Transcript Profile of the Response of Two Soybean Genotypes to Potassium Deficiency

Cheng Wang1,2,3, HaiFeng Chen2, QingNan Hao2, AiHua Sha2, ZhiHui Shan2, LiMiao Chen2, Rong Zhou2, HaiJian Zhi1, XinAn Zhou2*

1 National Center for Soybean Improvement, National Key Laboratory for Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing, China, 2 Key Laboratory of Oil Crop Biology of the Ministry of Agriculture, Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, Wuhan, China

Abstract
The macronutrient potassium (K) is essential to plant growth and development. Crop yield potential is often affected by lack of soluble K. The molecular regulation mechanism of physiological and biochemical responses to K starvation in soybean roots and shoots is not fully understood. In the present study, two soybean varieties were subjected to low-K stress conditions: a low-K-tolerant variety (You06-71) and a low-K-sensitive variety (HengChun04-11). Eight libraries were generated for analysis: 2 genotypes × 2 tissues (roots and shoots) × 2 time periods [short term (0.5 to 12 h) and long term (3 to 12 d)]. RNA derived from the roots and shoots of these two varieties across two periods (short term and long term) were sequenced and the transcriptomes were compared using high-throughput tag-sequencing. To this end, a large number of clean tags (tags used for analysis after removal of dirty tags) corresponding to distinct tags (all types of clean tags) were identified in eight libraries (L1, You06-71-root short term; L2, HengChun04-11-root short term; L3, You06-71-shoot short term; L4, HengChun04-11-root long term; L5, You06-71-root long term; L6, HengChun04-11-root long term; L7, You06-71-shoot long term; L8, HengChun04-11-shoot long term). All clean tags were mapped to the available soybean (Glycine max) transcript database (http://www.soybase.org). Many genes showed substantial differences in expression across the libraries. In total, 5,440 transcripts involved in 118 KEGG pathways were either up- or down-regulated. Fifteen genes were randomly selected and their expression levels were confirmed using quantitative RT-PCR. Our results provide preliminary information on the molecular mechanism of potassium absorption and transport under low-K stress conditions in different soybean tissues.

Introduction
Potassium (K) is one of the most abundant cations in plants. It is considered a vital soil resource. The fact that most plants can only absorb small amounts of soluble K from the soil frequently limits crop yield. Although concentrations of K+ in soil solution are in the range of only 0.1–6.0 mmol L−1, plants can still accumulate large amounts of this element, up to 2–6% of the total plant dry weight [1,2]. K is widely distributed within plants and plays essential roles in maintaining plasma membrane potential, ion homeostasis, enzyme activation, signal transduction, and many other physiological processes in plant tissues [3–5]. For these reasons, inorganic K fertilizers are used on crops.

K+ obtained from the external environment is primarily absorbed by plant roots, which grow toward nutrient-rich areas and transport the nutrient to the root surface through the soil [6]. In addition to long-term K deprivation, plant roots can tolerate transient shortages of potassium from spatial and temporal variations in the availability of the nutrients. When plants are subjected to K+ deprivation, root cells sense the change in [K+]ext and initiate a series of physiological reactions [7]. In recent years, the molecular mechanism of potassium uptake by plant roots and of the loading and transport within plants has become the focus of a great deal of research. Experimental data from various plants has demonstrated that there are both high- and low-affinity K+ transport systems [8,9]. In comparative analyses, K+ absorption from soil to living plant cells and K+ transport inside plants are mediated by plant K+ transporters (high-affinity) and channels (low-affinity) [10]. Low-K stress conditions trigger the expression of high-affinity K+ transporters and up-regulate certain K+ channel genes in plants. Although the details of the molecular mechanisms related to the regulation of the sensing, signaling, and associated relevant gene activity in soybeans under low-K stress conditions are poorly understood, recent studies have provided some valuable clues. Over the past few decades, most molecular research into the absorption, loading/unloading, and transport of K in roots and leaves has used the model plant, Arabidopsis thaliana. For example, homology-based cloning has allowed the isolation of a large number of genes encoding K+ transporters and channels and the
identification of their homologs in other plant species [11–13]. These transporters and channels make up the K\(^+\) absorption and transport system in plant cells. Seventy-one K\(^+\) transporter and channel genes have been isolated and identified in Arabidopsis. Some of these genes and their biological functions have been well described: several molecular and protein components have been shown to contribute to potassium uptake into plant roots. For example, AKT1 (shaker family) Arabidopsis thaliana K\(^+\) transporter 1/2) encodes an inward-rectifying K\(^+\) channel. When identified in heterologous systems, AKT2 appears as a weakly-rectifying K\(^+\) channel that primarily expresses in the phloem of plant tissue [14,15]. The K\(^+\) uptake permease (KUP) family is important for various physiological processes, and many of its members have been isolated and identified [16]. Arabidopsis thaliana high-affinity K\(^+\) transporter 1 (AtHKT1) has been shown to be a Na\(^+\)/K\(^+\) transporter localized to the xylem parenchyma cells of leaves, mediating the Na\(^+\)/K\(^+\) loading from xylem vessels into xylem parenchyma cells [17]. In addition to these transporters and channels, certain cytoplasmic enzymes in plant cells also have roles in K signal transduction, such as calcium signaling (CBL/ CIPK) and transcription factors (REST (repressor element-1 silencing transcription factor)) [18,19].

The genome of the soybean (Glycine max) has been completed sequenced. Schmutz et al. described the complete soybean genome sequence [20]. In that sequence a large number of genes were annotated using a variety of functional genomic methods. The availability of the genome sequences corresponding to these genes permit the characterization of their related functions. For example, the Solexa/Illumina platform permits comprehensive, accurate gene expression profiling [21]. There have been several recent studies of Solexa/Illumina sequencing technology in response to plant growth and development [22,23].

In the present study, we used Solexa sequencing system to perform deep sequencing analysis, producing a great quantity of tags at a reasonable expense. We then used a digital gene expression (DGE) (digital gene expression) profile to filter the differentially expressed genes in different plant tissues [24]. In this way, rare transcripts were more likely to be identified with more reproducible sensitivity (Pearson correlation coefficient with 0.9–0.95) [25]. The results of the Solexa sequencing technology tend to show strong correlations to those of quantitative PCR [26]. The objective of this study was to determine the expression levels of genes derived from low-K-tolerant and -sensitive varieties under low-K stress conditions to determine whether these genes had any effect on K utilization.

**Results**

Preliminary Screening of Soybean Varieties with Different Levels of Tolerance to Low-K Stress Conditions

To identify soybean varieties tolerant to low-K stress conditions, a total of 40 varieties (lines) were collected and subjected to preliminary screening under low-K stress conditions. Thirty low-K-tolerant varieties (You06-71, Xudou3, and ZhongDou33) and two low-K-sensitive varieties (HengChun04-11 and ZhongHuang13) were identified in this preliminary screening. The difference between the low-K-tolerant varieties and low-K-sensitive varieties mainly involved potassium content, dry weight, accumulation of potassium, and potassium utilization efficiency (KUE (potassium use efficiency)). These indices were used to further screen and evaluate the varieties under different K conditions.

Effect of Different Levels of K on the Biomass of Plants of Different Genotypes

Biomass is the final result of plant growth and development. It reflects the extent to which the plant has suffered from the low-nutrition stress. It has always been a measure of resistance to the low-nutrition stress conditions. The experimental results indicated that different levels of K had a significant effect on the dry matter weights for different soybean varieties (Table 1). Under the low-K stress conditions, all five soybean varieties grew more poorly than under ordinary conditions but to different extents. The growth of low-K-tolerant varieties was better than low-K-sensitive varieties. You06-71 showed the highest coefficient of relatively dry weight (0.67), and HengChun04-11 showed the lowest coefficient of relatively dry weight (0.25). The trend among coefficients of relatively dry weight in different soybean varieties was as follows: You06-71 (0.67) > ZhongDou33 (0.53) > Xudou3 (0.51) > ZhongHuang13 (0.39) > HengChun04-11 (0.25). Plant roots are key to nutritional uptake under the conditions of both –K and +K. The dry weights of the roots of low-K-tolerant variety You06-71 were 1.50 and 1.17 times those of low-K-tolerant variety HengChun04-11 under the conditions of both –K and +K. This means that You06-71 has a stronger root system than HengChun04-11 and therefore stronger nutritional uptake capability. The shoot is the organ responsible for the transportation of nutrients and the photosynthetic product. The dry weights of the shoots of low-K-tolerant variety You06-71 were 3.86 and 1.19 times those of the low-K-sensitive variety HengChun04-11 under both –K and +K conditions. These results indicated that You06-71 caused more matter to accumulate. The ratio of the dry weights of shoots and roots of reflected the ability of the plant transport and distribute nutrition to the ground. You06-71 showed the highest ratio of shoot to root dry mass under the low-K stress conditions of any variety; this means that You06-71 can maintain more shoot growth while cannibalizing less of its root material. This may be one reason why it is so highly K efficient.

