CRISPR/Cas9-Mediated Targeted Mutagenesis of Wild Soybean (*Glycine soja*) Hairy Roots Altered the Transcription Profile of the Mutant

Fengjuan Niu[1], Qiyan Jiang[1], Rui Cheng[2], Xianjun Sun[1], Zheng Hu[1], Lixia Wang[1] & Hui Zhang[1]

[1] Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, China

[2] Tianjin Institute of Agricultural Quality Standard and Testing Technology, Tianjin Academy of Agricultural Sciences, Tianjin, China

Correspondence: Qiyan Jiang, Institute of Crop Science, Chinese Academy of Agricultural Sciences, 80 Xueyuan Sourth Road, Haidian District, Beijing, China. Tel: 0086-10-6218-9045. E-mail: jiangqiyan@caas.cn

Zhang Hui, Institute of Crop Science, Chinese Academy of Agricultural Sciences, 80 Xueyuan Sourth Road, Haidian District, Beijing, China. Tel: 0086-10-6218-6649. E-mail: huizhang_06@caas.cn

Abstract

Clustered Regularly Interspaced Short Palindromic Repeat/CRISPR-associated protein 9 (CRISPR/Cas9) system has been regularly applied for genome editing and gene function identification in wild soybean (*Glycine max*) cultivars. However, till date no studies have demonstrated successful mutagenesis in wild soybean (*Glycine soja*) which is the ancestor of *Glycine max* and rich in stress tolerance genes. In the current study, we report the successful creation of mutations in the loci encoding plasma membrane Na^+/H^+ antiporter (*SOS1*) and nonselective cation channels (*NSCC*) in wild soybean hairy roots using the CRISPR/Cas9 system. Two genes, *GsSOS1* and *GsNSCC*, were mutagenized with frequencies of 28.5% and 39.9%, respectively. Biallelic mutations in *GsSOS1* were detected in transgenic hairy roots. *GsSOS1* mutants exhibited altered Na^+/K^+ ratios in the roots under both control and salt-treated conditions. However, no significant effects of *GsNSCC* mutation on Na^+/K^+ ratios were observed. RNA-Seq analysis revealed that both *GsSOS1* and *GsNSCC* mutation altered the transcription profiles in mutant roots. Many differentially expressed gene sets that are associated with various cellular functions were identified. Our results demonstrated that CRISPR/Cas9 systems as powerful tools for wild soybean genome editing and would significantly advance the gene mining and functional identification in wild soybean.

Keywords: CRISPR/Cas9, *Glycine soja*, transcription profile, Na^+ transport

1. Introduction

It is well known that soybean (*Glycine max*) is one of the most important crops of the world (Lam et al., 2010; Smil, 2000). Wild soybean (*Glycine soja*) is the known ancestor of *Glycine max*, possessing much greater adaptability to a variety of environmental stresses and is believed to be rich in stress tolerance genes. It has, therefore, been suggested as a potential source of germplasm to improve the agronomic traits of cultivated soybean (Ge et al., 2011; Wen et al., 2009). However, so far, only a few resistance genes have been successfully mined from wild soybean. In recent years, significant progress has been made in the utilization of CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeat/CRISPR-associated protein 9) system for genome editing and functional studies in various crops.

Multiple reports have described the generation of mutations using CRISPR/Cas9 in soybean (*Glycine max*), including somatic mutation and whole-plant soybean mutagenesis. In earlier years, the hairy soybean root system was used to detect the target gene editing efficiency with different sgRNA promoters (Sun et al., 2015; Du et al., 2016) or different target loci (Jacobs et al., 2015) and of endogenous and exogenous genes (Cai., 2015), which significantly improved the optimization of CRISPR/Cas9 system in soybean functional genomic research. Earlier, gene function investigations with CRISPR/Cas9 system combined with the hairy root transformation (due to its high efficiency and time-saving) was successfully implemented. For example, overexpression of *GmMYB118* in soybean hairy roots, significantly improved the plants' drought and salt tolerance when compared with

CRISPR-only transformed plants (Du et al., 2018). In recent years, before commencing the whole-plant soybean mutagenesis, the CRISPR/Cas9 constructs functionality are assessed by transient expression in the hairy roots (Do et al., 2018; Curtin et al., 2018), thus strengthening the realization of hairy roots as an excellent transgenic model system for soybean transformation/mutagenesis. Based upon above-mentioned information and other reports in *Glycine max*, we investigated for the existence of studies wherein successful transformation and mutagenesis in the wild soybean (*Glycine soja*) using root hair was reported. To our surprise, till date, no studies have reported successful utilization of the CRISPR/Cas9 for mutagenesis in wild soybean (*Glycine soja*).

Hence, the first question we want to address in this study is to see if we could perform mutagenesis in *Glycine soja* hairy roots in the same way as *Glycine max*. For this purpose, we selected two target genes *GsSOS1* and *GsNSCC*. Salt overly sensitive1 (*SOS1*), encodes a plasma membrane Na^+/H^+ antiporter, which exports Na^+ to the apoplast (Shi et al., 2000, 2002). Nonselective cation channels (*NSCC*) in the plasma membrane of higher plants form a large and diverse group of plant cation channels which are the major pathway for Na^+ influx into root cells (Amtamnn & Sanders, 1999; Tyerman & Skerrett, 1999; White, 1999). Because the unidirectional influx of Na^+ is rapid and greatly exceeds the rate of accumulation, efficient efflux of Na^+ to apoplast must function to minimize net uptake and achieve ion homeostasis in the plant cell (Tester & Davenport, 2003). Hence, *SOS1* is a crucial component of plants in the defence against sodium ions that have entered the cytoplasm. Mutants of Arabidopsis lacking *SOS1* are highly salt-sensitive and does not possess an effective Na^+ extrusion mechanism (Shi et al., 2002; Qiu et al., 2002). In addition to Na^+ efflux, some reports also highlighted the critical roles of *SOS1* in supporting vacuolar morphology, ion homeostasis, and membrane trafficking, thus mediating salt tolerance of root cells during the early stages of salt stress (Oh et al., 2010a). Further, *SOS1* mutants also exhibit altered plants pathogen responses and circadian rhythm (Oh et al., 2010b). Therefore, it is highly likely that the complex structure of the large *SOS1* protein is involved in more than one function. Like *SOS1*, *NSCC* also has multiple functions. Apart from Na^+ influx, *NSCC* was shown to be involved in the uptake of K^+, NH_4^+, Ca^{2+}, Mg^{2+}, micronutrients and trace elements, in ROS-, amino acid, purine- and cyclic nucleotide induced signaling, growth and development (Demidchik and Maathuis 2007). Both *SOS1* and *NSCC* are related to Na^+ transportation and have multiple functions. The second question we want to address is the alterations in the transcriptome as a result of the mutagenesis of *GsSOS1* and *GsNSCC*. In this study, we tried to perform CRISPR/Cas9 system in wild soybean to generate the mutagenesis of *GsSOS1* and *GsNSCC* genes and detect the effects of gene mutation on phenotype and transcriptome.

2. Material and Methods

2.1 Plant Material

The wild soybean seeds of *G. soja* (ZYD1239) were germinated in a growth chamber maintained at 25 °C with 16/8hrs of light/dark cycle. One-week old seedlings were selected for *A. rhizogenes*-mediated transformation (Sun et al., 2015; Ming et al., 2018). After the initiation of hairy root formation from the infection site, the hairy roots were covered with vermiculite to maintain high humidity. After 30 days, each hairy root was cut into two parts. One part was used for DNA extraction (for mutation detection), and the other part was used for total RNA extraction for RNA-seq analysis. The seedling, which does not exhibit any mutation, their hairy and the primary roots were excised and used as controls for those with the mutation. Both seedlings (with and without mutations) were subsequently treated for six hours with either water or 250 mM NaCl. Post-treatment, the roots, stems and leaves were collected to measure Na^+, K^+ concentrations. Further, plants with all hairy roots were also directly subjected to control and salt stress. All the roots (with gene mutation), stem, leaves were collected from each plant to determine the levels of Na^+, K^+ ions.

2.2 Vector Construction

A codon-optimized *Cas9* gene with an NLS was obtained from Professor Qu (Qu, State Key Laboratory for Protein and Plant Gene Research, Peking-Tsinghua Center for Life Sciences, College of Life Sciences, Peking University) and used for generation of *pCambia3301-Cas9* and *pUC57-GmU6-sgRNA* vectors as described earlier in Sun et al. (2015).

The pUC57-GmU6-sgRNA plasmid was digested using *Bsa*I (NEB, Massachusetts, USA) and purified with a TIANquick Midi purification kit (Tiangen, Beijing, China). Two target gene oligonucleotides (Table A1) were annealed to form sgRNA and were subsequently ligated into the *pUC57-GmU6-sgRNA* vector. Both vectors (*pCambia3301-Cas9* and *pUC57-GmU6-sgRNA*) were digested using *Eco*RI and *Hind*III, gel purified (TIANgel Midi purification kit, Tiangen, China) and ligated overnight using T4 DNA ligase (Fermentas) to obtain *pCas9-GmU6-sgRNA* vectors for different target genes.

2.3 Detection of Mutations in Target Genes

Genomic DNA was extracted using a DNAsecure Plant Kit (Tiangen, China) according to the manufacturer's instructions. The target genes were PCR amplified using with gene-specific primers (Table A2) using genomic DNA of each hairy root as a template. The PCR products were digested for 30 min with Bgl Ⅱ and $Apal$I, respectively. The undigested bands were purified and sequenced to detect gene mutation(s).

2.4 Determination of Ion Concentration

The harvested seedlings were separated into roots, stems and leaves and were initially oven-dried for 30 min at 105 ^{o}C and then at 65 ^{o}C until a constant weight is recorded. The fully dried tissues were weighted and grounded to a fine powder. The powdered material was digested with nitric acid and the total volume was made to 15 ml using ddH$_2$O (double distilled water). The solution without tissues samples was used as blank. K$^+$ and Na$^+$ concentrations were measured by ICP-OES spectrometer (ICAP6300, Thermo Fisher Scientific, US). The experiments were repeated three times. The ion content was determined using the following formula:

Ion content (mg/kg) = (The sample concentration – Blank control concentration) × Volume/Dry weight.

2.5 RNA-Seq Profiling Experiment

The hairy roots from the three sample groups (*CK*, *NSCC* and *SOS1*) were collected for RNA extraction using the RNAprep Pure Plant Kit (TIANGEN, China) according to the manufacturer's instructions. The qualified RNA samples were used for RNA-seq analysis. The strand-specific cDNA library was constructed as described before (Jiang et al., 2017). Suitable enriched fragments were sequenced using a HiSeq 2500 instrument (Illumina, USA).

