y. *et al.* (2014) *Dt2* is a gain-of-function MADS-Domain factor gene that specifies semideterminacy in soybean. *Plant Cell*, **26**, 2831–2842.
- Qi, X., Jiang, B., Wu, T., Sun, S., Wang, C., Song, W., Wu, C. *et al.* (2021) Genomic dissection of widely planted soybean cultivars leads to a new breeding strategy of crops in the post-genomic era. *Crop J.* <https://doi.org/10.1016/j.cj.2021.1001.1001>
- Qi, X., Li, M.W., Xie, M., Liu, X., Ni, M., Shao, G., Song, C. *et al.* (2014) Identification of a novel salt tolerance gene in wild soybean by whole-genome sequencing. *Nat. Commun.* **5**, 4340.
- Qin, L., Zhao, J., Tian, J., Chen, L., Sun, Z., Guo, Y., Lu, X. *et al.* (2012) The high-affinity phosphate transporter GmPT5 regulates phosphate transport to nodules and nodulation in soybean. *Plant Physiol.* **159**, 1634–1643.
- Qiu, J., Wang, Y., Wu, S., Wang, Y.Y., Ye, C.Y., Bai, X., Li, Z. *et al.* (2014) Genome re-sequencing of semi-wild soybean reveals a complex Soja population structure and deep introgression. *PLoS One*, **9**, e108479.
- Qu, Y., Guan, R., Bose, J., Henderson, S.W., Wege, S., Qiu, L. and Gilliam, M. (2021) Soybean CHX-type ion transport protein GmSALT3 confers leaf Na⁺ exclusion via a root derived mechanism, and Cl⁻ exclusion via a shoot derived process. *Plant Cell Environ.* **44**, 856–869.
- Quain, M.D., Makgopa, M.E., Márquez-García, B., Comadira, G., Fernandez-García, N., Olmos, E., Schnaubelt, D. *et al.* (2014) Ectopic phytoalexin expression leads to enhanced drought stress tolerance in soybean (*Glycine max*) and *Arabidopsis thaliana* through effects on strigolactone pathways and can also result in improved seed traits. *Plant Biotechnol. J.* **12**, 903–913.
- Radwan, O., Wu, X., Govindarajulu, M., Libault, M., Neece, D.J., Oh, M.H., Berg, R.H. *et al.* (2012) 14-3-3 proteins SGF14c and SGF14l play critical roles during soybean nodulation. *Plant Physiol.* **160**, 2125–2136.
- Rakocevic, M., Müller, M., Matsunaga, F.T., Neumaier, N., Farias, J.R.B., Nepomuceno, A.L. and Fuganti-Pagliarini, R. (2018) Daily heliotropic movements assist gas exchange and productive responses in DREB1A soybean plants under drought stress in the greenhouse. *Plant J.* **96**, 801–814.
- Rambani, A., Pantalone, V., Yang, S., Rice, J.H., Song, Q., Mazarei, M., Arelli, P.R. *et al.* (2020) Identification of introduced and stably inherited DNA methylation variants in soybean associated with soybean cyst nematode parasitism. *New Phytol.* **227**, 168–184.
- Rambani, A., Rice, J.H., Liu, J., Lane, T., Ranjan, P., Mazarei, M., Pantalone, V. *et al.* (2015) The methylome of soybean roots during the compatible interaction with the soybean cyst nematode. *Plant Physiol.* **168**, 1364–1377.
- Ramesh, S.V., Govindasamy, V., Rajesh, M.K., Sabana, A.A. and Praveen, S. (2019) Stress-responsive miRNAome of *Glycine max* (L.) Merrill: molecular insights and way forward. *Planta*, **249**, 1267–1284.
- Ray, D.K., Mueller, N.D., West, P.C. and Foley, J.A. (2013) Yield trends are insufficient to double global crop production by 2050. *PLoS One*, **8**, e66428.
- Rech, E.L., Vianna, G.R. and Aragao, F.J. (2008) High-efficiency transformation by biolistics of soybean, common bean and cotton transgenic plants. *Nat. Protoc.* **3**, 410–418.
- Redditt, T.J., Chung, E.H., Zand, K.H., Rodibaugh, N., Zhang, Y., Trinidad, J.C., Kim, J.H. *et al.* (2019) AvrRpm1 functions as an ADP-ribosyl transferase to modify NOI-domain containing proteins, including Arabidopsis and soybean RPM1-interacting protein 4. *Plant Cell*, **31**, 2664–2681.
- Rehman, N.U., Ali, M., Ahmad, M.Z., Liang, G. and Zhao, J. (2018) Strigolactones promote rhizobia interaction and increase nodulation in soybean (*Glycine max*). *Microb. Pathog.* **114**, 420–430.
- Reid, D.E., Ferguson, B.J. and Gresshoff, P.M. (2011) Inoculation- and nitrate-induced CLE peptides of soybean control NARK-dependent nodule formation. *Mol. Plant. Microbe Interact.* **24**, 606–618.
- Reid, D.E., Hayashi, S., Lorenc, M., Stiller, J., Edwards, D., Gresshoff, P.M. and Ferguson, B.J. (2012) Identification of systemic responses in soybean nodulation by xylem sap feeding and complete transcriptome sequencing reveal a novel component of the autoregulation pathway. *Plant Biotechnol. J.* **10**, 680–689.
- Reid, D.E., Li, D., Ferguson, B.J. and Gresshoff, P.M. (2013) Structure-function analysis of the GmRIC1 signal peptide and CLE domain required for nodulation control in soybean. *J. Exp. Bot.* **64**, 1575–1585.
- Reinprecht, Y., Luk-Labey, S.-y., Larsen, J., Poysa, V.w., Yu, K., Rajcan, I., Ablett, G.R. *et al.* (2009) Molecular basis of the low linolenic acid trait in soybean EMS mutant line RG10. *Plant Breed.* **128**, 253–258.
- Reis, P.A., Rosado, G.L., Silva, L.A., Oliveira, L.C., Oliveira, L.B., Costa, M.D., Alvim, F.C. *et al.* (2011) The binding protein BiP attenuates stress-induced cell death in soybean via modulation of the N-rich protein-mediated signaling pathway. *Plant Physiol.* **157**, 1853–1865.
- Ren, B., Wang, X., Duan, J. and Ma, J. (2019) Rhizobial tRNA-derived small RNAs are signal molecules regulating plant nodulation. *Science*, **365**, 919–922.
- Rendon-Anaya, M. and Herrera-Estrella, A. (2018) The advantage of parallel selection of domestication genes to accelerate crop improvement. *Genome Biol.* **19**, 147.
- Ribichich, K.F., Chiozza, M., Ávalos-Britez, S., Cabello, J.V., Arce, A.L., Watson, G., Arias, C. *et al.* (2020) Successful field performance in warm and dry environments of soybean expressing the sunflower transcription factor HB4. *J. Exp. Bot.* **71**, 3142–3156.
- Rincker, K., Nelson, R.L., Specht, J., Sleper, D., Cary, T., Cianzio, S., Casteel, S. *et al.* (2014) Genetic improvement of U.S. soybean in maturity groups II, III, and IV. *Crop Sci.* **54**, 1419–1432.
- Robison, J.D., Yamasaki, Y. and Randall, S.K. (2019) The ethylene signaling pathway negatively impacts CBF/DREB-regulated cold response in soybean (*Glycine max*). *Front. Plant Sci.* **10**, 121.
- Rodriguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E. and Lippman, Z.B. (2017) Engineering quantitative trait variation for crop improvement by genome editing. *Cell*, **171**, 470–480.
- Roesler, K., Shen, B., Bermudez, E., Li, C., Hunt, J., Damude, H.G., Ripp, K.G. *et al.* (2016) An improved variant of soybean type 1 diacylglycerol acyltransferase increases the oil content and decreases the soluble carbohydrate content of soybeans. *Plant Physiol.* **171**, 878–893.
- Román, Á., Andreu, V., Hernández, M.L., Lagunas, B., Picorel, R., Martínez-Rivas, J.M. and Alfonso, M. (2012) Contribution of the different omega-3 fatty acid desaturase genes to the cold response in soybean. *J. Exp. Bot.* **63**, 4973–4982.
- Roulin, A., Auer, P.L., Libault, M., Schlueter, J., Farmer, A., May, G., Stacey, G. *et al.* (2013) The fate of duplicated genes in a polyploid plant genome. *Plant J.* **73**, 143–153.
- Roy, S., Liu, W., Nandety, R.S., Crook, A., Mysore, K.S., Pislariu, C.I., Frugoli, J. *et al.* (2020) Celebrating 20 years of genetic discoveries in legume nodulation and symbiotic nitrogen fixation. *Plant Cell*, **32**, 15–41.
- Ruiz-Lozano, J.M., Porcel, R., Azcon, C. and Aroca, R. (2012) Regulation by arbuscular mycorrhizae of the integrated physiological response to salinity in

- plants: new challenges in physiological and molecular studies. *J. Exp. Bot.* **63**, 4033–4044.
- Ruiz-Vera, U.M., Siebers, M., Gray, S.B., Drag, D.W., Rosenthal, D.M., Kimball, B.A., Ort, D.R. et al. (2013) Global warming can negate the expected CO₂ stimulation in photosynthesis and productivity for soybean grown in the Midwestern United States. *Plant Physiol.* **162**, 410–423.
- Samanfar, B., Molnar, S.J., Charette, M., Schoenrock, A., Dehne, F., Golshani, A., Belzile, F. et al. (2017) Mapping and identification of a potential candidate gene for a novel maturity locus, *E10*, in soybean. *Theor. Appl. Genet.* **130**, 377–390.
- dos Santos, J.V.M., Valliyodan, B., Joshi, T., Khan, S.M., Liu, Y., Wang, J.X., Vuong, T.D. et al. (2016) Evaluation of genetic variation among Brazilian soybean cultivars through genome resequencing. *BMC Genom.* **17**, 110.
- Sato, S., Newell, C., Kolacz, K., Tredo, L., Finer, J. and Hinchey, M. (1993) Stable transformation via particle bombardment in two different soybean regeneration systems. *Plant Cell Rep.* **12**, 408–413.
- Schaarschmidt, S., Gresshoff, P.M. and Hause, B. (2013) Analyzing the soybean transcriptome during autoregulation of mycorrhization identifies the transcription factors GmNF-YA1a/b as positive regulators of arbuscular mycorrhization. *Genome Biol.* **14**, R62.
- Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L. et al. (2010) Genome sequence of the palaeopolyploid soybean. *Nature*, **463**, 178–183.
- Searle, I.R., Men, A.E., Laniya, T.S., Buzas, D.M., Iturbe-Ormaetxe, I., Carroll, B.J. and Gresshoff, P.M. (2003) Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. *Science*, **299**, 109–112.
- Sedivy, E.J., Wu, F. and Hanzawa, Y. (2017) Soybean domestication: the origin, genetic architecture and molecular bases. *New Phytol.* **214**, 539–553.
- Selote, D., Shine, M.B., Robin, G.P. and Kachroo, A. (2014) Soybean NDR1-like proteins bind pathogen effectors and regulate resistance signaling. *New Phytol.* **202**, 485–498.
- Severin, A.J., Cannon, S.B., Graham, M.M., Grant, D. and Shoemaker, R.C. (2011) Changes in twelve homoeologous genomic regions in soybean following three rounds of polyploidy. *Plant Cell*, **23**, 3129–3136.
- Severin, A.J., Woody, J.L., Bolon, Y.T., Joseph, B., Diers, B.W., Farmer, A.D., Muehlbauer, G.J. et al. (2010) RNA-Seq Atlas of *Glycine max*: a guide to the soybean transcriptome. *BMC Plant Biol.* **10**, 160.
- Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K. et al. (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 686–688.
- Shen, Y., Du, H., Liu, Y., Ni, L., Wang, Z., Liang, C. and Tian, Z. (2019) Update soybean Zhonghuang 13 genome to a golden reference. *Sci. China Life. Sci.* **62**, 1257–1260.
- Shen, Y., Liu, J., Geng, H., Zhang, J., Liu, Y., Zhang, H., Xing, S. et al. (2018a) *De novo* assembly of a Chinese soybean genome. *Sci. China Life. Sci.* **61**, 871–884.
- Shen, Y., Zhang, J., Liu, Y., Liu, S., Liu, Z., Duan, Z., Wang, Z. et al. (2018b) DNA methylation footprints during soybean domestication and improvement. *Genome Biol.* **19**, 128.
- Shen, Y.T., Zhou, Z.K., Wang, Z., Li, W.Y., Fang, C., Wu, M., Ma, Y.M. et al. (2014) Global dissection of alternative splicing in paleopolyploid soybean. *Plant Cell*, **26**, 996–1008.
- Shimomura, M., Kanamori, H., Komatsu, S., Namiki, N., Mukai, Y., Kurita, K., Kamatsuki, K. et al. (2015) The *Glycine max* cv. Enrei genome for improvement of Japanese soybean cultivars. *Int. J. Genomics*, **2015**, 358127.
- Shine, M.B., Gao, Q.M., Chowda-Reddy, R.V., Singh, A.K., Kachroo, P. and Kachroo, A. (2019) Glycerol-3-phosphate mediates rhizobia-induced systemic signaling in soybean. *Nat. Commun.* **10**, 5303.
- Shine, M.B., Yang, J.W., El-Habbak, M., Nagyabhyru, P., Fu, D.Q., Navarre, D., Ghabrial, S. et al. (2016) Cooperative functioning between phenylalanine ammonia lyase and isochorismate synthase activities contributes to salicylic acid biosynthesis in soybean. *New Phytol.* **212**, 627–636.
- Shu, Y., Zhou, Y., Mu, K., Hu, H., Chen, M., He, Q., Huang, S. et al. (2020) A transcriptomic analysis reveals soybean seed pre-harvest deterioration resistance pathways under high temperature and humidity stress. *Genome*, **63**, 115–124.
- Slattery, R.A., Grennan, A.K., Sivaguru, M., Sozzani, R. and Ort, D.R. (2016) Light sheet microscopy reveals more gradual light attenuation in light-green versus dark-green soybean leaves. *J. Exp. Bot.* **67**, 4697–4709.
- Sobhanian, H., Razavizadeh, R., Nanjo, Y., Ehsanpour, A.A., Jazii, F.R., Motamed, N. and Komatsu, S. (2010) Proteome analysis of soybean leaves, hypocotyls and roots under salt stress. *Proteome Sci.* **8**, 19.