Effect of Different Levels of K on the Relevant Indicators of K of Different Genotypes

Under low-K stress conditions, there were significant differences in the amount of K accumulated and the amount of KUE across different soybean varieties (Table 2). The accumulation of K is the product of dry weight and K content. The difference in dry weight is a major contributor to differences in K accumulation. The total K accumulation per plant of the low-K-tolerant You06-71 variety was 2.36 and 1.26 times of that of the low-K-sensitive variety HengChun04-11 under –K and +K conditions, respectively. In the shoots, K accumulation per You06-71 plant peaked under both –K and +K conditions. In addition, KUE in You06-71 plants were significantly higher under low-K stress conditions than +K conditions. These differences indicated that You06-71 had a pronounced ability to take up, distribute, and use K efficiently. In our experiment, under low-K stress conditions, the K content of HengChun04-11 was the highest out of all the soybean varieties studied, but it showed a weaker correlation to the plant’s total dry weight. The K content of You06-71 was the lowest out of all the soybean varieties studied, but showed the strongest correlation with total dry weight. This indicated that lower K content was sufficient to maintain the regular physiological functions of low-K-tolerant variety You06-71. The low-K-tolerant variety had two main characteristics, including strong ability to take up K and use it efficiently. You06-71 showed an obvious advantage in the KUE and absorbed K more efficiently than other varieties, and differed from each other the least with respect to losses in total dry weight.
under low-K stress conditions. Among the five soybean varieties, HengChun04-11 showed an obvious disadvantage on these indices. Combined with the date of phenotype (Figure 1), we select You06-71 as the relatively low-K-tolerant variety and HengChun04-11 as the relatively low-K-sensitive variety to perform the following sequencing experiment.

Solexa Sequencing Evaluation

In this study, to obtain a comprehensive view of the soybean gene expression profile under low-K stress conditions, we used the Solexa method to perform a high-throughput tag-sequencing analysis. Poly-(A)-enriched RNAs samples were prepared from You06-71(Y) and Hengchun04-11(HC) and used for low-K stress treatment lasting from 0.5 h to 12 d. The samples taken at 0.5, 2, 6, and 12 h were selected and used as the short-term library and those taken at 3, 6, 9, and 12 d were used as the long-term library. The following samples were used for Solexa sequencing: L1, You06-71-root short-term; L2, HengChun04-11-root short term; L3, You06-71-shoot short term; L4, HengChun04-11-shoot short term; L5, You06-71-root long term; L6, HengChun04-11-root long term; L7, You06-71-shoot long term; and L8, HengChun04-11-shoot long term. After sequencing, the number of tags for each library ranged from 5.50 to 5.99 million reads (raw tags) and the total number of tags exceeded two million tags (10^6) (Figure S1). This indicated that the number of detected genes plateaued when the number of sequences increased as the total number of tags increased. The number of tags producing distinct sequences ranged from 0.3 to 0.5 million (Table 3). To determine whether the sequencing depth of the Solexa technology was sufficient for transcriptome coverage, the sequencing saturation was analyzed for all eight libraries. The distribution of the various tag abundance categories between total and distinct tag counts showed different results for all libraries. These distributions can be used to evaluate the normality of the whole data (Figure 2). Among the total clean tags, more than 5% had 2–5 copies (3.28% on average), 23.48% of the tags had 5–100 copies, and more than 65% of the tags had more than 100 copies (69.24% on average). Among the distinct clean tags, more than 60% had 2–5 copies (60.01% on average), 34.31% of the tags had 5–100 copies, and fewer than 5% of the tags had more than 100 copies (4.95% on average).

Matching reads to genes is a critical step in the prediction of sequences and determination of the molecular events underlying gene expression. All clean tags were BLAST searched against the available genome (http://phytozome.net/). Then 66210 genes were used for the initial analysis, and 61,044 of these genes (92.2%) had CATG sites. Out of the eight data sets representing expressed sequences and transcriptomes from the eight libraries, an average of 81.43% of all sequence were matched by tags and an average of 13.48% of distinct tags were uniquely mappable to the reference sequences. The tags mapping to the database generated 26,840; 27,443; 27,399; 26,689; 28,300; 25,340; 28,237; and 26,752 tag-mapped transcripts for the L1, L2, L3, L4, L5, L6, L7, and L8 libraries (Table 3). The distribution of the various tag abundance categories between total and distinct tag counts showed different results for all libraries. These distributions can be used to evaluate the normality of the whole data (Figure 2). Among the total clean tags, more than 5% had 2–5 copies (3.28% on average), 23.48% of the tags had 5–100 copies, and more than 65% of the tags had more than 100 copies (69.24% on average). Among the distinct clean tags, more than 60% had 2–5 copies (60.01% on average), 34.31% of the tags had 5–100 copies, and fewer than 5% of the tags had more than 100 copies (4.95% on average).

For each K level, means within a column followed by the different letter are significantly different according to the Duncan’s multiple range, P<5%, n=5.

For each K level, means within a column followed by the different letter are significantly different according to the Duncan’s multiple range, P<5%, n=5.
In addition, sense and antisense genes play important roles in gene expression and regulation, sequencing tags mapping to the complementary strands of the sense gene would suggest that the antisense strand also had transcripts [27]. In our study, approximately 23.30% of the tags in the eight libraries mapped to the antisense strands (Table S2), indicating that those genes might be activated via sense or antisense regulation.

Functional Classification of Differentially Expressed Transcripts

GO (gene ontology) enrichment analysis of functional significance was used to evaluate the potential function of genes that showed significant differences between the two soybean varieties under low-K stress conditions. Gene ontology (GO) is an internationally standardized gene function classification system for comprehensively describing the properties of genes and their products in any organism. GO has three ontologies: molecular functions, cellular components, and biological processes. The basic unit of GO is the GO-term. Every GO-term belongs to a type of ontology. In gene expression profiling analysis, GO enrichment analyses of functional significance was performed using hypergeometric testing to map all differentially expressed genes to terms in the GO database (http://www.geneontology.org/) by looking for GO terms that were significantly enriched in a given DEG (differentially expressed gene) relative to the genome background.

In our study, predicted and known transcripts were divided into groups: molecular function transcripts, cellular component transcripts, and biological process transcripts. Of these, 4813

Figure 1. Comparison of the morphological changes of different soybean (Glycine max) varieties under low-K stress condition. You06-71 is the low-K tolerant variety; HengChun is the low-K sensitive variety. doi:10.1371/journal.pone.0039856.g001
Table 3. Categorization and abundance of tags.