Raw reads obtained from sequencing were filtered to exclude reads containing adapters, reads with more than 10% unknown nucleotides and low-quality reads containing more than 50% of bases with a quality score of ≤ 5 to obtain clean reads. The cleaned reads were mapped to the *G. max* references sequence using TopHat2 software with a tolerance of two mismatches (Kim et al., 2013). The soybean (*Glycine max* (Linn.) Merr.) genomic sequence available from the database (ftp://ftp.ensemblgenomes.org/pub/release-35/plants/fasta/glycine_max/ dna/Glycinemax.V1.0.dna. toplevel.fa.gz) was used for mapping of the reads. The basic sequencing results and assembly information are summarized in Table A3. The DEseq package was used to estimate differential gene expression after standardization of reading count (Anders and Huber 2012). The differentially expressed genes were considered to be significant at False Discovery Rate (FDR) < 0.01 and absolute fold change ≥ 2. GO annotation was carried out using Blast2GO software (Young et al., 2010). KEGG pathway annotation was performed using Path_finder software against the KEGG database (Kanehisa et al., 2008). After GO annotation of every unigene, WEGO was used to assign GO functions to all unigenes and to determine the distribution of gene functions of the species.

2.6 RT-PCR Assays

Sequencing results were validated by qRT-PCR analysis of a randomly selected set of genes. The total RNA was DNaseI-treated and used for cDNA synthesis. First-strand synthesis was carried out using Superscript III reverse transcriptase module (Invitrogen, USA) as the manufacturer's protocol. The comparative ΔΔCT method was used for relative quantitation of expression of genes (Schmittgen and Livak 2008). The *Gsr18SRNA* gene was amplified as an internal control. Primers used in this study were listed in Table A2.

3. Results

3.1 Targeted Mutagenesis in Soybean Hairy Roots

The CRISPR/Cas9-mediated genome-editing tool was utilized to edit the endogenous gene *GsSOS1* and *GsNSCC* in wild hairy soybean roots. PCR/RE assays were conducted to detect mutations in *GsSOS1* and *GsNSCC* target regions. The PCR-RE assay showed that gene mutations were induced (Figure1) and mutagenesis efficiencies for *GsSOS1* and *GsNSCC* were 28.5% and 39.9%, respectively (Table 1). We detected the biallelic and monoallelic mutations in *GsSOS1* and only the monoallelic mutations in *GsNSCC* (Figure 1). The undigested bands from the PCR-RE assay were cloned and sequenced to confirm the mutation. Most mutations in the two genes predominantly were multiple-nucleotide deletion. However, some rare nucleotide substitutions and insertions were also detected (Figure A1). Sequencing of gene clones from independent mutant roots revealed a variety of mutations for each root, suggesting that the introduced CRISPR-Cas9 system continued to modify the genes during hairy root development (Figure A1).

Table 1. Gene mutation in 2 target genes

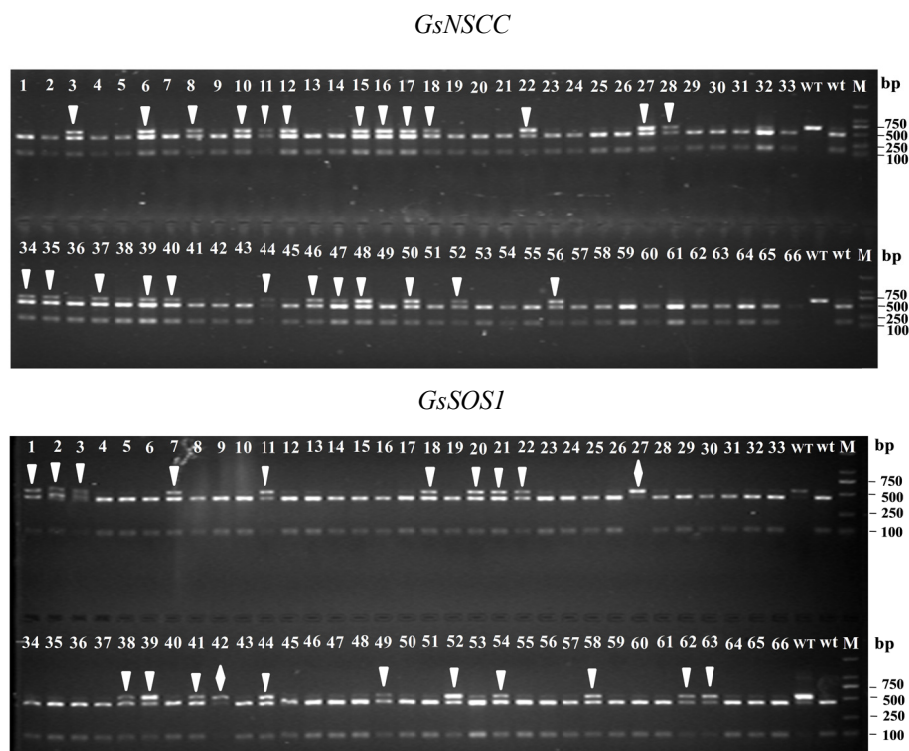| Target gene	Number of hair roots	Number of monoallelic mutation	Number of biallelic mutation	Mutation rate (%)
GsNSCC	168	67	0	39.9
GsSOS1	144	37	4	28.5

GsNSCC

GsSOS1

Figure 1. Detection of mutants using the PCR-restriction enzyme (PCR-RE) assay

Note. Lanes 1-66: Digested PCR products amplified from the independent hairy root samples indicating different mutation types; Rhombus: biallelic mutation, Triangle: monoallelic mutation, WT and wt: undigested and digested DNA of PCR products amplified from wild type controls respectively. M: D2000 marker.

3.2 Effects of Gene Mutation on Na+/K+ in Different Tissues

Na^+, K^+ concentration and Na^+/K^+ ratios were determined in all the hairy roots of each plant with a gene mutation. For this analysis, apart from hairy roots, stems and leaves of these plants were also included. Our results showed that there were no significant differences in Na^+/K^+ ratios between mutants and the wild type in roots, stems and leaves under normal and salt treatment conditions (Figure A2).

The root mutation rates of each plant broadly differ from zero to 100% (Table A4). Hence, hairy roots with no mutation were excised out. The salt treatment was done for six hours and the mutant roots, as well as the stems and leaves were collected to measure Na^+, K^+ concentration and Na^+/K^+ ratio. Under control condition, Na^+/K^+ ratios in *GsSOS1* mutant roots and stems were significantly higher than that in wild type. After salt treatment, a significant difference in Na^+/K^+ ratios was observed between the roots of *GsSOS1* mutant and wild type. However, no significant effects of *GsNSCC* mutation on Na^+/K^+ ratios in different tissues under normal and salt treatment conditions were detected (Figure 2).
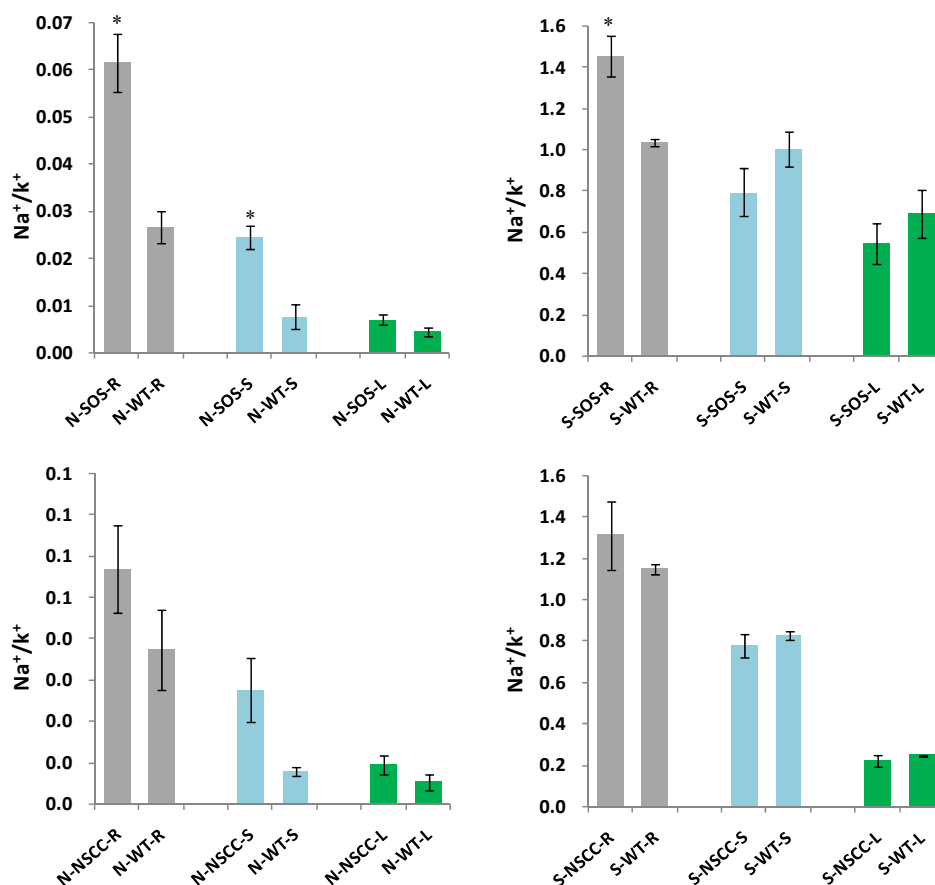
Figure 2. Na$^+$/K$^+$ ratio in different tissues of *GsSOS1* and *GsNSCC* mutants and wild type plants

Note. After detection of mutation in the roots (and subsequently removing the roots without mutation), leaves and stems tissues were collected to determine Na$^+$ and K$^+$ levels. S and N: Salt and Normal; *SOS*/*NSCC* and WT: *SOS*/*NSCC* mutation and Wild type; R, S and L: Root, Stem and Leaf. * indicate a significant difference at *P* < 0.05 compared with the corresponding controls.

3.3 Transcriptome Analysis of GsSOS1 and GsNSCC Mutant Roots

3.3.1 Effects of Gene Mutation on the Expression of target and Non-target Genes

Nine mRNA libraries of *GsmSOS1* (*SOS*), *GsNSCC* (*NSCC*) mutants and wild type (CK) soybean roots, each with three replicates, generated a total of 140.39 Gb raw reads via the Illumina/Solexa sequencing platform and obtained 138.42 Gb of clean reads after analysis. Mapping of reads indicated, ~81.03 to 83.84% clean reads, ~79.05 to 81.95% unique reads and one perfectly matching locus in the soybean genome (Table A3).