- Somers, D.A., Samac, D.A. and Olhoft, P.M. (2003) Recent advances in legume transformation. *Plant Physiol.* **131**, 892–899.
- Song, H., Yin, Z., Chao, M., Ning, L., Zhang, D. and Yu, D. (2014) Functional properties and expression quantitative trait loci for phosphate transporter GmPT1 in soybean. *Plant Cell Environ.* **37**, 462–472.
- Song, Q., Hyten, D.L., Jia, G., Quigley, C.V., Fickus, E.W., Nelson, R.L. and Cregan, P.B. (2015) Fingerprinting soybean germplasm and its utility in genomic research. *G3: Genes - Genomes - Genetics*, **5**, 1999–2006.
- Song, Q.X., Li, Q.T., Liu, Y.F., Zhang, F.X., Ma, B., Zhang, W.K., Man, W.Q. et al. (2013a) Soybean *GmbZIP123* gene enhances lipid content in the seeds of transgenic *Arabidopsis* plants. *J. Exp. Bot.* **64**, 4329–4341.
- Song, Q.X., Lu, X., Li, Q.T., Chen, H., Hu, X.Y., Ma, B., Zhang, W.K. et al. (2013b) Genome-wide analysis of DNA methylation in soybean. *Mol. Plant*, **6**, 1961–1974.
- Song, Q.J., Marek, L.F., Shoemaker, R.C., Lark, K.G., Concibido, V.C., Delannay, X., Specht, J.E. et al. (2004) A new integrated genetic linkage map of the soybean. *Theor. Appl. Genet.* **109**, 122–128.
- Song, Q., Yan, L., Quigley, C., Fickus, E., Wei, H., Chen, L., Dong, F. et al. (2020) Soybean BARCSoySNP6K: An assay for soybean genetics and breeding research. *Plant J.* **104**, 800–811.
- Song, Y., Ji, D., Li, S., Wang, P., Li, Q. and Xiang, F. (2012) The dynamic changes of DNA methylation and histone modifications of salt responsive transcription factor genes in soybean. *PLoS One*, **7**, e41274.
- Song, Z.Y., Tian, J.L., Fu, W.Z., Li, L., Lu, L.H., Zhou, L., Shan, Z.H. et al. (2013c) Screening Chinese soybean genotypes for *Agrobacterium*-mediated genetic transformation suitability. *J. Zhejiang Univ. Sci. B*, **14**, 289–298.
- Specht, J.E., Hume, D.J. and Kumudini, S.V. (1999) Soybean yield potential - A genetic and physiological perspective. *Crop Sci.* **39**, 1560–1570.
- Subramanian, S., Hu, X., Lu, G., Odelland, J.T. and Yu, O. (2004) The promoters of two isoflavone synthase genes respond differentially to nodulation and defense signals in transgenic soybean roots. *Plant Mol. Biol.* **54**, 623–639.
- Subramanian, S., Stacey, G. and Yu, O. (2006) Endogenous isoflavones are essential for the establishment of symbiosis between soybean and *Bradyrhizobium japonicum*. *Plant J.* **48**, 261–273.
- Sukumaran, A., McDowell, T., Chen, L., Renaud, J. and Dhaubhadel, S. (2018) Isoflavonoid-specific prenyltransferase gene family in soybean: *GmPTO1*, a pterocarpan 2-dimethylallyltransferase involved in glyceollin biosynthesis. *Plant J.* **96**, 966–981.
- Sun, L., Miao, Z., Cai, C., Zhang, D., Zhao, M., Wu, Y., Zhang, X. et al. (2015a) *GmHs1-1*, encoding a calcineurin-like protein, controls hard-seededness in soybean. *Nat. Genet.* **47**, 939–943.
- Sun, X.J., Hu, Z., Chen, R., Jiang, Q.Y., Song, G.H., Zhang, H. and Xi, Y.J. (2015b) Targeted mutagenesis in soybean using the CRISPR-Cas9 system. *Sci Rep-Uk*, **5**, 10342.
- Sun, Z., Su, C., Yun, J., Jiang, Q., Wang, L., Wang, Y., Cao, D. et al. (2019) Genetic improvement of the shoot architecture and yield in soya bean plants via the manipulation of GmMiR156b. *Plant Biotechnol. J.* **17**, 50–62.
- Svitashev, S., Young, J.K., Schwartz, C., Gao, H., Falco, S.C. and Cigan, A.M. (2015) Targeted mutagenesis, precise gene editing, and site-specific gene insertion in Maize using Cas9 and guide RNA. *Plant Physiol.* **169**, 931–945.
- Syed, N.H., Prince, S.J., Mutava, R.N., Patil, G., Li, S., Chen, W., Babu, V. et al. (2015) Core clock, *SUB1*, and *ABAR* genes mediate flooding and drought responses via alternative splicing in soybean. *J. Exp. Bot.* **66**, 7129–7149.
- Takeshima, R., Hayashi, T., Zhu, J., Zhao, C., Xu, M., Yamaguchi, N., Sayama, T. et al. (2016) A soybean quantitative trait locus that promotes flowering under long days is identified as *FT5a*, a *FLOWERING LOCUS T* ortholog. *J. Exp. Bot.* **67**, 5247–5258.
- Takeshima, R., Nan, H., Harigai, K., Dong, L., Zhu, J., Lu, S., Xu, M. et al. (2019) Functional divergence between soybean *FLOWERING LOCUS T* orthologues *FT2a* and *FT5a* in post-flowering stem growth. *J. Exp. Bot.* **70**, 3941–3953.
- Tanaka, A., Fujita, K. and Terasawa, H. (1985) Growth and dinitrogen fixation of soybean root system affected by partial exposure to nitrate. *Soil Sci. Plant Nutr.* **31**, 637–645.

- Tanaka, K., Nguyen, C.T., Libault, M., Cheng, J. and Stacey, G. (2011) Enzymatic activity of the soybean ecto-apyrase GS52 is essential for stimulation of nodulation. *Plant Physiol.* **155**, 1988–1998.
- Tang, F., Yang, S., Liu, J. and Zhu, H. (2016) *Rj4*, a gene controlling nodulation specificity in soybeans, encodes a thaumatin-like protein but not the one previously reported. *Plant Physiol.* **170**, 26–32.
- Tang, X., Su, T., Han, M., Wei, L., Wang, W., Yu, Z., Xue, Y. et al. (2017) Suppression of extracellular invertase inhibitor gene expression improves seed weight in soybean (*Glycine max*). *J. Exp. Bot.* **68**, 469–482.
- Tao, Y., Jordan, D.R. and Mace, E.S. (2020) A graph-based pan-genome guides biological discovery. *Mol. Plant*, **13**, 1247–1249.
- Tavakolan, M., Alkharouf, N.W., Khan, F.H. and Natarajan, S. (2013) SoyProDB: A database for the identification of soybean seed proteins. *Bioinformatics*, **9**, 165–167.
- Tavakolan, M., Alkharouf, N.W., Matthews, B.F. and Natarajan, S.S. (2014) SoyProLow: a protein database enriched in low abundant soybean proteins. *Bioinformatics*, **10**, 599–601.
