Pak. J. Bot., 51(4): 1291-1296, 2019. DOI: http://dx.doi.org/10.30848/PJB2019-4(9)

ANALYSIS OF THE TRANSCRIPTOME OF NORMAL AND MUTANT WILLOW-SHAPED LEAVES OF *RICINUS COMMUNIS* USING HIGH-THROUGHPUT RNA SEQUENCING

WEI ZHOU[1], YAXING ZHOU[2], ZHENSHENG SHI[1,*], YUN WANG[2], GUOLI ZHU[3], ZHIBIAO HE[3] AND YONGSHENG CHEN[2]

[1]*Maize Research Institute, Shenyang Agricultural University, Shenyang, Liaoning Province, China*
[2]*Inner Mongolia University for Nationalities, Tongliao, Inner Mongolia, China*
[3]*Tongliao Academy of Agricultural Science, Inner Mongolia, Tongliao, Inner Mongolia, China*
[*]*Corresponding author's email: shi7nu@163.com*

Abstract

Ricinus communis is the source of castor oil and widely used. Lm type female strains of castor were mutants of castor which showed some phenotypes including typical willow-shaped leaves morphological characteristics, while leaves on the normal bisexual strain in *Ricinus communis* are palmar. To date, the molecular mechanisms responsible for the growth of the typical willow-shaped leaves have not yet been clarified. In this study, we performed transcriptome analysis to explore differentially expressed genes (DEGs) between normal palm-shaped leaves and mutant willow-shaped ones. Through high-throughput RNA sequencing, a total of 155,862 unigenes were finally annotated after assembly, and 1,530 DEGs in the mutants were identified, in which 1,205 DEGs were up-regulated and 325 DEGs were down-regulated. Gene Ontology (GO) Classification divided the unigenes and DEGs into 3 main categories, i.e. molecular function, biological process and cell component. According to the COG analysis, most of the unigenes were classified into the 'general function prediction only'. Some interesting candidate genes that may contribute to the formation of willow-shaped leaves such as C179712, C271005, C99773 and C241332 were verified using quantitative real-time PCR (qRT-PCR) analysis, and the results were in agreement with that in the RNA-Seq which confirmed the reliability of the RNA-seq data. These results laid the foundation for studying the mechanism responsible for the formation of the willow-shaped leaves.

Key words: Transcriptome, *Ricinus communis*, RNA sequencing.

Introduction

Ricinus communis, commonly known as castor bean, has great economic values, indicating the importance of castor bean breeding. Lm type female strains of castor were obtained by exposing castor to 60Coγ (Shen *et al.*, 1994). The appearance of this mutant isofemale line provided a new, promising approach for castor heterosis utilization and population improvement (Zhao *et al.*, 2016). Willow-shaped leaves are one of the typical morphological characteristics of this female strain without male flowers. The willow-shaped leaves appear with the emergence of buds and shed until the fruit ripening. So far, the molecular mechanisms responsible for the growth of willow-shaped functional leaves have not yet been clarified.

The transcriptome is the basis for structural and functional gene studies (Zhou *et al.*, 2012). Transcriptomics resolves biological issues by taking advantage of the expression regulation of all genes and the functions of proteins (Mahan *et al.*, 1993). Transcriptomics not only investigates the expression changes of each gene in transcriptomic level, but also predicts their transcriptomic location and annotation, and the functional structure of each gene in the genome (Costa *et al.*, 2010; Ruan *et al.*, 2004). The continuous development and modification of high-throughput transcriptome sequencing provides a new idea and method for functional genomic research in different organisms, and has been applied to study multiple species including humans, yeast, licorice, arabidopsis, and soybean (Wang *et al.*, 2009; Li *et al.*, 2012).

This study was aimed to explain the reasons for the formation of willow-shaped leaves on gene transcription level by analyzing the transcriptome of leaves on single female castor and normal strains (amphiprotic strains) through high-throughput sequencing. Changes in the expression of genes that related to leaf morphogenesis and male flowers formation were verified using quantitative real-time PCR (qRT-PCR) and discussed. This study provided valuable genetic resources for further studying the molecular mechanism of castor leaf mutant. Furthermore, due to the willow-shaped leaves are typical morphological characteristics which appear with the emergence of buds and shed until the fruit ripening, DEGs verified in this type of mutant may have more significance to help explore candidate genes that can be used in the genetic breeding.

Materials and Methods

Materials: Lm type female strains with willow-shaped leaves and normal amphiprotic strains (Fig. 1) were obtained from the east of the Inner Mongolia Autonomous Region (China), located between north latitude 42°15′-45°41′ and east longitude 119°15′-123°43′ in August 2, 2015. Ten plants of each strain were selected when the willow shaped functional leaves grew to 1-6 cm and the normal leaves grew to 2-15 cm. The Lm type female strains and normal amphiprotic strains were denoted as RL and RD, respectively.