| Summary                                      | L1   | L2   | L3   | L4   | L5   | L6   | L7   | L8   |
|----------------------------------------------|------|------|------|------|------|------|------|------|
| Raw tags                                     | 6249661 | 6082210 | 6025833 | 5777397 | 6170727 | 602146 | 5940203 | 5909627 |
| Distinct tags                                | 440455 | 574060 | 444939 | 426457 | 645685 | 602146 | 439324 | 363528 |
| Clean tags Total number                      | 5985559 | 5704575 | 5741830 | 5503142 | 5736751 | 5496158 | 567587 | 5683891 |
| Number of distinct tags                      | 440455 | 574060 | 444939 | 426457 | 645685 | 602146 | 439324 | 363528 |
| Clean tags Total number                      | 5985559 | 5704575 | 5741830 | 5503142 | 5736751 | 5496158 | 567587 | 5683891 |
| Number of distinct tags                      | 440455 | 574060 | 444939 | 426457 | 645685 | 602146 | 439324 | 363528 |
| All tags mapping to gene                     | 4743425 | 4433325 | 4755583 | 4582123 | 453668 | 447895 | 4788454 | 4767238 |
| Clean tags as % of total                     | 79.25% | 77.72% | 82.82% | 83.26% | 79.08% | 80.93% | 84.49% | 83.87% |
| Number of distinct tags                      | 96697 | 97305 | 91574 | 85365 | 101664 | 106076 | 91523 | 84341 |
| Clean tags as % of distinct                  | 54.39% | 48.81% | 54.67% | 53.86% | 47.46% | 50.36% | 52.87% | 58.61% |
| Unambiguous tags mapping to gene             | 3650621 | 3425562 | 3442014 | 3331664 | 3455437 | 3266134 | 3315656 | 3443689 |
| Total % of clean tags                        | 60.99% | 60.05% | 59.95% | 60.54% | 60.23% | 59.43% | 58.50% | 60.59% |
| Number of distinct tags                      | 76207 | 76010 | 72175 | 67777 | 79268 | 81635 | 71463 | 67372 |
| Clean tags as % of distinct                  | 42.87% | 38.13% | 43.09% | 42.76% | 37.00% | 38.75% | 41.29% | 46.82% |
| All Tag-mapped Genes                         | 33717 | 34993 | 34771 | 34291 | 35967 | 36353 | 35999 | 33739 |
| % of ref genes                               | 50.92% | 52.85% | 52.52% | 51.79% | 54.32% | 54.91% | 54.37% | 50.96% |
| Unambiguous tag-mapped genes                 | 26848 | 27443 | 27339 | 26889 | 28300 | 28540 | 28237 | 26752 |
| % of ref genes                               | 40.55% | 41.45% | 41.29% | 40.61% | 42.74% | 43.11% | 42.65% | 40.40% |
| Mapping to mitochondrion                     | 201 | 297 | 161 | 191 | 267 | 267 | 216 | 140 |
| Clean tags as % of total                     | 0.00% | 0.01% | 0.00% | 0.00% | 0.00% | 0.01% | 0.00% | 0.00% |
| Number of distinct tags                      | 25 | 33 | 29 | 22 | 39 | 33 | 32 | 16 |
| Clean tags as % of distinct                  | 0.01% | 0.02% | 0.02% | 0.01% | 0.02% | 0.02% | 0.02% | 0.01% |
| Mapping to chloroplast                       | 4185 | 6981 | 2982 | 22891 | 8466 | 9592 | 36984 | 12778 |
| Clean tags as % of total                     | 0.07% | 0.12% | 0.52% | 0.42% | 0.15% | 0.17% | 0.65% | 0.22% |
| Number of distinct tags                      | 244 | 341 | 481 | 422 | 386 | 360 | 487 | 362 |
| Clean tags as % of distinct                  | 0.14% | 0.17% | 0.29% | 0.27% | 0.18% | 0.17% | 0.28% | 0.25% |
| Mapping to genome                            | 633696 | 598340 | 570702 | 564597 | 613028 | 518917 | 531827 | 578709 |
| Clean tags as % of total                     | 10.59% | 10.49% | 9.94% | 10.26% | 10.69% | 9.44% | 9.38% | 10.18% |
| Number of distinct tags                      | 38961 | 47412 | 46550 | 45832 | 55250 | 47841 | 51542 | 36698 |
| Clean tags as % of distinct                  | 21.92% | 23.78% | 27.79% | 28.92% | 25.79% | 22.71% | 29.78% | 25.50% |
| Unknown tags Total number                    | 603779 | 665632 | 385564 | 333340 | 578322 | 519435 | 310106 | 325026 |
| Clean tags as % of total                     | 10.09% | 11.67% | 6.72% | 6.06% | 10.08% | 9.45% | 5.47% | 5.72% |
differentially expressed transcripts were divided into 50 functional groups (Figure 3). For molecular function, the most significant enrichment was observed among the various binding genes (GO: 0005488), including 40 DEGs encoding calcium ion binding proteins (GO: 0005509), 66 DEGs encoding iron ion binding proteins (GO: 0005506), 55 DEGs encoding zinc ion binding proteins (GO: 0008270), and 4 DEGs encoding potassium ion binding proteins (GO: 0030955). Two hundred and eighty DEGs were found to encode transferase activity (GO: 0016740), these included kinase activity genes (GO: 0016301). Ninety-four DEGs were found to encode transporter activity (GO: 0005215). These included transmembrane transporter activity genes (GO: 0022857). For cellular components, 99.29% of DEGs were found to be involved in the “cell” structure. These included genes involved in the plasma membrane and external encapsulating structures such as the cell wall and cell envelope. There were 524 DEGs encoded in membranes (GO: 0016020). These included plastid membranes (GO: 0005886) and mitochondrial membranes (GO: 0031966). For biological processes, 813 DEGs were found to be involved in metabolic process (GO: 0008152) and 62 DEGs were found to be involved in stress response (GO: 0006950), including responses to salt stress (GO: 0006970), osmotic stress (GO: 0006970), ethylene stimulus (GO: 0009723), and oxidative stress (GO: 0034599).

We also matched differentially expressed genes to terms in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. KEGG is the major public pathway-related database. Different genes usually cooperate with each other to exercise their biological functions. Pathway-based analysis helps the user to further understand the biological functions of specific genes. Pathway enrichment analysis identifies significantly enriched metabolic pathways and signal transduction pathways in DEGs by comparing them to the whole-genome background [28]. The formula used to determine DEGs is the same as that used in GO analysis. We arranged 5440 differentially expressed genes to 118 KEGG pathways (Figure 4). The pathways with the greatest numbers of unique sequences were all metabolic pathways (1435 members) (ko01100). Metabolic pathways are large complexes comprising several networks. These networks can involve biosynthesis of secondary metabolites, carbohydrate metabolism, lipid metabolism, or amino acid metabolism [29–32]. Other pathways, such as flavonoid biosynthesis (150 members) (ko00941), purine metabolism (118 members) (ko00230), ABC transporters (44 members) (ko02010), and ubiquitin mediated proteolysis (140 members) (ko06320).
Genes Found to be Differentially Expressed in the Two Soybean Varieties

Schmid et al., Lieng Ma et al., and Aceituno et al. concluded that the Arabidopsis transcriptome varied strongly from one organ to another, and showed organ-specific expression during development [37–39]. These studies suggest that specific plant organs had their own independent transcriptomes. In soybeans, by comparing our eight Solexa libraries, we identified a large number of differentially expressed transcripts. The existences of differentially expressed genes in plants hint at the molecular basis of root and shoot development under low-K stress conditions. With reference to the significance of digital gene expression profiles, we adopted a rigorous algorithm to identify differentially expressed genes in their two corresponding libraries (eg: L1 vs L2).

Figure 3. Histogram presentation of gene ontology classification. It includes three main categories: biological processes, cellular components, and molecular functions. The y-axes on the left and right indicate the percentage of a specific category of genes in that main category and the number of genes in a category, respectively.
doi:10.1371/journal.pone.0039856.g003
regulated genes were those with a $|\log_2 \text{ratio}| > 1$, FDR $< 0.001$, $P < 0.01$) [24]. Analysis of the eight libraries revealed 24,534; 24,831; 24,767; 24,325; 25,593; 25,384; and 24,544, tag-mapped sense transcripts for L1, L2, L3, L4, L5, L6, L7, and L8. We found that many genes were differentially expressed in the two libraries (comparing the transcripts originating from low-K-tolerant and low-K-sensitive plants). The results showed 13,231; 13,963; 15,091; and 17,285 differentially expressed genes for L1 vs. L2, L3 vs. L4, L5 vs. L6, and L7 vs. L8, respectively. After filtering with FDR $< 0.001$, $P < 0.01$ and $|\log_2 \text{Ratio}| > 1$ and excluding genes with unknown function (average 41.89%) (Table S4), a total of 1094 differentially expressed genes were detected in L1 vs. L2. These included both up-regulated genes (690) and down-regulated genes (404). A total of 1289 differentially expressed genes were detected in L3 vs. L4. These included both up-regulated genes (752) and down-regulated genes (537). A total of 1370 differentially expressed genes were detected in L5 vs. L6. These included both up-regulated genes (693) and down-regulated genes (677). A total of 2490 differentially expressed genes were detected in L7 vs. L8. These included both up-regulated genes (1314) and downregulated genes (1176) (Table S5) (Figure 5). The predicted genes in different libraries are shown in Table S6.

In addition, in data sets from roots and shoots, a large number of genes showed differential expression between short term (0–12 h) and long term (3–12 d). This suggests that many genes were regulated in growing plants under low-K stress conditions. The DGE results revealed 279 genes that were differentially expressed in both the short term and long term in roots and 532 genes that were differentially expressed in the both short term and long term in shoots. There were 53 genes that were continuously expressed in all data sets (Table S7, Figure 6). Meanwhile, we presented the 10 most differentially expressed genes in the L1/L2, L3/L4, L5/L6, L7/L8 libraries, for a total of 40. The relative abundance is expressed as a TPM (number of transcripts per million clean tags; when one tag matched one gene perfectly. TPM represented the level of specific gene expression) ratio of the target group to that of the control group. As shown in Table 4, these genes included transcription factors, protein kinases, and other gene products.