- Teshima, T., Yamada, N., Yokota, Y., Sayama, T., Inagaki, K., Koeduka, T., Uefune, M. et al. (2020) Suppressed Methionine γ -Lyase expression causes hyperaccumulation of S-methylmethionine in soybean seeds. *Plant Physiol.* **183**, 943–956.
- Tian, Z., Wang, J.W., Li, J. and Han, B. (2020) Designing Future Crops: challenges and strategies for sustainable agriculture. *Plant J.* **105**, 1165–1178.
- Tian, Z., Wang, X., Lee, R., Li, Y., Specht, J.E., Nelson, R.L., McClean, P.E. et al. (2010) Artificial selection for determinate growth habit in soybean. *Proc. Natl Acad. Sci. USA*, **107**, 8563–8568.
- Tian, Z., Zhao, M., She, M., Du, J., Cannon, S.B., Liu, X., Xu, X. et al. (2012) Genome-wide characterization of nonreference transposons reveals evolutionary propensities of transposons in soybean. *Plant Cell*, **24**, 4422–4436.
- Torkamaneh, D., Laroche, J., Tardivel, A., O'Donoghue, L., Cober, E., Rajcan, I. and Belzile, F. (2018) Comprehensive description of genomewide nucleotide and structural variation in short-season soya bean. *Plant Biotechnol. J.* **16**, 749–759.
- Torkamaneh, D., Laroche, J., Valliyodan, B., O'Donoghue, L., Cober, E., Rajcan, I., Vilela, A.R. et al. (2021) Soybean (*Glycine max*) Haplotype Map (GmHapMap): a universal resource for soybean translational and functional genomics. *Plant Biotechnol. J.* **19**, 324–334.
- Trick, H.N. and Finer, J.J. (1998) Sonication-assisted *Agrobacterium*-mediated transformation of soybean [*Glycine max* (L.) Merrill] embryogenic suspension culture tissue. *Plant Cell Rep.* **17**, 482–488.
- Tsuda, M., Kaga, A., Anai, T., Shimizu, T., Sayama, T., Takagi, K., Machita, K. et al. (2015) Construction of a high-density mutant library in soybean and development of a mutant retrieval method using amplicon sequencing. *BMC Genom.* **16**, 1014.
- Tucker, M.L., Murphy, C.A. and Yang, R. (2011) Gene expression profiling and shared promoter motif for cell wall-modifying proteins expressed in soybean cyst nematode-infected roots. *Plant Physiol.* **156**, 319–329.
- Tucker, M.L., Xue, P. and Yang, R. (2010) 1-Aminocyclopropane-1-carboxylic acid (ACC) concentration and ACC synthase expression in soybean roots, root tips, and soybean cyst nematode (*Heterodera glycines*)-infected roots. *J. Exp. Bot.* **61**, 463–472.
- Turck, F., Fornara, F. and Coupland, G. (2008) Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. *Annu. Rev. Plant Biol.* **59**, 573–594.
- Turner, M., Nizampatnam, N.R., Baron, M., Coppin, S., Damodaran, S., Adhikari, S., Arunachalam, S.P. et al. (2013) Ectopic expression of miR160 results in auxin hypersensitivity, cytokinin hyposensitivity, and inhibition of symbiotic nodule development in soybean. *Plant Physiol.* **162**, 2042–2055.
- Um, J.H., Kim, S., Kim, Y.K., Song, S.B., Lee, S.H., Verma, D.P. and Cheon, C.I. (2013) RNA interference-mediated repression of S6 kinase 1 impairs root nodule development in soybean. *Mol. Cells*, **35**, 243–248.
- Upadhyay, S.K., Kumar, J., Alok, A. and Tuli, R. (2013) RNA-guided genome editing for target gene mutations in wheat. *G3: Genes - Genomes - Genetics*, **3**, 2233–2238.
- Valdés-López, O., Thibivilliers, S., Qiu, J., Xu, W.W., Nguyen, T.H., Libault, M., Le, B.H. et al. (2011) Identification of quantitative trait loci controlling gene expression during the innate immunity response of soybean. *Plant Physiol.* **157**, 1975–1986.
- Valliyodan, B., Cannon, S.B., Bayer, P.E., Shu, S.Q., Brown, A.V., Ren, L.H., Jenkins, J. et al. (2019) Construction and comparison of three reference-quality genome assemblies for soybean. *Plant J.* **100**, 1066–1082.
- Valliyodan, B., Dan, Q., Patil, G., Zeng, P., Huang, J., Dai, L., Chen, C. et al. (2016) Landscape of genomic diversity and trait discovery in soybean. *Sci. Rep.* **6**, 23598.
- Wagner, N., Mroczka, A., Roberts, P.D., Schreckengost, W. and Voelker, T. (2011) RNAi trigger fragment truncation attenuates soybean FAD2-1 transcript suppression and yields intermediate oil phenotypes. *Plant Biotechnol. J.* **9**, 723–728.
- Wan, X., Hontelez, J., Lillo, A., Guarnerio, C., van de Peut, D., Fedorova, E., Bisseling, T. et al. (2007) *Medicago truncatula* *ENOD40-1* and *ENOD40-2* are both involved in nodule initiation and bacteroid development. *J. Exp. Bot.* **58**, 2033–2041.
- Wang, D., Liang, X., Bao, Y., Yang, S., Zhang, X., Yu, H., Zhang, Q. et al. (2020a) A malectin-like receptor kinase regulates cell death and pattern-triggered immunity in soybean. *EMBO Rep.* **21**, e50442.
- Wang, D., Liu, Y.X., Yu, Q., Zhao, S.P., Zhao, J.Y., Ru, J.N., Cao, X.Y. et al. (2019a) Functional analysis of the soybean *GmCDPK3* gene responding to drought and salt stresses. *Int. J. Mol. Sci.* **20**, 5909.
- Wang, F., Chen, H.W., Li, Q.T., Wei, W., Li, W., Zhang, W.K., Ma, B. et al. (2015a) *GmWRKY27* interacts with *GmMYB174* to reduce expression of *GmNAC29* for stress tolerance in soybean plants. *Plant J.* **83**, 224–236.
- Wang, G. and Xu, Y. (2008) Hypocotyl-based *Agrobacterium*-mediated transformation of soybean (*Glycine max*) and application for RNA interference. *Plant Cell Rep.* **27**, 1177–1184.
- Wang, H., Ji, F., Zhang, Y., Hou, J., Liu, W., Huang, J. and Liang, W. (2019b) Interactions between hydrogen sulphide and nitric oxide regulate two soybean citrate transporters during the alleviation of aluminium toxicity. *Plant Cell Environ.* **42**, 2340–2356.
- Wang, H. and Wang, H. (2015) The miR156/SPL module, a regulatory hub and versatile toolbox, gears up crops for enhanced agronomic traits. *Mol. Plant*, **8**, 677–688.
- Wang, H., Zhou, L., Fu, Y., Cheung, M.Y., Wong, F.L., Phang, T.H., Sun, Z. et al. (2012a) Expression of an apoplast-localized BURP-domain protein from soybean (*GmRD22*) enhances tolerance towards abiotic stress. *Plant Cell Environ.* **35**, 1932–1947.
- Wang, J., Chu, S., Zhang, H., Zhu, Y., Cheng, H. and Yu, D. (2016a) Development and application of a novel genome-wide SNP array reveals domestication history in soybean. *Sci. Rep.* **6**, 20728.