Transcriptome sequencing data processing and analysis: The reads with joint sequences, data with excessive N content and short sequences were eliminated from the raw sequencing data. Next, the sequences were extended to contigs through overlaps between sequences

using Trinity (http://trinityrnaseq.sourceforge.net). Thereafter the contigs were connected into transcript sequences according to the sequential paired-end information. All transcripts were clustered ground on similarity to obtain single genes. Finally, the unigene sequences were subjected to BLASTX searches against public databases such as Nr, Nt, SwissProt, Gene Ontology (GO), COG and KEGG, and annotation information regarding these unigenes was obtained (http://soap.genomics.org. cn/SOAPdenovo-Trans.html).

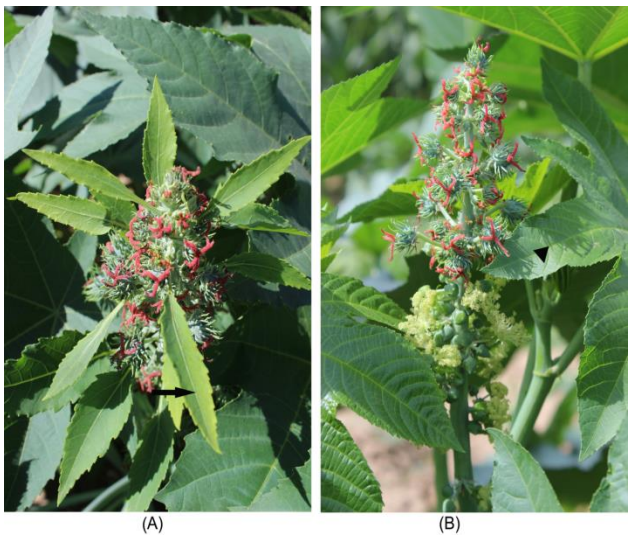

Fig. 1. (A) Lm-type female strain plants (disappeared stamens) with willow shaped leaves. (B) Normal bisexual plants with palm-shaped leaves. Arrow: willow shaped leaf; Arrowhead: palm-shaped leaves.

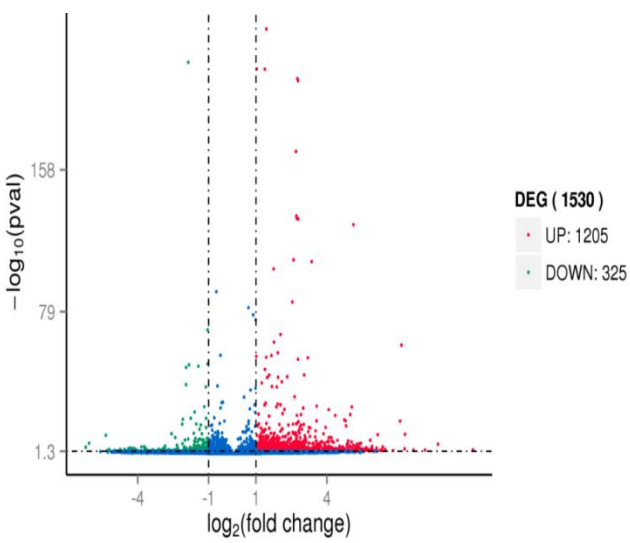

Fig. 2. Differential expression of genes from two samples. Note: Red spots represent up-regulated DEGs and green spots indicate down-regulated DEGs. Those shown in blue are unigenes that do not show obvious changes.

qRT-PCR validation: The candidate genes were validated and quantified via quantitative real-time PCR (qRT-PCR) according to the methods described in the published reference (Qi *et al.*, 2015). The primers used for qRT-PCR are listed in Table 1.

Results

Sequencing and assembly of transcriptome samples: RNA-seq was performed and data was analyzed according to previously published methods (Hu *et al.*, 2014; Wilhelm & Landry, 2009), finally as summarized in Table 2, the total sequencing reads from RL and RD library were 55, 498, 354 and 54, 864, 992, respectively. The total bases were 5.13G bp and 5.18G bp, respectively. The total reads from RL and RD were then assembled and produced 124, 572 contigs, with an average length of 477 bp, and an N_{50} of 649 bp (Table 3). Then, all the contigs were assembled to generate 59, 500, 144 scaffolds, with an average length of 501 bp and an N_{50} of 1,373 bp (Table 3).

Functional annotation of unigenes: In order to verify and annotate the unigenes assembled, BLAST software was used to compare the unigenes sequences generated with the sequences in NR, NT, SwissProt, TrimBL, GO, COG, and KEGG databases. As shown in Table 4, a total of 155,862 unigenes were finally annotated, among which 6,352 had homologous sequences in Swiss Prot, 42,764 in Nr, 65, 521 in Nt, 12,029 in COG, and 29,196 in KEGG.

DEG analysis: Overall, a total of 1,530 DEGs between RL and RD were identified in the present study under the control condition of log2 fold change > 1, p-value < 0.05. Among these genes, 1,205 genes were up-regulated and 325 were down-regulated (Fig. 2). In addition, 186 genes were found to be expressed only in the RL strain.

Gene Ontology (GO) classification: In this study, a total of 65,535 unigenes and 1,530 DEGs were divided into 3 main categories, i.e. cell component, molecular function and biological process, and 54 subclasses (Fig. 3). Most of the DEGs in cell components were classified as 'organelles' and 'cell parts', and the majority of the DEGs in molecular functions were 'catalytic activity' and 'binding', while the largest group in biological processes was 'metabolic process', 'Response to stimulus' and 'biological regulation'.

