Introduction

Rhus (Anacardiaceae) contains more than 250 species found in temperate and tropical regions worldwide that are characterized by the presence of several compound groups, including tannins, flavonoids, triterpenoids, phenolics, and aromatic alkanes (Djakpo and Yao 2010). Of these, Rhus chinensis is an important plant that has long been used in folk medicine for the treatment of cold, fever, and malaria in China, Japan, and Korea (Rayne and Mazza 2007). Notably, the extract of galls induced by infection with Schlechtendalia chinensis (Bell) is known to be effective in bacterial control since galls are rich in gallotannin (50–70%) and phenolic compounds, such as gallic acid (Djakpo and Yao 2010). The tannin-rich horned galls have extensive use in medicine, chemical engineering, and the food and textiles industries and account for the majority of the yield of gallnut in China (Chen 2000).

In nature, species-specific gall is abnormal plant tissue, and is generally induced by insect parasitism, which then evolves into a relationship between the insect and the host plant akin to mutualism (Cooper and Rieske 2010; Tokuda 2012; Nabity et al. 2013). To disclose the adaptive significance of galls, several hypotheses have been proposed, including defense from natural enemies, protection from the external environment, expansion of the surface area available for consumption, and enhancement of the nutritional quality of plant tissue (Price et al. 1987). Based on the nutrition hypothesis for the adaptive nature, gall insects control the nutrient levels in galls for their own benefit (Hartley and Lawton 1992). Mapes and Davies (1998) reported that the gall-fly larva is needed to maintain early gall development. When the gall inducer dies in a young developing gall, cell division of gall tissue slows and cell differentiation is altered back to more normal stem tissue. Gall-fly mortality and even the timing of gall-fly mortality can have substantial effects on the quality of gall tissue (Sarahe et al. 2008). There is a dynamic biochemical interaction that is usually well balanced for most of the Aphididae between aphids and plants. Gall-forming Aphidoidea trigger and control abnormal growth in the plant to the insects’ advantage, possibly by eliciting vigorous oxidation in selective meristematic tissues, thereby limiting supply of molecular oxygen and inhibiting oxygen-dependent growth-controls (Miles 1999). However, the molecular mechanisms that underlie gall induction and gall development remain unknown. RNA-seq is a powerful and efficient tool to identify differentially expressed genes (DEGs) at the whole-genome level and is especially effective for screening the transcriptome of non-sequenced species (Wang et al. 2009).

In the present study, we investigated the molecular processes involved in gall and leaf development in R. chinensis by high throughput RNA-seq. A range of analysis related to transcriptome analysis, including de novo assembly, gene characterization, gene classification, and gene enrichment, were performed. In order to identify the molecular features that distinguish gall and leaf and to disclose what makes the gall unique, the gall and three types of leaf were compared. Two major questions were addressed: What is the main molecular difference between gall and leaf and which factors are needed to involve in the gall initiation and development?
Materials and methods

Plant material

R. chinensis and Plagiommium maximoviczii (Lindb) T. Kop. are the primary and secondary hosts to which S. chinensis migrates. In mid-April, fundatrices appear at the R. chinensis and induce horned gallnuts that remain small for several months and then rapidly expand in the autumn while the number of female aphids in a gall drastically increases. In the late October, the galls crack open and alate aphids disperse, and their offspring develop on the secondary hosts in winter. In the next spring, alate sexuparae are formed on the secondary hosts and migrate to R. chinensis where they produce sexual males and females. After sexual reproduction, each female produces only one fundatrix that can induce a new gall (Takada 1991).

As reported previously (Wang et al. 2016), galls and leaves were collected at 100 days from gall formation in eight-year-old trees at a natural plantation located in Yanjin county (Yunnan province) of China (28°04′~28°11′ N, 104°14′~104°27′ E). Four types of samples were randomly harvested from different locations of the tree crown and mixed. Of these, three types of tissue samples were collected from trees with galls, which contain galls, leaves that grew at the same branch as the gall taken (GL), and leaves from branches without any gall (LW). In addition, control leaves were collected from trees without any galls (CL) (Figure 1). After collection, galls were immediately cut into two halves, and aphids as well as impurities on the inner gall surface were removed, followed by washing with deionized water and wiping by filter paper. Next, samples were immediately frozen in liquid nitrogen and stored at −80°C. Ten randomly selected galls or leaves per plant were pooled, and measurements were based on three biological and technical replicates, respectively.

RNA extraction, library construction, and RNA-seq

Total RNA was extracted by TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer’s protocol. The concentration of the purified RNA was determined using a UV–VIS spectrophotometer (Quawell Q5000, San Jose, CA, USA), and the purity and degradation of the total RNA were checked with 1% agarose gels. Using Illumina’s TruSeq RNA Sample Preparation Kit (Illumina, Inc., San Diego, CA, USA), the mRNA-seq library was constructed, and mRNA was purified using oligo (dT) magnetic beads. The mRNA was fragmented into small pieces using divalent cations after purification, and the cleaved RNA fragments were used for first strand cDNA synthesis using reverse transcriptase and random primers, followed by second strand cDNA synthesis using DNA polymerase I and RNaseH. Subsequently, the cDNA fragments were subjected to an end repair process and ligation of adapters. The product was purified and enriched by PCR to create the final cDNA library, which was sequenced using an Illumina HiSeq™ 2000 sequencing platform.

Analysis of sequencing results

After sequencing, raw reads were filtered by discarding the reads with adaptor contamination, masking low-quality reads with ambiguous ‘N’ bases and removing the reads in which more than 10% bases had a Qvalue <20 (Tang et al. 2011). Next, the clean reads obtained were assembled into contigs by the Trinity program, in which an optimized k-mer length of 25 was used for de novo assembly (Xie et al. 2012). The contigs were linked into transcripts according to the paired-end information of the sequences. Furthermore, the transcripts were clustered by nucleotide sequence identity, and the longest transcripts were defined as unigenes to eliminate redundant sequences.