- Wang, J., Kuang, H.Q., Zhang, Z.H., Yang, Y.Q., Yan, L., Zhang, M.C., Song, S.K. et al. (2020b) Generation of seed lipoxygenase-free soybean using CRISPR-Cas9. *Crop J.* **8**, 432–439.
- Wang, J.W., Schwab, R., Czech, B., Mica, E. and Weigel, D. (2008) Dual effects of miR156-targeted *SPL* genes and *CYP78A5/KLUH* on plastochron length and organ size in *Arabidopsis thaliana*. *Plant Cell*, **20**, 1231–1243.
- Wang, J., Shine, M.B., Gao, Q.M., Navarre, D., Jiang, W., Liu, C., Chen, Q. et al. (2014a) Enhanced disease susceptibility1 mediates pathogen resistance and virulence function of a bacterial effector in soybean. *Plant Physiol.* **165**, 1269–1284.
- Wang, K., Bu, T., Cheng, Q., Dong, L., Su, T., Chen, Z., Kong, F. et al. (2021a) Two homologous *LHY* pairs negatively control soybean drought tolerance by repressing the ABA responses. *New Phytol.* **229**, 2660–2675.
- Wang, L., Cao, C., Ma, Q., Zeng, Q., Wang, H., Cheng, Z., Zhu, G. et al. (2014b) RNA-seq analyses of multiple meristems of soybean: novel and alternative transcripts, evolutionary and functional implications. *BMC Plant Biol.* **14**, 169.
- Wang, L., Liu, L., Ma, Y., Li, S., Dong, S. and Zu, W. (2018a) Transcriptome profiling analysis characterized the gene expression patterns responded to combined drought and heat stresses in soybean. *Comput. Biol. Chem.* **77**, 413–429.
- Wang, L., Sun, S., Wu, T., Liu, L., Sun, X., Cai, Y., Li, J. et al. (2020c) Natural variation and CRISPR/Cas9-mediated mutation in *GmPRR37* affect photoperiodic flowering and contribute to regional adaptation of soybean. *Plant Biotechnol. J.* **18**, 1869–1881.

- Wang, L., Sun, Z., Su, C., Wang, Y., Yan, Q., Chen, J., Ott, T. et al. (2019c) A GmNINA-miR172c-NNC1 regulatory network coordinates the nodulation and autoregulation of nodulation pathways in soybean. *Mol. Plant*, **12**, 1211–1226.
- Wang, M., Li, W., Fang, C., Xu, F., Liu, Y., Wang, Z., Yang, R. et al. (2018b) Parallel selection on a dormancy gene during domestication of crops from multiple families. *Nat. Genet.* **50**, 1435–1441.
- Wang, N., Liu, Y., Cai, Y., Tang, J., Li, Y. and Gai, J. (2020d) The soybean U-box gene *GmPUB6* regulates drought tolerance in *Arabidopsis thaliana*. *Plant Physiol. Biochem.* **155**, 284–296.
- Wang, S., Liu, S., Wang, J., Yokosho, K., Zhou, B., Yu, Y.-C., Liu, Z. et al. (2020e) Simultaneous changes in seed size, oil content, and protein content driven by selection of SWEET homologues during soybean domestication. *Natl Sci. Rev.* **7**, 1776–1786.
- Wang, S., Yokosho, K., Guo, R., Whelan, J., Ruan, Y.L., Ma, J.F. and Shou, H. (2019d) The soybean sugar transporter GmSWEET15 mediates sucrose export from endosperm to early embryo. *Plant Physiol.* **180**, 2133–2141.
- Wang, T., Sun, M.Y., Wang, X.S., Li, W.B. and Li, Y.G. (2016b) Over-expression of GmGla-regulated soybean miR172a confers early flowering in transgenic *Arabidopsis thaliana*. *Int. J. Mol. Sci.* **17**, 645.
- Wang, X., Chen, L. and Ma, J. (2019e) Genomic introgression through interspecific hybridization counteracts genetic bottleneck during soybean domestication. *Genome Biol.* **20**, 22.
- Wang, X. and Komatsu, S. (2018) Proteomic approaches to uncover the flooding and drought stress response mechanisms in soybean. *J. Proteomics.* **172**, 201–215.
- Wang, X., Li, Y., Zhang, H., Sun, G., Zhang, W. and Qiu, L. (2015b) Evolution and association analysis of *GmCYP78A10* gene with seed size/weight and pod number in soybean. *Mol. Biol. Rep.* **42**, 489–496.
- Wang, X., Zhao, J., Fang, Q., Chang, X., Sun, M., Li, W. and Li, Y. (2021b) GmAKT1 is involved in K⁺ uptake and Na⁺/K⁺ homeostasis in *Arabidopsis* and soybean plants. *Plant Sci.* **304**, 110736.
- Wang, Y., Gu, Y., Gao, H., Qiu, L., Chang, R., Chen, S. and He, C. (2016c) Molecular and geographic evolutionary support for the essential role of *GIGANTEA* in soybean domestication of flowering time. *BMC Evol. Biol.* **16**, 79.
- Wang, Y., Li, K., Chen, L., Zou, Y., Liu, H., Tian, Y., Li, D. et al. (2015c) MicroRNA167-directed regulation of the auxin response factors *GmARF8a* and *GmARF8b* is required for soybean nodulation and lateral root development. *Plant Physiol.* **168**, 984–999.
- Wang, Y., Liu, M., Ge, D., Akhter, B.J., Li, Y., Kong, J., Liu, K. et al. (2020f) Hydroperoxide lyase modulates defense response and confers lesion-mimic leaf phenotype in soybean (*Glycine max* (L.) Merr.). *Plant J.* **104**, 1315–1333.
- Wang, Y., Suo, H., Zheng, Y., Liu, K., Zhuang, C., Kahle, K.T., Ma, H. et al. (2010) The soybean root-specific protein kinase GmWINK1 regulates stress-responsive ABA signaling on the root system architecture. *Plant J.* **64**, 230–242.
- Wang, Y., Wang, L., Zou, Y., Chen, L., Cai, Z., Zhang, S., Zhao, F. et al. (2014c) Soybean miR172c targets the repressive AP2 transcription factor NNC1 to activate *ENOD40* expression and regulate nodule initiation. *Plant Cell*, **26**, 4782–4801.
- Wang, Y., Yang, W., Zuo, Y., Zhu, L., Hastwell, A.H., Chen, L., Tian, Y. et al. (2019f) *GmYUC2a* mediates auxin biosynthesis during root development and nodulation in soybean. *J. Exp. Bot.* **70**, 3165–3176.
- Wang, Y., Yu, K.F., Poysa, V., Shi, C. and Zhou, Y.H. (2012b) A single point mutation in *GmHMA3* affects Cadmium (Cd) translocation and accumulation in soybean seeds. *Mol. Plant*, **5**, 1154–1156.
- Wang, Z. and Tian, Z. (2015) Genomics progress will facilitate molecular breeding in soybean. *Sci. China Life Sci.* **58**, 813–815.
- Wang, Z., Zhou, Z., Liu, Y., Liu, T., Li, Q., Ji, Y., Li, C. et al. (2015d) Functional evolution of phosphatidylethanolamine binding proteins in soybean and *Arabidopsis*. *Plant Cell*, **27**, 323–336.
