



Article

GmWRI1c Increases Palmitic Acid Content to Regulate Seed Oil Content and Nodulation in Soybean (*Glycine max*)

Haowei Zheng, Duo Zhao, Wentao Shao, Yun Lu, Wenhui Wang, Yanjiao Hu, Jiajia Li, Shangshang Zhu and Xiaobo Wang *

School of Agronomy, Anhui Agricultural University, Hefei 230036, China

* Correspondence: wangxiaobo@ahau.edu.cn; Tel./Fax: +86-0551-65786209

Abstract: Soybean (*Glycine max*) is an important oil crop, but the regulatory mechanisms underlying seed oil accumulation remain unclear. We identified a member of the *GmWRI1s* transcription factor family, *GmWRI1c*, that is involved in regulating soybean oil content and nodulation. Overexpression of *GmWRI1c* in soybean hairy roots increased the expression of genes involved in glycolysis and de novo lipogenesis, the proportion of palmitic acid (16:0), and the number of root nodules. The effect of *GmWRI1c* in increasing the number of root nodules via regulating the proportion of palmitic acid was confirmed in a recombinant inbred line (RIL) population. *GmWRI1c* shows abundant sequence diversity and has likely undergone artificial selection during domestication. An association analysis revealed a correlation between seed oil content and five linked natural variations (Hap1/Hap2) in the *GmWRI1c* promoter region. Natural variations in the *GmWRI1c* promoter were strongly associated with the *GmWRI1c* transcript level, with higher *GmWRI1c* transcript levels in lines carrying *GmWRI1c*^{Hap1} than in those carrying *GmWRI1c*^{Hap2}. The effects of *GmWRI1c* alleles on seed oil content were confirmed in natural and RIL populations. We identified a favourable *GmWRI1c* allele that can be used to breed new varieties with increased seed oil content and nodulation.



Citation: Zheng, H.; Zhao, D.; Shao, W.; Lu, Y.; Wang, W.; Hu, Y.; Li, J.; Zhu, S.; Wang, X. *GmWRI1c* Increases Palmitic Acid Content to Regulate Seed Oil Content and Nodulation in Soybean (*Glycine max*). *Int. J. Mol. Sci.* **2022**, *23*, 13793. <https://doi.org/10.3390/ijms232213793>

Academic Editor: Setsuko Komatsu

Received: 20 October 2022

Accepted: 7 November 2022

Published: 9 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: *GmWRI1c*; natural variation; nodule number; palmitic acid (16:0); seed oil content; soybean (*Glycine max*)

1. Introduction

Soybean is widely used in modern society. It provides health benefits for humans through its oil, protein, isoflavones, phospholipids, and vitamins [1–3]. With rapid population growth, there has been a significant increase in the demand for soybean oil. Therefore, it is of great significance to identify genes involved in the synthesis of oil, and in the regulation of this process, in order to breed new, high-quality soybean varieties.

Soybean oil consists of five main fatty acids; palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1), linoleic acid (18:2), and linolenic acid (18:3). Palmitic acid and stearic acid are saturated fatty acids, while oleic acid, linoleic acid, and linolenic acid are unsaturated fatty acids [4]. Most studies on fatty acid components focus on food storage and safety. Saturated fatty acids can improve the oxidative stability of soybean oil, which is beneficial for storage. Unsaturated fatty acids can reduce the cholesterol content in human blood [5,6]. However, the roles of these five fatty acids in plant growth and development are unclear.

WRI1 belongs to the ANT group of the AP2 transcription factor supergene family and promotes fatty acid synthesis by binding to the conserved AW-box in the promoters of its target genes [7,8]. *WRI1* was first discovered in Arabidopsis wrinkled seed coat mutants [9]. Previous studies have shown that *AtWRI1* and its homologs *BnWRI1*, *HaWRI1*, *CiWRI1*, and *GmWRI1a* can significantly increase the oil content in seeds of various species [10–15]; for example, *BnWRI1* regulates expression of glycolysis gene *BnBCCP* to increase seed oil content, and *GmWRI1a* regulates expression of de novo lipogenesis gene *GmACP* to increase seed oil content [11,12]. In soybean, *GmWRI1b* can improve plant structure and increase yield under field conditions [16]. A recent study proposed that *WRI1* is also involved in

the regulation of soybean nodulation. In previous studies, the number of nodules in hairy roots was increased by the overexpression of *GmWRI1a/b* but decreased by the silencing of *GmWRI1a/b* [17]. These findings suggest that *WRI1* plays a vital role in plant growth and development. Thus, studies on soybean *WRI1* genes have important implications for crop health, growth, and yield.

The root nodule is a unique structure in legumes. Rhizobia fix nitrogen from the air into ammonia, thereby supplying nitrogen for the growth and development of their host plants. The roots and nodules of the host supply carbon and nutrients to rhizobia [18]. Symbiotic nodulation is a complex trait affected by nodulation factors, endogenous hormones, and the external environment. Previous studies have found that nodulation factors directly regulate nodulation in legume plants. Nodulation factors include nodule receptors (encoded by *NINs*), nodule factor receptors (encoded by *NFR1s*), symbiotic receptor kinases (encoded by *SYMRK/DMI2*), calmodulin kinases (encoded by *CCaMK/DMI3*), nodule signalling channel protein (encoded by *NSP*), and nodule marker oligopeptides (encoded by *ENOD40*) [19,20]. Plant hormones also play an important role in the nodulation process [21]. Nodulation is promoted by auxin and cytokinin [22–24] and inhibited by abscisic acid [25,26]. Legume nodulation is also affected by environmental factors. For example, light, nitrogen, and salt stress inhibit legume nodulation [27–29]. Recent studies have also shown that lipid metabolism genes are involved in regulating soybean nodulation [17,30,31], but it is still unclear how particular fatty acids affect the number of soybean nodules. Therefore, studies on the interaction between lipid metabolism and nodulation are of great practical significance.

WRI1 encodes a transcription factor with critical biological functions in plants, such as regulating oil synthesis, seed development, nodulation, and yield [10–16]. However, only *GmWRI1a* and *GmWRI1b* have been studied in soybean so far. Exploring the biological functions of transcription factors encoded by other *GmWRI1s* and determining their selection patterns during domestication is very important for soybean breeding.

2. Results

2.1. Identification of WRI Gene Family Members in Soybean

WRI belongs to the AP2 transcription factor family. We downloaded the original HMM file of AP2 (PF00847) from the Pfam database, identified genes containing AP2 domains in the Williams 82 soybean reference genome, and then constructed a soybean-specific AP2 HMM. A total of 166 candidate genes containing AP2 domains were identified in the Williams 82 soybean reference genome using the AP2-specific HMM. The conserved structural domains of the candidate genes were further analysed using Pfam, NCBI-CDD, and Smart databases to remove genes lacking AP2 structural domains and those with severe deletions of AP2 structural domains. Finally, we identified 162 genes encoding AP2 transcription factors in soybean (Supplementary Dataset S2).

A phylogenetic tree was constructed with the neighbour-joining (NJ) method using the 162 soybean genes containing AP2 domains and *WRI* genes in *Arabidopsis* (*AtWRI1*, *AtWRI2*, *AtWRI3*, *AtWRI4*) (Supplementary Figure S1). Fifteen soybean genes with AP2 domains and *Arabidopsis* *WRI* genes were in the same evolutionary clade. An evolutionary analysis was conducted for these 15 soybean *WRI* genes and those from other species (*Arabidopsis*, *M. sativa*, and *B. napus*). Six soybean genes belonged to the *WRI1* group, three belonged to the *WRI2* group, and six belonged to the *WRI3* and *WRI4* groups (Figure 1A). The six *GmWRI1s* showed the highest homology with *AtWRI1*, and they all had two AP2 domains (Figure 1B). They were named *GmWRI1a*, *GmWRI1b*, *GmWRI1c*, *GmWRI1d*, *GmWRI1e*, and *GmWRI1f*. *GmWRI1a* and *GmWRI1b* have been reported to be involved in fatty acid synthesis and regulation of soybean nodulation [17].