Functional annotation

To annotate the unigenes, a BLASTX search (E ≤ 10^-5) with unigenes was performed against protein databases, including non-redundant protein (Nr), Swiss-Prot, Kyoto Encyclopedia of Genes and Genome (KEGG) and Cluster of Orthologous Groups (COG) of proteins, and the best aligning results were used to determine the sequence direction of

![Figure 1. Photo showing sample types. Two types of trees were sampled. The first was a galled tree that was infected with gall-forming aphids (A and B), and the other was an ungalled tree that was uninfected (C). From the two types of trees, four types of tissues were harvested, including gall, GL, LW, and CL. CL represents control leaves that grow on ungalled trees; GL signifies galled leaves that grow in the same branches as galls and are adjacent to galls; LW represents leaves without gall that grow in the gall-free branches in galled trees. Panels A, B, and C represent 100, 155, and 100 days from gall initiation, respectively.](image-url)
the unigenes. ESTScan was used to predict the coding regions and sequence direction when a unigene could not be aligned to sequences in any of the above databases (Grabherr et al. 2011). The expression abundance of the unigenes was represented by the number of fragments per kilobase of transcript per million fragments mapped (FPKM) (Iseli et al. 1999). Additionally, Blast2GO (Grabherr et al. 2011) and WEGO (Conesa et al. 2005) were used for the functional annotation of genes. The KEGG pathway annotation was performed using Blastall software and the KEGG database.

**Detection of SSR and single-nucleotide polymorphism markers**

The assembled sequences were screened to search for potential simple sequence repeat (SSR) markers using MISA (http://pgrc.ipk-gatersleben.de/misa/). The principle requirement for the identification of perfect dinucleotide motifs was a minimum of six repeats, and for tri-, tetra-, penta-, and hexa-nucleotide motifs, a minimum of five repeats (Ye et al. 2006; Wei et al. 2011). In addition, we used SOAPsnp to detect single-nucleotide polymorphisms (SNPs) in the transcripts, which can be identified on the consensus sequence by comparison with the unigenes (Zeng et al. 2010).

**Analysis of DEGs**

To determine the expression levels of unigenes, the ratio of FPKM values was defined as the fold-changes in the expression of each gene, and thus DEGs between two groups were identified (Mortazavi et al. 2008). The false discovery rate (FDR) control method was used to identify the threshold of the P-value for transcript abundance (Reiner et al. 2003). In this study, significant DEGs were identified when the fold change values were above log2 (FPKM of sample/FPKM of control) ≥ 1, and FDR ≤ 0.001.

**Results**

**Transcriptome sequencing and de novo assembly**

RNA from different developmental tissues was pooled to obtain a broad gene library related to gall development; a total of 86.17 million raw reads and 6.71 gigabase pairs were obtained. With quality checks, the average guanine and cytosine (GC) content reached 43.73%. High-quality reads were defined as reads with Q ≥ 20 and no ambiguous ‘N’ by SOAPdenovo (Li et al. 2009). Thus, after removal of adaptor sequences and contaminated or short reads, 43,027,436 high-quality reads were assembled into 88,592 contigs. Furthermore, short-read sequences were assembled into 59,522 unigenes with a mean length of 889 bp using the Trinity de novo assembly program, including 40,588 unigenes (68.19%) with lengths of 100–1000 bp, 11,855 unigenes (19.92%) with lengths of 1000–2000 bp, and 7079 unigenes (11.89%) with lengths >2000 bp (Table 1 and Figure 2). The N50 values of contigs and unigenes were 1224 and 1675 bp, respectively. The number of reads based on the relative position in a gene (5′–3′) exhibited a normal distribution (Figure 3), which indicates a randomly fragmented transcriptome. To facilitate access to the *R. chinensis* transcriptome data, raw paired-end sequence data in FASTQ format were deposited in the NCBI Sequence Read Archive (SRA) database with the accession number SRA302347.

**Functional annotation and classification**

For annotating the assembled unigenes, a threshold of 10^-5 was used when performing a BLASTX search against diverse databases, consisting of the NCBI non-redundant protein (Nr) database, NCBI non-redundant nucleotide sequence (Nt) database, UniProt/Swiss-Prot, KEGG, C OG of proteins...
and Gene Ontology (GO). Next, 63.39% (37,732) of the unigenes were successfully annotated. Based on the BLASTX results, 36,153 (60.74%), 31,128 (52.30%), and 24,856 (41.76%) unigenes displayed significant similarity to Nr, Nt, and Swiss-Prot database, respectively (Table 2). With regard to the species annotation for the unigenes, 8345 (23%) unigenes displayed significant homology with sequences of *Ricinus communis*, *Populus balsamifera subsp. trichocarpa*, and *Amygdalus persica*, respectively. *R. chinensis* and the four species mapped are dicotyledonous, implying that these species may share similar genetic and genomic characteristics. A total of 38,943 coding sequences (CDSs) were successfully identified, of which, 35,833 CDSs were based on BLASTX, and another 3110 gene relied on ESTscan software prediction. The remaining 20,579 unigenes (34.57%) was not identified, possibly because the length of the CDSs in some transcripts is too short, the background database does not have sufficient information, and some transcripts correspond to non-coding RNAs.

Based on GO classification, 42,508 unigenes were divided into three categories, which includes 51.76% in biological processes, 29.16% in cellular components, and 19.08% in molecular functions. The three largest biological processes were cellular processes, metabolic processes, and single-organism process. For the cellular components, the major classifications were cell, cell part, and organelle. Under the molecular functions category, most of the classifications were of the catalytic activity, binding, and transporter activity type. The above-mentioned data imply that gall development is an extremely complicated biological process.

Likewise, 14,409 (38.19%) of the 37,732 unigenes that were successfully annotated were assigned to the COG database based on Nr annotation to predict and classify possible functions, and 25 different functional classes were designated. Of these, the largest group is general function prediction (18.20%), followed by transcription (9.38%), replication, recombination and repair (8.51%), posttranslational modification, protein turnover, chaperones (7.53%), signal transduction mechanisms (7.44%), carbohydrate transport and metabolism (5.53%), and translation, ribosomal structure and biogenesis (5.36%). In this study, 21,880 unigenes (57.99%) had significant matches and were assigned to 128 KEGG pathways (Table S1) through a BLASTX with \( E < 10^{-3} \) against the KEGG database. Of all the categories, the pathways with the highest unigene representation were metabolic pathways (ko01100, 4830 unigenes, 22.08%), followed by biosynthesis of secondary metabolites (ko01110, 2585 unigenes, 11.81%), plant–aphid interactions (ko04626, 1894 unigenes, 8.66%), plant hormone signal transduction (ko04075, 1254 unigenes, 5.73%), and spliceosome (ko03040, 850 unigenes, 3.88%).