Differentially Expressed Genes Encoding Transcription Factors

Transcription factors are involved in the regulation of gene expression in response to stress in higher plants [40]. Changes in gene transcription are associated with changes in the expression of transcription factors. Our DEG results showed that 144 genes encoding transcription factors were induced 2.02 to 968-fold, including 77 up-regulated and 67 down-regulated genes. Among the 144 genes, 47 were from the MYB gene family, 15 were bZIP (basic region/leucine zipper motif) transcription factors, 8 were AP2-EREBP transcription factors, 7 were GRAS (GAI, RGA, SCR domain transcriptional regulators) family transcription factors, 6 were WRKY transcription factors, 6 were heat shock
transcription factors, 4 were CPP (cystein-rich polycomb-like protein) transcription factors, 3 were ethylene-responsive transcription factors, 3 were ERF domain-containing transcription factors, 2 were bHLH transcription factors, 2 were AN1-like transcription factors, 1 was a BEL1-like transcription factor, 1 was a GATA transcription factor, and one was a TGA10 transcription factor [41–47]. The others were mainly putative transcription factors (Table S8).

Differentially Expressed Genes Encoding Kinases

Kinases play important roles in the development of eukaryotic cells. They operate in a large number of distinct perception and signaling pathways, such as the signal transduction pathways controlling cell growth, differentiation, and death [48]. The response of cells to stress is mediated by a large number of protein kinases [49]. Protein kinases also act as molecular switches to regulate potassium channels [50]. Among our DEGs, a total of 248 genes encoding kinases were identified, including 147 up-regulated and 101 down-regulated genes. Among these genes, 29 were serine or threonine protein kinases, 13 were calcium-dependent protein kinases (CDPKs), 6 were CBL-interacting protein kinases, 5 were tyrosine kinases, 4 were leucine-rich repeat transmembrane protein kinases, 3 were wall-associated kinases, 2 were MAP-kinases, 1 was a stress-induced receptor-like kinase, and 1 was a kinase-associated protein phosphatase [51]. Another 40 were different types of receptor-like kinases, and the remaining genes coded for other types of kinases (Table S9).

Potassium Absorption- and Transportation-related Genes

Potassium absorption and transportation are essential physiological and biochemical processes for plants. Among our DEGs, a total of 12 genes encoding potassium channels and ion transporters showed significant differential expression in this study, including eight up-regulated genes (Glyma13g23960, Glyma07g30360, Glyma06g09720, Glyma12g29190, Glyma07g08030, Glyma15g10140, Glyma08g09720, and Glyma08g09720), three down-regulated genes (Glyma08g01900, Glyma10g24620 and Glyma05g33660), and one bidirectionally regulated genes (Glyma03g31130).

Oxidative Stress-related Genes

Reactive oxygen species (ROS) are known to be involved in signaling pathways specific to low-K stress potassium conditions [52]. Peroxidase, cytochrome P450, and glutathione S-transferase play important roles during oxidative stress in plants [53]. In our study, a total of 70 genes encoding oxidative stress-related products were differentially expressed under low-K stress conditions. These included 39 up-regulated and 31 down-regulated genes. Among these, 34 cytochrome P450 genes and 19...
lipoxigenase genes were identified. Genes encoding peroxidase and glutathione S-transferase were not found among the DEGs (Table S10).

**Genes Involved in Carbohydrate and Energy Metabolism**

A large number of differentially expressed genes were found to be involved in carbohydrate and energy metabolism. These included glycolysis and the citrate cycle (TCA cycle), oxidative phosphorylation, photosynthesis, and nitrogen metabolism. Glycolysis metabolism is the process of converting glucose into pyruvate and generating small amounts of ATP (energy) and NADH (reducing power). This process has been reported to play an important role in potassium nutrition in plants [54]. Among our DEGs, 17 genes involved in glycolysis showed increased transcript abundance. These genes encoded, among other things, glyceraldehyde 3-phosphate dehydrogenase (Glyma12g18190), a glyceraldehyde 3-phosphate dehydrogenase (Glyma11g13110), and others. In addition, 16 genes related to oxidoreductase activity were identified, those encoding alcohol dehydrogenase (Glyma07g07100), NADPH-specific isocitrate dehydrogenase (Glyma16g27990-3.32-fold), and others.

**Other Differentially Regulated Genes**

One other type of high-level differentially expressed gene related to low-K stress conditions was shown. After analysis of the DEGs, five genes encoding the ATP binding cassette (ABC) transporter family were found to be differentially expressed (Glyma09g04980, Glyma10g08280, Glyma10g08280, Glyma10g08280, and Glyma10g08280). Four defense response genes were also identified, those encoding abscisic stress response protein (Glyma16g03770). NBS-LRR type disease resistance proteins (Glyma12g36510), wound-induced basic proteins (Glyma12g36510), and salt-tolerance protein (Glyma12g36510). In addition, 16 genes related to oxidoreductase activity were identified, those encoding alcohol dehydrogenase (Glyma09g04980), NADPH-specific isocitrate dehydrogenase (Glyma14g39160), a short-chain dehydrogenase (Glyma16g04630), fatty acid desaturase (Glyma14g39160), and others. In addition, two gibberellin-responsive protein genes were found to be up-regulated under low-K stress conditions (Glyma17g36220-4.11-fold, Glyma16g04630, Glyma12g36510). One gene encoding a BURP

---

### Table 4. The most differentially expressed annotated genes in L1 vs. L2, L3 vs. L4, L5 vs. L6, and L7 vs. L8 libraries based on expressed tag frequency (TPM: transcripts per million).

| Gene ID          | TPM ratio | Functional annotation                      | Gene ID          | TPM ratio | Functional annotation                      |
|------------------|-----------|--------------------------------------------|------------------|-----------|--------------------------------------------|
| Glyma08g28800    | 6900      | Full = 605 ribosomal protein L30            | Glyma03g33620    | 1738      | ADP-glucose small subunit PvAGPS1           |
| Glyma19g38570    | 3859      | Syntactic, plant, putative                  | Glyma18g09290    | 1160      | Disease resistance protein                  |
| Glyma20g03230    | 2924      | Matrix metalloproteinase                    | Glyma13g23060    | 1325      | Sucrose-phosphate synthase                  |
| Glyma13g23060    | 1119      | Sucrose-phosphate synthase                  | Glyma18g51930    | 1238      | NBS-LRR disease-resistance protein         |
| Glyma14g09540    | 986       | MYB transcription factor MYB81              | Glyma20g27410    | 697       | Receptor-like kinase homolog RK20-1         |
| Glyma16g30630    | 2384      | Leucine-rich repeat                         | Glyma09g30360    | 418       | Purple acid phosphatase                     |
| Glyma12g18190    | 1946      | TNP1                                        | Glyma06g37150    | 309       | S-adenosylmethionine decarboxylase          |
| Glyma07g07100    | 1508      | Rpp4 candidate 3                            | Glyma11g13110    | 273       | F-box and WD40 domain protein              |
| Glyma09g33850    | 1052      | Pyrophosphate isomerase                     | Glyma18g09290    | 237       | Cytochrome P450 enzyme                     |
| Glyma13g25800    | 701       | Protein tyrosine kinase                     | Glyma18g50690    | 17.15     | Serine-threonine kinase                     |
| Glyma08g28800    | 2996      | Full 605 ribosomal protein L30              | Glyma02g47460    | 8787      | ADP-glucose small subunit PvAGPS1           |
| Glyma04g40720    | 1498      | 605 acid ribosomal protein P1               | Glyma19g38570    | 5752      | Syntactic, plant, putative                  |
| Glyma10g02690    | 940       | FIE                                         | Glyma02g03230    | 5364      | Matrix metalloproteinase                    |
| Glyma19g03570    | 819       | Cell division protease                      | Glyma12g01260    | 2982      | Serine carboxypeptidase                     |
| Glyma02g03420    | 801       | UDP-glucosyltransferase                     | Glyma16g33910    | 1994      | R 8 protein                                |
| Glyma09g31740    | 927       | Dehydrin-like protein                       | Glyma07g07100    | 5542      | Rpp4 candidate 3                           |
| Glyma18g04520    | 672       | High mobility group family                  | Glyma12g33150    | 1196      | Structural constituent of cell wall         |
| Glyma12g11390    | 509       | MYB transcription factor MYB117             | Glyma16g06870    | 1108      | Ribonucleoside-diphosphate reductase chain  |
| Glyma10g07770    | 491       | Dehydration-responsive protein              | Glyma06g35630    | 1003      | Tyrosine aminotransfer                     |
| Glyma13g07530    | 418       | Ubiquitin-protein ligase                    | Glyma01g41280    | 968       | Polyprotein                                |

**Identification of Related Response to Low K Stress**

doi:10.1371/journal.pone.0039856.t004
domain protein (Glyma14g20440) and ten encoding dehydration-responsive binding proteins were differentially expressed. The expression of Glyma09g08330 and Glyma15g19910 was found to be up-regulated. The others were down-regulated.