GsSOS1 and *GsNSCC* mutation affected the expression of many other genes in wild soybean roots. There were 571 up-regulated and 1246 genes down-regulated in *GsNSCC* mutant roots when compared with the wild type. In *GsSOS1* mutant roots, 908 and 1031 genes were up- and down-regulated were observed (Figure 3A, Supplementary File 1). The expression of the 887-common set of genes was significantly affected either of the mutations (Figure 3B). These results indicate a) both *GsSOS1* and *GsNSCC* are associated with common pathways or b) be regulating genes with a similar pathway. To verify the expression profiles of differentially expressed genes, about forty genes were randomly selected, and their expressions were validated using qRT-PCR. Among these, thirty-five genes expression profiles matched with that observed in RNAseq data (Figure 4).
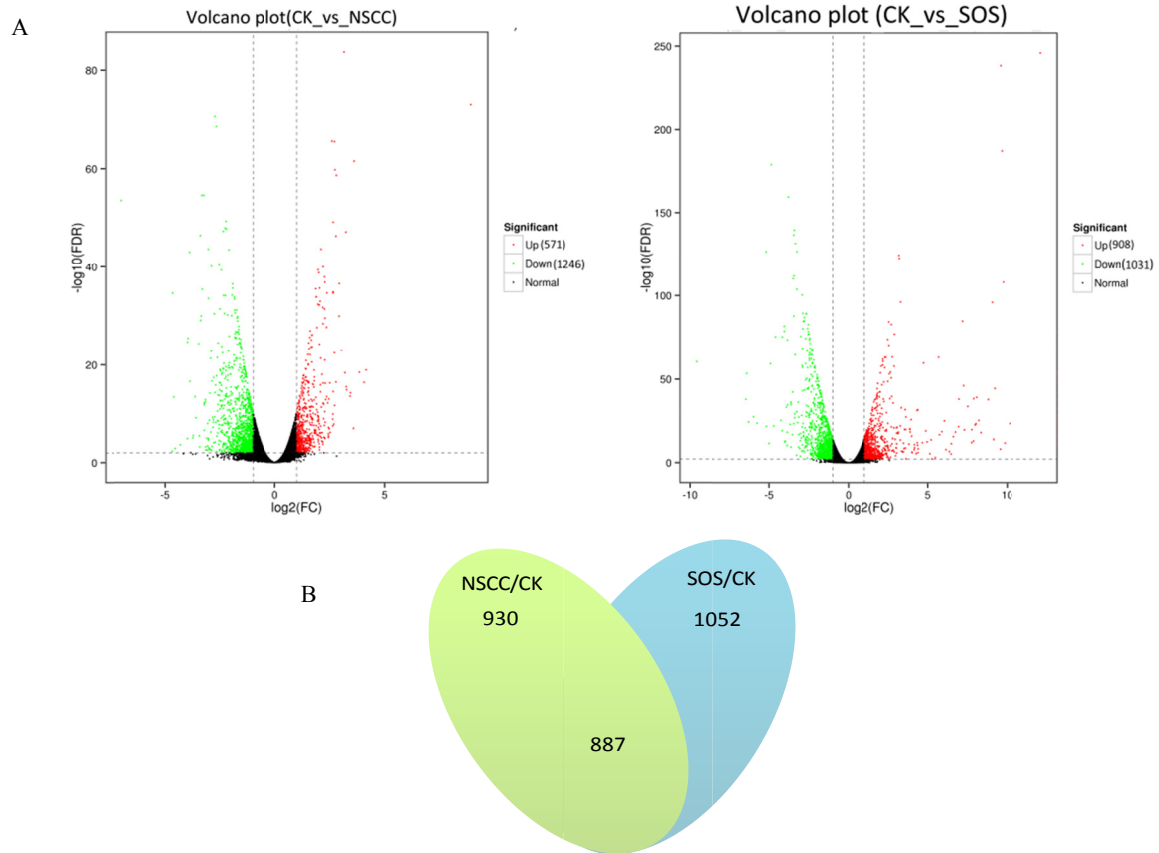
Figure 3. Differentially expressed genes (DEGs) in the roots of wild type and *GsSOS1*/*GsNSCC* mutant plants

Note. A: Volcano plot of all expressed genes. The red dots represent up-regulated genes with log2 (FC) > 1 and FDR < 0.01, the green dots represent down-regulated genes with log2 (FC) < -1 and FDR < 0.01, and the black dots represent genes with no significant difference.

B: Overlap of differentially expressed genes affected by the *GsSOS1* and *GsNSCC* mutation. CK, wild-type control; *SOS*, *GsSOS1* mutation; *NSCC*, *GsNSCC* mutation. FDR: False Discovery Rate. FC: Fold Change.
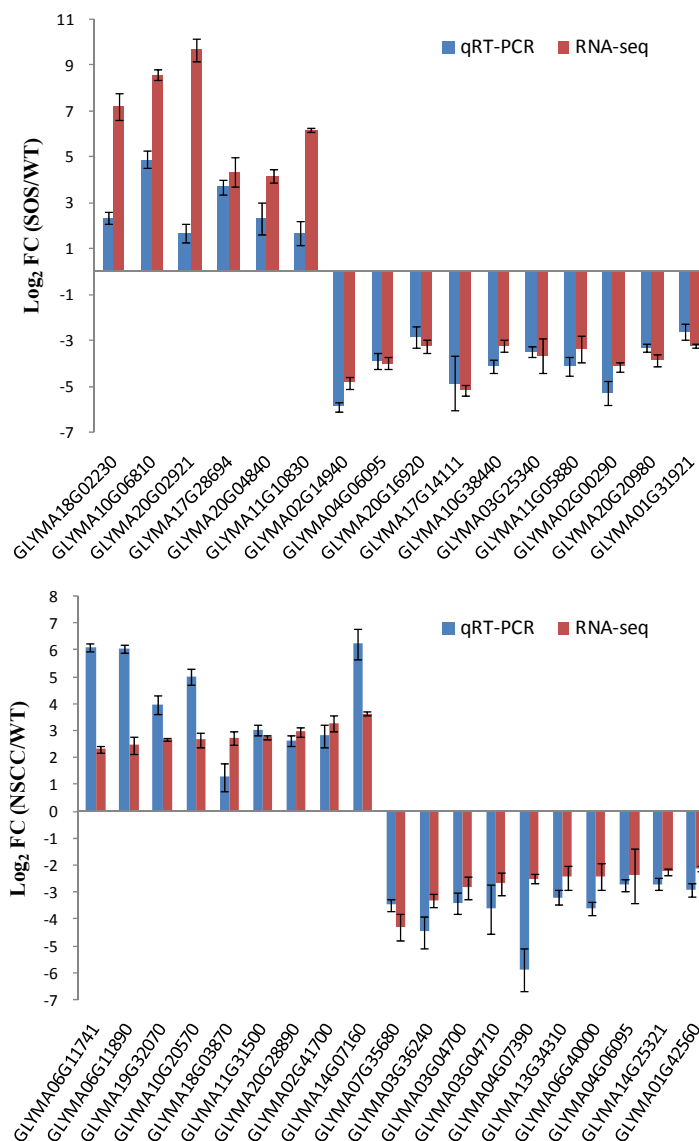
Figure 4. Experimental validation of differentially expressed genes by qRT-PCR

Note. Gene ID is shown on the x-axis. The comparative ΔΔCT method was used for the qRT-PCR experiments, and *GsRNAr-18S* was selected as the reference.

3.3.2 The Function Prediction of the DEGs Affected by the Gene Mutation

Blast2GO software was used to assign GO and KEGG functional classifications the DEGs. As a result, The DEGs were successfully classified into the three main GO categories of biological process, cellular component, and molecular function. The DEGs affected by *GsSOS1* and *GsNSCC* mutation were further categorized into 40, and 38 GO functional groups respectively (Figure A3). Further, the DEGs affected by *GsSOS1* and *GsNSCC* mutation were mapped to 94 and 95 KEGG pathways respectively (Supplementary File 2). The pathways with the highest unigene representations were those associated with plant hormone signal transduction, phenylpropanoid biosynthesis, starch-sucrose metabolism, and plant-pathogen interactions (Figure A4).

The results of GO enrichment analysis showed that *GsSOS1* and *GsNSCC* mutation had significant effects on the expression of many genes (Figure 5). For example, the biological process, the number of genes related to the protein phosphorylation in *GsNSCC* mutant, and transcription regulation and nodulation in *GsSOS1* mutant respectively, were most affected (Figure 5). Interestingly, few genes that exhibited up-regulation in *GsSOS1* mutant roots were also associated with the nodulation process (Figure A5). Significantly, both *GsSOS1* and *GsNSCC* mutations altered the expression of a set of the gene that is associated with transporter activity,