### Detection of SSR and SNP markers

Based on the screening of 59,522 unigenes, 7526 sequences containing 8926 SSRs were identified. Of these, 1172 unigenes sequences contain more than one SSR, and 391 unigene sequences contain SSRs in a compound formation. The largest numbers of motifs are of the mononucleotide and trinucleotide type, which account for 40.43% and 29.01% of the SSR sequences, respectively (Table 3). Dinucleotide (1956) is the most abundant repeat type in all SSRs, followed by AG/CT (1139), AAG/CTT (916), AT/AT (670), and ATC/ATG (355), respectively. The number of quad-, penta-, and hexa-nucleotide motifs is relatively less. Furthermore, 73,226 SNPs were identified with a frequency of 1/722, and contain 45,648 transitions and 27,578 transversions. Of these, the majority are of the A/G and C/T type, which account for 31.09% and 31.25% of all SNPs, respectively. The four other types of nucleotide variations, A/C, A/T, C/G, and G/T, account for 9.22%, 11.39%, 7.75%, and 9.31% of the SNPs, respectively (Table 4).

### Sequencing of digital gene expression

To identify the digital gene expression (DGE) profiles involved in gall development, four DGE libraries (Gall, GL, LW, and CL) were generated and sequenced using an Illumina platform. Quality control and filtering of raw data were implemented to decrease data noise, which included removing the sequences of adapters, sequences with a high content of unknown bases and low-quality reads before further analysis. The percentages of clean reads from the four tissues are 99.58%, 99.61%, 99.69%, and 99.87%, respectively. In this study, 73,226 SNPs were identified with a frequency of 1/722, and contain 45,648 transitions and 27,578 transversions. Of these, the majority are of the A/G and C/T type, which account for 31.09% and 31.25% of all SNPs, respectively. The four other types of nucleotide variations, A/C, A/T, C/G, and G/T, account for 9.22%, 11.39%, 7.75%, and 9.31% of the SNPs, respectively (Table 4).
reaches a certain value, which implies that the number of detected genes tends to saturation (Figure 6).

Comparison of gene expression between tissues

As shown in Figure 7, 39,845, 36,389, 35,796, and 39,526 genes were annotated from Gall, GL, LW, and CL tissues, respectively. Of these, 24,951 genes overlapped in all four tissues. In addition, 5985, 1141, 1194, and 2633 genes specifically expressed in only one tissue type. As for gene expression, those genes that satisfied the criteria of more than a twofold difference between tissues and an FDR $\leq 0.001$ were defined as DEGs. Thus, 7979, 1634, and 3190 up-regulated genes, and 6792, 2889, and 6474 down-regulated genes were found when Gall, GL, and LW were compared with CL, respectively (Figures 8 and 9). These DEGs were further analyzed by GO classification (Figure 10). Three GO classifications, cellular processes, metabolic processes, and single-organism processes, are the three largest biological processes. In the category of cellular components, the major classifications were cell, cell part, and organelle. Catalytic activity, binding, and transporter activity formed the biggest group in molecular function. Interestingly, the GO classification for different tissues showed the same pattern, which indicates that these tissues (Gall, GL, and LW) possibly have similar metabolic processes. To verify this conclusion, a pathway enrichment analysis was performed (Figure 11). For all groups of comparisons, the three largest classifications are biosynthesis of secondary metabolites, plant–aphid interactions, and plant hormone signal transduction. Interestingly, starch and sucrose metabolism was enriched both in the Gall vs. GL and GL vs. CL comparisons, which implies that gall-related tissues are more involved in starch and sucrose metabolism, compared to the leaf.

DEGs involved in plant hormone signal transduction

Generally, IAA, GA, CTK, and brassinosteroid (BR) are classified as growth hormones, while ABA, ETH, JA and salicylic acid (SA) are classified as defense hormones. A total of 28 DEGs involved in IAA metabolism were identified, including 3 auxin influx carrier proteins, 5 auxin-responsive proteins, 16 auxin response factors, and 4 SAUR family proteins. Additionally, four isoforms of the two-component response regulator ARR, which plays an important role in CTK signal transduction, were annotated. Of these, the gene expression ratio of three genes (Unigene10325, CL3428.Contig1 and Unigene18453) in Gall vs. CL is noticeably higher than in YL vs. CL and LW vs. CL. Moreover, four gibberellin receptor (GID1) and two DELLA proteins were identified. In the ABA signal transduction category, 19 protein phosphatase 2c and 2 serine/threonine-protein kinase SRK2E-like isoform 1 proteins were found. Five ETH-related genes were identified. Interestingly, the majority of the ABA and ETH-related genes were over expressed in Gall tissues (Table 5).

Functional genes involved in secondary metabolism, plant–aphid interactions, and starch and sucrose metabolism

Overall, plant hormones regulate primary and secondary metabolism. To determine the metabolic differences among tissues, the three largest pathways (secondary metabolism, plant–aphid interactions, and starch and sucrose metabolism) were selected, and 50 unigenes with the highest expression in gall tissues for each pathway were screened. Interestingly, the majority of these 150 unigenes have higher expression levels in Gall than in GL and LW. After further analysis by hierarchical clustering (Figure 13), three clustering patterns were found. Cluster ‘a’ consists of 53 genes that have the highest expression in Gall/CL, the lowest expression in GL/CL, and tempered expression in LW/CL. Cluster ‘b’ consists of 93
genes whose expression is dramatically decreased going from Gall/CL to GL/CL, and then LW/CL. Cluster 'c' only includes four genes, of which three genes (beta-glucosidase 46-like, endo-1, 4-beta-glucanase, polygalacturonase precursor) are only expressed in Gall/CL, and the WRKY transcription factor was significantly increased going from Gall/CL, GL/CL, and through to LW/CL.

In addition, to confirm the reliability of the transcriptome data, the expression levels of 20 unigenes, including ARP, ARF, TCRR, FPG, DPGI, SPK, PP2C, EIP, ER, and LMYC2 (hormone signal transduction); PP1A, BIRK, SPKB, and PPA (plant aphid); ASAS, CBC, and CP (secondary metabolism); and SAI, CSCS, and G8E (starch and sucrose metabolism) were detected by real-time quantitative PCR.

Figure 5. Read distribution mapped on reference genes for each sample.

Figure 6. Sequencing saturation analysis for each sample.
(Figure 14). A significantly positive correlation between qRT-PCR and DEseq transcripts ($R^2 = 0.64$) was found.