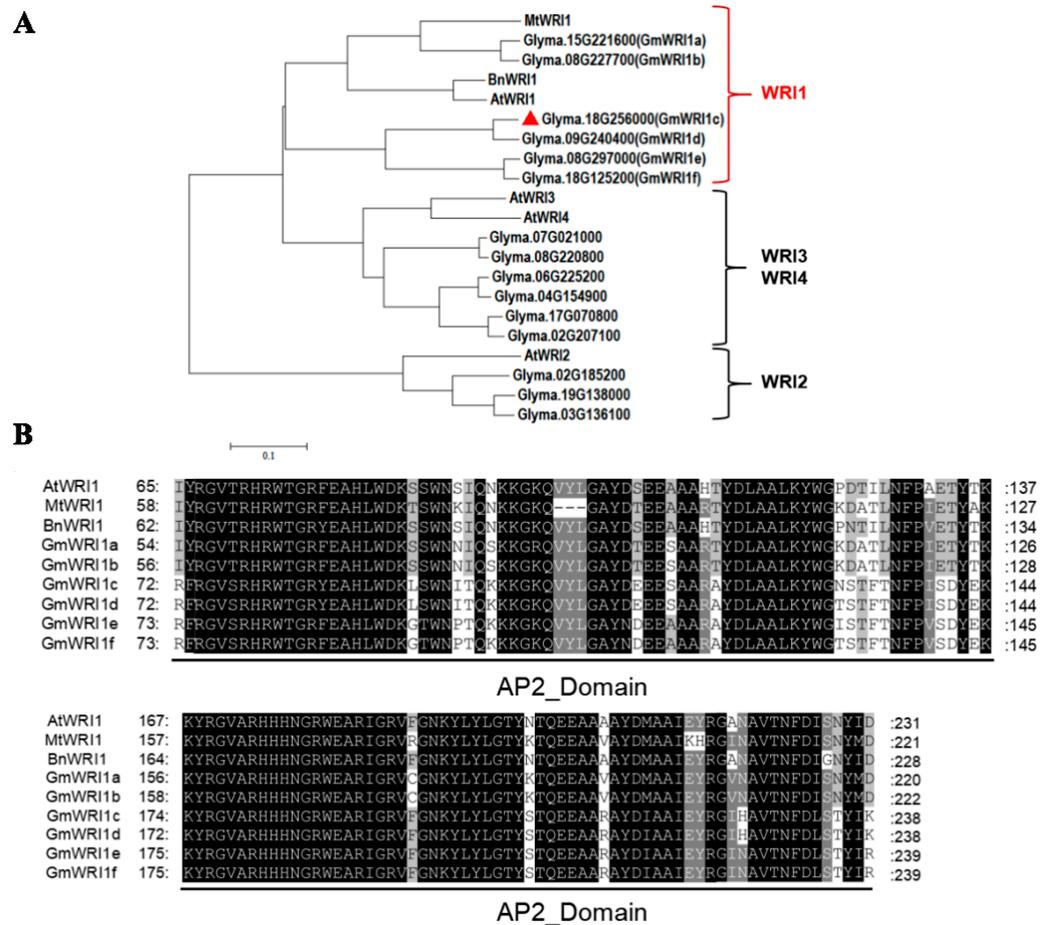


Figure 1. Phylogenetic and conserved domain analyses of GmWRI1s. (A) Phylogenetic tree of WRIs from *Arabidopsis*, soybean, *Medicago sativa*, and *Brassica napus*. (B) Conserved domain analysis of representative WRI1 proteins in *Arabidopsis*, soybean, *M. sativa*, and *B. napus*.

The tissue expression patterns of the six *GmWRI1s* were determined on the basis of data from the Phytozome database. *GmWRI1a* and *GmWRI1b* were expressed in the root, nodule, and seed; *GmWRI1c* was expressed in the nodule; *GmWRI1d* and *GmWRI1f* were barely expressed in soybean; and *GmWRI1e* was expressed in the leaf and seed (Supplementary Figure S2). Searches of the Soybase database revealed that *GmWRI1c*, *GmWRI1d*, and *GmWRI1f* are located in soybean oil QTL intervals (*GmWRI1c* is located in the QTL intervals of seed oils 6–4 and seed oils 24–27; *GmWRI1d* is located in the QTL intervals of seed oils 4–11; and *GmWRI1f* is located in the QTL intervals of seed oils 27–10) [32–35]. As *GmWRI1d* and *GmWRI1f* are hardly expressed in soybean, we focused on *GmWRI1c*.

2.2. Conserved Function of GmWRI1c in Fatty Acid Synthesis

The 35S::*GmWRI1c*::GFP fusion expression vector was transiently expressed in *N. benthamiana* leaves. The localisation analysis revealed that that *GmWRI1c* was located in the nucleus (Figure 2A–C), consistent with its role as a transcription factor.

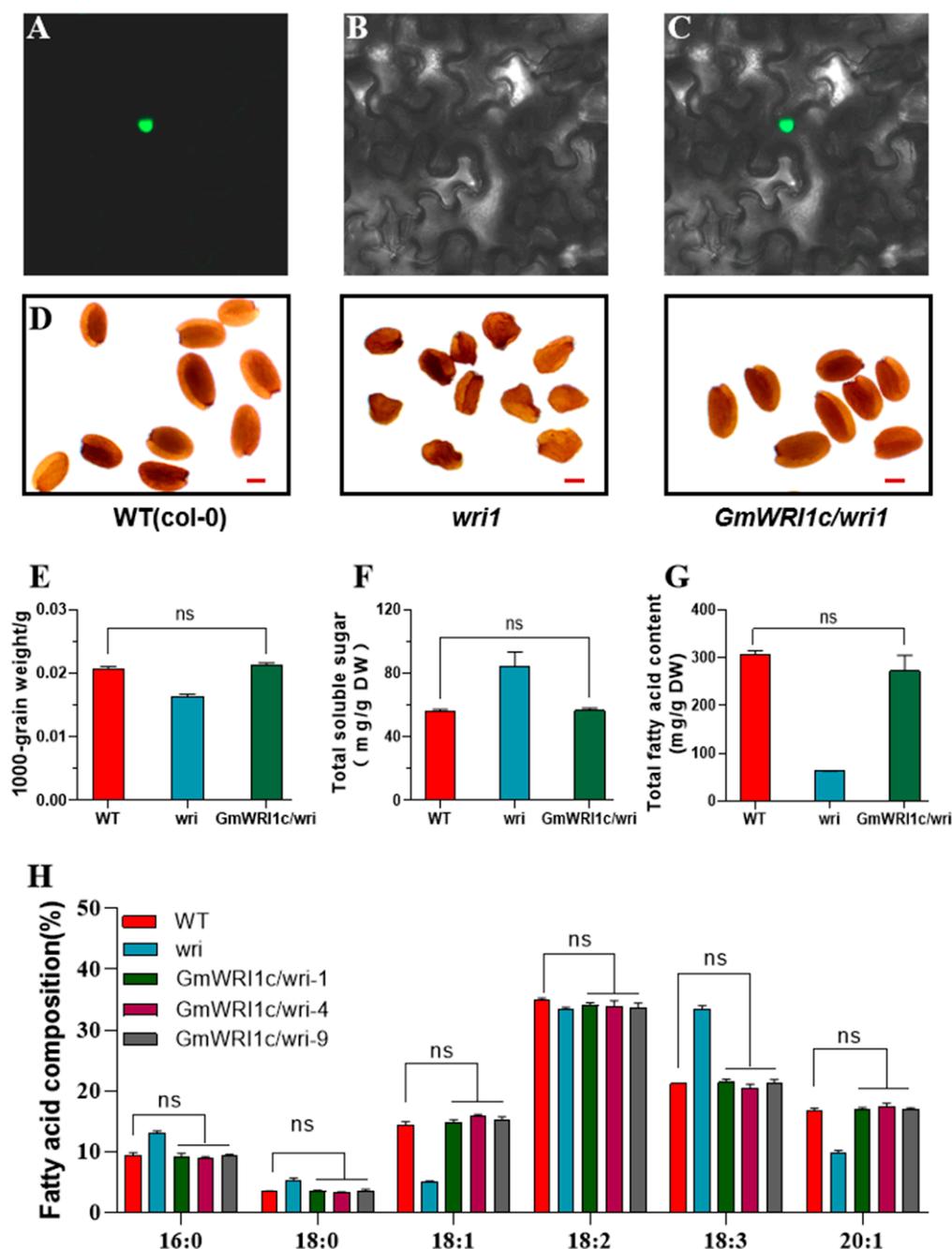


Figure 2. *GmWRI1c* restores the *Arabidopsis wri1* mutant phenotype. (A–C) Subcellular localisation of *GmWRI1c*. Fluorescence (A), bright-field (B), and merged (C) images of *GmWRI1c*-GFP. The magnification of microscope is 20 times. (D) Mature seeds of the wild type (Col-0), the *wri1* mutant, and *wri1* complemented with *35S::GmWRI1c/wri1*. Scale bar, 0.5 mm. (E–H) 1000-grain weight, total soluble sugar, total fatty acid content, and fatty acid composition of mature seeds of the wild type (Col-0), the *wri1* mutant, and *wri1* complemented with *35S::GmWRI1c/wri1*. *p*-values were determined by two-tailed two-sample Welch's *t*-tests.

We constructed a stably inherited *GmWRI1c* overexpression line in the *Arabidopsis wri1* mutant background (Supplementary Figure S3). *GmWRI1c* completely restored the wrinkle phenotype, thousand-grain weight, soluble sugar content, total fatty acid content, and fatty acid composition of *wri1* mutant seeds (Figure 2D–H). These results indicate that *GmWRI1c* has a conserved function in fatty acid synthesis.

2.3. *GmWRI1c* Specifically Increases Palmitic Acid Content by Regulating Genes Related to Glycolysis and De Novo Lipogenesis

To explore the function of *GmWRI1c* in soybean, we analysed lipids from both the roots and nodules of *GmWRI1c*-OE transgenic roots. Compared with the wild type, the line overexpressing *GmWRI1c* (Figure 3A,D) showed total fatty acid content in the roots and root nodules that was significantly increased by 38.22% and 32.8%, respectively (Figure 3B,E). Moreover, overexpression of *GmWRI1c* specifically increased the proportion of palmitic acid in roots and root nodules by 8.99% and 7.25%, respectively (Figure 3C,F). These results suggest that *GmWRI1c* promotes the biosynthesis of fatty acids, especially palmitic acid.