- Watanabe, S., Hideshima, R., Xia, Z., Tsubokura, Y., Sato, S., Nakamoto, Y., Yamanaka, N. et al. (2009) Map-based cloning of the gene associated with the soybean maturity locus *E3*. *Genetics*, **182**, 1251–1262.
- Watanabe, S., Xia, Z.J., Hideshima, R., Tsubokura, Y., Sato, S., Yamanaka, N., Takahashi, R. et al. (2011) A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. *Genetics*, **188**, 395–407.
- Wei, W., Liang, D.W., Bian, X.H., Shen, M., Xiao, J.H., Zhang, W.K., Ma, B. et al. (2019) GmWRKY54 improves drought tolerance through activating genes in abscisic acid and Ca²⁺ signaling pathways in transgenic soybean. *Plant J.* **100**, 384–398.
- Wei, W., Tao, J.J., Chen, H.W., Li, Q.T., Zhang, W.K., Ma, B., Lin, Q. et al. (2017) A histone code reader and a transcriptional activator interact to regulate genes for salt tolerance. *Plant Physiol.* **175**, 1304–1320.
- Wei, X. and Huang, X. (2018) Identification of a seed dormancy gene in soybean sheds light on crop domestication. *Sci. China Life Sci.* **61**, 1439–1441.
- Wei, Y., Balaceanu, A., Rufian, J.S., Segonzac, C., Zhao, A., Morcillo, R.J.L. and Macho, A.P. (2020a) An immune receptor complex evolved in soybean to perceive a polymorphic bacterial flagellin. *Nat. Commun.* **11**, 3763.
- Wei, Z., Pan, T., Zhao, Y., Su, B., Ren, Y. and Qiu, L. (2020b) The small GTPase Rab5a and its guanine nucleotide exchange factors are involved in post-Golgi trafficking of storage proteins in developing soybean cotyledon. *J. Exp. Bot.* **71**, 808–822.
- Whitham, S.A., Qi, M., Innes, R.W., Ma, W., Lopes-Caitar, V. and Hewezi, T. (2016) Molecular soybean-pathogen interactions. *Annu. Rev. Phytopathol.* **54**, 443–468.
- Widiasari, K., Alazem, M. and Kim, K.H. (2020) Soybean resistance to soybean mosaic virus. *Plants*, **9**, 219.
- Willson, J. (2020) Resolving the roles of structural variants. *Nat. Rev. Genet.* **21**, 507.
- Wilson, R.F. (2008) *Soybean: Market Driven Research Needs in Genetics and Genomics of Soybean*. New York, NY: Springer Science Press.
- Wolfgang, G. and Charles, Y.Q. (2017) Genetic separation of southern and northern soybean breeding programs in North America and their associated allelic variation at four maturity loci. *Mol. Breed.* **37**, 8.
- Wong, C.E., Zhao, Y.T., Wang, X.J., Croft, L., Wang, Z.H., Haerizadeh, F., Mattick, J.S. et al. (2011) MicroRNAs in the shoot apical meristem of soybean. *J. Exp. Bot.* **62**, 2495–2506.
- Wong, J., Gao, L., Yang, Y., Zhai, J., Arikat, S., Yu, Y., Duan, S. et al. (2014) Roles of small RNAs in soybean defense against *Phytophthora sojae* infection. *Plant J.* **79**, 928–940.
- Wrather, J.A. and Koenning, S.R. (2006) Estimates of disease effects on soybean yields in the United States 2003 to 2005. *J. Nematol.* **38**, 173–180.
- Wu, D., Li, D., Zhao, X., Zhan, Y., Teng, W., Qiu, L., Zheng, H. et al. (2020) Identification of a candidate gene associated with isoflavone content in soybean seeds using genome-wide association and linkage mapping. *Plant J.* **104**, 950–963.
- Wu, F., Price, B.W., Haider, W., Seufferheld, G., Nelson, R. and Hanzawa, Y. (2014) Functional and evolutionary characterization of the *CONSTANS* gene family in short-day photoperiodic flowering in soybean. *PLoS One*, **9**, e85754.
- Wu, F., Sedivy, E.J., Price, W.B., Haider, W. and Hanzawa, Y. (2017) Evolutionary trajectories of duplicated *FT* homologues and their roles in soybean domestication. *Plant J.* **90**, 941–953.
- Wu, G., Park, M.Y., Conway, S.R., Wang, J.W., Weigel, D. and Poethig, R.S. (2009) The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell*, **138**, 750–759.
- Wu, W., Lin, Y., Liu, P., Chen, Q., Tian, J. and Liang, C. (2018) Association of extracellular dNTP utilization with a GmPAP1-like protein identified in cell wall proteomic analysis of soybean roots. *J. Exp. Bot.* **69**, 603–617.
- Wu, Z., Wang, M., Yang, S., Chen, S., Chen, X., Liu, C., Wang, S. et al. (2019) A global coexpression network of soybean genes gives insights into the evolution of nodulation in nonlegumes and legumes. *New Phytol.* **223**, 2104–2119.
- Xia, Z.J., Watanabe, S., Yamada, T., Tsubokura, Y., Nakashima, H., Zhai, H., Anai, T. et al. (2012) Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. *Proc. Natl Acad. Sci. USA*, **109**, E2155–E2164.
- Xia, Z., Zhai, H., Lu, S., Wu, H. and Zhang, Y. (2013) Recent achievement in gene cloning and functional genomics in soybean. *Sci. World J.* **2013**, 281367.
- Xie, M., Chung, C.Y., Li, M.W., Wong, F.L., Wang, X., Liu, A., Wang, Z. et al. (2019) A reference-grade wild soybean genome. *Nat. Commun.* **10**, 1216.

- Xiong, L., Li, C., Li, H., Lyu, X., Zhao, T., Liu, J., Zuo, Z. *et al.* (2019) A transient expression system in soybean mesophyll protoplasts reveals the formation of cytoplasmic GmCRY1 photobody-like structures. *Sci. China Life Sci.* **62**, 1070–1077.
- Xu, C., Xia, C., Xia, Z., Zhou, X., Huang, J., Huang, Z., Liu, Y. *et al.* (2018) Physiological and transcriptomic responses of reproductive stage soybean to drought stress. *Plant Cell Rep.* **37**, 1611–1624.
- Xu, H., Zhang, L., Zhang, K. and Ran, Y. (2020) Progresses, challenges, and prospects of genome editing in soybean (*Glycine max*). *Front. Plant Sci.* **11**, 571138.
- Xu, M., Xu, Z., Liu, B., Kong, F., Tsubokura, Y., Watanabe, S., Xia, Z. *et al.* (2013a) Genetic variation in four maturity genes affects photoperiod insensitivity and PHYA-regulated post-flowering responses of soybean. *BMC Plant Biol.* **13**, 91.
- Xu, M.L., Yamagishi, N., Zhao, C., Takeshima, R., Kasai, M., Watanabe, S., Kanazawa, A. *et al.* (2015) The soybean-specific maturity gene *E1* family of floral repressors controls night-break responses through down-regulation of *FLOWERING LOCUS T* orthologs. *Plant Physiol.* **168**, 1735–1746.
- Xu, X., Zeng, L., Tao, Y., Vuong, T., Wan, J., Boerma, R., Noe, J. *et al.* (2013b) Pinpointing genes underlying the quantitative trait loci for root-knot nematode resistance in palaeopolyploid soybean by whole genome resequencing. *Proc. Natl Acad. Sci. USA*, **110**, 13469–13474.