oxidoreductase activity. The most differentially expressed genes in *GsSOS1* and *GsNSCC* mutant roots were associated with the same metabolic pathways, such as plant hormone signaling system, phenylpropanoid biosynthesis, starch and sucrose metabolism, circadian rhythm and plant-pathogen interaction (Figure A4).
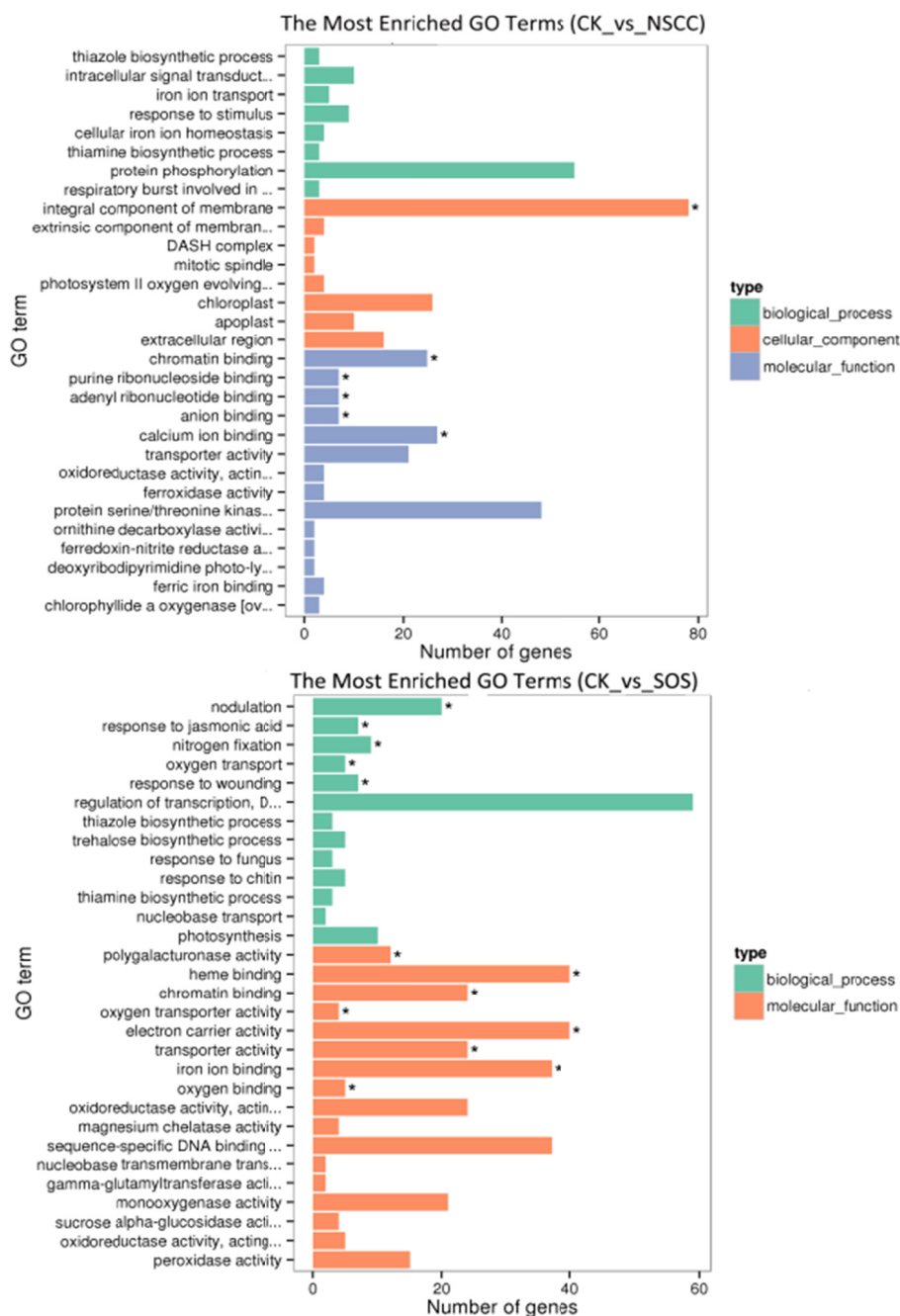

Figure 5. GO enrichment analysis of differentially expressed genes in *GsNSCC* and *GsSOS1* mutant roots compared with the CK

Note. GO annotation was carried out using Blast2GO software. Kolmogorov-Smirnow test was used for GO statistical analysis.

4. Discussion

CRISPR-Cas9 system has been extensively used in targeted genome editing in different crop species, including soybean (*Glycine max*) (Curtin et al., 2018; Li et al., 2015; Cai et al., 2018a, 2018b, 2019; Bao et al., 2019; Admin et al., 2019). Because of low transformation efficiency and long protocol of *A. tumefaciens* mediated

whole-plant transformation in soybean (Taylor et al., 2006), many researchers carried out the soybean gene editing research in hairy roots (Sun et al., 2015; Du et al., 2016; Jacobs et al., 2015; Cai et al., 2015; Du et al., 2018; Ming et al., 2018). Hairy soybean roots are an excellent model system for transformation and mutagenesis, for carrying out a functional study of the genes. For example, overexpression of *GmMYB118* in hairy soybean roots, led to improved drought and salt tolerance of the plants. Furthermore, CRISPR-transformed plants exhibited reduced stresses tolerance (Du et al., 2018). *GmNAC15* overexpression in hairy roots of soybean enhances their salt tolerance (Ming et al., 2018). However, the hairy root system cannot regenerate whole transgenic plants, and therefore successful heritable mutations cannot be achieved. Furthermore, the mutation frequencies in hairy roots are hard to be 100% and exhibit high variability, due to design of sgRNAs, different promoters driving sgRNA cassette and the target gene (Sun et al., 2015; Du et al., 2016; Cai et al., 2015). The mutant and the normal roots grow together which might cause some deviation in phenotypic and physiological data of CRISPR-transformed plants. The mutation efficiencies for *GsSOS1* and *GsNSCC* in this study were 28.5 and 39.9%, respectively. When all the hairy roots of each CRISPR-transformed plant were together collected, the Na^+/K^+ ration between mutants and the wild type exhibited no significant difference under normal and salt stress conditions. Exclusion of hairy roots with no mutation, the significant difference of Na^+/K^+ ratios in roots was detected between *GsSOS1* mutant and wild type. No significant effects of *GsNSCC* mutation on Na^+/K^+ ratios in different tissues under normal and salt treatment conditions were detected, perhaps because *GsNSCC* has two copies in the genome, while only one copy of *GsSOS1* (Supplementary File 3 and 4). Hence, some researchers study the sgRNA efficiencies with a hairy root system and identify the gene function with whole-plant transformation method (Do et al., 2019; Curtin et al., 2018). Therefore, our transcriptome analysis was based on the mutant hairy roots of target genes to study the influence of the target gene mutation on other gene expressions.

SOS1, a plasma membrane Na^+/H^+-antiporter, is known as a crucial component in the defence of plants against sodium ions that have entered the cytoplasm. The mutations in Arabidopsis *SOS1* results in (*i*) enhanced sensitivity to higher levels of NaCl (Wu et al., 1996; Zhu et al., 1998), (*ii*) accumulation of significantly higher levels of Na^+ in mutants than that of wild type under salt stress (Shi et al., 2000), and (*iii*) K^+ acquisition impairment and alterations in Na^+ to K^+ ratios (Wu et al., 1996). In present study, *GsSOS1* was edited in the roots with CRISPR-cas9 system into different mutant types. Under 250 mM NaCl treatment, Na^+/K^+ ratios in *GsSOS1* mutant roots were significantly higher than that of the wild type. Many studies have previously reported that *SOS1* appears to have many functions, that includes supporting vacuolar morphology, ion homeostasis and membrane trafficking. Many of these functions are critically associated with tolerance of root tissue during the early stages of salt stress. In this study, with RNA-seq analysis, *GsSOS1* mutation affected the expression of many other genes in wild soybean roots. Oh et al. (2010a) also outlined the function of genes and pathways that are affected when *SOS1* is either mutated by T-DNA insertion or by the reduction of *SOS1* transcripts through RNAi-interference (Oh, et al., 2010a). Oh et al. (2010a) also reported that non-availability of *SOS1*, altered the expression of genes related to pH homeostasis, membrane trafficking and ion transportation during salt stress. The similar results were obtained in this study. Eight aquaporin genes, including PIPs, TIPs, SIPs and NIPs, were differentially expressed in *GsSOS1* edited roots compared with wild type. The expression differences of some the K^+, Na^+ and Ca^+ transporters between *GsSOS1* mutant and wild type were also observed, Some of the K^+ channel protein-coding genes *SKOR* and *AKT1*, sodium/calcium exchanger, sodium transporter *HKT1*, as well as 11 other intracellular traffic-related genes (Supplementary File 3). In Arabidopsis *SOS1* mutant, expression of CNGC19 gene, encoding a calmodulin-binding cyclic nucleotide-gated channel, is strongly upregulated (Oh et al., 2010b). In this study, the alteration in the expression levels of four *CNGC* genes in *GsSOS1* mutant roots was also detected. In addition, KEGG analysis of DEGs in *GsSOS1* mutant roots showed that *GsSOS1* mutation significantly influenced the expression of many other genes that involved in different pathway, including hormone signal transduction, phenylpropanoid biosynthesis, starch-sucrose metabolism, circadian rhythm and plant-pathogen interactions (Figure A4). Earlier, a few of the reports also suggest biological functions that are affected in *SOS1* mutants of Arabidopsis, includes plants pathogen responses and circadian rhythm (Oh et al., 2010b). Hence, it is highly likely that the *SOS1* is involved in more than one functions such as those observed in the current study.

Nonselective cation channels (*NSCC*) in the plasma membrane of higher plants form a large and diverse group of plant cation channels by which the bulk of toxic Na^+ influx into plant roots (Amtmann & Sanders, 2002; Tyerman & Skerrett, 1999; White, 1999). In this study, *GsNSCC* mutation altered the expression of many genes that are integral components of the membrane (Figure 6). Unlike *GsSOS* mutants, *GsNSCC* mutation did not cause any change in the Na^+/K^+ ratios under normal condition and salt treatment (Figure 3). This raises the possibility of involvement of other types of *NSCC*s for Na^+ influx in the root cells (Tyerman, 2002). In the

Arabidopsis, 20 Glu receptor family genes may form nonselective ion channels (Lacombe et al., 2001). Apart from toxic Na^+ influx, *NSCC*s were also involved in nutritional uptake of K^+, NH_4^+, Ca^{2+}, Mg^{2+}, micronutrients including trace elements, in ROS-, amino acid, purine- and cyclic nucleotide induced signaling, growth and development (Demidchik et al., 2002). In this study, *GsNSCC* mutation altered the expression of 1817 genes that are involved in different metabolism pathways (Figures 6 and A4). One notable effect was detected on the genes associated with protein phosphorylation of biological process or protein serine/threonine kinase of molecular function. In other studies, it was reported that the activity of some of these nonselective cation channels is modulated by phosphorylation (Kaupp & Seifert, 2002). Hence, it can be postulated that *GsNSCC* may be involved in protein phosphorylation or *GsNSCC* mutant may need the functional complementation of other *NSCC* protein and the activity of *NSCC* modulated by phosphorylation. As *SOS1*, *NSCC* may also be involved in more than one function. Both *NSCC* and *SOS* are also associated with biological processes (albeit in different ways) such as Na^+ transportation, that may be the reason of sharing 887 differentially expressed genes between two mutants of *GsNSCC* and *GsSOS*. We also found many genes whose expression were significantly altered uniquely in a mutant specific manner. For example, from the results of hierarchical clustering analysis of differentially expressed genes, some genes were up-regulated only in *GsSOS1* mutant but not in *GsNSCC* mutant. Majority of these genes are involved in nodulation (Figure A5).