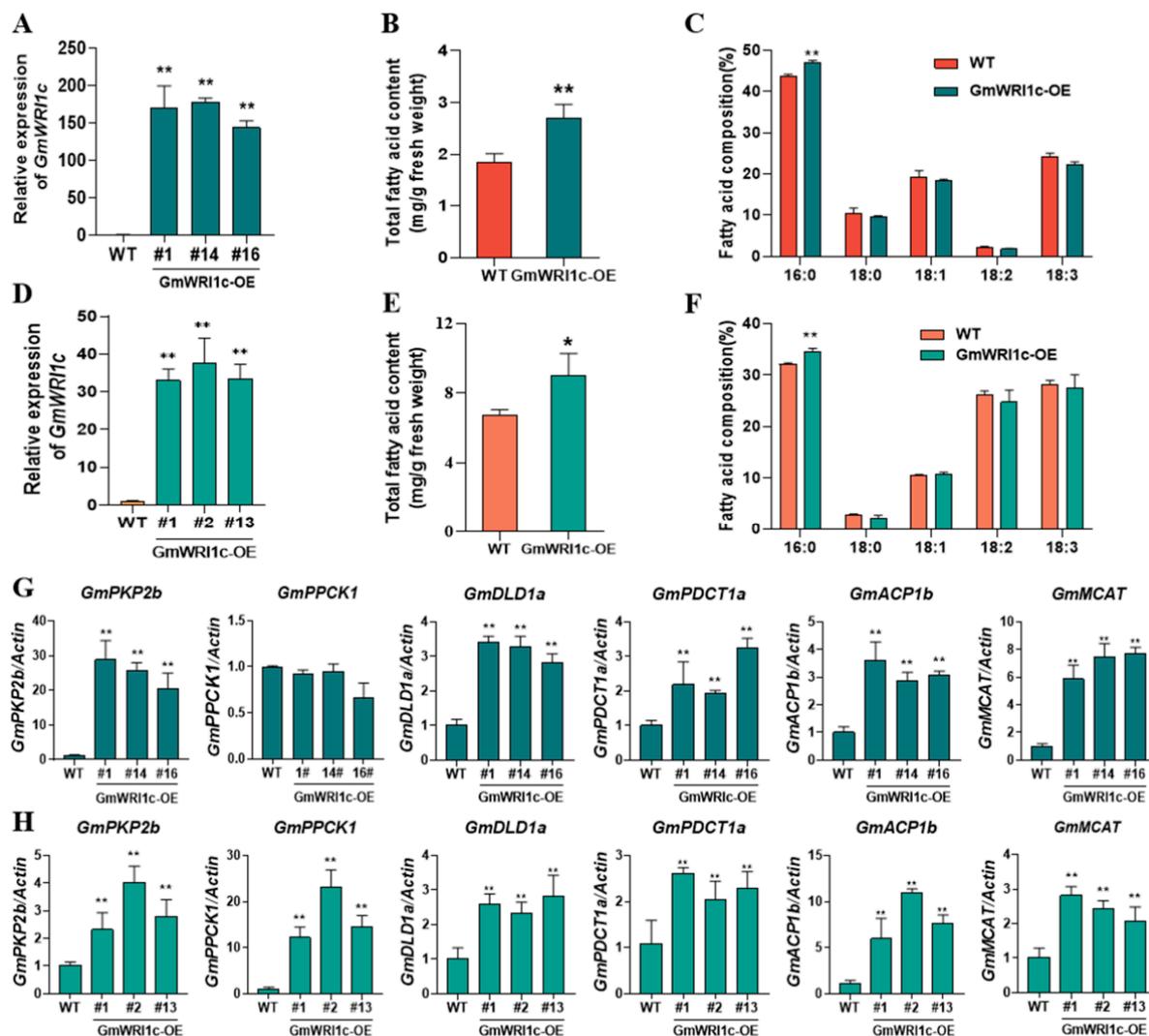


Figure 3. Analyses of *GmWRI1c* function in soybean. (A) Identification of transgenic hairy roots. (B,C) Total fatty acid content and fatty acid composition in *GmWRI1c*-OE transgenic hairy roots compared with control (values are means \pm SD from three biological replicates). (D) Identification of transgenic root nodules. (E,F) Total fatty acid content and fatty acid composition in *GmWRI1c*-OE transgenic root nodules compared with control (values are means \pm SD from three biological replicates). (G) Transcript levels of genes involved in glycolysis and de novo fatty acid synthesis in hairy roots of transgenic plants. (H) Transcript levels of genes involved in glycolysis and de novo fatty acid synthesis in root nodules of transgenic plants. * $p < 0.05$ and ** $p < 0.01$ (Student's *t*-test).

We then conducted qRT-PCR analyses to detect the transcript levels of genes involved in glycolysis and de novo lipogenesis in the *GmWRI1c*-OE transgenic hairy roots and nodules. Compared with the line expressing the empty vector, the *GmWRI1c*-OE line

showed significantly ($p < 0.01$) increased transcript levels of genes involved in glycolysis (e.g., *GmPDHE*, *GmPKP*, *GmDLD*) and de novo lipogenesis (e.g., *GmPDCT*, *GmACP*, *GmMCAT*) (Figure 3G,H; Supplementary Figure S4). These results suggest that *GmWRI1c* specifically increases palmitic acid content by regulating genes involved in glycolysis and de novo lipogenesis.

2.4. Overexpression of *GmWRI1c* Promotes Nodulation Genes' Expression and Increases Nodule Numbers in Soybean

The transcript levels of *GmWRI1c* in different soybean tissues were detected by qRT-PCR. The transcript levels of *GmWRI1c* were high in the root nodules and low in the seeds at 28 and 32 days after flowering (Figure 4A). Moreover, the transcript levels of *GmWRI1c* were significantly ($p < 0.01$) increased at 6 h after rhizobia infection (Figure 4B). Previous studies have shown that rhizobia invade root hair cells at 6–8 h after infection [36], so we hypothesised that *GmWRI1c* plays a role in the nodulation process of soybean. To test this hypothesis, we generated transgenic soybean hairy roots overexpressing *GmWRI1c* and then counted the number of root nodules. Compared with hairy roots expressing the empty vector, those of the *GmWRI1c*-OE line had significantly ($p < 0.01$) more root nodules. The average number of nodules was 69.8 on the transgenic roots compared with 43.4 on the roots of the control (Figure 4C,D), indicating that overexpression of *GmWRI1c* results in increased nodule numbers in soybean.

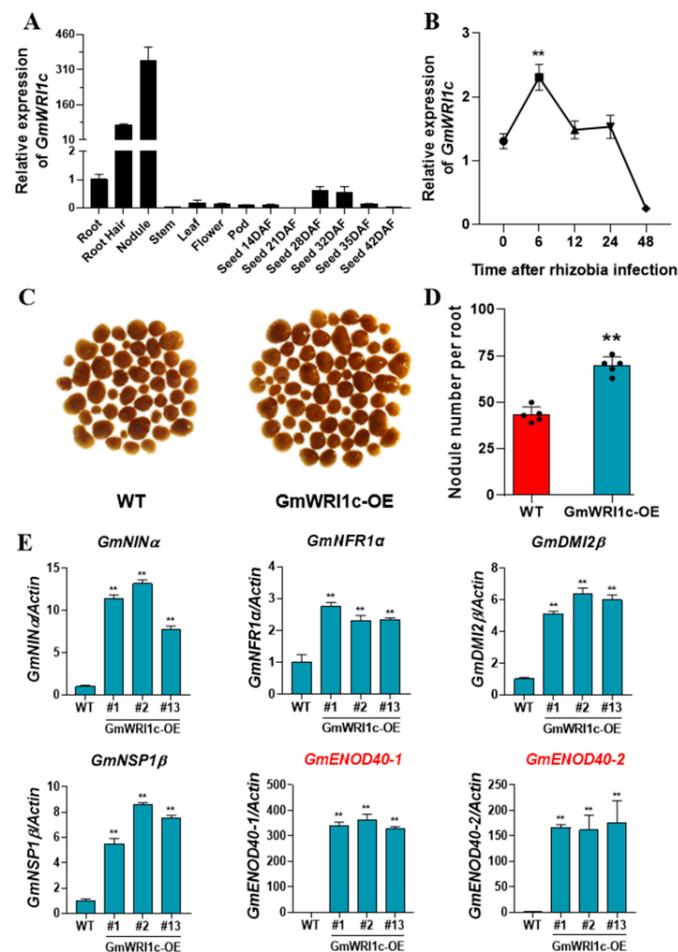


Figure 4. Effects of *GmWRI1c* on nodulation in soybean. (A) Transcript profiles of *GmWRI1c* in different tissues. (B) Transcript levels of *GmWRI1c* in roots after rhizobial infection. (C,D) Number of root nodules in *GmWRI1c*-OE transgenic hairy roots compared with control ($n = 5$). (E) Transcript level of gene encoding nodule factor in transgenic nodules. Values are means \pm SD from three technical replicates. ** $p < 0.01$ (Student's t -test).

Previous studies have detected rapid up-regulation of early nodulation genes (e.g., *ENOD40*, *NIN*) in soybean roots after rhizobial infection [37,38]. *ENOD40* significantly influences nodulation and has been widely used as a marker gene for nodule formation [39,40]. For example, knocking down *ENOD40* in *Lotus japonicus* suppresses root nodule formation [41], and an AP2 transcription factor, *GmNNCI*, was found to suppress *GmENOD40-1* expression to regulate root nodule formation in soybean roots [42]. In this study, we also found that rhizobial infection induced *GmWRI1c* expression, so we speculated that *GmWRI1c* may regulate the expression of genes related to nodulation. To test this idea, we determined the transcript levels of nodulation-related genes in *GmWRI1c*-OE transgenic root nodules and found that *GmNIN*, *GmNFR1*, *GmDMI*, *GmNSP1*, and *GmENOD40* transcript levels were elevated. In particular, the transcript level of *GmENOD40* was increased by several hundred-fold (Figure 4E; Supplementary Figure S5). The increased transcript levels of nodulation-related genes in transgenic root nodules of *GmWRI1c*-OE coincided with increased numbers of root nodules.