**Discussion**

**General characteristics of the transcriptome of R. chinensis**

The significance of the BLAST comparison depends partly on the length of the query sequence; short sequencing reads would rarely be matched to known genes (Novaes et al. 2008). In the present study, 63.39% (37,732) of the unigenes were annotated, which is on the lower end when compared to other uncharacterized plants, such as *Litsea cubeba* (56.00%) (Han et al. 2013), *Siraitia grosvenorii* (59.90%) (Tang et al. 2011), *Phyllostachys heterocycla* (69.75%) (He et al. 2013), *Dendrocalamus sinicus* (78.71%) (Cui et al. 2016), *Dendrocalamus latiflorus* (78.90%) (Liu et al. 2012), and *Mangifera indica* (80%) (Azim et al. 2014). The large number of unmapped unigenes might be due to the short sequence reads from sequencing and the relatively short sequences of the assembled unigenes, most of which probably lack conserved functional domains (Hou et al. 2011). Other possible reasons include the fact that some of these unigenes belong to non-coding RNAs (Liu et al. 2012) and insufficient sequence information in public databases (Hou et al. 2011).

**DEGs involved in plant–aphid interactions**

Galls are products of parasitism by aphids upon the Rhus tree; their development is an extremely complicated physiological process and is implicated in active attack by aphids, passive defense of trees, confrontation between the host and the aphid, and mutual coordination between the two. It is expected that massive transcripts linked to plant–aphid interactions would be enriched in the four tissues, especially in galls. Serine/threonine-protein kinases phosphorylate the OH group of serine or threonine residues, leading to a functional change in the target protein; protein phosphorylation has been identified as one of the most important events in the disease resistance pathway (Cao et al. 2011). For example, the tomato bacterial speck disease resistance gene *Pto* encodes a serine/threonine-protein kinase (Martin et al. 1993a); the bacterial blight disease resistance gene *Xa21* (Song et al. 2012).

![Image of Venn diagram showing the number of DEGs in different tissues.](image1)

**Figure 7.** Venn diagram showing the number of DEGs in different tissues.

![Image of comparison of number of DEGs in different tissues.](image2)

**Figure 8.** Comparison of number of DEGs in different tissues.

![Image of scatter plots showing DEGs in different tissues.](image3)

**Figure 9.** Scatter plots showing DEGs in different tissues. A DEG has to satisfy the following criteria: FDR $\leq 0.001$ and abs ($\log2(Y/X)$) $\geq 1$. Orange, blue and grey dots represent up-regulation, down-regulation and no change, respectively.
In rice, flagellin perception gene *FLS2* (Gómez-Gómez et al. 2001) in Arabidopsis, stem rust resistance gene *Rpg5* in barley (Brueggeman et al. 2008), and stripe rust resistance gene *Yr36* in wheat (Fu et al. 2009) also code for a serine/threonine-protein kinase domain. The network of serine/threonine-protein kinases in plant cells appears to act as a ‘central processor unit’, accepting input information from receptors that sense environmental conditions, phytohormones, and other external factors, and converting it into appropriate outputs, such as changes in metabolism, gene expression, and cell growth and division (Hardie 1999).

![Figure 10. Functional annotation of DEGs based on GO categories.](image)

![Figure 11. Pathway enrichment analysis of DEGs.](image)

The Rich factor represents the ratio of the number of DEGs, accounting for all annotated genes that share the same pathway as the DEGs. Larger values of the Rich factor imply a greater degree of enrichment. The Qvalue is the adjusted P-value by multiple hypothesis testing. The Qvalue ranges from 0 to 1, and values closer to 0 imply more significant enrichment.
| Transcript ID | Log2 ratio (Gall vs. CL) | Log2 ratio (Gl vs. CL) | Log2 ratio (LV vs. CL) | Blast annotation |
|---------------|--------------------------|------------------------|------------------------|------------------|
| CL3586.Contig1 | 1.88 | 0.61 | Null | | |
| CL5074.Contig2 | −0.18 | −2.02 | 0.64 | 0.61 | |
| CL11399 | −0.14 | 0.08 | 0.91 | PREDICTED: auxin-responsive protein IA14-like isoform 1 | |
| CL11202 | 5.35 | 5.35 | −0.32 | 0.27 | |
| Unigene20138 | −0.68 | 1.07 | −2.54 | 0.26 | |
| Unigene19895 | 5.47 | 0.37 | 0.10 | 0.26 | |
| CL3860.Contig3 | 3.19 | 0.49 | 0.36 | 0.36 | |
| CL89.Contig1 | 4.21 | 1.99 | 1.51 | 0.27 | |
| CL5334.Contig1 | 1.53 | 0.77 | 2.20 | 0.77 | |
| Unigene18995 | 1.22 | 0.64 | 2.20 | 0.64 | |
| CL2119.Contig1 | 1.84 | 0.97 | 0.35 | 0.35 | |
| Unigene13899 | 1.11 | 1.69 | 0.27 | 0.27 | |
| Unigene162674 | 0.31 | 1.91 | 0.79 | 0.79 | |
| Unigene14582 | 1.49 | 0.49 | 1.51 | 0.51 | |
| CL725.Contig1 | 4.21 | 1.99 | 0.50 | 0.50 | |
| CL18338 | −4.66 | −1.22 | −0.15 | 0.15 | |
| CL4747.Contig4 | −0.32 | 0.73 | 0.37 | 0.37 | |
| Unigene16436 | −1.71 | −2.01 | 0.04 | 0.04 | |
| CL725.Contig2 | 1.29 | 0.37 | −0.66 | 0.66 | |
| Unigene16506 | 0.63 | 0.56 | −0.48 | 0.48 | |

Notes: The Blast annotation includes the gi number, e-value, gene description, and species for the gene. Null means that value does not exist.
protein kinases were identified, indicating their indispensable role in gall formation.

In the present study, 24 WRKY transcription factors were identified, including WRKY1, WRKY 2, WRKY17, WRKY22, WRKY26, WRKY44, WRKY47, WRKY57, WRKY60, WRKY72, and WRKY74. The large family of WRKY transcription factors exerts an influence over plant growth, development, and plant environment interactions (Ülker and Somssich 2004; Rushton et al. 2010). In Arabidopsis thaliana, AtWRKY44 is involved in trichome development (Johnson et al. 2002), while AtWRKY6, AtWRKY53, and AtWRKY70 are involved in leaf senescence (Robatzek and Somssich 2002; Ülker et al. 2007). AtWRKY3, AtWRKY4, AtWRKY52, AtWRKY53, and AtWRKY70 all participate in the plant’s response to pathogen attack (Pandey and Somssich 2009). AtWRKY22 and AtWRKY29 are two downstream components in a MAPK pathway that is responsible for conferring resistance against both bacterial and fungal pathogens (Asai et al. 2002). A transcript (CL1111.Contig2_Rhuscm) was identified as WRKY transcription factor 22-like, which was highly expressed in LW, mildly expressed in GL, and slightly expressed in Gall. This implies that both leaves and galls show a significant defense response. A transcript (Unigene11387_Rhuscm) identified as WRKY1 was highly expressed in Gall, mildly expressed in LW, and slightly expressed in GL. This is consistent with evidence that the order of withering is GL, LW, and Gall (Figure 1(B)). According to a previous study (Qiao et al. 2016), WRKY1 is associated with stomatal movement in drought stress. Thus, it is possible that WRKY1 regulates water management during gall development.