Quantitative Real-time PCR (qRT-PCR) Confirmation

To confirm the reliability of Solexa/Illumina sequencing technology, 15 candidate genes were confirmed by qRT-PCR using RT-primers (Table S11). We used the geNorm software to evaluate the stable of housekeeping genes. TUA5 and UBQ10 were ranked the most stable of all samples. UKN2, ACT2/7 and HDC consistently ranked poorly. The optimal numbers of housekeeping genes required for RT-PCR data normalization were determined by the geNorm (Table S12). Finally, three housekeeping genes (TUA5, UBQ10, and UKN2) were selected to normalize the level of gene expression (V3/4 = 0.133<0.15) (Table S13). Differential expression was observed for all the candidate genes, indicating that they are involved in the regulatory networks that are active during low-K stress condition. The results of RT-PCR showed that the qRT-PCR assessments (relative expressed level) of these 15 genes (tag-mapped genes) were consistent with those of Solexa sequencing analysis (Figure 7).

Discussion

Soybeans require a great deal of soluble potassium. The mechanism of high K use efficiency of plant was complicated; it can be affected by different factors, such as plant’s ability to take up, transport, distribute, and utilize K. Therefore, screening the soybean germplasm for highly K-efficient varieties and varieties with improved of K-nutritional traits may be an effective way to relieve the plant’s demand for potassium. In our experiment, the growth of five soybean varieties with different low-K tolerance were reduced to different degrees as the concentration of K decreased. Low-K-sensitive varieties were affected more than the low-K-tolerant varieties. You06-71 has a high dry weight, high potassium accumulation, high KUE, and low potassium content under the low-K stress conditions, but the variety HengChun04-11 does not. These results indicate that variety You06-71 has a greater capability to take up K and use it efficiently than HengChun04-11. The selection of a suitable variety was critical to the following sequencing experiment.

The objective of the present study was to perform a preliminary exploration of transcripts involved in the development of soybean roots and shoots under low-K stress conditions and to provide background work for investigations of the mechanism behind their regulation. There are few existing reports showing the transcriptional changes associated with the absorption and transportation of potassium in soybean root and shoot transcript profiles under low-K stress conditions, and those that are available do not include deep sequencing results. For this reason, we used the Solexa system to analyze differential transcript abundance and regulations in response to low-K stress conditions in two soybean varieties (You06-71 and HengChun04-11). In our experiments, the differentially expressed genes were screened using transcriptome annotation. We discovered that several were involved in the regulation of the absorption and transportation of potassium and of metabolic pathways. Several other biological processes that had not previously been linked to K stress, such as natural killer-cell-mediated cytotoxicity, thiamine metabolism, monoterpenoid biosynthesis, zeatin biosynthesis, the phosphatidylinositol signaling system, and the biosynthesis of flavone and flavonol, were dramatically altered during K-stress [56–59]. These may be novel genes relevant to potassium use efficiency (KUE) in soybeans.

In the present study, many potential potassium-regulated genes were identified in roots and shoots. In plants, potassium processing...
involves several steps, including uptake, translocation, and recycling [60,61]. Some of the candidate genes identified in this study may play important roles in these processes. These include putative potassium transporters, potassium channel, and potassium channel regulatory factors. The K⁺ uptake permeases (KUPs) encode plasma membrane K⁺/H⁺ symporters [62]. They have been found in plasma membranes of various cell types and are thought to catalyze K⁺ influx in to cells under low-apoplastic-K⁺ concentrations [63]. The transcript representing the KUP-related potassium transporter was Glyma08g09720, which showed up-regulated expression (4.95-fold) in the low-K-tolerant variety, implying that this gene might assist the KUP gene family in K⁺ absorption under low-K stress conditions. AKT2 encoded voltage-independent K⁺-channels expressed in the phloem and xylem parenchyma. It has been implicated in both the loading and unloading of the phloem and is responsible for the recirculation of K⁺ from the shoots to the roots [64]. A potassium channel transcript (Glyma08g20030) that was determined by a sequence analysis to represent the AKT2 gene was hypothesized to be linked to the ion transport process. AtKC1 is a shaker-like channel family gene. It has been identified as an important, dual-affinity, inward-rectifying K⁺ channel that participates in K⁺ uptake from soil into cells in Arabidopsis roots [65]. Recent studies using electrophysiological techniques demonstrated that AtKC1 could not form a functional K⁺ channel by itself but that it might form one with AKT1. The inhibition of AKT1-mediated inward K⁺ currents was significantly enhanced by the injection of AtKC1 cRNA [66]. In our study, the transcript representing the AtKC1 gene (Glyma05g33660) was down-regulated in the low-K-tolerant variety. In the short-term, the expression level of this gene was higher in HengChun04-11 than in You06-71. This might explain why soybean You06-71 is tolerant to low-K conditions and soybean HengChun04-11 is not. KAT1 is a highly selective inward-rectifying potassium channel. This voltage-gated channel mediates long-term potassium influx into guard cells, leading to stomatal openings [67]. Expression of the transcript representing the KAT1 gene, Glyma15g10140, was found to be up-regulated (1.63-fold, 3.39-fold) in the low-K-tolerant variety during both periods, suggesting that this variety was better able to perceive potassium deficiency, open the stoma, and stimulate potassium transporta-
tion. Another putative potassium transporter may play a key role in K⁺ assimilation and translocation and could improve stress tolerance in soybeans.

We also noticed a significant phenomenon involving several transcription factors. A single transcription factor can regulate the expression of many genes. Transcriptional control of the expression of stress-response genes is a crucial part of the plant response to a range of abiotic and biotic stresses [68]. This means that transcription factors can show some complexity and overlap in their responses to different stresses and are likely to lead to new ways of enhancing crop tolerance to environmental stress [69]. Some transcription factors are related to the development of potassium-related metabolic pathways [70]. For example, the expression level of the Glyma09g08060 (WRKY50) (2.97-fold) gene was up-regulated in the low-K-tolerant variety over the short term. WRKY50 is a defense-related transcription factor belonging to the WRKY family of proteins [46]. Armengaud et al. pointed out that increased jasmonic acid (JA) levels in K⁺-starved plants might repress otherwise inducible defense responses [71]. Although WRKY50 proteins have not been reported to regulate the expression of JA signaling components, the absence of these proteins relieves this repression and restores JA-derived signaling in plants [72]. In this way, WRKY50 may have positive effects under low-K stress conditions. In our study, several AP2/EREBP (apetala 2-ethylene-responsive element binding proteins) transcription factors were found to be among our DEG. The AP2/EREBP family proteins are unique to plants and share a highly conserved region of about 60 to 70 amino acids [44]. Many AP2/EREBP transcription factors can activate genes in response to abiotic stress and are involved in the ABA response [43,73]. The expression levels of these genes were up-regulated in the roots and shoots of the low-K-sensitive plant varieties. We assume that these genes can indirectly regulate the potassium channel involved in K-metabolic pathways in soybeans. The function of this gene will be studied by RNA-interference or by over-expression in transgenic plants in the future. In addition, we found many members of the MYB transcription factor family to be differentially expressed between two varieties (i.e., Glyma11g22960, Glyma07g05960, etc.). Some reports have suggested that these genes regulate a wide array of processes in plants, including the synthesis of secondary metabolites, different aspects of plant development, abiotic stress responses, and other processes [41]. This indicates that the soybean MYB family genes may be responsive to K-stress. One gene encoding GATA transcription factors (Glyma1lg04060) was found to be up-regulated (3.51-fold) in both soybean varieties. Many GATA factors can activate or inactivate genes in response to environmental deficiencies [74]. For this reason, we assume that this gene is involved in low K- resistance in soybeans.

We also found many protein kinases to be differentially expressed. Among these, plant Ca²⁺ sensors. These included calmodulin (CaM), calcineurin B-like proteins (CBLs), CBL-interacting protein kinases (CIPKs), and Ca²⁺-dependent protein kinases (CDPKs). CDPKs are Ca²⁺ sensors unique to higher plants. They are structurally different from Ca²⁺-regulated protein kinases in other species [75]. The CDPKs have been proven to be involved in regulating plant responses to environmental stresses in different species [76,77]. CIPK23 has been shown to directly phosphorylate the K⁺ transporter AtAKT1. CIPK23 is activated by the binding of two calcineurin B-like proteins, CBL1 and CBL9 [78]. The CBL1/9-CIPK23 pathway efficiently improves the activity of AKT1 and enhances K⁺ uptake under low-K stress conditions. The tag representing the CBL1/9 gene (Glyma05g05580-2.10-fold), CIPK23 gene (Glyma14g04430-1.83-fold), and CDPKs (Glyma14g02680-16.23-fold, Glyma04g38150-2.31-fold, and Glyma14g00320-2.47-fold) was found to be up-regulated in both varieties. We hypothesize that these genes might make an important contribution to adaptation to low-K stress conditions in soybeans. In addition, wall-associated kinases are also involved in various processes in plants, including pathogen resistance, heavy-metal tolerance, and organ development [79]. However, their role in tolerance to potassium deficiency is only minimally understood. In our DGE results, three genes encoding the wall-associated kinases were found to be up-regulated under low-K stress conditions. Meanwhile, two genes encoding MAPK-kine

(lyma17g06020, Glyma15g18860) were up-regulated in the low-K-tolerant variety. Recent investigations have confirmed major roles for defined MAPK pathways in plant development, cell proliferation, and hormone physiology, and both biotic and abiotic stress signaling [80]. These genes might play an important role in plant growth and development as well as in hormone and stress responses, and might be important for adaptation to low-K stress condition in soybeans.