2.5. Superior Allele of *GmWRI1c*^{Hap1} Increases Seed Oil Content and Nodule Numbers in Soybean

Sequence variations in *GmWRI1c* were investigated in 20 high-oil cultivated soybeans and 20 low-oil cultivated soybeans (Supplementary Dataset S3). For each sample, an approximately 4.0 kb region of *GmWRI1c* was sequenced, including its promoter, intron, and exon. In total, 99 single nucleotide polymorphisms (SNPs) and insertions and deletions (InDels) were detected in this 4.0 kb region (Supplementary Dataset S4). An association analysis using the mixed model method found that the *GmWRI1c* promoter had five polymorphic sites significantly ($p < 0.01$) associated with variations in seed oil content in two environments (Supplementary Table S2; Supplementary Dataset S5). We divided the five polymorphic sites into Hap1 and Hap2 (Figure 5A). A cleaved amplified polymorphic sequence (CAPS) marker (Supplementary Figure S6) was developed for one of the C/G mutations (chr18-54487077), and the genetic effects of the two haplotypes (Hap1/Hap2) were analysed. Analyses of 644 individuals in a natural soybean population revealed 383 carrying the Hap1 allele and 261 carrying the Hap2 allele (Supplementary Dataset S1). Correlation analyses were conducted between the two haplotypes of *GmWRI1c* and the oil content of soybean seeds in the two environments. The results showed that *GmWRI1c*^{Hap1} significantly ($p < 0.01$) increased the oil content of soybean seeds in both environments, increasing the seed oil content by 3.73% and 4.09%, respectively (Figure 5B). The 644 soybean lines were divided into four class intervals on the basis of their seed oil content (<18%, 18–20%, 20–23%, and >23%). The proportion of *GmWRI1c*^{Hap1} increased with increasing oil content. The proportion of *GmWRI1c*^{Hap1} was much higher than that of *GmWRI1c*^{Hap2} in the high-oil cultivated soybeans (Figure 5C). We investigated the frequency of the alleles and found that *GmWRI1c*^{Hap1} was present in 7%, 53.29%, and 72.91% of wild soybeans, landraces, and elite lines, respectively (Figure 5D). This result indicates that *GmWRI1c*^{Hap1} has undergone artificial selection during early soybean domestication. These results suggest that *GmWRI1c*^{Hap1} is a superior allele of *GmWRI1c* that improves soybean seed oil content.

To further understand the mechanism by which variations in the *GmWRI1c* promoter affect the oil content of soybean seeds, we analysed gene expression driven by the two haplotypes' promoters (1600 bp upstream of the *GmWRI1c* transcription start site) in soybean hairy root transformation assays. The GUS activity was significantly higher ($p < 0.01$) in the transgenic nodules harbouring the *GmWRI1c*^{Hap1} promoter than in those harbouring the *GmWRI1c*^{Hap2} promoter (Figure 5E,F). Next, we determined the transcript levels of *GmWRI1c* in the nodules in soybean varieties with different haplotypes using qRT-PCR. The transcript levels of *GmWRI1c* were significantly ($p < 0.01$) higher in five high-oil cultivars with Hap1 than in five low-oil cultivars with Hap2 (Figure 5G; Supplementary Table S3). These data indicate that the differential expression of *GmWRI1c* in soybean nodules is the reason for differences in the oil contents of the soybean seeds.

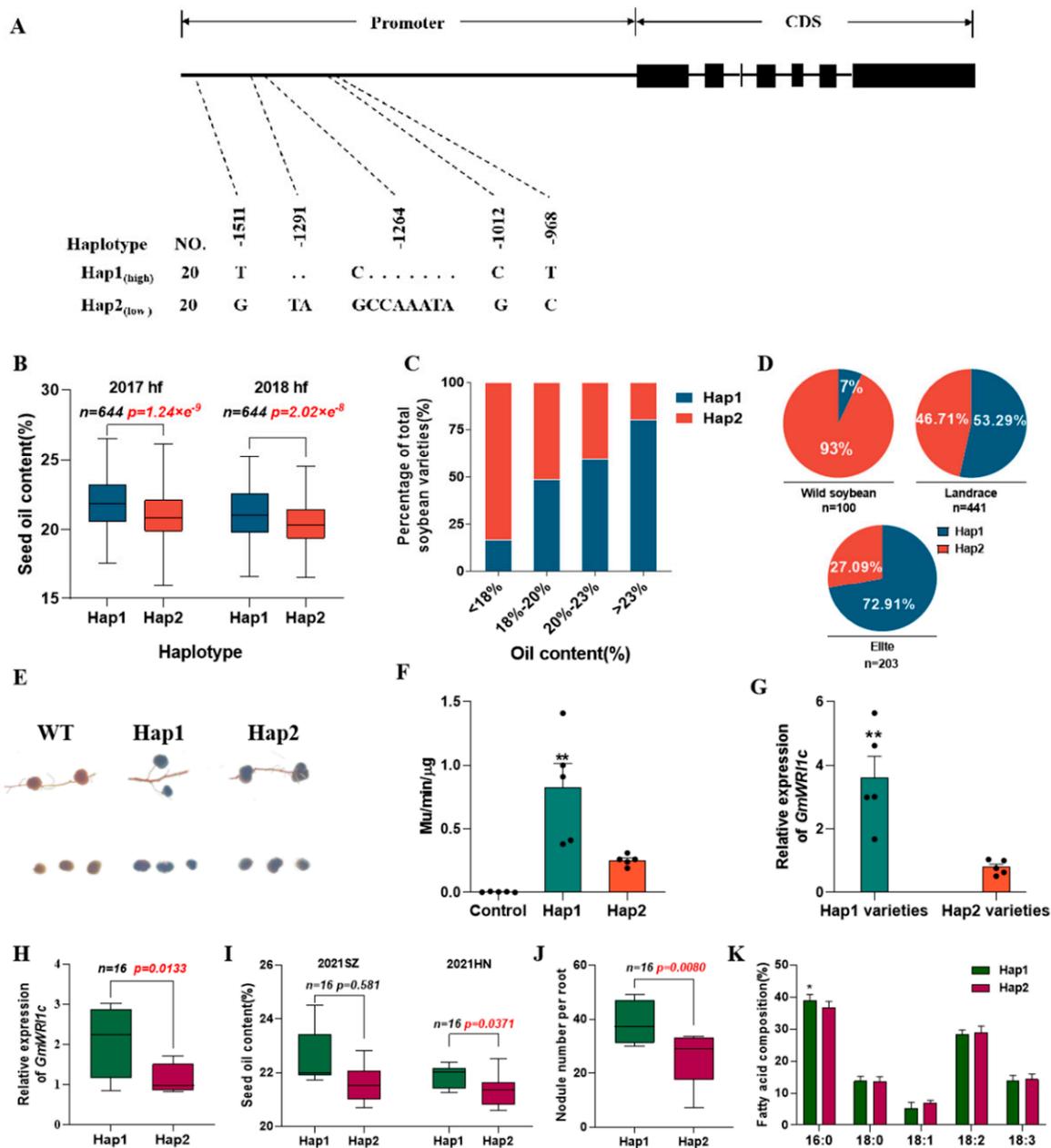


Figure 5. Effect of *GmWRI1c* on seed oil content and number of root nodules. (A) Haplotype analysis of *GmWRI1c* in 40 cultivated soybeans. (B) Boxplots of seed oil content for two haplotypes in 644 cultivated soybeans. (C) Percentages of two haplotypes of *GmWRI1c* in soybean lines grouped according to oil contents. (D) Distribution of two haplotypes of *GmWRI1c* in cultivated and wild soybeans. (E) GUS staining of transgenic nodules with gene expression driven by promoters of two *GmWRI1c* haplotypes. (F) Quantitative analysis of GUS activity in transgenic nodules ($n = 5$). (G) *GmWRI1c* transcript levels in root nodules of different soybean cultivars ($n = 5$). (H–K) Boxplots of gene transcript levels (H), seed oil contents (I), nodule numbers (J), and fatty acid composition (K) in soybean RILs with different haplotypes of *GmWRI1c*. 2017hf: grown in Hefei, China, in 2017. 2018hf: grown in Hefei, China, in 2018. 2021SZ: grown in Suzhou, China, in 2021. 2021HN: grown in Hainan, China, in 2021. Data are means \pm SD from three biological replicates. p -values were determined with two-tailed two-sample Welch’s t -tests. * $p < 0.05$ and ** $p < 0.01$.

To confirm the effect of the *GmWRI1c* allele type on seed oil content and the number of root nodules in soybean, the seed oil content, number of root nodules, and *GmWRI1c* sequences were analysed in a soybean RIL population (18 lines) derived from a cross

between Jin dou 21 (JD21) and Fu shen dou (FSD). JD21 carries the Hap1 type of *GmWRI1c*, whereas FSD carries Hap2. Consistent with observations in natural populations, in the RIL population the transcript levels of *GmWRI1c* and the seed oil contents were higher in the lines with Hap1 than in those with Hap2 (Figure 5H,I). We also found that, compared with the RILs carrying Hap2, those carrying Hap1 had more root nodules (Figure 5J) and higher palmitic acid contents in the nodules (Figure 5K). These results confirm that the superior *GmWRI1c* allele is associated with higher seed oil content and a higher number of root nodules in soybean and, thus, will be useful for breeding new varieties with increased oil content.