**DEGs involved in secondary metabolism**

Two shikimate dehydrogenase genes from grapevine, VvSDH3 and VvSDH4, were mainly expressed in immature berry tissues in which galloylated flavan-3-ols were accumulated, and the content of gallic acid, β-glucogallin, and galloylated flavan-3-ols was increased (Bontpart et al. 2016). In contrast, two transcripts encoding shikimate dehydrogenase (CL2974.Contig1 and CL3909.Contig1) that expressed highly in galls were identified in this study, attesting to the fact that shikimate dehydrogenase has an impact on gallic acid metabolism.

In plants, anthranilate synthases are monofunctional, and consist of two non-identical subunits, the alpha (ASA) and the beta subunit (ASB) (Crawford 1989). ASA alone can convert chorismate to anthranilate with ammonia acting as an amino group donor, while ASB acts as an accessory component to ASA. In most plants, there are two non-identical copies of ASA genes that are regulated by different mechanisms (Bohmann et al. 1995; Tozawa et al. 2001). In Arabidopsis, the mRNA expression of ASA1 was approximately 10 times higher than that of ASA2 and was induced by pathogen infection as well as wounding (Niyogi and Fink 1992). At the protein level, the enzymatic activities of ASA are also differentially regulated through feedback inhibition by Trp (Bohmann et al. 1996; Kreps et al. 1996). The regulated expression of ASA genes suggests that the control of anthranilate flux is an important step in regulating the production of Trp (Bohmann et al. 1996). In this study, two transcripts encoding the anthranilate synthase alpha subunit (CL3745.Contig1 and Unigene19101) were identified with abundant expression in galls compared with leaf tissues, suggesting that the anthranilate synthase alpha subunit is possibly associated with gallic acid synthesis.

**DEGs involved in starch and sucrose metabolism**

Fructokinase directly regulates sucrose degradation after the conversion of β-D-fructofuranose and D-fructose-6-phosphate, and the latter is the starting material for glycolysis (Martin et al. 1993b). Additionally, fructokinase performs a regulatory role in photosynthetic tissues, regulating photosynthesis, growth, and senescence (Dai et al. 1999). During hypoxia, overexpression of fructokinase can improve adenylate energy status in tomato root (Gharbi et al. 2007). Additionally, fructokinase significantly is up-regulated during fiber elongation in cotton (Yang et al. 2008). In this study, a fructokinase-2 gene (Unigene15517) was abundantly expressed in gall compared with the other three types of tissues, which implies that up-regulation of the fructokinase-2 gene might be associated with gall formation.

Invertase, also called beta-fructofuranosidase, is a glycoprotein that cleaves the terminal non-reducing beta-fructofuranoside residues. Based on solubility, optimum pH,
isoelectric point, and subcellular localization, plant invertases can be classified into three subgroups: vacuolar (soluble acid), cytoplasmic (soluble alkaline), and cell-wall-bound invertase. Soluble acid invertase has important biological functions in sucrose metabolism and predominantly hydrolyses sucrose for growth and developmental processes (Kulshrestha et al. 2013). In addition, sucrose hydrolysis by soluble acid invertase helps in the regulation of osmotic pressure, which controls cell expansion and depends on the size of vacuole (Ji et al. 2005). Sucrose synthase can catalyze the cleavage of sucrose and produce uridine diphosphoglucose, which is the substrate for cellulose synthesis (Nairn et al. 2008). Three transcripts with abundant expression in gall encoding soluble acid invertase (Unigene4744), sucrose synthase (CL127.Contig1), and cellulose synthase A catalytic subunit 3 (CL1316.Contig1) were identified.

Pectinesterase is a ubiquitous cell-wall-associated enzyme thought to be responsible for the demethylation of galacturonic residues in high-molecular-weight pectin (Phan et al. 2007), which has been implicated in many developmental processes, including cellular adhesion, stem elongation, pollen tube development, abscission, and fruit ripening (Brummell and Harpster 2001; Micheli 2001; Bosch and Hepler 2005). Twelve transcripts encoding pectinesterase were enriched in four types of tissues. Of these, eight transcripts were highly expressed in gall, and four transcripts were substantially expressed in the other three tissues. Given that galls experience a substantial accumulation of biomass when compared with leaves, we speculate that the above-mentioned DEGs, which are involved in sugar and starch metabolism, play a key role in primary substance metabolism.

Endogenous hormones coordinate primary and secondary metabolism

Plants reorganize their metabolism as they establish vegetative structures, acquire nutrients, produce viable offspring, survive drought, and other abiotic stresses, as well as navigate the challenges of maintaining mutualistic relationships while thwarting the advances of parasitic or herbivorous organisms. Phytohormones are important regulators of plant growth and development, in addition to playing an important role in their adaptation to their respective environments. Often, phytohormones are classified by their most prominent function, such as ‘growth hormones’ or ‘defense hormones’ (Schäfer et al. 2016). However, these hormones have also been shown to participate in adaptations different from their classical functions. For example, JAs can regulate flower development (Stitz et al. 2014) and senescence induction (Ueda and Kato 1980), while cytokinins also participate in interactions with pathogens and herbivores (Smigocki et al. 1993; Choi et al. 2010). In addition, many phytohormone pathways interact with each other.

Phytohormone interactions can occur at the signaling level. A classic example is the GA pathway amplifying JA signaling by the binding of the GA-regulated DELLA proteins to JASMONATE ZIM-DOMAIN (JAZ)-suppressors, negative regulators of JA signaling (Hou et al. 2010). Phytohormones can also influence each other at the metabolic level, as is
reported for AXs that can reduce CK levels by promoting cytokinin oxidase/reductase-mediated degradation of CKs (Werner et al. 2006; Carabelli et al. 2007). However, phytohormone pathways are also known to interact at the signaling and metabolic levels. Additionally, phytohormones can simultaneously affect nutritional value and the abundance of defense compounds.