- Xu, Z.L., Ali, Z., Xu, L., He, X.L., Huang, Y.H., Yi, J.X., Shao, H.B. *et al.* (2016) The nuclear protein GmbZIP110 has transcription activation activity and plays important roles in the response to salinity stress in soybean. *Sci. Rep.* **6**, 20366.
- Xue, Y.B., Xiao, B.X., Zhu, S.N., Mo, X.H., Liang, C.Y., Tian, J., Liao, H. *et al.* (2017) *GmPHR25*, a *GmPHR* member up-regulated by phosphate starvation, controls phosphate homeostasis in soybean. *J. Exp. Bot.* **68**, 4951–4967.
- Xun, H.W., Yang, X.D., He, H.L., Wang, M., Guo, P., Wang, Y., Pang, J.S. *et al.* (2019) Over-expression of *GmKR3*, a TIR-NBS-LRR type *R* gene, confers resistance to multiple viruses in soybean. *Plant Mol. Biol.* **99**, 95–111.
- Yamaguchi, Y., Barona, G., Ryan, C.A. and Pearce, G. (2011) GmPep914, an eight-amino acid peptide isolated from soybean leaves, activates defense-related genes. *Plant Physiol.* **156**, 932–942.
- Yan, Q., Wang, L. and Li, X. (2018) GmBEHL1, a BES1/BZR1 family protein, negatively regulates soybean nodulation. *Sci. Rep.* **8**, 7614.
- Yan, Z., Hossain, M.S., Wang, J., Valdes-Lopez, O., Liang, Y., Libault, M., Qiu, L. *et al.* (2013) miR172 regulates soybean nodulation. *Mol. Plant. Microbe Interact.* **26**, 1371–1377.
- Yan, Z., Hossain, M.S., Arikiti, S., Valdés-López, O., Zhai, J., Wang, J., Libault, M. *et al.* (2015) Identification of microRNAs and their mRNA targets during soybean nodule development: functional analysis of the role of miR393j-3p in soybean nodulation. *New Phytol.* **207**, 748–759.
- Yan, Z., Hossain, M.S., Valdés-López, O., Hoang, N.T., Zhai, J., Wang, J., Libault, M. *et al.* (2016) Identification and functional characterization of soybean root hair microRNAs expressed in response to Bradyrhizobium japonicum infection. *Plant Biotechnol. J.* **14**, 332–341.
- Yang, A., He, C. and Zhang, K. (2006) Improvement of *Agrobacterium*-mediated transformation of embryogenic calluses from maize elite inbred lines. *Vitro Cell. Dev. Biol. Plant*, **42**, 215–219.
- Yang, C.F., Huang, Y.Z., Lv, W.H., Zhang, Y.Y., Bhat, J.A., Kong, J.J., Xing, H. *et al.* (2020) GmNAC8 acts as a positive regulator in soybean drought stress. *Plant Sci.* **293**, 110442.
- Yang, J. and Huang, X. (2018) A new high-quality genome sequence in soybean. *Sci. China Life Sci.* **61**, 1604–1605.
- Yang, S., Tang, F., Gao, M., Krishnan, H.B. and Zhu, H. (2010) *R* gene-controlled host specificity in the legume-rhizobia symbiosis. *Proc. Natl Acad. Sci. USA*, **107**, 18735–18740.
- Yang, Y., Chi, Y., Wang, Z., Zhou, Y., Fan, B. and Chen, Z. (2016) Functional analysis of structurally related soybean GmWRKY58 and GmWRKY76 in plant growth and development. *J. Exp. Bot.* **67**, 4727–4742.
- Yang, X., Kim, M.Y., Ha, J. and Lee, S.H. (2019) Overexpression of the soybean NAC gene *GmNAC109* increases lateral root formation and abiotic stress tolerance in transgenic *Arabidopsis* plants. *Front. Plant Sci.* **10**, 1036.
- Yasutaka, T., Hisakazu, M., Meilan, X., Baohui, L., Hiroko, N., Toyoaki, A., Fanjiang, K. *et al.* (2013) Genetic variation in soybean at the maturity locus *E4* is involved in adaptation to long days at high latitudes. *Agronomy*, **3**, 117–134.
- Yi, J., Derynck, M.R., Li, X., Telmer, P., Marsolais, F. and Dhaubhadel, S. (2010) A single-repeat MYB transcription factor, GmMYB176, regulates *CHS8* gene expression and affects isoflavonoid biosynthesis in soybean. *Plant J.* **62**, 1019–1034.
- Yu, H., Lin, T., Meng, X., Du, H., Zhang, J., Liu, G., Chen, M. *et al.* (2021) A route to *de novo* domestication of wild allotetraploid rice. *Cell*, **184**, 1156–1170 e1114.
- Yu, L., Liu, Y., Zeng, S., Yan, J., Wang, E. and Luo, L. (2019) Expression of a novel PSK-encoding gene from soybean improves seed growth and yield in transgenic plants. *Planta*, **249**, 1239–1250.
- Yu, Y.G., Maroof, M.A.S., Buss, G.R., Maughan, P.J. and Tolin, S.A. (1994) Rflp and microsatellite mapping of a gene for soybean mosaic-virus resistance. *Phytopathology*, **84**, 60–64.
- Yuan, L., Xie, G.Z., Zhang, S., Li, B., Wang, X., Li, Y., Liu, T. *et al.* (2021) GmLCLs negatively regulate ABA perception and signalling genes in soybean leaf dehydration response. *Plant Cell Environ.* **44**, 412–424.
- Yue, Y.L., Liu, N.X., Jiang, B.J., Li, M., Wang, H.J., Jiang, Z., Pan, H.T. *et al.* (2017) A single nucleotide deletion in *J* encoding GmELF3 confers long juvenility and is associated with adaption of tropic soybean. *Mol. Plant*, **10**, 656–658.
- Zeng, X., Liu, H., Du, H., Wang, S., Yang, W., Chi, Y., Wang, J. *et al.* (2018) Soybean MADS-box gene *GmAGL1* promotes flowering via the photoperiod pathway. *BMC Genom.* **19**, 51.
- Zhai, H., Lu, S.X., Liang, S., Wu, H.Y., Zhang, X.Z., Liu, B.H., Kong, F.J. *et al.* (2014a) *GmFT4*, a homolog of *FLOWERING LOCUS T*, is positively regulated by *E1* and functions as a flowering repressor in soybean. *PLoS One*, **9**, e89030.
- Zhai, H., Lu, S.X., Wang, Y.Q., Chen, X., Ren, H.X., Yang, J.Y., Cheng, W. *et al.* (2014b) Allelic variations at four major maturity *E* genes and transcriptional abundance of the *E1* gene are associated with flowering time and maturity of soybean cultivars. *PLoS One*, **9**, e97636.
- Zhang, B., Wang, M., Sun, Y., Zhao, P., Liu, C., Qing, K., Hu, X. *et al.* (2021) *Glycine max* *NNL1* restricts symbiotic compatibility with widely distributed bradyrhizobia via root hair infection. *Nat. Plants*, **7**, 73–86.