3. Discussion

Since the discovery of *WRI1* in *Arabidopsis* wrinkled seed coat mutants, its homologs have been identified in diverse plant species [43]. Previous studies have found that *WRI1* can directly bind to the AW box in the promoters of genes involved in glycolysis and de novo lipogenesis to regulate their transcript levels [17]. In this study, we identified a *WRI1* homolog that is expressed in root nodules and enhances the expression of genes related to glycolysis and de novo lipogenesis. This results in an increase in the proportion of palmitic acid out of total fatty acids. A recent study found that transient expression of *PaWRI1* in tobacco leaves specifically increased the proportion of palmitic acid [44], consistent with the results of our study. As palmitic acid is the main product of de novo lipogenesis [45,46], it is likely that *GmWRI1c* plays a main role in regulating de novo lipogenesis.

Symbiotic rhizobia–legume interactions are energy-demanding processes, and recent studies have shown that fatty acid composition and content strongly affect soybean nodule development and nodulation [17,30,31]. *GmWRI1a/b* directly regulates glycolysis and fatty acid biosynthesis, resulting in increased total fatty acid content and soybean nodulation [17]. Silencing of *GmACP* reduced the proportion of palmitic acid in transgenic hairy roots and inhibited soybean nodulation [30]. In the present study, overexpression of *GmWRI1c* significantly increased the transcript level of *GmACP*, leading to an increase in the ratio of palmitic acid in roots and nodules and more abundant root nodules. The results of our study and other studies suggest that palmitic acid has an important effect on soybean nodulation. Moreover, we detected a relationship between *GmWRI1c* alleles and the number of root nodules in the RIL population. The proportion of palmitic acid in the root nodules was higher in the RILs carrying the Hap1 type of *GmWRI1c* than in those carrying the Hap2 type. Therefore, we suggest that *GmWRI1c* may affect soybean nodulation by regulating the proportion of palmitic acid.

Legume domestication is one of the earliest technological innovations. From wild soybean to cultivated soybean, the seed size and seed oil content have undergone significant changes [47]. During soybean domestication, the seed size of *GmSWEET10a*, *POWR1*, *GmST105*, and *GmSS1* have significantly increased [48–51], and the seed oil content of *GmZF351*, *GmZF392*, and *GmST1* have significantly increased [52–54]. In this study, a nucleotide diversity analysis of *GmWRI1c* revealed a strong selective molecular footprint. *GmWRI1c*^{Hap1} was detected in 7%, 52.39%, and 72.91% of wild soybeans, landraces, and elite, respectively. The seed oil content was found to be significantly higher in lines carrying *GmWRI1c*^{Hap1} than in those carrying *GmWRI1c*^{Hap2} in both natural and RIL populations. These data indicate that *GmWRI1c*^{Hap1} is a beneficial haplotype that has been strongly favoured during soybean selection.

Recent studies have reported that divergence of gene expression driving trait evolution may be a common genetic mechanism in soybean domestication [55]. For example, natural variations in the promoters of *GmGBP* and *GsERD15B* affect gene expression, which alters the flowering time and salt tolerance in soybean [56,57]. The results of our transgenic experiments show that the promoter activity of *GmWRI1c*^{Hap1} is significantly higher than that of *GmWRI1c*^{Hap2}. Further analyses of natural and RIL populations confirmed that *GmWRI1c*^{Hap1} is expressed at higher levels than is *GmWRI1c*^{Hap2}. These data indicate that

differential expression of *GmWRI1c* in soybean nodules is responsible for differences in seed oil contents among lines and varieties.

In conclusion, we identified a crucial AP2 family gene, *GmWRI1c*, that is closely associated with nodule number and seed oil content in soybean. *GmWRI1c* promotes glycolysis and fatty acid biosynthesis and increases the proportion of palmitic acid to regulate soybean nodulation (Figure 6). Our analysis indicates that *GmWRI1c* has undergone artificial selection during early soybean domestication and identified a beneficial haplotype of *GmWRI1c*. The superior allele of *GmWRI1c*^{Hap1} is associated with higher seed oil content and an increased number of root nodules. Thus, the superior allele will be useful for breeding new high-oil soybean varieties.

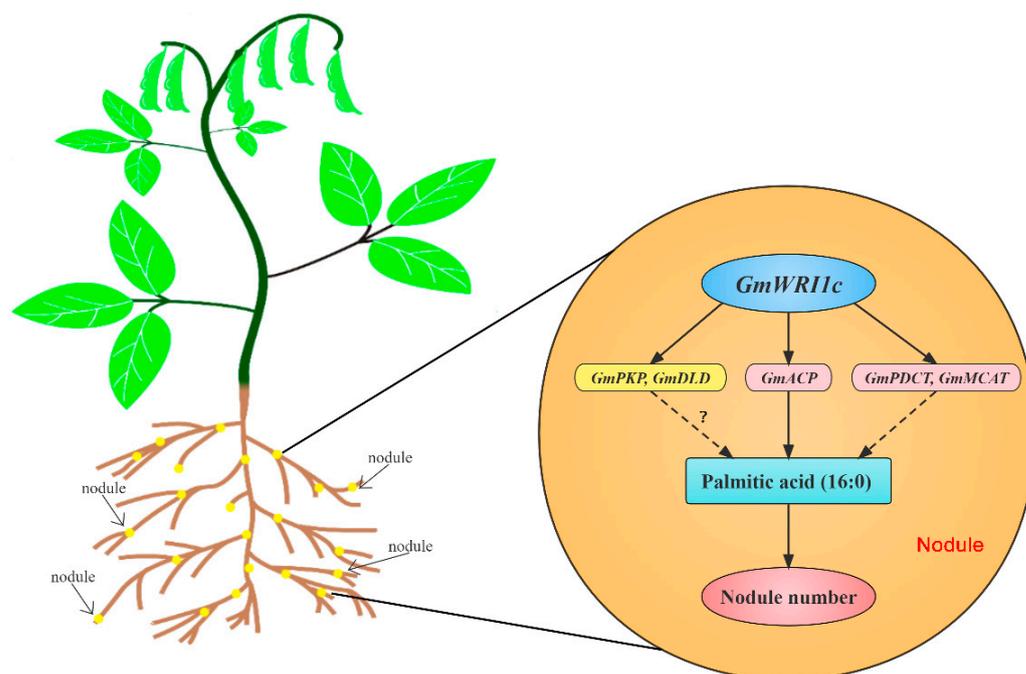


Figure 6. A working model for *GmWRI1c* regulation of soybean nodulation. ?—The mechanism of *GmPKP* and *GmDLD* increase palmitic acid is unclear.

4. Materials and Methods

4.1. Plant Materials and Growth Conditions

In total, 744 soybean germplasm resources were grown in Hefei, China. The experiment used a random block design, with 2 m long rows and 0.4 m between-row spacing. The population consisted of 100 wild soybeans, 441 landraces, and 203 elite lines from China and foreign soybean eco-regions (Supplementary Dataset S1). The *Arabidopsis thaliana* *wri1* mutant (SALK_008559C) was used for transformation; all *Arabidopsis* and *Nicotiana benthamiana* (tobacco) plants were grown under standard growth conditions [58]. The transformation recipient, soybean variety Tianlong NO.1, and other soybean materials used for qRT-PCR were grown in a greenhouse under a 16 h/8 h (light/dark) photoperiod at 23/28 °C (night/day).

4.2. Identification and Bioinformatics Analysis of *GmWRI1c*

The latest reference genome of soybean and the hidden Markov model (HMM) profile of AP2 (PF00847) were downloaded from the Phytozome soybean genome and PFAM databases. The *hmmsearch* program in HMMER software [59] was used to search the AP2 HMM profile against the soybean genome, and the reliable results were screened based on an E-value less than or equal to 1×10^{-10} . The soybean-specific AP2 HMM profile was constructed using the *hmmbuild* program and then the soybean reference genome was searched again with an E-value threshold 1×10^{-10} . To avoid missing other members of

the AP2 gene family, the AP2-type protein sequences from soybean reported on the NCBI website were used as query sequences for BLASTP alignment with the E-value threshold of 1×10^{-5} . We combined the genes predicted using the two methods and removed the duplicates. If multiple transcripts corresponding to a single AP2 gene were obtained, only the most reliable one was retained. The protein sequences of the putative genes were submitted to NCBI-CDD3 and PFAM online databases for verification, and the putative genes without AP2 domains were excluded. The genes with an E value greater than or equal to 1×10^{-20} predicted by the Pfam database were also excluded. Finally, members of the soybean AP2 gene family were obtained with high confidence. For multiple alignment analysis, the amino acid sequences of WRI proteins from soybean, Arabidopsis, Medicago sativa, and Brassica napus were downloaded from the NCBI website. The phylogenetic tree was constructed with MEGA7.0 software using the NJ method [60]. The bootstrap tests were conducted with 1000 replicates, and other parameters were set to default.

4.3. Subcellular Localisation Analysis

The coding sequence (CDS) of *GmWRI1c* was amplified from the soybean cultivar Williams 82 using gene-specific primers (Supplementary Table S1). The CDS of *GmWRI1c* was cloned into pCAMBIA3301-GFP in-frame at the N-terminus of GFP using a one-step cloning kit (VAZYME, Nanjing, China). This construct was transferred into *Agrobacterium tumefaciens* GV3101, which was then injected into *Nicotiana benthamiana* leaves. After 36–48 h of incubation, the subcellular localization of *GmWRI1c* was observed in the epidermal cells of inoculated tobacco leaves under a laser scanning confocal microscope (Zeiss LSM880, Carl Zeiss, Jena, Germany) [61].