In the present study, abundant enrichment of signal transduction pathways involving seven types of hormones was identified, which is in accordance with our previous study, wherein concentrations for the seven types of hormones showed significant differences. However, the PCA results between hormone concentration and signal transduction do not completely overlap (Figure 12), possibly due to differences between hormone synthesis and signal transduction. Interestingly, ABA has high weight in the PCA results for both hormone concentration and signal transduction, suggesting that it plays a key role in gall formation. In this study, 22 transcripts were related to ABA signal transduction, of which 19 transcripts were annotated as protein phosphatase 2c, and the majority of them retain higher expression in galls compared with other tissues.

According to our previous study (Wang et al. 2016), exogenous GA application can effectively promote gall development. In this study, six DEGs involved in GA signal transduction were found, including gibberellin receptor GID1 (Unigene11910, Unigene16383, Unigene11215, and Unigene11024), DELLA protein GAI1-like isoform 1 (Unigene22181), and DELLA protein SLR1-like (Unigene15152). Both gibberellin receptor GID1 and DELLA protein are components involved in the signal transduction of GA. The gibberellin receptor GID1 binds in a complex with the DELLA protein to form a GA–GID1–DELLA trimer after receiving the GA signal, which induces degradation of the DELLA

Figure 14. RT-qPCR validation of 20 genes involved in hormone signal transduction, secondary metabolism, plant–aphid interactions, and starch and sucrose pathways. Different letters indicate differences at \( P \leq 0.05 \) based on the least significant difference test. Data are the means of three replicates, and each replicate was measured three times. The gene names and the primers used for RT-qPCR analysis are shown in Table 6. ARF: PREDICTED: auxin response factor 5-like isoform 2; ARP: Auxin-responsive protein IAA1, putative; ASAS: anthranilate synthase alpha subunit; BIRK: BRASSINOSTEROID INSENSITIVE 1-associated receptor kinase 1 precursor, putative; CBC: chloroplast biotin carboxylase; CP: cytochrome P450; CSCS: Cellulose synthase A catalytic subunit 3; G8E: glycosyltransferase 8E; DPGI: DELLA protein GAI1-like isoform 1; EP: ethylene-insensitive protein 2-like; ER: ethylene receptor; FPG: F-box protein GID2-like; PP1A: Pathogenesis-related protein 1a; PP2C: protein phosphatase 2c, putative; PPA: hypothetical protein PRUPE_ppa002088mg; SAI: soluble acid invertase; SPK: serine/threonine-protein kinase SRK2E-like isoform 1; SPKB: serine/threonine-protein kinase br1, putative; TCRR: two-component response regulator ARR12-like.
protein by the ubiquitin 26S proteasome. Therefore, the inhibitory effect of DELLA protein on plant growth is reversed (Vierstra 2003; Ueguchi-Tanaka et al. 2007). Of the four unigenes encoding the gibberellin receptor GID1, only Unigene11024 is consistent with the majority of DEGs in being related to secondary metabolism and plant–aphid interactions (Figure 13); therefore, we speculate that Unigene11024 has an important role in GA signal transduction regulating gall development, and its specific function needs to be determined.

Currently, there are no mRNA expression data for *R. chinensis*. In this study, transcriptome data based on RNA-seq technology were obtained and gene expression profiling of four tissues was performed. We identified thousands of DEGs enriched in the biosynthesis of secondary metabolites, plant–aphid interactions, and plant hormone signal transduction, which were abundantly expressed in galls compared with other leaf tissues. Of the seven endogenous hormones, IAA and ABA dominated hormone signal transduction according to PCA analysis. Additionally, the transcriptome data were verified by quantitative real-time PCR (Figure 14). In conclusion, due to aphid attack, hormone signal transduction coordinates primary and secondary metabolism and the defense response, leading to gall induction (Figure 15).
These data lay the groundwork for further genomic research on gallnuts.

Acknowledgements

XMC and KKJ designed the research; HYW and KC conducted the research; JL, HC, ZXY, CW, and QL analyzed the data; HYW, HXW, and SXS wrote the paper. All authors have read and approved the final manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the grant for Essential Scientific Research of Chinese National Non-profit Institute [grant number CAFYBB2014ZD005], the National Natural Science Foundation of China [grant numbers U1402263, 31370651 and 31372266], and the National High-tech RandD Program of China [grant number 2014AA021801].

References

Asai T, Tena G, Plotnikova J, Willmann MR, Chiu WL, Gomez-Gomez L, Boller T, Ausubel FM, Sheen J. 2002. MAP kinase signalling cascade in Arabidopsis innate immunity. Nature. 415(6875):977–983.

Azim MK, Khan IA, Zhang Y. 2014. Characterization of mango (Mangifera indica L.) transcriptome and chloroplast genome. Plant Mol Biol. 85(1–2):193–208.

Bolhmann J, DeLuca V, Eilert U, Martin W. 1996. Anthranilate synthase in pollen tubes. Plant Cell. 17(12):3219–3226.

Bolhmann J, Lins T, Martin W, Eilert U. 1996. Anthranilate synthase from Ruta graveolens (duplicated AS [alpha] genes encode tryptophan-sensitive and tryptophan-insensitive isoenzymes specific to amino acid and alkaloid biosynthesis). Plant Physiol. 111(2):507–514.

Bontpart T, Marlin T, Vialet S, Guiraud JL, Pinasseau L, Meudec E, Sommerer N, Cheynier V, Terrier N. 2016. Two shikimate dehydrogenases, VvSDH3 and VvSDH4, are involved in gallic acid biosynthesis in grapevine. J Exp Bot. 67(11):3357–3350.

Bosh M, Hepler PK. 2005. Pectin methyl transferases and pectin dynamics in pollen tubes. Plant Cell. 17(12):3219–3226.

Bruggeman R, Druka A, Nirmala J, Cavileer T, Drader T, Rostoks N, Mirdologi A, Bennypaul H, Gill U, Kudrna D. 2008. The stem rust resistance gene Rpg5 encodes a protein with nucleotide-binding-site, leucine-rich, and protein kinase domains. Proc Natl Acad Sci U S A. 105(39):14970–14975.

Brummell DA, Harpster MH. 2001. Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. Plant Mol Biol. 47(1–2):311–339.