In conclusion, the mutants with CRISPR/Cas9 gene-editing system in the hairy roots of wild soybean were successfully generated. Two genes, *GsSOS1* and *GsNSCC*, were mutagenized with frequencies of 28.5% and 39.9%, respectively. *GsSOS1* and *GsNSCC* mutations significantly altered the transcriptome of mutant roots. Further, many differentially expressed genes are associated with various cellular functions according to the multiple functions of *SOS1* and *NSCC* proteins. The wild soybean genome editing would advance the gene mining and functional identification in wild soybean for improving the agronomic traits of cultivated soybean.

Acknowledgements

References

Amin, N., Ahmad, N., Wu, N., Pu, X., Ma, T., Du, Y., … Wang, P. (2019). CRISPR-Cas9 mediated targeted disruption of *FAD2-2* microsomal omega-6 desaturase in soybean (*Glycine max* L.). *BMC Biotechnol., 19*, 9. https://doi.org/10.1186/s12896-019-0501-2

Amtmann, A., & Sanders, D. (1999). Mechanisms of Na^+ uptake by plant cells. *Adv. Bot. Res., 29*, 75-112. https://doi.org/10.1016/S0065-2296(08)60310-9

Anders, S., & Huber, W. (2012). *Differential expression of RNA-Seq data at the gene level-the DESeq package*. European Molecular Biology Laboratory (EMBL), Heidelberg, Germany.

Bao, A., Chen, H., Chen, L., Chen, S., Hao, Q., Guo, W., … Cao, D. (2019). CRISPR/Cas9-mediated targeted mutagenesis of *GmSPL9* genes alters plant architecture in soybean. *BMC Plant Biol., 19*, 131. https://doi.org/10.1186/s12870-019-1746-6

Cai, Y., Chen, L., Liu, X., Guo, C., Sun, S., Wu, C., … Hou, W. (2018b). CRISPR/Cas9-mediated targeted mutagenesis of *GmFT2a* delays flowering time in soya bean. *Plant Biotechnol. J., 16*, 176-185. https://doi.org/10.1111/pbi.12758

Cai, Y., Chen, L., Liu, X., Sun, S., Wu, C., Jiang, B., … Hou, W. (2015). CRISPR/Cas9-mediated genome editing in soybean hairy roots. *PLoS One, 10*, e0136064. https://doi.org/10.1371/journal.pone.0136064

Cai, Y., Chen, L., Sun, S., Wu, C., Yao, W., Jiang, B., … Hou, W. (2018a). CRISPR/Cas9-mediated deletion of large genomic fragments in soybean. *Int. J. Mol. Sci., 19*, 3835. https://doi.org/10.3390/ijms19123835

Cai, Y., Wang, L., Chen, L., Wu, T., Liu, L., Sun, S., … Hou, W. (2019). Mutagenesis of *GmFT2a* and *GmFT5a* mediated by CRISPR/Cas9 contribute for expanding the regional adaptability of soybean. *Plant Biotechnol. J., 18*, 1-12. https://doi.org/10.1111/pbi.13199

Curtin, S. J., Xiong, Y., Michno, J. M., Campbell, B. W., Stec, A. O., Čermák, T., … Stupar, R. M. (2018). CRISPR/Cas9 and TALENS generate heritable mutations for genes involved in small RNA processing of

Glycine max and *Medicago truncatula*. *Plant Biotechnol. J., 16*, 1125-1137. https://doi.org/10.1111/pbi.12857

Demidchik, V., & Maathuis, F. J. M. (2007). Physiological roles of nonselective cation channels in plants: from salt stress to signalling and development. *New Phytol., 175*, 387-404. https://doi.org/10.1111/j.1469-8137.2007.02128.x

Demidchik, V., Davenport, R. J., & Tester, M. (2002). Nonselective cation channels in plants. *Annu. Rev. Plant Biol., 53*, 67-107. https://doi.org/10.1146/annurev.arplant.53.091901.161540

Do, P. T., Nguyen, C. X., Bui, H. T., Tran, L. T. N., Stacey, G., Gillman, J. D., … Stacey, M. G. (2019). Demonstration of highly efficient dual gRNA CRISPR/Cas9 editing of the homeologous *GmFAD2-1A* and *GmFAD2-1B* genes to yield a high oleic, low linoleic and α-linolenic acid phenotype in soybean. *BMC Plant Biol., 19*, 311. https://doi.org/10.1186/s12870-019-1906-8

Du, H., Zeng, X., Zhao, M., Cui, X., Wang, Q., Yang, H., … Yu, D. (2016). Efficient targeted mutagenesis in soybean by TALENs and CRISPR/Cas9. *J. Biotechnol., 217*, 90-97. https://doi.org/10.1016/j.jbiotec.2015.11.005

Du, Y. T., Zhao, M. J., Wang, C. T., Gao, Y., Wang, Y. X., Liu, Y. W., … Ma, Y. Z. (2018). Identification and characterization of *GmMYB118* responses to drought and salt stress. *BMC Plant Biol., 18*, 320. https://doi.org/10.1186/s12870-018-1551-7

Ge, Y., Li, Y., Lv, D. K., Bai, X., Ji, W., Cai, H., … Zhu, Y. M. (2011). Alkaline-stress response in *Glycine soja* leaf identifies specific transcription factors and ABA-mediated signaling factors. *Funct. Integr. Genomics., 11*, 369-379. https://doi.org/10.1007/s10142-010-0191-2

Jacobs, T. B., LaFayette, P. R., Schmitz, R. J., & Parrott, W. A. (2015). Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnol, 15*, 16. https://doi.org/10.1186/s12896-015-0131-2

Jiang, Q. Y., Niu, F. J., Sun, X. J., Hu, Z., Li, X. H., Ma, Y. Z., & Zhang, H. (2017). RNA-seq analysis of unintended effects in transgenic wheat overexpressing the transcription factor *GmDREB1*. *The Crop Journal, 5*, 207-218. https://doi.org/10.1016/j.cj.2016.12.001

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., … Yamanishi, Y. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res., 36*, D480-D484. https://doi.org/10.1093/nar/gkm882

Kaupp, U. B., & Seifert, R. (2002). Cyclic nucleotide-gated ion channels. *Physiol. Rev., 82*, 769-824. https://doi.org/10.1152/physrev.00008.2002

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol., 14*, R36. https://doi.org/10.1186/gb-2013-14-4-r36

Lacombe, B., Becker, D., Hedrich, R., DeSalle, R., Hollmann, M., Kwak, J. M., … Coruzzi, G. (2001). The identity of plant glutamate receptors. *Science, 292*, 1486-1487. https://doi.org/10.1126/science.292.5521.1486b

Lam, H. M., Xu, X., Liu, X., Chen, W., Yang, G., Wong, F. L., … Zhang, G. (2010). Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection, *Nat. Genet., 42*, 1053-1059. https://doi.org/10.1038/ng.715

Li, M., Hu, Z., Jiang, Q., Sun, X. J., Guo, Y., Qi, J. C., & Zhang, H. (2018). *GmNAC15* overexpression in hairy roots enhances salt tolerance in soybean. *J. Integr. Agric., 17*, 530-538. https://doi.org/10.1016/S2095-3119(17)61721-0

Li, Z., Liu, Z. B., Xing, A., Moon, B. P., Koellhoffer, J. P., Huang, L., … Cigan, A. M. (2015). Cas9-guide RNA directed genome editing in soybean. *Plant Physiol., 169*, 960-970. https://doi.org/10.1104/pp.15.00783

Oh, D. H., Ali, Z., Hyeong, C. P., Bressan, R. A., Dae, J. Y., & Bohnert, H. J. (2010b). Consequences of *SOS1* deficiency: Intracellular physiology and transcription. *Plant Signal. Behav., 5*, 766-768. https://doi.org/10.4161/psb.5.6.11777

Oh, D. H., Lee, S. Y., Bressan, R. A., Yun, D. J., & Bohnert, H. J. (2010a). Intracellular consequences of *SOS1* deficiency during salt stress. *J. Exp. Bot., 61*, 1205-1213. https://doi.org/10.1093/jxb/erp391

Qiu, Q. S., Guo, Y., Dietrich, M. A., Schumaker, K. S., & Zhu, J. K. (2002). Regulation of *SOS1*, a plasma membrane Na$^+$/H$^+$ exchanger in *Arabidopsis thaliana*, by *SOS*2 and *SOS*3. *Proc. Natl. Acad. Sci. USA*. *99*, 8436-8441. https://doi.org/10.1073/pnas.122224699

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc., 3*, 1101-1108. https://doi.org/10.1038/nprot.2008.73

Shi, H., Ishitani, M., Kim, C., & Zhu, J. K. (2000). The *Arabidopsis thaliana* salt tolerance gene *SOS1* encodes a putative Na$^+$/H$^+$ antiporter. *Proc. Natl. Acad. Sci., 97*, 6896-6901. https://doi.org/10.1073/pnas.120170197

Shi, H., Quintero, F. J., Pardo, J. M., & Zhu, J. K. (2002). The putative plasma membrane Na$^+$/H$^+$ antiporter *SOS1* controls long-distance Na$^+$ transport in plants. *Plant Cell, 14*, 465-477. https://doi.org/10.1105/tpc.010371

Smil, V. (2000). Magic beans. *Nature, 407*, 567. https://doi.org/10.1038/35036653

Sun, X., Hu, Z., Chen, R., Jiang, Q., Song, G., Zhang, H., & Xi, Y. (2015). Targeted mutagenesis in soybean using the CRISPR-Cas9 system. *Sci. Rep., 5*, 10342. https://doi.org/10.1038/srep10342

Taylor, C. G., Fuchs, B., Collier, R., & Lutke, W. K. (2006). Generation of composite plants using *Agrobacterium rhizogenes*, *Methods Mol. Biol., 343*, 155-167. https://doi.org/10.1385/1-59745-130-4:155

Tester, M., & Davenport, R. (2003). Na$^+$ tolerance and Na$^+$ transport in higher plants. *Ann. Bot., 91*, 503-527. https://doi.org/10.1093/aob/mcg058

Tyerman, S. D. (2002). Nonselective cation channels. Multiple functions and commonalities. *Plant Physiol., 128*, 327-328. https://doi.org/10.1104/pp.900021

Tyerman, S. D., & Skerrett, I. M. (1999). Root ion channels and salinity. *Sci. Hortic., 78*, 175-235. https://doi.org/10.1016/S0304-4238(98)00194-0

Wen, Z., Ding, Y., Zhao, T., & Gai, J. (2009). Genetic diversity and peculiarity of annual wild soybean (*G. soja* Sieb. et Zucc.) from various eco-regions in China. *Theor. Appl. Genet., 119*, 371-381. https://doi.org/10.1007/s00122-009-1045-y