4.4. Generation of Transgenic Arabidopsis

The CDS of *GmWRI1c* was cloned into the pCAMBIA3301 vector harbouring the CaMV 35S promoter to generate 35S:*GmWRI1c*. The empty vector served as the control. Subsequently, the vectors in the pCAMBIA3301 backbone were transferred into GV3101 and then used to transform Arabidopsis *wri1* mutant plants using the floral-dip method [62].

4.5. RNA Extraction and cDNA Synthesis

Total RNA was isolated from soybean tissues using an RNA prep Pure Plant Kit (TaKaRa, Otsu, Japan), and first-strand cDNA was synthesised using a PrimeScript™ 1st Strand cDNA Synthesis Kit (TaKaRa).

4.6. Quantitative Real-Time PCR Analysis

Quantitative real-time PCR (qRT-PCR) analysis was performed with a CFX96™ Real-time System (BIO-RAD, Hercules, USA) with 10 µL of SYBR Mix (VAZYME, Nanjing, China), 6 µL of ddH₂O, 1 µL of each primer, and 2 µL of cDNA template to a final volume of 20 µL. The implemented reaction procedure was 95 °C for 60 s, followed by 40 cycles of 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s. The relative gene expression levels were calculated using the $2^{-\Delta\Delta CT}$ method [63] by normalization to *GmACTIN* (Glyma.19G147900) [64] or *AtACTIN7* (AT5G09810) in soybean and Arabidopsis, respectively [65]. All the primers used in this experiment are listed in Supplementary Table S1.

4.7. Soybean Hairy Root Transformation and Nodulation Assay

The CDS of *GmWRI1c* was cloned into the pCAMBIA3301 vector harbouring the CaMV 35S promoter to generate 35S:*GmWRI1c*. The 1.6 kb promoter region of *GmWRI1c* was amplified from Tian Long No. 1 and Ji Lin Cai Li Xiang using specific primers (Supplementary Table S1). The promoter of *GmWRI1c* was cloned into the pCAMBIA1301 vector in-frame at the N-terminus of GUS using a one-step cloning kit (VAZYME). Subsequently, the freeze–thaw method was used to transform the pCAMBIA3301 vector containing *GmWRI1c*-CDS or pro-*GmWRI1c* into the soybean cultivar K599 (Tian Long No. 1) to obtain transgenic hairy roots according to methods published previously [66,67]. The

transgenic soybean hairy roots and nodules were identified followed by the statistics for the transgenic hairy roots' nodule numbers, and the RNA of transgenic roots and nodules were finally extracted.

4.8. Determination of Soluble Sugars Content, Total Fatty Acid Content, and Fatty Acid Composition

The soluble sugars content was determined using a plant soluble sugars content test kit, according to the manufacturer's instructions (Jiancheng Bioengineering Institute, Nanjing, China). To determine total fatty acid content and fatty acid composition in soybean seeds, transgenic Arabidopsis seeds, soybean hairy roots and nodules, dry seeds (10 mg) of Arabidopsis, fresh soybean hairy roots (100 mg), and nodules (100 mg) were used to extract fatty acids. The fatty acid fraction from each sample was analysed by gas chromatography (SHIMADZU GC-2010 Plus, Japan) using heptadecanoic acid as an internal standard, according to the method described previously [68].

4.9. Determination of GUS Activity

Transgenic root nodules expressing two haplotypes of *GmWRI1c* and wild-type root nodules were stained with GUS staining buffer [69] containing 1 mM X-gluc (5-bromo-4-chloro-3-indolyl- β -D-glucuronide) (Gold BioTechnology, St. Louis, MO, USA), 100 mM sodium phosphate (pH 7.5), 0.5 mM potassium ferricyanide, 0.5 mM potassium ferrocyanide, 10 mM EDTA, and 0.1% (*v/v*) Triton X-100. The transgenic root nodules were incubated at 37 °C for 4 h and cleared with 70% (*v/v*) ethanol. For the quantitative assay, 100 mg transgenic fresh root nodules expressing each of the two haplotypes was ground in liquid nitrogen, and then GUS activity was detected according to the instructions of the GUS reporter quantitative assay kit (Coolaber, Beijing, China).

4.10. Soybean Oil Content Determination

The seed oil content of soybean grains was determined using an NIR grain quality analyser (DA7200, Perten Instruments, Springfield, IL, USA). The determination of seed oil content was conducted for each sample in triplicate.

4.11. Tassel Analysis and Genotyping

Association analysis was carried out between sites with genetic variations and oil content data using TASSEL 5.0 software. A haplotype analysis was performed to identify sites with significant effects on seed oil content ($p < 0.05$).

4.12. Statistical Analysis

All experimental data were obtained from three or more independent experiments with replicates and were analysed using Student's *t*-test. Significant differences between the two datasets represented 95% confidence limits.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms232213793/s1>.

Author Contributions: H.Z. and X.W. conceived and designed the contents. H.Z., D.Z., W.S., Y.L., W.W., Y.H., J.L. and S.Z. conducted the experiment; H.Z., D.Z. and X.W. wrote and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key Research and Development Program of China (2021YFD1201605), the Key Research and Development Program of Anhui Province, China (202204c06020026), and the Natural Science Foundation of Anhui Province, China (2208085MC61).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data supporting the findings of this study are available within the paper and within its supplementary data published online.

Acknowledgments: We would like to express our special appreciation and thanks to Penghui Li from Anhui Agricultural University for his help revising the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

Cleaved amplified polymorphic sequence	CAPS
Coding sequence	CDS
Haplotype	Hap
Hidden Markov model	HMM
Quantitative real-time PCR	qRT-PCR
Recombinant inbred line	RIL

References

1. Tyug, T.S.; Prasad, K.N.; Ismail, A. Antioxidant capacity, phenolics and isoflavones in soybean by-products. *Food Chem.* **2010**, *123*, 583–589. [\[CrossRef\]](#)
2. Ademiluyi, A.O.; Oboh, G. Soybean phenolic-rich extracts inhibit key-enzymes linked to type 2 diabetes (alpha-amylase and alpha-glucosidase) and hypertension (angiotensin I converting enzyme) in vitro. *Exp. Toxicol. Pathol.* **2013**, *65*, 305–309. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Silva, F.D.; Perrone, D. Characterization and stability of bioactive compounds from soybean meal. *LWT-Food Sci. Technol.* **2015**, *63*, 992–1000. [\[CrossRef\]](#)
4. Li, H.; Peng, Z.Y.; Yang, X.H.; Wang, W.D.; Fu, J.J.; Wang, J.H.; Han, Y.J.; Chai, Y.C.; Guo, T.T.; Yang, N.; et al. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat. Genet.* **2013**, *45*, 43–50. [\[CrossRef\]](#)
5. Simopoulos, A.P.; Faergeman, O.; Bourne, P.G. Action plan for a healthy agriculture, healthy nutrition, healthy people. *World Rev. Nutr. Diet.* **2011**, *102*, 1–5.
6. de Lorgeril, M.; Salen, P. New insights into the health effects of dietary saturated and omega-6 and omega-3 polyunsaturated fatty acids. *BMC Med.* **2012**, *10*, 50. [\[CrossRef\]](#)
7. Cernac, A.; Benning, C. WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis. *Plant J.* **2004**, *40*, 575–585. [\[CrossRef\]](#)
8. Baud, S.; Mendoza, M.S.; To, A.; Harscoet, E.; Lepiniec, L.; Dubreucq, B. WRINKLED1 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during seed maturation in Arabidopsis. *Plant J.* **2007**, *50*, 825–838. [\[CrossRef\]](#)
9. Focks, N.; Benning, C. *wrinkled1*: A novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. *Plant Physiol.* **1998**, *118*, 91–101. [\[CrossRef\]](#)
10. Liu, J.; Hua, W.; Zhan, G.M.; Wei, F.; Wang, X.F.; Liu, G.H.; Wang, H.Z. Increasing seed mass and oil content in transgenic *Arabidopsis* by the overexpression of *wri1*-like gene from *Brassica napus*. *Plant Physiol. Biochem.* **2010**, *48*, 9–15. [\[CrossRef\]](#)
11. Wu, X.L.; Liu, Z.H.; Hu, Z.H.; Huang, R.Z. BnWRI1 coordinates fatty acid biosynthesis and photosynthesis pathways during oil accumulation in rapeseed. *J. Integr. Plant Biol.* **2014**, *56*, 582–593. [\[CrossRef\]](#)
12. Chen, L.; Zheng, Y.H.; Dong, Z.M.; Meng, F.F.; Sun, X.M.; Fan, X.H.; Zhang, Y.F.; Wang, M.L.; Wang, S.M. Soybean (*Glycine max*) WRINKLED1 transcription factor, GmWRI1a, positively regulates seed oil accumulation. *Mol. Genet. Genom.* **2018**, *293*, 401–415. [\[CrossRef\]](#)
13. Zhou, X.F.; Dai, Y.Q.; Wu, H.J.; Zhong, P.Q.; Luo, L.J.; Shang, Y.J.; Tan, P.P.; Peng, F.R.; Tian, Z.X. Characterization of *CiWRI1* from *Carya illinoensis* in Seed Oil Biosynthesis. *Forests* **2020**, *10*, 818. [\[CrossRef\]](#)
14. Lim, A.R.Q.; Kong, Q.; Singh, S.K.; Guo, L.; Yuan, L.; Ma, W. Sunflower WRINKLED1 Plays a Key Role in Transcriptional Regulation of Oil Biosynthesis. *Int. J. Mol. Sci.* **2022**, *23*, 3054. [\[CrossRef\]](#)
15. Wang, Z.K.; Wang, Y.Z.; Shang, P.; Yang, C.; Yang, M.M.; Huang, J.X.; Ren, B.Z.; Zuo, Z.H.; Zhang, Q.Y.; Li, W.B.; et al. Overexpression of Soybean *GmWRI1a* Stably Increases the Seed Oil Content in Soybean. *Int. J. Mol. Sci.* **2022**, *23*, 5084. [\[CrossRef\]](#)
16. Guo, W.; Chen, L.M.; Chen, H.F.; Yang, H.L.; You, Q.B.; Bao, A.L.; Chen, S.L.; Hao, Q.N.; Huang, Y.; Qiu, D.Z.; et al. Overexpression of *GmWRI1b* in soybean stably improves plant architecture and associated yield parameters, and increases total seed oil production under field conditions. *Plant Biotechnol. J.* **2020**, *18*, 1639–1641. [\[CrossRef\]](#)
17. Chen, B.B.; Zhang, G.Y.; Li, P.H.; Yang, J.H.; Guo, L.; Benning, C.; Wang, X.M.; Zhao, J. 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.* **2020**, *18*, 155–171. [\[CrossRef\]](#)
18. Udvardi, M.; Poole, P.S. Transport and Metabolism in Legume-Rhizobia Symbioses. *Annu. Rev. Plant Biol.* **2013**, *64*, 781–805. [\[CrossRef\]](#)
19. Yang, W.C.; Katinakis, P.; Hendriks, P.; Smolders, A.; de Vries, F.; Spee, J.; van Kammen, A.; Bisseling, T.; Franssen, H. Characterization of *GmENOD40*, a gene showing novel patterns of cell-specific expression during soybean nodule development. *Plant J. Cell Mol. Biol.* **1993**, *3*, 573–585. [\[CrossRef\]](#)