Cao A, Xing L, Wang X, Yang X, Wang W, Sun Y, Qian C, Ni J, Chen Y, Liu D. 2011. Serine/threonine kinase gene Stk-V, a key member of powdery mildew resistance gene Pm21, confers powdery mildew resistance in wheat. P Natl Acad Sci USA. 108(19):7727–7732.

Carabelli M, Possenti M, Sessa G, Ciolfi A, SassI M, Morelli G, Ruberti I. 2007. Canopy shade causes a rapid and transient arrest in leaf development through auxin-induced cytokinin oxidase activity. Gene Dev. 21(15):1863–1868.

Chen JH. 2000. Situation and prospect of chemical utilitization of gall tannins in China. Chem Ind Forest Prod. 20(2):70–82.

Choi J, Huh SU, Kojima M, Sakakibara H, Paek KH, Hwang I. 2010. The cytokinin-activated transcription factor AR2 promotes plant immunity via TGA3/NPR1-dependent salicylic acid signaling in Arabidopsis. Dev Cell. 19(2):284–295.

Conea A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 21(18): 3674–3676.

Cooper WR, Rieske LK. 2010. Gall structure affects ecological associations of Dryococcus kuriphilus (Hymenoptera: Cynipidae). Environ Entomol. 39(3):787–797.

Crawford IP. 1989. Evolution of a biosynthetic pathway: the tryptophan paradigm. Annu Rev Microbiol. 43(1):567–600.

Cui K, Wang H, Liao S, Tang Q, Li L, Cui Y, He Y. 2016. Transcriptome sequencing and analysis for culm elongation of the world’s largest bamboo (Dendrocalamus sinuosus). PloS one. 11(6):e0157362.

Dai N, Schaffer A, Petreikov M, Shahak Y, Giller Y, Ratner E, Levine A, Granot D. 1999. Overexpression of Arabidopsis hexokinase in tomato plants inhibits growth, reduces photosynthesis, and induces rapid senescence. Plant Cell. 11(7):1253–1266.

Djakko O, Yao W. 2010. Rhus chinensis and Galla Chinensis—folklore to modern evidence: review. Phytother Res. 24(12):1739–1747.

Dubouzet JG, Ishihara A, Matsuda F, Miyagawa H, Iwata H, Wakasa K. 2007. Integrated metabolomic and transcriptional analyses of high-yield tryptophan rice expressing a mutant anthranilate synthase alpha subunit. J Exp Bot. 58(12):3309–3321.

Fu D, Uauy C, Distelfeld A, Blech A, Epstein L, Chen X, Sela H, Fahima T, Dubcovsky J. 2009. A kinase-START gene confers temperature-dependent resistance to wheat stripe rust. Science. 323(5919):1357–1360.

Gharbi I, Ricard B, Rodin D, Maucourt M, Andreieu MH, Bizid E, Smiti S, Brouquisse R. 2007. Effect of hexokinase activity on tomato root metabolism during prolonged hypoxia. Plant Cell Environ. 30(4):508–517.

Gómez-Gómez L, Bauer Z, Boller T. 2001. Both the extracellular leucine-rich repeat domain and the kinase activity of FLS2 are required for flagellin binding and signaling in Arabidopsis. Plant Cell. 13(5):1155–1163.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 29(7):644–652.

Han XJ, Wang YD, Chen YC, Lin LY, Wu QK. 2013. Transcriptome sequencing and expression analysis of terpenoid biosynthesis genes in Litsaea cubeba. PloS one. 8(10):e76890.

Hardie D. 1999. Plant protein serine/threonine kinases: classification and functions. Annu Rev Plant Biol. 50(1):97–131.

Hartley SE, Lawton JH. 1992. Host-plant manipulation by gallinsects: a test of the nutrition hypothesis. J Anim Ecol. 61(1):113–119.

He CY, Cui K, Zhang JG, Duan AG, Zeng YF. 2013. Next-generation sequencing-based mRNA and microRNA expression profiling analyses of sequenced pathways involved in the rapid growth of developing culms in Moso bamboo. BMC Plant Biol. 13:119.

Hou R, Bao Z, Wang S, Su H, Li Y, Du H, Hu J, Wang S, Hu X. 2011. Transcriptome sequencing and de novo analysis for Yesso scallop (Pariapenaeus yessoensis) using 454 GS FLX. PloS one. 6(6):e21560.

Hou X, Lee LCY, Xia K, Yan Y, Yu H. 2010. DELLAs modulate jasmonate signaling via competitive binding to JAZs. Dev Cell. 19(6):884–894.

Iseli C, Jongeneel CV, Bucher P. 1999. ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. Proc Int Conf Intell Syst Mol Biol. 99:138–148.

Ji X, Van den Ende W, Van Laere A, Cheng S, Bennett J. 2005. Structure, evolution, and expression of the two invertase gene families of rice. J Mol Evol. 60(5):615–634.

Johnson CS, Kolveski B, Smyth DR. 2002. TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of Arabidopsis, encodes a WRKY transcription factor. Plant Cell. 14(6):1539–1559.

Kreps IA, Panappa T, Dong W, Town CD. 1996. Molecular basis of [alpha]-methyltryptophan resistance in amt-1, a mutant of Arabidopsis thaliana with altered tryptophan metabolism. Plant Physiol. 110(4):1159–1165.

Kulshreshtha S, Tyagi P, SinhI V, Yadavalli KS. 2013. Invertase and its applications—a brief review. J Pharmacy Res. 7(9):792–797.

Li R, Li Y, Fang X, Yang H, Wang J, Kristiansen K, Wang J. 2009. SNP detection for massively parallel whole-genome resequencing. Genome Res. 19(6):1124–1132.

Liu M, Qiao G, Jiang J, Yang H, Xie L, Xie J, Zhuo R. 2012. Transcriptome sequencing and de novo analysis for ma bamboo (Dendrocalamus latiflorus Munro) using the Illumina platform. PloS one. 7(10):e46766.

Lu H, Gorman E, McNight TD. 2005. Molecular characterization of two anthranilate synthase alpha subunit genes in Camptotheca acuminata. Planta. 221(3):352–360.
Maps CC, Davies PG. 1998. Plant hormones and gall formation by *Eurosta solidaginis* on *Solidago altissima*. In: Čoška G, Mattson WJ, Stone GN, Price PW, editors. The biology of gallinducing arthropods. St. Paul (MN): USDA Forest Service North Central Research Station General Technical Report NC-199: p. 161–172.

Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganal MW, Spivey R, Wu T, Earle ED, Tanksley SD. 1993a. Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science. 262:1432–1432.