Several other regulated genes were observed under low-K stress conditions. The ethylene response protein (Glyma16g01500-2.24-fold) showed short-term up-regulation of expression levels in the roots. Ethylene regulates many cellular and developmental processes in plants through a signaling pathway conserved in monocots and dicots [81]. Reactive oxygen species (ROS) are
known to be involved in low-potassium signaling pathways [52]. Some reports have shown that ethylene acts upstream of reactive oxygen species in response to potassium deprivation [52]. Low-K⁺ inducible HAK5 expression is dependent on the production of ROS [52]. HAK5 also has an important role in the inducible high-affinity K⁺ uptake system in Arabidopsis roots [83]. These observations suggest the possibility that the ethylene-ethylene response protein-ROS-HAK5-dependent pathways are involved in low-K⁺-induced plant responses and that the ethylene response protein is a key element in the integration of signals for the regulation of defense response genes.

Five genes were predicted to be members of the ATP binding cassette (ABC) transporter family. All five were differentially expressed between the two varieties. The ABC superfamily is a large, diverse group of ubiquitous proteins, most of which mediate transport across biological membranes. ABC transporters have been shown to function not only as ATP-dependent pumps, but also as ion channels and channel regulators [84]. Several plant ABC genes have been found to participate in the abiotic stress response [85]. These genes might be helpful in investigating the mechanism of the response to low-K stress conditions.

One gene encoding BURP domain protein (Glyma14g20440) was found to be differentially expressed in the two varieties. We identified all 23 members of the BURP family in soybeans and found 17 to be stress-responsive, responding to at least one of the three stress treatments (ABA, NaCl, PEG) [86]. Previous research revealed that BURP-domain protein family genes might be important for plants responding to stress conditions. The soybean BURP gene may be K-responsive to low-K stress condition. Plants have evolved diverse adaptive physiological and biochemical mechanisms to resist various stressors. Consequently, that stress can alter the expression levels of related genes.

The present study investigated a large number of transcripts and identified a large number of differentially expressed genes related to the mechanisms of potassium regulation in two soybean genotypes under low-K stress conditions. We used Solexa/Illumina deep-sequencing technology. However, many ambiguous tags and unannotated genes were discounted. Future improvements to the knowledge base of soybean genome sequences may resolve some of these ambiguities, leading to the discovery of large numbers of genes related to potassium regulation. Further verification of these differentially expressed genes via transgenic technology could lead to improvement of the soybean’s resistance to low-K stress condition and its KUE.

Materials and Methods

Selection of Plant Materials and Growth Conditions

The identification of a low-K tolerant variety was an important step in our study. Forty soybean varieties (lines) were collected from the Oil Crops Research Institute at the Chinese Academy of Agricultural Science (Table S1). Sterilized seeds were pregerminated on moistened filter paper in a plant growth chamber at 60% humidity and 28°C, under a 16/8 h (light/dark) cycle for 3–4 days. Then the seedlings were transferred into plastic pots with 1/2 strength modified Hoagland nutrient solution (2 mM Ca(NO₃)₂·4H₂O, 2.5 mM KNO₃, 0.5 mM NH₄NO₃, 0.5 mM KH₂PO₄, 1 mM MgSO₄·7H₂O, 0.05 mM Fe-EDTA, 0.005 mM KI, 0.1 mM H₃BO₃, 0.1 mM MnSO₄·H₂O, 0.03 mM ZnSO₄·7H₂O, 0.0001 mM CuSO₄·5H₂O, 0.001 mM Na₂MnO₄·2H₂O, 0.0001 mM CoCl₂·6H₂O). Their cotyledons were removed to eliminate any additional supply of nutrition. When the plants reached the 2-leaf stage, they were treated with 1/2 Hoagland nutrient solution with different concentrations of K⁺ (K1 level: 0.2 mmol L⁻¹, K2 level: 0.5 mmol L⁻¹, K3 level: 1.0 mmol L⁻¹, K4 level: normal K concentration). The low K nutrient solution was made by a modification of 1/2 Hoagland nutrient solution: the KNO₃ and KH₂PO₄ in the 1/2 Hoagland nutrient solution were replaced by NH₄NO₃ and NH₄H₂PO₄ respectively; the final [K⁺] was adjusted by adding KCl [87]. The culture solution was refreshed every 4 days. We sampled the plant tissue at 6-day intervals. Stress symptoms (yellow leaves and small plants) were observed within 12 days. Different parts of the plant were sampled and dried to assess plant dry weight and potassium content using an inductively coupled plasma emission spectrometer for preliminary evaluation of low-K-tolerant soybean varieties. This experiment was conducted once. The relevant formulas are as follows [88]: (1) K accumulation = K content × dry weight; (2) K use efficiency (KUE): dry weight (g) ×1000/K content (mg/g).

In our experimental treatment, four K⁺ concentrations (K1, K2, K3 and K4) were selected for our nutrient solution. Among these concentrations, K4 represented a level sufficient for normal plant growth [89]. K1, K2 and K3 here represent the low-K⁺ level for plant growth. We found that an excessively low K⁺ concentration (K1 level) could lead to abnormal growth and plant death. We used K2 and K4 to represent low-K (-K) and high-K (+K) conditions.

After preliminary screening experiments, three low-K-tolerant varieties and two low-K-sensitive varieties were observed. We selected seeds from You06-71 (low-K-tolerant) and HengChun04-11 (low-K-sensitive) soybean varieties and grew them in 1/2 Hoagland solution. When plants reached the 2-leaf stage, they were treated with 1/2 Hoagland nutrient solution at K2 and K4 levels. The roots and shoots were harvested separately after 0.5, 2, 6, and 12 h and after 3, 6, 9, and 12 days of this treatment. Each sample was derived from at least three independent plants and three biological repeats. The plant tissues were frozen in liquid nitrogen and kept at -80°C until RNA isolation. Total RNA was extracted from the root and shoot samples using the method described by Liu in 2006 [90].

Library Construction and Sequencing

At least 10 µg of total RNA (≥300 ng/µl) was sent to the Beijing Genomics Institute for Solexa sequencing (commercial service). The experimental process included sample preparation and sequencing. Ten micrograms of total RNA was extracted, and oligo (dT) magnetic adsorption beads were used to purify mRNA. These were then used as primers to synthesize the first and second-strand cDNA. The 5’ ends of the tags were generated using two types of endonuclease: NlaIII and DpnII. The bead-bound cDNA was then digested with restriction enzyme NlaIII, which recognizes and cuts off the CATG sites. All non-3’ cDNA fragments connected to oligo beads were washed away and an Illumina adaptor1 was ligated to the sticky 5’ end of each digested bead-bound cDNA fragment. The junction of the Illumina adaptor1 and the CATG site is the recognition site of Mmel, which is a type of endonuclease with separate recognition and digestion sites. It cuts 17 bp downstream of the CATG site, producing tags with adaptor1. After the removal of the 3’ fragments through magnetic bead precipitation, an Illumina adaptor2 was ligated to the 3’ ends of the tags, producing tags with different adaptors at both ends to form a tag library. After 15 cycles of linear PCR amplification, 95 bp fragments were purified by 6% TBE PAGE gel electrophoresis. After denaturation, the single-chain molecules were fixed onto the Illumina Sequencing Chip (flow cell). Each molecule grew into a single-molecule cluster-sequencing template through in situ amplification. Then four types of nucleotides, each labeled with...
four colors, were added and sequencing was performed by synthesis (SBS). Each tunnel generated millions of raw reads with a length of 35 bp. Then the raw data (tag sequences and counts) were produced. The raw sequences had 3’ adaptor fragments as well as a few low-quality sequences and several types of impurities. Raw sequences were transformed into clean tags after certain data-processing steps: (1) Removal of 3’ adaptor sequence: Because tags are only 21 bp long and the sequencing reads are 35 bp long, raw sequences include 3’ adaptor sequences. (2) Removal of empty reads (reads with only 3’ adaptor sequences but no tags). (3) Low-quality tags removal (tags with unknown sequences “N”). (4) Removal of tags that are too long or too short, leaving tags of exactly 21 bp. (5) Removal of tags with a copy number of 1 (likely sequencing error). (6) Generation of clean tags. “Clean tags” here refers to the tags used for analysis after filtering the dirty tags. “Distinct tags” include all types of clean tags. The underlying principles and specific step of tag preparation are shown in Table S14.