20. Murakami, Y.; Miwa, H.; Imaizumi-Anraku, H.; Kouchi, H.; Downie, J.A.; Kawaguchi, M.; Kawasaki, S. Positional cloning identifies *Lotus japonicus* NSP2, a putative transcription factor of the GRAS family, required for *NIN* and *ENOD40* gene expression in nodule initiation. *DNA Res.* **2006**, *13*, 255–265. [[CrossRef](#)]
21. Liu, H.; Zhang, C.; Yang, J.; Yu, N.; Wang, E.T. Hormone modulation of legume-rhizobial symbiosis. *J. Integr. Plant Biol.* **2018**, *60*, 632–648. [[CrossRef](#)] [[PubMed](#)]
22. Ng, J.L.P.; Hassan, S.; Truong, T.T.; Hocart, C.H.; Laffont, C.; Frugier, F.; Mathesius, U. Flavonoids and Auxin Transport Inhibitors Rescue Symbiotic Nodulation in the *Medicago truncatula* Cytokinin Perception Mutant *cre1*. *Plant Cell* **2015**, *27*, 2210–2226. [[CrossRef](#)] [[PubMed](#)]
23. Cai, Z.M.; Wang, Y.N.; Zhu, L.; Tian, Y.P.; Chen, L.; Sun, Z.X.; Ullah, I.; Li, X. GmTIR1/GmAFB3-based auxin perception regulated by miR393 modulates soybean nodulation. *New Phytol.* **2017**, *215*, 672–686. [[CrossRef](#)] [[PubMed](#)]
24. Tan, S.; Sanchez, M.; Laffont, C.; Boivin, S.; Le Signor, C.; Thompson, R.; Frugier, F.; Brault, M. A Cytokinin Signaling Type-B Response Regulator Transcription Factor Acting in Early Nodulation. *Plant Physiol.* **2020**, *183*, 1319–1330. [[CrossRef](#)]
25. Suzuki, A.; Akune, M.; Kogiso, M.; Imagama, Y.; Osuki, K.; Uchiumi, T.; Higashi, S.; Han, S.Y.; Yoshida, S.; Asami, T.; et al. Control of nodule number by the phytohormone abscisic acid in the roots of two leguminous species. *Plant Cell Physiol.* **2004**, *45*, 914–922. [[CrossRef](#)]
26. Ding, Y.L.; Kalo, P.; Yendrek, C.; Sun, J.H.; Liang, Y.; Marsh, J.F.; Harris, J.M.; Oldroyd, G.E.D. Abscisic Acid Coordinates Nod Factor and Cytokinin Signaling during the Regulation of Nodulation in *Medicago truncatula*. *Plant Cell* **2008**, *20*, 2681–2695. [[CrossRef](#)]
27. Taylor, B.N.; Menge, D.N.L. Light regulates tropical symbiotic nitrogen fixation more strongly than soil nitrogen. *Nat. Plants* **2018**, *4*, 655–661. [[CrossRef](#)]
28. Wang, X.D.; Chen, K.; Zhou, M.M.; Gao, Y.K.; Huang, H.M.; Liu, C.; Fan, Y.Y.; Fan, Z.H.; Wang, Y.N.; Li, X. GmNAC181 promotes symbiotic nodulation and salt tolerance of nodulation by directly regulating GmNINA expression in soybean. *New Phytol.* **2022**, *236*, 656–670. [[CrossRef](#)]
29. Guan, R.X.; Qu, Y.; Guo, Y.; Yu, L.L.; Liu, Y.; Jiang, J.H.; Chen, J.G.; Ren, Y.L.; Liu, G.Y.; Tian, L.; et al. Salinity tolerance in soybean is modulated by natural variation in *GmSALT3*. *Plant J.* **2014**, *80*, 937–950. [[CrossRef](#)]
30. Wang, J.; Toth, K.; Tanaka, K.; Nguyen, C.T.; Yan, Z.; Brechenmacher, L.; Dahmen, J.; Chen, M.J.; Thelen, J.J.; Qiu, L.J.; et al. A Soybean Acyl Carrier Protein, GmACP, Is Important for Root Nodule Symbiosis. *Mol. Plant-Microbe Interact.* **2014**, *27*, 415–423. [[CrossRef](#)]
31. Zhang, G.Y.; Ahmad, M.Z.; Chen, B.B.; Manan, S.; Zhang, Y.L.; Jin, H.N.; Wang, X.M.; Zhao, J. 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.* **2020**, *103*, 1351–1371. [[CrossRef](#)]
32. Lee, S.H.; Bailey, M.A.; Mian, M.A.; Carter, T.E., Jr.; Shipe, E.R.; Ashley, D.A.; Parrott, W.A.; Hussey, R.S.; Boerma, H.R. RFLP loci associated with soybean seed protein and oil content across populations and locations. *TAG. Theor. Appl. Genet. Theor. Und Angew. Genet.* **1996**, *93*, 649–657. [[CrossRef](#)]
33. Brummer, E.C.; Graef, G.L.; Orf, J.H.; Wilcox, J.R.; Shoemaker, R.C. Mapping QTL for seed protein and oil content in eight soybean populations. *Crop Sci.* **1997**, *37*, 370–378. [[CrossRef](#)]
34. Reinprecht, Y.; Poysa, V.W.; Yu, K.F.; Rajcan, I.; Ablett, G.R.; Pauls, K.P. Seed and agronomic QTL in low linolenic acid, lipoxygenase-free soybean (*Glycine max* (L.) Merrill) germplasm. *Genome* **2006**, *49*, 1510–1527. [[CrossRef](#)]
35. Qi, Z.M.; Wu, Q.; Han, X.; Sun, Y.N.; Du, X.Y.; Liu, C.Y.; Jiang, H.W.; Hu, G.H.; Chen, Q.S. Soybean oil content QTL mapping and integrating with meta-analysis method for mining genes. *Euphytica* **2011**, *179*, 499–514. [[CrossRef](#)]
36. Oldroyd, G.E.D.; Downie, J.M. Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annu. Rev. Plant Biol.* **2008**, *59*, 519–546. [[CrossRef](#)]
37. Minami, E.; Kouchi, H.; Cohn, J.R.; Ogawa, T.; Stacey, G. Expression of the early nodulin, ENOD40, in soybean roots in response to various lipo-chitin signal molecules. *Plant J. Cell Mol. Biol.* **1996**, *10*, 23–32. [[CrossRef](#)]
38. Liu, M.; Soyano, T.; Yano, K.; Hayashi, M.; Kawaguchi, M. ERN1 and CYCLOPS coordinately activate NIN signaling to promote infection thread formation in *Lotus japonicus*. *J. Plant Res.* **2019**, *132*, 641–653. [[CrossRef](#)]
39. Charon; Sousa; Crespi; Kondorosi, Alteration of *enod40* expression modifies *medicago truncatula* root nodule development induced by *sinorhizobium meliloti*. *Plant Cell* **1999**, *11*, 1953–1966.
40. Xu, H.Y.; Li, Y.J.; Zhang, K.F.; Li, M.J.; Fu, S.Y.; Tian, Y.Z.; Qin, T.F.; Li, X.X.; Zhong, Y.J.; Liao, H. *miR169c-NFYA-C-ENOD40* modulates nitrogen inhibitory effects in soybean nodulation. *New Phytol.* **2021**, *229*, 3377–3392. [[CrossRef](#)]
41. Kouchi, H.; Kinoshita, E.; Ridge, R.W.; Kumagai, H. RNAi knockdown of *ENOD40s* leads to significant suppression of nodule formation in *Lotus japonicus*. *Plant Cell Physiol.* **2005**, *46*, S91.
42. Wang, L.X.; Sun, Z.X.; Su, C.; Wang, Y.L.; Yan, Q.Q.; Chen, J.H.; Ott, T.; Li, X. A GmNINA-miR172c-NNC1 Regulatory Network Coordinates the Nodulation and Autoregulation of Nodulation Pathways in Soybean. *Mol. Plant* **2019**, *12*, 1211–1226. [[CrossRef](#)] [[PubMed](#)]
43. Tang, T.; Du, C.; Song, H.; Aziz, U.; Wang, L.L.; Zhao, C.Z.; Zhang, M. Genome-wide analysis reveals the evolution and structural features of WRINKLED1 in plants. *Mol. Genet. Genom.* **2019**, *294*, 329–341. [[CrossRef](#)] [[PubMed](#)]
44. Behera, J.R.; Rahman, M.M.; Bhatia, S.; Shockey, J.; Kilaru, A. Functional and Predictive Structural Characterization of WRINKLED2, A Unique Oil Biosynthesis Regulator in Avocado. *Front. Plant Sci.* **2021**, *12*, 648494. [[CrossRef](#)] [[PubMed](#)]