White, P. J. (1999). The molecular mechanism of sodium influx to root cells. *Trends Plant Sci., 4*, 245-246. https://doi.org/10.1016/S1360-1385(99)01435-1

Wu, S. J., Ding, L., & Zhu, J. K. (1996). *SOS1*, a genetic locus essential for salt tolerance and potassium acquisition. *Plant Cell, 8*, 617-627. https://doi.org/10.2307/3870339

Young, M. D., Wakefield, M. J., Smyth, G. K., & Oshlack, A. (2010). Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol., 11*, R14. https://doi.org/10.1186/gb-2010-11-2-r14

Zhu, J. K., Liu, J., & Xiong, L. (1998). Genetic analysis of salt tolerance in Arabidopsis: Evidence for a critical role of potassium nutrition. *Plant Cell, 10*, 1181-1191. https://doi.org/10.1105/tpc.10.7.1181

Supplementary Files

Supplementary Files of this paper are retrieved from http://www.ccsenet.org/journal/index.php/jas/article/download/0/0/43446/45847

Appendix

Table A1. Target sequences of two genes in wild soybean and oligonucleotides used to express sgRNA in the vectors

Target gene	Target site sequences	Oligo forward (5'-3')	Oligo reverse (5'-3')	Enzyme	Gene annotation
GsNSCC	CCACACGTGCAA GAGGTGCACGG	attgCCACACGTG CAAGAGGTGCA	aaacTGCACCTCT TGCACGTGTGG	*ApaL* I	cyclic nucleotide-gated ion channel (NSCC)
GsSOS1	GTGGCTTTGTTG AAAGATCTTGG	attgGTGGCTTTG TTGAAAGATCT	aaacAGATCTTTC AACAAAGCCAC	*Bgl* II	Sodium/hydrogen exchanger protein (SOS)

The restriction enzyme sites are showed in green color. The PAMs are highlighted with red.

Table A2. Primers used in this study

Items	Primer name	Forward Primer	Reverse Primer
qRT-PCR experiments			
	GLYMA06G11741	GCAGTCGGATTTGTGGA	AAAGGTGGCAGCGTGTT
	GLYMA06G11890	AACCCACTTCCACCAT	ATAGCAACACGACCCA
	GLYMA19G32070	AGGATGCCAGGCTCTACA	GCAAAACCAAACCAATGAC
	GLYMA10G20570	TGAAATGTGCTGCTATGG	TTGTAGGTGGACTCATCG
NSCC/WT UP	*GLYMA18G03870*	CAGTGAGCCCTCGTTCT	CCTTACACCAGGATTTAGT
	GLYMA11G31500	TACGACCCATTTGGAAGA	AGAAACCGTGATGAACCTC
	GLYMA20G28890	TCAAGGGTTGTGAAAGC	GCATAGCATAAGAAGAGGC
	GLYMA02G41700	CAAGAAGGGAAAGATGTG	TTGTAGGTGCCCAGTAGA
	GLYMA14G07160	CCAGGGATAAGGATAGC	GATGAGGAGTGGAAACC
	GLYMA06G14880	GCTAAAATGGGACTTGG	TTTGGCTGAAACTCTGA
	GLYMA07G35680	GGTTCATAGATGCTGCTG	GTCCCATAGAAAGTAAAGAG
	GLYMA03G36240	GTCTCACAAGGGTCGTT	AAGAGGTGTAAAGGGATG
	GLYMA03G04700	AAGTCGTGAAGCAGCAG	CACCACCGTTTGTAGGG
	GLYMA03G04710	CTCTTCAATGGTGGTTCT	GTTCGCAAAGTCCTCGT
NSCC/WT DOWN	*GLYMA04G07390*	TTTCGATCACAGCGTAG	AAACATCATAGTGGTAGGC
	GLYMA13G34310	GACTATTGGTTATGCTCCTC	TCTGTGGGTCTCCTTCC
	GLYMA06G40000	GGACAAGCGACGATGA	CTGTGCCTGAAGGACTC
	GLYMA04G06095	TCCCAATAGCCGTGAAC	GCAGCAACCTCTTGAATC
	GLYMA14G25321	GTCGGATAACAGGGTCG	TGGAGGCTGCTGAGTGT
	GLYMA01G42560	GGCTTGTGGCTTGGTTTA	GATGATGTGAGTTCTTTGG
	GLYMA18G02230	GGCGAGTTCAAGGGAG	CCACGGAGCACTGTCT
	GLYMA10G06810	CAATTCCTGAGGCTGTG	GGTTGTGGGTTCGTGT
SOS/WT UP	*GLYMA20G02921*	CCATCACCACCCATAGC	GGCAACATTTGGAGCAG
	GLYMA17G28694	AGGCAAGGTCAAAGTGT	GGTCCTGCTTGGTAGAT
	GLYMA20G04840	GGAGGGATTGACTACTGC	TTACCTTGCCAAACTGG
	GLYMA11G10830	GCCATCAACTTTGTAGC	AGTGGGTCAACTCTTCC
	GLYMA02G14940	CCCGAGGAACAAAGGAA	TGTACCGAGCCAGACCC
	GLYMA04G06095	TCAAGAGGTTGCTGCTAA	CACCGTTCCATCACATT
	GLYMA20G16920	GGAAAGGGACAAGGGTT	GAGGGAAGTTCAGAATAGC
	GLYMA17G14111	CCTCTAACCTTCCGACCAC	AACCTCGCAGACCCACTT
	GLYMA10G38440	GAGCAATCCCAAGAAACGC	GGTCCCGAGCCAAATCC
SOS/WT DOWN	*GLYMA03G25340*	GAGCCGTTGCCGTTAC	GCTTGTCCAGCCATCC
	GLYMA11G05880	GCGTTGCTTCCGTCTT	GCTTGTCCAGCCATCC
	GLYMA02G00290	GAGACCTACCCAGAAGAA	ATGTGCCTGAAACTACG
	GLYMA20G20980	AGCAGCAGAAGGCACA	CACAGACCATCCCACAA
	GLYMA01G31921	TACAAATGCGTAGCCC	TGAGTTTCCTCGTCCT
Reference	*GsrRNA18S*	AAACGGCTACCACATCCAAG	CCTTCAATGGATCCATCGTTA
Normal PCR			
gene-specific primers	*NSCC-JC*	AAAAGTTGCGACGAAGGAA	TGGGGGACTGGGAGGAT
	SOS-JC	AATACAGCTTACTTTTCCG	AGCAAGACCCATTCCTA

Table A3. Composition of mRNA libraries from wild soybean roots of mutant and wild type plants

Sample ID	Raw Reads	Raw Bases(bp)	Clean Reads	Clean Bases(bp)	Mapped Reads	Uniq Mapped Reads	Multiple Map Reads
CK-1	97,578,216	14,636,732,400	95,670,324	14,350,548,600	79,563,455 (83.16%)	78,097,028 (81.63%)	1,466,427 (1.53%)
CK-2	103,213,950	15,482,092,500	101,918,250	15,287,737,500	84,979,791 (83.38%)	83,525,357 (81.95%)	1,454,434 (1.43%)
CK-3	107,604,634	16,140,695,100	105,533,340	15,830,001,000	88,483,435 (83.84%)	85,448,097 (80.97%)	3,035,338 (2.88%)
SOS-1	117,348,420	17,602,263,000	114,819,574	17,222,936,100	94,664,154 (82.45%)	90,764,263 (79.05%)	3,899,891 (3.40%)
SOS-2	103,396,196	15,509,429,400	102,085,584	15,312,837,600	82,716,351 (81.03%)	81,597,540 (79.93%)	1,118,811 (1.10%)
SOS-3	105,391,056	15,808,658,400	103,888,360	15,583,254,000	85,960,041 (82.74%)	84,528,630 (81.36%)	1,431,411 (1.38%)
NSCC-1	91,553,034	13,732,955,100	91,333,464	13,700,019,600	75,750,326 (82.94%)	74,671,597 (81.76%)	1,078,729 (1.18%)
NSCC-2	105,314,898	15,797,234,700	104,215,410	15,632,311,500	86,575,785 (83.07%)	85,404,962 (81.95%)	1,170,823 (1.12%)
NSCC-3	104,582,400	15,687,360,000	103,334,614	15,500,192,100	85,210,271 (82.46%)	83,867,517 (81.16%)	1,342,754 (1.30%)

Table A4. Mutation rates of each plant

Plant No.	Glyma08g09730 (SOS1)			Glyma18g49890 (NSCC)		
	Number of hair roots	Number of mutation	Mutation rate (%)	Number of hair roots	Number of mutation	Mutation rate (%)
1	13	6	46.2	6	6	100.0
2	6	4	66.7	12	6	50.0
3	9	1	11.1	6	2	33.3
4	9	3	33.3	11	2	18.2
5	5	1	20.0	7	4	57.1
6	5	3	60.0	6	0	0.0
7	6	1	16.7	10	5	50.0
8	2	0	0.0	7	0	0.0
9	3	1	33.3	8	1	12.5
10	1	0	0.0	8	1	12.5
11	4	0	0.0	8	2	25.0
12	9	3	33.3			
13	6	2	33.3			
14	5	1	20.0			
15	7	1	14.3			
16	5	1	20.0			

GsNSCC

```
    ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAGGTGCACGGCGTCCACGCACGTGCGCGCC    WT
```