Data Processing and Digital Tag Profiling
Sequencing-received raw image data was transformed by base calling into the sequence data, which is called raw data or raw reads, and it is stored in FASTQ format. These types of files store information about read sequences and quality. Each read is described in four lines in FASTQ files. Raw sequence reads were filtered by the Illumina pipeline. Raw sequences have 3’ adaptor fragments, a few low-quality sequences, and several types of impurities. All low-quality tags, such as short tags (<21 nt), empty reads, overly long tags, and tags with only a single copy were removed. Raw sequences were transformed into clean tags after data processing. To map DGE tags to the soybean genome (Glycine max), we created virtual libraries containing all the possible sequences 17 bases in length that were located next to a NlaIII restriction site (CATG). These high-quality sequences were mapped to genome reference sequences (http://phytozome.net/) (Phytozome V6.0). To monitor mapping events on both strands, the sense and complementary anti-sense sequences were included in the data collection process. During mapping, any transcript detected more than once during these experiments was excluded.

Gene Annotation
Currently, the full genome sequence for the soybean (Glycine max) is available (http://www.soybase.org/). We used it to align and identify clean sequencing reads. Virtual libraries were built containing all possible CATG+17-base sequences in reference gene sequences. All clean tags were mapped to the reference sequences and only mismatches of no more than 1 bp were considered. Clean tags mapped to reference sequences from multiple genes were filtered. The remaining clean tags were designed to be unambiguous. The number of unambiguous clean tags per gene was calculated and then normalized to TPM. In this way, we have provided a more accurate and scientifically sound means of measuring the level of gene expression in these two varieties under low-K stress conditions.

Statistical Evaluation of Differential Gene Expression
Statistical comparison was performed to identify differentially expressed genes between the libraries using the rigorous algorithm method described by Audic [24]. Gene expression was normalized to transcripts per million (TPM) clean tags. For gene expression variance, T-testing was performed to determine significant differences. P values corresponding to differential gene expression were assessed. The false discovery rate (FDR) method was used to determine the threshold P value through a multiple test and analysis as per Benjamini [91]. We used FDR ≤ 0.001 and the absolute value of $|\log2\text{Ratio}| ≥ 1$ as a threshold to judge the significance of differences in gene expression. More stringent criteria, using smaller FDR values and bigger fold-change values, were used to identify DEGs.

Quantitative RT-PCR Analysis
Analysis of the gene expression was performed via real-time PCR as described previously [92]. Evidence shows that transcription levels of housekeeping genes can vary considerably in response to changes in experimental conditions and across different types of tissue types [93,94]. For this reason, we selected five housekeeping genes to evaluate the levels of gene expression. These five genes were TUA5, ACT2/7, UBQ10, UKN2, and HDC. The stability of housekeeping gene expression was analyzed with geNorm software (V3.50). We determined the value of stability measure (M) via a stepwise exclusion of the least stable housekeeping gene, created a stability ranking, and then used the relatively stable housekeeping gene to normalize the level of selected gene expression [95].

Selected gene sequences were obtained from the soybean sequence database. Primers were designed using Primer Premier 5.0 software (Premier Biosoft International, Palo Alto, CA, U.S.) and were synthesized commercially (Shanghai Sangon Biological Engineering Technology & Services Co., Ltd., Shanghai, China). RT-PCR was conducted with gene-specific primers and housekeeping gene primers (Table S11). real-time PCR reactions were performed with a Bio-Rad real-time thermal cycling system (Bio-Rad, Hercules, CA, U.S.) using SYBR-Green to assess gene expression. The reaction mixture (20 µl) contained 0.5 µl of each primer and the appropriate amounts of enzymes, cDNA, and fluorescent dyes. All runs used a negative control without any additional target cDNA, resulting in a lack of detectable fluorescence. A range of five dilutions of the total cDNA was tested under the same conditions. Amplification reactions were initiated with a pre-denaturing step at 95°C for 10 s followed by 40 cycles of denaturing (95°C for 5 s), annealing (55°C for 10 s), and extension (72°C for 15 s) during the second stage, and a final stage of 55°C to 95°C to determine the dissociation curves of the amplified products. All reactions were performed with at least three replicates. The relative quantitation of gene expression was calculated and normalized to housekeeping genes. Data were analyzed using Bio-Rad CFX Manager software and MS Excel 2003.

Supporting Information
Figure S1 Sequencing saturation analysis of the eight libraries in (A) HengChun04-11 and (B) You06-71. The number of detected genes was found to increase as the total number of tags increased.

Table S1 Soybean varieties used for screening.

Table S2 Expression of anti-sense genes detected from the eight libraries.

Table S3 Parts of pathways used for the DEGs.

Table S4 Unknown genes differentially expressed in the roots and shoots of two varieties of soybeans under low-K stress conditions.
Table S5 Differentially expressed genes involved in the soybean roots (Glycine max) and shoot development under low K stress conditions. (XLS)

Table S6 Predicted genes differentially expressed in the roots and shoots of two varieties of soybeans under low-K stress conditions. (XLS)

Table S7 Differentially expressed genes found in both roots and shoots. (XLS)

Table S8 Transcription factors involved in genes differentially expressed in the roots and shoots of two varieties of soybeans under low-K stress conditions. (XLS)

Table S9 Protein kinases involved in genes differentially expressed in the roots and shoots of two varieties of soybeans under low-K stress conditions. (XLS)

Table S10 Oxidative stress-related genes in differentially expressed in the roots and shoots of two varieties of soybeans under low K stress conditions. (XLS)