45. Zakim, D.; Herman, R.H. Regulation of fatty acid synthesis. *Am. J. Clin. Nutr.* **1969**, *22*, 200–213. [[CrossRef](#)]
46. Liu, Q.; Wu, M.; Zhang, B.L.; Shrestha, P.; Petrie, J.; Green, A.G.; Singh, S.P. Genetic enhancement of palmitic acid accumulation in cotton seed oil through RNAi down-regulation of *ghKAS2* encoding beta-ketoacyl-ACP synthase II (KASII). *Plant Biotechnol. J.* **2017**, *15*, 132–143. [[CrossRef](#)]
47. Sedivy, E.J.; Wu, F.Q.; Hanzawa, Y. Soybean domestication: The origin, genetic architecture and molecular bases. *New Phytol.* **2017**, *214*, 539–553. [[CrossRef](#)]
48. Wang, S.D.; Liu, S.L.; Wang, J.; Yokosho, K.; Zhou, B.; Yu, Y.C.; Liu, Z.; Frommer, W.B.; Ma, J.F.; Chen, L.Q.; et al. Simultaneous changes in seed size, oil content and protein content driven by selection of *SWEET* homologues during soybean domestication. *Natl. Sci. Rev.* **2020**, *7*, 1776–1786. [[CrossRef](#)]
49. Duan, Z.B.A.; Zhang, M.; Zhang, Z.F.; Liang, S.; Fan, L.; Yang, X.; Yuan, Y.Q.; Pan, Y.; Zhou, G.A.; Liu, S.L.; et al. Natural allelic variation of *GmSTO5* controlling seed size and quality in soybean. *Plant Biotechnol. J.* **2022**, *20*, 1807–1818. [[CrossRef](#)]
50. Goettel, W.; Zhang, H.Y.; Li, Y.; Qiao, Z.Z.; Jiang, H.; Hou, D.Y.; Song, Q.J.; Pantalone, V.R.; Song, B.H.; Yu, D.Y.; et al. *POWR1* is a domestication gene pleiotropically regulating seed quality and yield in soybean. *Nat. Commun.* **2022**, *13*, 3051. [[CrossRef](#)]
51. Zhu, W.W.; Yang, C.; Yong, B.; Wang, Y.; Li, B.B.; Gu, Y.Z.; Wei, S.M.; An, Z.H.; Sun, W.K.; Qiu, L.J.; et al. An enhancing effect attributed to a nonsynonymous mutation in *SOYBEAN SEED SIZE 1*, a *SPINDLY*-like gene, is exploited in soybean domestication and improvement. *New Phytol.* **2022**, *236*, 1375–1392. [[CrossRef](#)]
52. Li, Q.T.; Lu, X.; Song, Q.X.; Chen, H.W.; Wei, W.; Tao, J.J.; Bian, X.H.; Shen, M.; Ma, B.A.; Zhang, W.K.; et al. Selection for a Zinc-Finger Protein Contributes to Seed Oil Increase during Soybean Domestication. *Plant Physiol.* **2017**, *173*, 2208–2224. [[CrossRef](#)]
53. Lu, L.; Wei, W.; Li, Q.T.; Bian, X.H.; Lu, X.; Hu, Y.; Cheng, T.; Wang, Z.Y.; Jin, M.; Tao, J.J.; et al. A transcriptional regulatory module controls lipid accumulation in soybean. *New Phytol.* **2021**, *231*, 661–678. [[CrossRef](#)]
54. Li, J.; Zhang, Y.H.; Ma, R.R.; Huang, W.X.; Hou, J.J.; Fang, C.; Wang, L.S.; Yuan, Z.H.; Sun, Q.; Dong, X.H.; et al. Identification of *ST1* reveals a selection involving hitchhiking of seed morphology and oil content during soybean domestication. *Plant Biotechnol. J.* **2022**, *20*, 1110–1121. [[CrossRef](#)]
55. Lu, X.; Li, Q.T.; Xiong, Q.; Li, W.; Bi, Y.D.; Lai, Y.C.; Liu, X.L.; Man, W.Q.; Zhang, W.K.; Ma, B.; et al. The transcriptomic signature of developing soybean seeds reveals the genetic basis of seed trait adaptation during domestication. *Plant J.* **2016**, *86*, 530–544. [[CrossRef](#)]
56. Zhao, L.; Li, M.M.; Xu, C.J.; Yang, X.; Li, D.M.; Zhao, X.; Wang, K.; Li, Y.H.; Zhang, X.M.; Liu, L.X.; et al. Natural variation in *GmGBP1* promoter affects photoperiod control of flowering time and maturity in soybean. *Plant J.* **2018**, *96*, 147–162. [[CrossRef](#)]
57. Jin, T.; Sun, Y.Y.; Shan, Z.; He, J.B.; Wang, N.; Gai, J.Y.; Li, Y. Natural variation in the promoter of *GsERD15B* affects salt tolerance in soybean. *Plant Biotechnol. J.* **2021**, *19*, 1155–1169. [[CrossRef](#)]
58. Li, D.; Jin, C.Y.; Duan, S.W.; Zhu, Y.N.; Qi, S.H.; Liu, K.G.; Gao, C.H.; Ma, H.L.; Zhang, M.; Liao, Y.C.; et al. MYB89 Transcription Factor Represses Seed Oil Accumulation. *Plant Physiol.* **2017**, *173*, 1211–1225. [[CrossRef](#)]
59. Mistry, J.; Finn, R.D.; Eddy, S.R.; Bateman, A.; Punta, M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* **2013**, *41*, 12. [[CrossRef](#)]
60. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)]
61. Li, P.H.; Dong, Q.; Ge, S.J.; He, X.Z.; Verdier, J.; Li, D.Q.; Zhao, J. Metabolic engineering of proanthocyanidin production by repressing the isoflavone pathways and redirecting anthocyanidin precursor flux in legume. *Plant Biotechnol. J.* **2016**, *14*, 1604–1618. [[CrossRef](#)] [[PubMed](#)]
62. Clough, S.J.; Bent, A.F. Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J. Cell Mol. Biol.* **1998**, *16*, 735–743. [[CrossRef](#)] [[PubMed](#)]
63. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)] [[PubMed](#)]
64. Ahmad, M.Z.; Rehman, N.U.; Yu, S.W.; Zhou, Y.Z.; ul Haq, B.; Wang, J.J.; Li, P.H.; Zeng, Z.X.; Zhao, J. GmMAX2-D14 and -KAI interaction-mediated SL and KAR signaling play essential roles in soybean root nodulation. *Plant J.* **2020**, *101*, 334–351. [[CrossRef](#)]
65. Zhang, D.; Yang, H.F.; Wang, X.C.; Qiu, Y.J.; Tian, L.H.; Qi, X.Q.; Qu, L. Cytochrome P450 family member CYP96B5 hydroxylates alkanes to primary alcohols and is involved in rice leaf cuticular wax synthesis. *New Phytol.* **2020**, *225*, 2094–2107. [[CrossRef](#)]
66. Kereszt, A.; Li, D.X.; Indrasumunar, A.; Nguyen, C.D.T.; Nontachaiyapoom, S.; Kinkema, M.; Gresshoff, P.M. *Agrobacterium rhizogenes*-mediated transformation of soybean to study root biology. *Nat. Protoc.* **2007**, *2*, 948–952. [[CrossRef](#)]
67. Ahmad, M.Z.; Zhang, Y.; Zeng, X.; Li, P.; Wang, X.; Benedito, V.A.; Zhao, J. Isoflavone malonyl-CoA acyltransferase *GmMaT2* is involved in nodulation of soybean by modifying synthesis and secretion of isoflavones. *J. Exp. Bot.* **2021**, *72*, 1349–1369. [[CrossRef](#)]
68. Liu, Y.F.; Li, Q.T.; Lu, X.; Song, Q.X.; Lam, S.M.; Zhang, W.K.; Ma, B.; Lin, Q.; Man, W.Q.; Du, W.G.; et al. Soybean GmMYB73 promotes lipid accumulation in transgenic plants. *BMC Plant Biol.* **2014**, *14*, 73. [[CrossRef](#)]
69. Jefferson, R.A.; Kavanagh, T.A.; Bevan, M.W. GUS fusions: Beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **1987**, *6*, 3901–3907. [[CrossRef](#)]