Martin T, Frommer WB, Salanoubat M, Willmitzer L. 1993b. Expression of an Arabidopsis sucrose synthase gene indicates a role in metabolization of sucrose both during phloem loading and in sink organs. Plant J. 4(2):367–377.

Micheli F. 2001. Pectin methyltransferases: wall cell enzymes with important roles in plant physiology. Trends Plant Sci. 6(9):414–419.

Miles PW. 1999. Aphid saliva. Biol Rev. 74:81–85.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 5(7):621–628.

Nabity PD, Haus MJ, Berenbaum MR, DeLucia EH. 2013. Leaf-galling insects perform better in galls. Ecol Entomol. 33(3):385–393.

Nairn CJ, Lennon DM, Wood-Jones A, Nairn AV, Dean JFD. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 5(7):621–628.

Nabity PD, Haus MJ, Berenbaum MR, DeLucia EH. 2013. Leaf-galling insects perform better in galls. Ecol Entomol. 33(3):385–393.

Price PW, Fernandes GW, Waring GL. 1987. The adaptive nature of insect galls. Environ Entomol. 16(1):15–24.

Qiao Z, Li CL, Zhang W. 2016. WRKY1 regulates stomatal movement in drought-stressed Arabidopsis thaliana. Plant Mol Biol 91(1–2):53–65.

Rayne S, Mazza G. 2007. Biological activities of extracts from sumac (Rhus spp.): a review. Plant Food Hum Nutr. 62(4):163–175.

Reiner A, Yekutieli D, Benjamini Y. 2003. Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics. 19(3):368–375.

Robatsek S, Somssich IE. 2002. Targets of ARWKR6Y regulation during plant senescence and pathogen defense. Gene Dev. 16(9):1139–1149.

Rushton PJ, Somssich IE, Ringler P, Shen QJ. 2010. WRKY transcription factors. Trends Plant Sci. 15(5):247–258.

Satish D, Catherine B, Warreng A. 2008. Testing the nutrition hypothesis for the adaptive nature of insect galls: does a non-adapted herbivore perform better in galls. Ecol Entomol. 33(3):385–393.

Schäfer M, Brüttig C, Baldwin IT, Kallenbach M. 2016. High-throughput quantification of more than 100 primary and secondary metabolites, and phytohormones by a single solid-phase extraction based sample preparation with analysis by UHPLC–HESI–MS/MS. Plant Methods. 12(1):30.

Sinigaglia C, Neal J, McCanna I, Douglass L. 1993. Cytokininin-mediated insect resistance in *Nicotiana* plants transformed with the ipt gene. Plant Mol Biol. 23(2):325–335.

Song W, Wang G, Chen L, Kim H. 1995. A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. Science. 270(5243):1804.

Stitz M, Hartl M, Baldwin IT, Gaquerel E. 2014. Jasmonoyl-l-isoleucine coordinates metabolic networks required for anthesis and floral attractant emission in wild tobacco (*Nicotiana attenuata*). Plant Cell. 26(10):3964–3983.

Takada H. 1991. Does the sexual female of *Schlechtendalia chinensis* (Bell) (Homoptera: Pemphigidae) “viviparously” produce the Fundatrix? Appl Entomol Zool. 26(1):117–121.

Tang Q, Ma X, Mo C, Wilson IW, Song C, Zhao H, Yang Y, Fu W, Qiu D. 2011. An efficient approach to finding *Serratia* *grosvenori* triterpene biosynthetic genes by RNA-seq and digital gene expression analysis. BMC Genomics. 12(1):343.

Tokuda M. 2012. Biology of Asphondyliini (Diptera: Cecidomyiidae). Entomol Sci. 15(4):361–383.

Tozawa Y, Hasegawa H, Terakawa T, Wakasa K. 2001. Characterization of rice anthranilate synthase a-subunit genes OASA1 and OASA2, tryptophan accumulation in transgenic rice expressing a feedback-sensitive mutant of OASA1. Plant Physiol. 126(4):1493–1506.

Ueda J, Kato J. 1980. Isolation and identification of a senescence-promoting substance from wormwood (*Artemisia absinthium* L.). Plant Physiol. 66(2):246–249.

Ueguchi-Tanaka M, Nakajima M, Katoh E, Ohmiya H, Asano K, Saji S, Hongyu X, Ashikari M, Kitano H, Yamaguchi I. 2007. Molecular interactions of a soluble gibberelin receptor, GID1, with a rice DELLA protein, SLR1, and gibberellin. Plant Cell. 19(7):2140–2155.

Ulker B, Mukhtar MS, Somssich IE. 2007. The WRKY70 transcription factor of Arabidopsis influences both the plant senescence and defense signaling pathways. Planta. 226(1):125–137.

Ulker B, Somssich IE. 2004. WRKY transcription factors: from DNA binding towards biological function. Curr Opin Plant Biol. 7(5):491–498.

Vierstra RD. 2003. The ubiquitin/26S proteasome pathway, the complex last chapter in the life of many plant proteins. Trends Plant Sci. 8(3):135–142.

Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 10(1):57–63.

Wang H, Liu J, Cui K, Chen H, Yang Z, Wu H, Shao S, King-Jones K, Chen X. 2016. Gibberellic acid is selectively downregulated in response to aphid-induced gall formation. Acta Physiol Plant. 38(9):214.

Wei W, Qi X, Wang L, Zhang Y, Hua W, Li D, Lv H, Zhang X. 2011. Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. BMC Genomics. 12(1):451.

Werner T, Köllmer I, Bartrina I, Holst K, Schmülling T. 2006. An efficient approach to finding EST-SSR markers. BMC Genomics. 12(1):451.

Xie F, Burklew CE, Yang Y, Liu M, Xiao P, Zhang B, Qui D. 2012. De novo sequencing and a comprehensive analysis of purple sweet potato (*Ipomoea batatas*) transcriptome. Planta. 236(1):101–113.

Yang YW, Bian SM, Yao Y, Liu JY. 2008. Comparative proteomic analysis provides new insights into the fiber elongating process in cotton. J Proteome Res. 7(11):4623–4637.

Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L. 2006. WEGO: a web tool for plotting GO annotations. Nucleic Acids Res. 34(Suppl. 2):W293–W297.

Zeng S, Xiao G, Guo J, Fei Z, Xu Y, Roe BA, Wang Y. 2010. Development of an EST dataset and characterization of EST-SSRs in a traditional Chinese medicinal plant, *Epimedium sagittatum* (Sieb. Et Zucc.) Maxim. BMC Genomics. 11(1):94.