References
1. Adams F (1974) Soil solution. The plant root and its environment Univ Press of Virginia, Charlottesville: 441–481.
2. Tisdale S, Nelson W, Beaton J, Hartlin J (1993) Soil and fertilizer nitrogen. Soil fertility and fertilizers 5: 109–175.
3. Schroeder JI, Kwap JM, Allen GJ (2001) Guard cell abscisic acid signalling and engineering drought hardiness in plants. Nature 410: 327–330.
4. Assmann SM (2003) Open stomatal opens the door to ABA signaling in Arabidopsis guard cells. Trends Plant Sci 8: 151–153.
5. Sutter JU, Sieben C, Hartel A, Eisenach C, Thiel G, et al. (2007) Abscisic acid triggers the endocytosis of the Arabidopsis KAT1 K+ channel and its recycling to the plasma membrane. Curr Biol 17: 1396–1402.
6. Jungk A, Claassen N (1997) Ion diffusion in the soil-root system. Adv Agron 61: 53–110.
7. Akáraz R, Marco F, Cuevas JC, Patron M, Ferrando A, et al. (2006) Involvement of polyanines in plant response to abiotic stress. Biotechnol lett 20: 1867–1876.
8. Ashley M, Grant M, Grabov A (2006) Plant responses to potassium deficiencies: a role for potassium transport proteins. J Exp Bot 57: 425–436.
9. Ward JM, M ser P, Schroeder JI (2006) Plant ion channel: gene families, physiology, and functional genomics analyses. Annu Rev Physiol 71: 59–82.
10. Pyo YJ, Gieth M, Schroeder JI, Cho MH (2010) High-affinity K+ transport in Arabidopsis: AtHAK5 and AKT1 are vital for seedling establishment and postgermination growth under low-potassium conditions. Plant Physiol 153: 863–875.
11. Lebandy A, Very AA, Sentenac H (2007) K+ channel activity in plants: genes, regulations and functions. FEBS lett 581: 2357–2366.
12. Gieth M (2007) Potassium transporters in plants-Involvement in K+ acquisition, redistribution and homeostasis. FEBS Lett 581: 2348–2356.
13. Chen YF, Wang Y, Wu WH (2006) Membrane transporters for nitrogen, potassium and sodium: a review of the literature. Am J Med 98: 75–84.
14. Winkel-Shirley B (2002) Biosynthesis of flavonoids and effects of stress. Curr Opin Plant Biol 5: 218–223.
15. Libault M, Farmer A, Brechenmacher L, Drnevich J, Langley RJ, et al. (2010) Complete transcriptome of the soybean root hair cell, a single-cell model, and its alteration in response to Bradyrhizobium japonicum infection. Plant physiol 152: 454–462.
16. Bryan J, Aguilar-Bryan L (1999) Sulfonylurea receptors: ABC transporters that control ATP-sensitive K+ channels. Bsa-Biomembranes 1461: 285–303.
36. Fragu G, Cluwa NH (2002) Ubiquitin-mediated proteolysis in plant hormone signal transduction. Trends Cell Biol 12: 308–311.
37. Ma L, Sun N, Liu X, Jiao Y, Zhao H, et al. (2005) Organ-specific expression of Arabidopsis genome during development. Plant Physiol 138(1): 80–91.
38. Schmidt M, Davin T, Shen SR, Page TJ, Demar M, et al. (2005) A gene expression map of Arabidopsis thaliana development. Nat Genet 37: 501–506.
39. Ache P, Fromm J, Koroleva O, Ache M, et al. (2002) A gene expression map of Arabidopsis thaliana development. Nat Genet 37: 501–506.
40. Veřy AA, Sentenac H (2003) Molecular mechanisms and regulation of K+ transport in higher plants. Annu Rev Plant Biol 54: 575–603.
41. Stracke R, Werber M, Weisshaar B (2001) The R2R3-MYB gene family in Arabidopsis thaliana.Curr Opin Plant Biol 4: 447–456.
42. Karam B, Rhonda C (2002) Transcription factors in plant defense and stress responses. Curr Opin Plant Biol 5(3): 218–223.
43. Ricarda D, Roman M, Jochen H, Mathias B (2001) Reduction of potassium nutrition for glycolysis and nitrogen assimilation in Arabidopsis roots. Multilevel analysis of primary metabolism provides new insights into the role of potassium nutrition for glycolysis and nitrogen assimilation in Arabidopsis roots. Plant physiol 150: 772–785.
44. Lee SC, Lan WJ, Kim BG, Li, Cheong YH, et al. (2007) A protein phosphorylation/dephosphorylation network regulates a plant potassium channel. P Natl Acad Sci USA 104: 15959–15964.
45. Berger H, Pachlinger R, Morozov I, Goller S, Narendja F, et al. (2006) The GATA factor AraGATA regulates localization and in vivo binding site occupancy of the nitrate activator NiaR. Mol Microbiol 59: 433–446.
46. Hunter T (1987) A thousand and one protein kinases. Cell 50: 823–9.
47. Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Trends Cell Biol 12: 308–311.
48. Paul A, Wilson S, Belham CM, Robinson CJM, Scott PH, et al. (1997) Stress-activated protein kinases: activation, regulation and function. Cell Signal 9: 403–406.
49. Jonas EA, Kaczmarek LK (1996) Regulation of potassium channels by protein kinases. Curr Opin Neurobiol 6: 318–323.
50. Eulgem T, Somischa IE (2007) Networks of WRKY transcription factors in defense signaling. Curr Opin Plant Biol 5(3): 218–223.
51. Schrader LA, Berbaum SG, Nadin BM, Ren Y, Bui D, et al. (2006) ERK/ MAPK regulates the Kv1.2 potassium channel by direct phosphorylation of the pore-forming subunit. Am J Physiol Cell Physiol 290: C532-C540.
52. Shin R, Berg RH, Schachtman DP (2005) Reactive oxygen species and root hairs in Arabidopsis root response to nitrogen, phosphorus and potassium deficiency. Plant Cell Physiol 46: 1350–1357.
53. Ricarda D, Roman M, Jochen H, Mathias B (2001) Reduction of potassium nutrition for glycolysis and nitrogen assimilation in Arabidopsis roots. Plant physiol 150: 772–785.
54. Lee SC, Lan WJ, Kim BG, Li, Cheong YH, et al. (2007) A protein phosphorylation/dephosphorylation network regulates a plant potassium channel. P Natl Acad Sci USA 104: 15959–15964.
55. Lee SC, Lan WJ, Kim BG, Li, Cheong YH, et al. (2007) A protein phosphorylation/dephosphorylation network regulates a plant potassium channel. P Natl Acad Sci USA 104: 15959–15964.
56. Veřy AA, Sentenac H (2003) Molecular mechanisms and regulation of K+ transport in higher plants. Annu Rev Plant Biol 54: 575–603.
57. Moreatta A, Bottino C, Vitale M, Pende D, Cantonii C, et al. (2001) Activating receptors and coreceptors involved in human natural killer cell-mediated cytolysis. Annu Rev Immunol 19(1): 197–230.
58. Ruchook M, Nowakova E, Ostrouchova K (2000) Modulation of thiamine metabolism in Zea mays seedlings under conditions of abiotic stress. J Exp Bot 51(15): 4133–4143.
59. Ricarda D, Roman M, Jochen H, Mathias B (2004) Reduction of potassium content and phosphatidylinositol 3-kinase-dependent Akt phosphorylation by tumor necrosis factor-a rescues axotomized retinal ganglion cells from retrograde cell death in vivo. J Neurosci 21(6): 2058–2066.
60. Winkel-Shirley B (2002) Biosynthesis of flavonoids and effects of stress. Curr Opin Plant Biol 5(3): 218–223.
61. Very AA, Sentenac H (2003) Molecular mechanisms and regulation of K+ transport in higher plants. Annu Rev Plant Biol 54: 575–603.
62. Masar M, Gierth M, Schroeder JI (2005) The potassium transporter AtHAK5 functions in K+ deprivation-induced high-affinity K+ uptake and AKT1 K+ channel contribution to K+ uptake kinetics in Arabidopsis roots. Plant Physiol 137: 1103–1114.
63. Apel K, Hirt H (2004) Reactive oxygen species, mitochondrial oxidative stress, and signal transduction. Annu Rev Plant Biol 55: 373–399.
64. Armengaud P, Sulpcze R, Miller AJ, Stitt M, Amtmann A, et al. (2009) Multilevel analysis of primary metabolism provides new insights into the role of potassium nutrition for glycolysis and nitrogen assimilation in Arabidopsis roots. Plant physiol 150: 772–785.
65. Lee SC, Lan WJ, Kim BG, Li, Cheong YH, et al. (2007) A protein phosphorylation/dephosphorylation network regulates a plant potassium channel. P Natl Acad Sci USA 104: 15959–15964.
66. Veřy AA, Sentenac H (2003) Molecular mechanisms and regulation of K+ transport in higher plants. Annu Rev Plant Biol 54: 575–603.
67. Nakamura RL, McKeendrew Jr WL, Hirsch RE, Selbybrook JC, Gaber RF, et al. (1995) Expression of an Arabidopsis potassium channel gene in guard cells. Plant physiol 109: 371–374.
68. Singh KB, Foley RC, Sanchez LO (2002) Transcription factors in plant defense and stress responses. Curr Opin Plant Biol 5(3): 218–223.
69. Cheng NH, Motes CM, Blancaflor EB, Moore M, et al. (2008) Identifying of Related Response to Low K Stress
70. Tran N, Do Van Khoa L, Tran Thi Ba Chanh T, Tran Thanh M, et al. (2003) Regulation of the ABA-sensitive Arabidopsis potassium channel gene GORK in response to water stress. FEBS lett 554: 119–126.
71. Armengaud P, Breitling R, Ammann A (2004) The potassium-dependent transcription of Arabidopsis reveals a prominent role of jasmonic acid in signalling. Plant Physiol 136: 2565–2576.
72. Gao QM, Venugopal S, Navarde D, Kachoo (2010) A Low Oleic Acid-Derived Repression of Jasmionic Acid-Inducible Defense Responses Requires the WRKY50 and WRKY51 Proteins. Plant Physiol155: 464–476.
73. Han J, Hosh S, Arce P, Wenske S, Reesela M, et al. (2003) Regulation of the ABA-sensitive Arabidopsis potassium channel gene GORK in response to water stress. FEBS lett 554: 119–126.
74. Crespo JL, Daacho K, Ushimaru T, Hall MN (2001) The GAT C factor transcription GLN3 and GAT1 in response to salt stress in Saccharomyces cerevisiae. J Biol Chem 276: 34441–34444.
75. Wan B, Lin Y, Tsai T (2007) Expression of rice Ca2+-dependent protein kinases (CDPKs) genes under different environmental stresses. FEBS lett 581: 1179–1184.
76. Itakura T, Ishihara K, Mizuno T, Ezura Y, Vazquez Y, et al. (2001) A Ca2+-dependent protein kinase that endorses rice plants with cold-and salt-stress tolerance functions in vascular bundles. Plant Cell Physiol 42: 1228–1233.
77. Cheng NH, Motes CM, Moore M, et al. (2008) Identifying of Related Response to Low K Stress
78. Scarselli M, Nogueras MM, Valverde P, Martinez-Garcia J, et al. (2004) Ca2+-dependent protein kinase gene from rice, is induced by cold and gibberelin in rice leaf sheath. Plant Mol Biol 55: 541–552.
79. Xiu J, Li JD, Chen LQ, Wang Y, Liu LL, et al. (2006) A protein kinase, interacting with two calciumin B-like proteins, regulates K+ transporter AKT1 in Arabidopsis. Cell 125: 1347–1360.
80. Young ET, Dombek KM, Tachibana C, Ideker T (2003) Multiple pathways are activated in response to low potassium. Biochem J 370: 159–167.