Mutants

```
1.  ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAGGCGCACGGCGTCCACGCACGTGCGCGCC    S1(X1)
    ACGTACGCCAGCCTGAACTGCAGCCACACGTG---TA-----ACGGCGTCCACGCACGTGCGCGCC    -8,S1(X1)
    ACGTACGCCAGCCTGAACTGCAGCCACACGTGCA-------ACGGCGTCCACGCACGTGCGCGCC    -8(X1)
2.  ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGA--TGCACGGCGTCCACGCACGTGCGCGCC    -2(X1)
    ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAA--------------------------GCC    -28(X1)
    ACGTACGCCAGCCTGAACTGCAGCCACACGTGCA-G--CCACACGAGTTTCCACGCACGTGCGCGCC    -3,+1,S5(X1)
3.  ACGTACGCCAGCCTGAACTGCAGCCACAC----------GCACGGCGTCCACGCACGTGCGCGCC    -11(X1)
    ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAA-----GCACGGCGTCCACGCACGTGCGCGCC    -5(X1)
    ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAGGTTGCACGGCGTCCACGCACGTGCGCGCC    +1(X1)
4.  ACGT---------------------------------------ACGCACGTGCGCGCC    -47(X1)
    ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAGGTTGCACGGCGTCCACGCACGTGCGCGCC    +1(X1)
5.  ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGA---GCACGGCGTCCACGCACGTGCGCGCC    -3(X1)
    ACGTACGCCAGCCTGAACTGCA------------------------------------GCC    -41(X1)
6.  ACGTACGCCAGCCTGAACTGCAGCCA-------------------------CACGTGCGCGCC    -28(X2)
    ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAGGTTGCACGGCGTCCACGCACGTGCGCGCC    +1(X1)
7.  ACGTACGCCAGCCTGAACT-------------------------GCGTCCACGCACGTGCGCGCC    -26(X1)
    ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAGGTGCGCGGCGTCCACGCACGTGCGCGCC    S1(X1)
8.  ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAGGCGCACCGCGTCCACGCACGTCCGCGCC    S3(X1)
    ACGTACGCCAGCCTGAACTACAGCCACACGTGCAAGAGATGCACGGCGTCCACGCACGTGCGCGCC    S3(X1)
9.  ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAGATGCACCGCGTCCACGCACGTCCGCGCC    S3(X1)
10. ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAGGTGCGCGGCGTCCACGCACGTGCGCGCC    S1(X2)
11. ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAA----TGCACGGCGTCCACGCACGTGCGCGCC    -4(X1)
12. ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAA----TGCACGGCGTCCACGCACGTGCGCGCC    -4(X1)
13. ACGTACGCCAGCCTGAACTGCAGCCACACGTG-----------GGCGTCCACGCACGTGCGCGCC    -12(X2)
14. ACGTACGCCAGCCTGAACT--------------AGCCA--CA---CGTCCACGCACGTGCGCGCC    -18,S3(X1)
15. ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAG-----CACGGCGTCCACGCACGTGCGCGCC    -5(X1)
16. ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAGGCGCACGGCGTCCACGCACGTGCGCGCC    S1(X1)
17. ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAGGTGCGCGGCGTCCACGCACGTGCGCGCC    S1(X2)
18. ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGA---GCACGGCGTCCACGCACGTGCGCGCC    -3(X1)
    ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAG---CACGGCGTCCACGCACGTGCGCGCC    -3(X1)
    ACGTACGCCAGC-------------------------------------GT------C    -51(X1)
    ACGTACGCCAGCCTGAACTGCAGCCACACGTGCAAGAAGTTGCACGGCGTCCACGCACGTGCGCGCC    +1,S2(X1)
    TC----------------------------(80bp)----------------------------TCC    -80(X1)
```

GsSOS1

```
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGATCTTGGTGCCAGCAAAAAGCTAAGCACGATAAT    WT
```

Mutants

```
1.  TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGA--TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -2(X4)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGAGTCTTGGTGCCAGCAAAAAGCTAAGCACGATAAT    +1(X1)
```

```
2.  TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGATTCTTGGTGCCAGCAAAAAGCTAAGCACGATAAT    +1(X1)
    TGATCCTGTTGCTGTTGTGGCTTTGT---------TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -9(X1)
3.  TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGAT--TGGTGCCAGCAAAAAGCTAAGCACGATAAT    -2(X2)
4.  TGATCCTGTTGCTGTTGTGGCTTTG--------TCTTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -8(X1)
    TGATCCTGTTGCTGTTGTGGCTTTGTT----GATCTTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -4(X2)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAAGA--TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -2(X1)
    TGATCCTGTTGCTGTTGTGGCTTTG---------TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -10(X1)
5.  TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGA----GGTGCCAGCAAAAAGCTAAGCACGATAAT    -4(X1)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGA--TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -2(X8)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGATC----TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -4,S2(X1)
    TGATC--------------------------TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -30(X1)
6.  TGATCCTGTTGCTGTTGTGGCTTTGTTGAAA------GGTGCCAGCAAAAAGCTAAGCACGATAAT    -6(X1)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAA-----TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -5(X1)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGA--TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -2(X3)
    TGATCCTGTTGCTGTTGTGGCTTTGT---------TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -9(X1)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGAGTCTTGGTGCCAGCAAAAAGCTAAGCACGATAAT    +1(X1)
7.  TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGA--TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -2(X6)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGA--------CCAGCAAAAAGCTAAGCACGATAAT    -8(X1)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGACTCTTGGTGCCAGCAAAAAGCTAAGCACGATAAT    +1(X1)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAA-------GGTGCCAGCAAAAAGCTAAGCACGATAAT    -7(X1)
8.  TGATCCTGTTGCTGTTGTGGCTT----------TCTTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -10(X1)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGA--TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -2(X1)
    TGATCCTGTTGCTGTTGTGGCTT----------TCTTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -10(X2)
    TGATCCTGTTGCTGTTGTGGCTTTGTT------TCTTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -6(X1)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAA------TGGTGCCAGCAAAAAGCTAAGCACGATAAT    -6(X1)
    TGATCCTGTTGCTTGGAAAACATCACTGTTGCTTGGAGGACTTCTTGGTGCCAGCAAAAAGCTAAG
    CACGATAAT                                                         +9,S20(X1)
9.  TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGAC-TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -1(X1)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGA--TTGGTGCCAGCAAAAAGCTAAGCACGATAAT    -2(X2)
    TGATCCTGTTGCTGTTGTGGCTTTGTTGAAAGAAAAAGCATTGAAGCTGCAGGATCCACGATGTCG
    GATACGATGTCCAGGACATCTGGCCCGAAAATACTGGACATATAAATCTGTTATATCTTTAACAGA
    TTATTGTGCAGTTAGCAAGAGATAAGATGATCTATCTTTAGGAACGAATTCTTGGTGCCAGCAAAA
    AGCTAAGCACGATAAT                                                  +148(X1)
```

Figure A1. Gene sequences from independent mutant lines

Note. Wild-type sequences of the target genes were shown with the protospacer adjacent motif sequence highlighted in red. The numbers of the changed nucleotides were shown to the right of each sequence. +, - and S indicate insertion, deletionand substitution. Inserted and substituted nucleotides were shown in green.
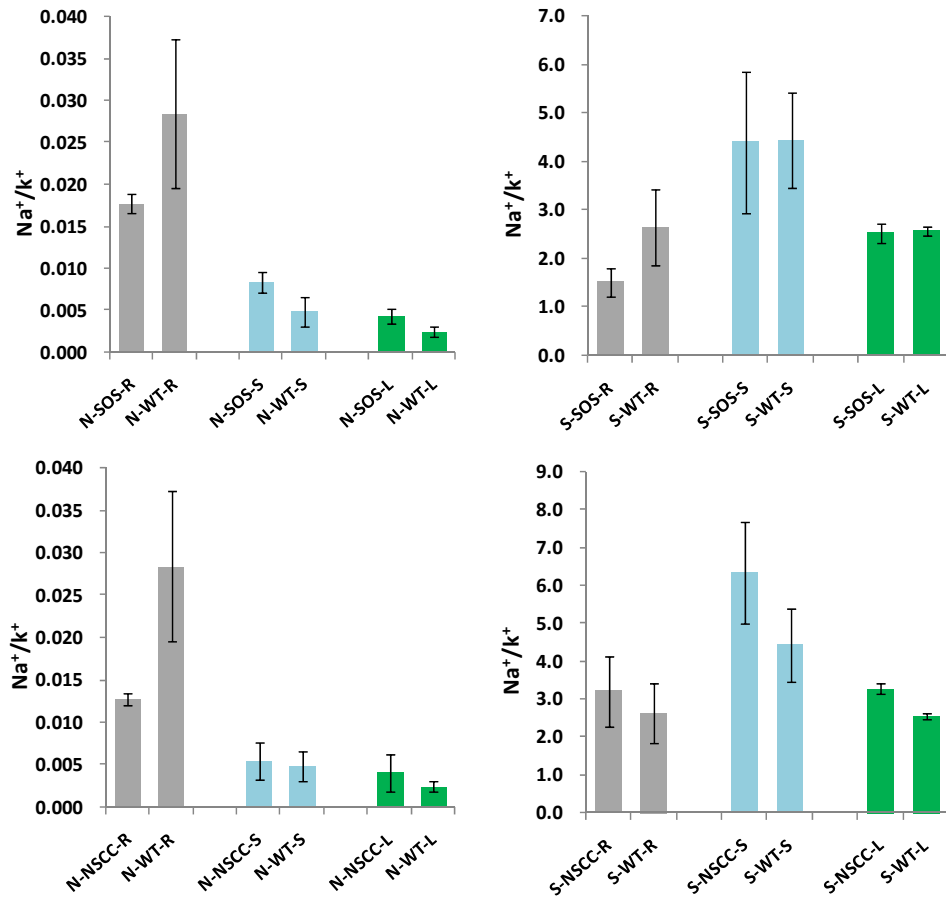
Figure A2. Na$^+$/K$^+$ rations in different tissues of *GsSOS1* and *GsNSCC* mutant composite plant and wild type