- Zhang, C., Qi, M., Zhang, X., Wang, Q., Yu, Y., Zhang, Y. and Kong, Z. (2020a) Rhizobial infection triggers systemic transport of endogenous RNAs between shoots and roots in soybean. *Sci. China Life Sci.* **63**, 1213–1226.
- Zhang, D., Song, H., Cheng, H., Hao, D., Wang, H., Kan, G., Jin, H. *et al.* (2014) The acid phosphatase-encoding gene *GmACP1* contributes to soybean tolerance to low-phosphorus stress. *PLoS Genet.* **10**, e1004061.
- Zhang, D., Wang, X., Li, S., Wang, C., Gosney, M.J., Mickelbart, M.V. and Ma, J. (2019a) A post-domestication mutation, *Dt2*, triggers systemic modification of divergent and convergent pathways modulating multiple agronomic traits in soybean. *Mol. Plant*, **12**, 1366–1382.
- Zhang, D., Zhang, H., Hu, Z., Chu, S., Yu, K., Lv, L., Yang, Y. *et al.* (2019b) Artificial selection on *GmOLEO1* contributes to the increase in seed oil during soybean domestication. *PLoS Genet.* **15**, e1008267.
- Zhang, D.J., Zhao, M.X., Li, S., Sun, L.J., Wang, W.D., Cai, C.M., Dierking, E.C. *et al.* (2017) Plasticity and innovation of regulatory mechanisms underlying seed oil content mediated by duplicated genes in the palaeopolyploid soybean. *Plant J.* **90**, 1120–1133.
- Zhang, G., Ahmad, M.Z., Chen, B., Manan, S., Zhang, Y., Jin, H., Wang, X. *et al.* (2020b) Lipidomic and transcriptomic profiling of developing nodules reveals the essential roles of active glycolysis and fatty acid and membrane lipid biosynthesis in soybean nodulation. *Plant J.* **103**, 1351–1371.
- Zhang, G., Chen, M., Li, L., Xu, Z., Chen, X., Guo, J. and Ma, Y. (2009) Overexpression of the soybean *GmERF3* gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. *J. Exp. Bot.* **60**, 3781–3796.
- Zhang, H., Goettel, W., Song, Q., Jiang, H., Hu, Z., Wang, M.L. and An, Y.C. (2020c) Selection of *GmSWEET39* for oil and protein improvement in soybean. *PLoS Genet.* **16**, e1009114.
- Zhang, H., Yang, Y., Sun, C., Liu, X., Lv, L., Hu, Z., Yu, D. *et al.* (2020d) Up-regulating *GmETO1* improves phosphorus uptake and use efficiency by promoting root growth in soybean. *Plant Cell Environ.* **43**, 2080–2094.

- Zhang, J., Singh, A., Mueller, D.S. and Singh, A.K. (2015) Genome-wide association and epistasis studies unravel the genetic architecture of sudden death syndrome resistance in soybean. *Plant J.* **84**, 1124–1136.
- Zhang, J., Wang, X., Lu, Y., Bhushal, S.J., Song, Q., Cregan, P.B., Yen, Y. et al. (2018) Genome-wide scan for seed composition provides insights into soybean quality improvement and the impacts of domestication and breeding. *Mol. Plant*, **11**, 460–472.
- Zhang, N.N., Zou, H., Lin, X.Y., Pan, Q., Zhang, W.Q., Zhang, J.H., Wei, G.H. et al. (2020e) Hydrogen sulfide and rhizobia synergistically regulate nitrogen (N) assimilation and remobilization during N deficiency-induced senescence in soybean. *Plant Cell Environ.* **43**, 1130–1147.
- Zhang, P., Du, H., Wang, J., Pu, Y., Yang, C., Yan, R., Yang, H. et al. (2020f) Multiplex CRISPR/Cas9-mediated metabolic engineering increases soya bean isoflavone content and resistance to soya bean mosaic virus. *Plant Biotechnol. J.* **18**, 1384–1395.
- Zhang, Q., Li, H., Li, R., Hu, R., Fan, C., Chen, F., Wang, Z. et al. (2008) Association of the circadian rhythmic expression of GmCRY1a with a latitudinal cline in photoperiodic flowering of soybean. *Proc. Natl Acad. Sci. USA*, **105**, 21028–21033.
- Zhang, W., Liao, X., Cui, Y., Ma, W., Zhang, X., Du, H., Ma, Y. et al. (2019c) A cation diffusion facilitator, GmCDF1, negatively regulates salt tolerance in soybean. *PLoS Genet.* **15**, e1007798.
- Zhao, B., Dai, A., Wei, H., Yang, S., Wang, B., Jiang, N. and Feng, X. (2016) *Arabidopsis KLU* homologue *GmCYP78A72* regulates seed size in soybean. *Plant Mol. Biol.* **90**, 33–47.
- Zhao, J., Chen, L., Zhao, T. and Gai, J. (2017) Chicken Toes-Like Leaf and Petaloid Flower (CTP) is a novel regulator that controls leaf and flower development in soybean. *J. Exp. Bot.* **68**, 5565–5581.
- Zhao, L., Li, M., Xu, C., Yang, X., Li, D., Zhao, X., Wang, K. et al. (2018) Natural variation in *GmGBP1* promoter affects photoperiod control of flowering time and maturity in soybean. *Plant J.* **96**, 147–162.
- Zhao, M., Meyers, B.C., Cai, C., Xu, W. and Ma, J. (2015a) Evolutionary patterns and coevolutionary consequences of MIRNA genes and microRNA targets triggered by multiple mechanisms of genomic duplications in soybean. *Plant Cell*, **27**, 546–562.
- Zhao, X., Han, Y., Li, Y., Liu, D., Sun, M., Zhao, Y., Lv, C. et al. (2015b) Loci and candidate gene identification for resistance to *Sclerotinia sclerotiorum* in soybean (*Glycine max* L. Merr.) via association and linkage maps. *Plant J.* **82**, 245–255.
- Zhao, X., Jing, Y., Luo, Z.H., Gao, S.N., Teng, W.L., Zhan, Y.H., Qiu, L.J. et al. (2021) GmST1, which encodes a sulfotransferase, confers resistance to soybean mosaic virus strains G2 and G3. *Plant Cell Environ.* **44**, 2777–2779.
- Zheng, F., Wu, H., Zhang, R., Li, S., He, W., Wong, F.L., Li, G. et al. (2016) Molecular phylogeny and dynamic evolution of disease resistance genes in the legume family. *BMC Genom.* **17**, 402.
- Zhou, Q.Y., Tian, A.G., Zou, H.F., Xie, Z.M., Lei, G., Huang, J., Wang, C.M. et al. (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., Gou, Z., Lyu, J., Li, W., Yu, Y. et al. (2015) Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat. Biotechnol.* **33**, 408–414.
- Zhou, Z., Wang, Z., Li, W., Fang, C., Shen, Y., Li, C., Wu, Y. et al. (2013) Comprehensive analyses of microRNA gene evolution in paleopolyploid soybean genome. *Plant J.* **76**, 332–344.
- Zsogon, A., Cermak, T., Naves, E.R., Notini, M.M., Edel, K.H., Weinl, S., Freschi, L. et al. (2018) *De novo* domestication of wild tomato using genome editing. *Nat. Biotechnol.* **36**, 1211–1216.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1 Representative genes related to agronomically important traits in soybean.