Note. We collected all the roots with mutation and non-mutation as well as stems and leaves to measure the content of Na$^+$ and K$^+$. S and N: Salt and Normal; SOS/NSCC and WT: SOS/NSCC mutation and Wild type; R, S and L: Root, Shoot and Leaf.
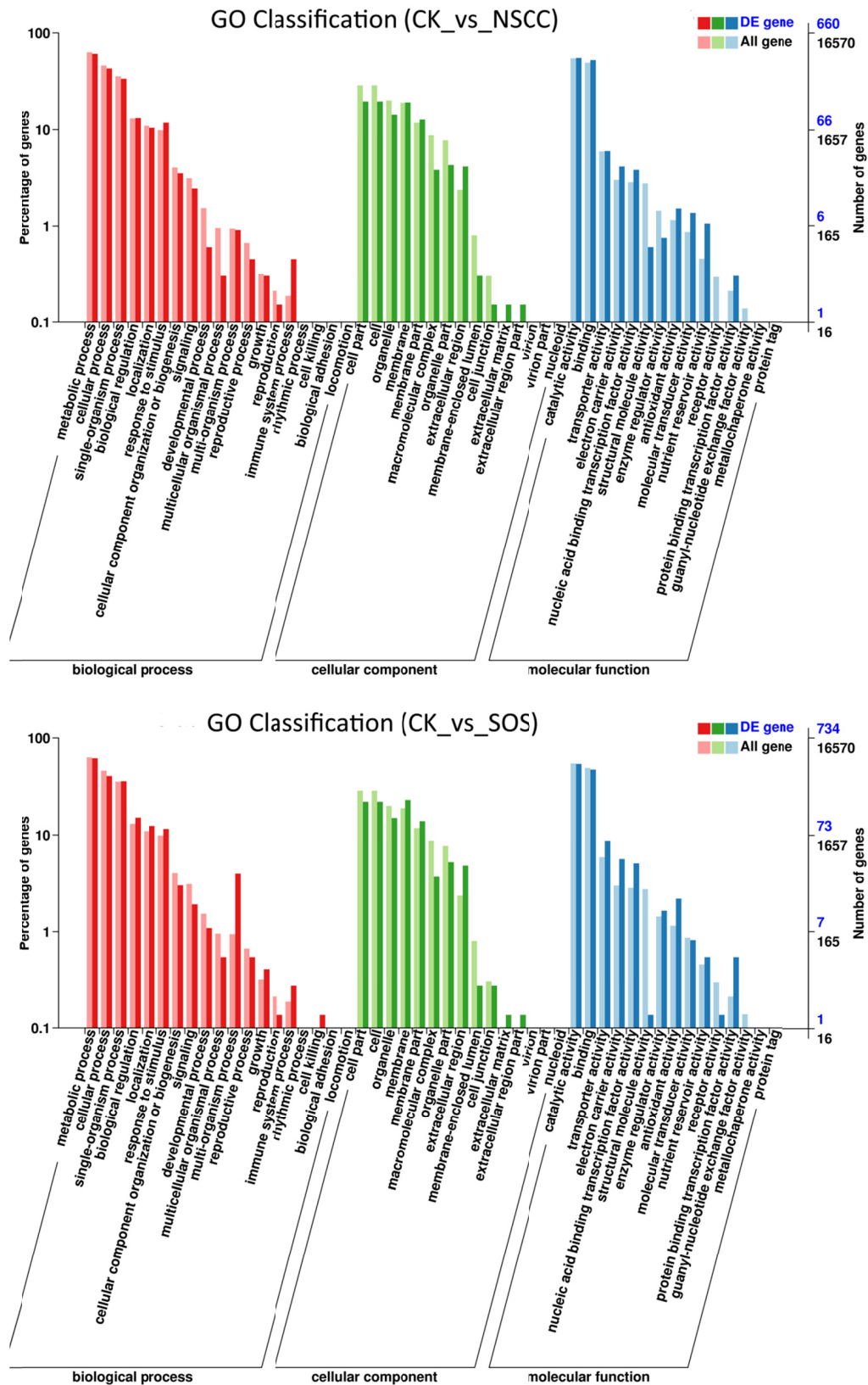
Figure A3. Gene ontology analysis. Expressed and differentially expressed genes were classified in the three main gene ontology categories: Molecular Functions, Biological Processes and Cellular Components
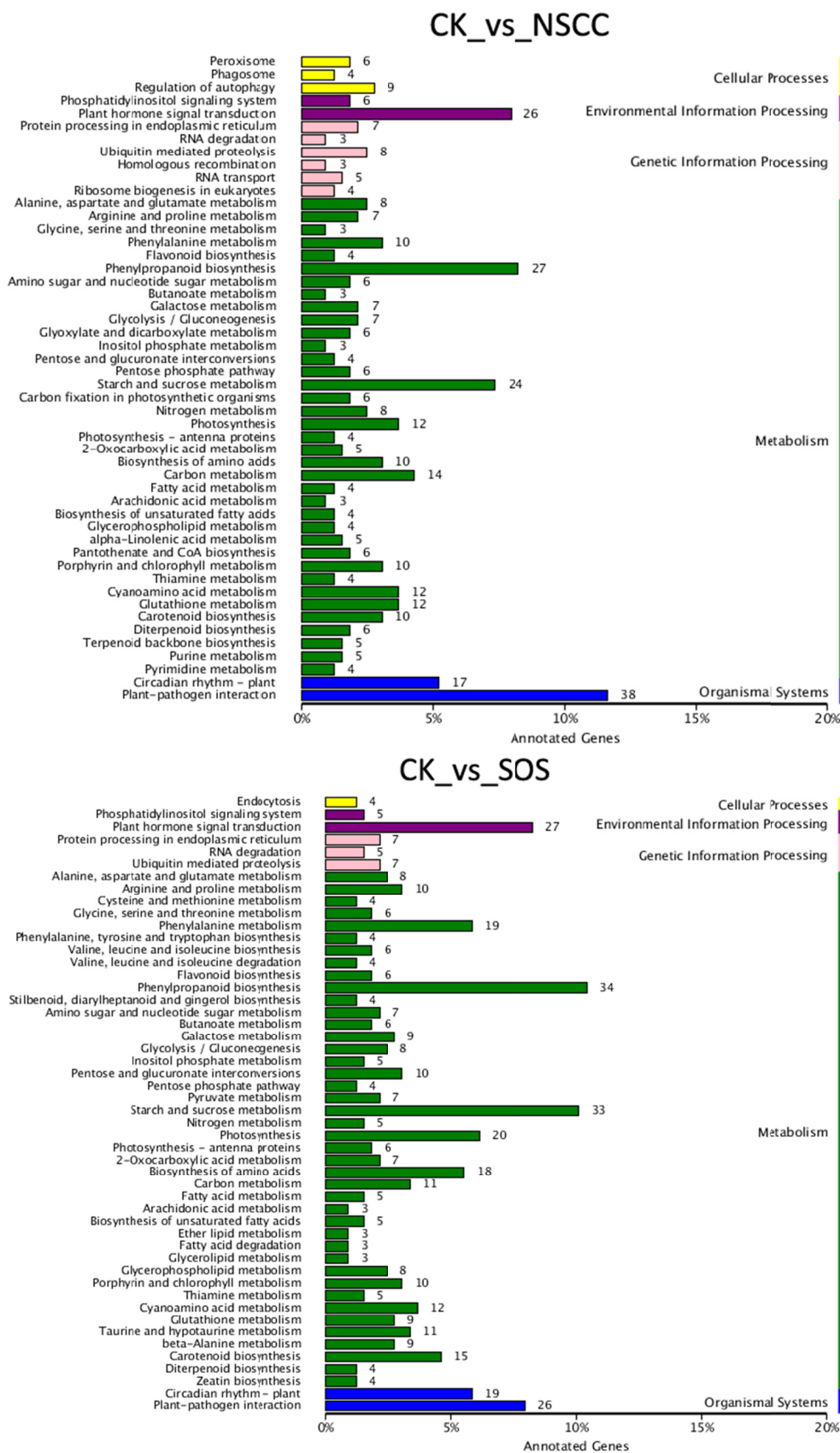
Figure A4. KEGG analysis of differentially expressed genes in *GsNSCC* and *GsSOS1* mutant roots
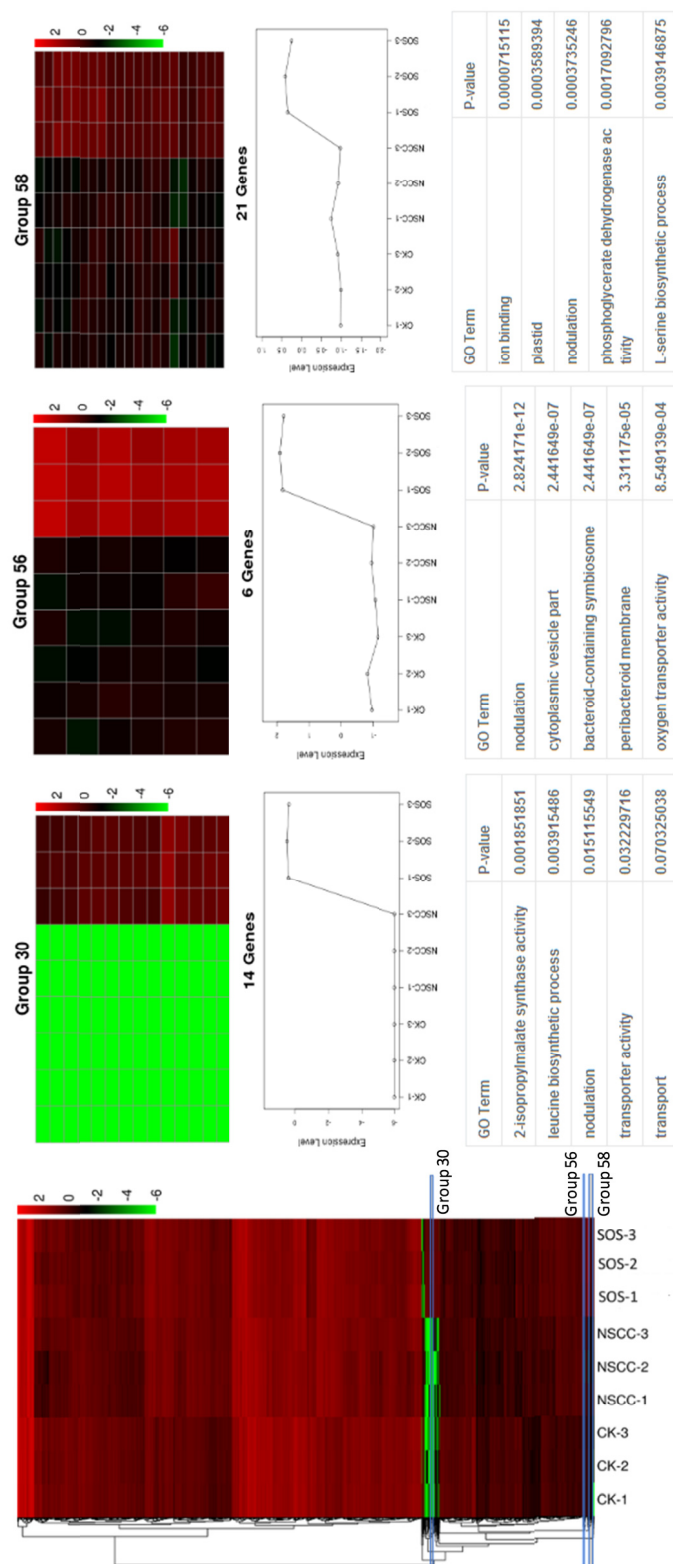
Figure A5. Hierarchical clustering analysis of differentially expressed genes from *GsSOS1* or *GsNSCC* mutation roots (SOS, NSCC) and wild type control (CK), each with 3 replicates (-1, -2 and -3). Rows represent individual genes. Genes that increased and decreased in abundance are indicated in red and green, respectively

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).