COG classification: A total of 12,028 unigenes in the transcriptome database were classified into 25 COG functional classifications (Fig. 4), among which, 1,628 unigenes were classified into the 'general function prediction only', 766 unigenes were classified into the 'transcription', and 711 were classified into the 'replication, recombination and repair' function. The least represented categories were 'Extracellular structure'.

DEGs involved in leaf morphogenesis and formation of male flowers: The primary purpose of this study was to discover genes that may contribute to the formation of willow-shaped functional leaves on Lm type female strains of castor (mutant). Thus, DEGs that have been related to leaf morphogenesis were selected. In addition, as the willow-shaped functional leaves were the typical morphological characteristics of this single female strain without male flowers, we speculated that the formation of willow-shaped functional leaves and the disappearance of male flowers may be controlled by some common genes. Thus, DEGs that have been implied in male flower formation were also screened. The gene ID, value of log2, and up- or down-regulation in the RNA-Seq for these selected genes were presented in Table 5.

Table 1. Primers used to fluorescent quantitative PCR.

Gene ID	Forward primer sequence	Reverse primer sequence
C179712	AGGACGAGATCCAACGCATC	TTCATATGTCACCTGGCGCA
C271005	GATCAGGGCTTCTCGTACCG	CCACGTTTGACCCGAGATGA
C99773	ACAAGCTTCCAGATTCGGCA	TGTGTTGTTGGGTGACTGCT
C241332	AGCACGGATGGTTGCTATTTG	AATAATAAGAGGCTCGATCCCAG
Actin	CCCAGCACACAGCAG	CTTGAAGAGGAAGAG

Table 2. Summary of the sequencing and assembly data for RL and RD.

Library	RL	RD
No. of raw reads	55,498,354	54,864,992
No. of high-quality reads after removing low-quality reads	51,334,388	51,868,586
GC contents (%)	46.92	47.58
Q30 percentage #	93.05	93.32

Q30 percentage is a standard for the quality assessment of sequencing data, with a larger value associated with a smaller sequencing error (Hu *et al.*, 2014; Wilhelm & Landry 2009). It refers to the average value which is greater than, or equal to 30%

Table 3. Assembly of the sequences obtained for two genotypes (RL and RD).

Contig			Scaffold		
Length range (bp)	**Total number**	**Percentage**	**Length range (bp)**	**Total number**	**Percentage**
100-150	47,134	37.84	100-500	92,688	78.07
150-200	16,137	12.95	500-1000	9,042	7.62
200-500	35,056	28.14	1000-1500	5,694	4.8
≥ 500	26,245	21.07	1500-2000	4,438	3.74
			≥ 2000	6,856	5.78%
Total number	124,572	100.00%		118,718	100.00%
Total length	59,378,739			59,500,144	
N_{50} length*	649			1,373	
Mean length	477			501	

*N_{50}, the median length (in nt) of contigs, transcripts, and unigenes, respectively

Table 4. The number of annotation distribution in the database of unigenes.

Database name	Annotated number	300 bp ≤ length ≤ 1000 bp	Length ≥ 1000 bp
COG	12,029	7,074	2,353
KEGG	29,196	16,939	3,298
Nr	42,764	25,593	4,078
Nt	65,521	6,683	58,207
Swissport	6,352	3,587	2,326
Annotate total number	155,862	59,876	70,262

Table 5. Log2, Direction, Description, FDR and KEGG gene name of a portion of differentially expressed genes.

Gene ID	Log2	Direction	Description	FDR	KEGG gene name
C243706	1.23625951	Up	agamous-like MADS-box protein AGL80-like [Citrus sinensis] gb |KDO36881.1| hypothetical protein CISIN_1g040635mg [Citrus sinensis]	$4.51E^{-12}$	cit:102617847
C179712	1.897457596	Up	*Ricinus communis* Axial regulator YABBY1,	$5.35E^{-11}$	--
C271005	-2.70245524	Down	*Ricinus communis* DELLA protein GAIP-B, putative, mRNA	$3.10E^{-22}$	rcu:RCOM_1586710
C99773	-10.9780682	Down	*Ricinus communis* Chitin-inducible gibberellin-responsive protein, putative, mRNA	$9.75E^{-30}$	--
scaffold16544	-2.82434221	Down	*Ricinus communis* Gibberellin receptor GID1, putative, mRNA microsatellite sequence	0	rcu:RCOM_0814010
C241332	-2.31199576	Down	*Ricinus communis* S-adenosyl methionine- dependent methyltransferase, putative, mRNA	$7.64E^{-14}$	rcu:RCOM_0432090
C224878	-1.15912593	Down	Theobroma cacao Phytochrome interacting factor 3, putative isoform 3 (TCM_026851) mRNA, complete cds	$6.15E^{-09}$	--
C83856	-3.90989732	Down	*N. tabacum* Rca mRNA for ribulose bisphosphate carboxylase activase (pTA1.1)	$6.61E^{-23}$	--

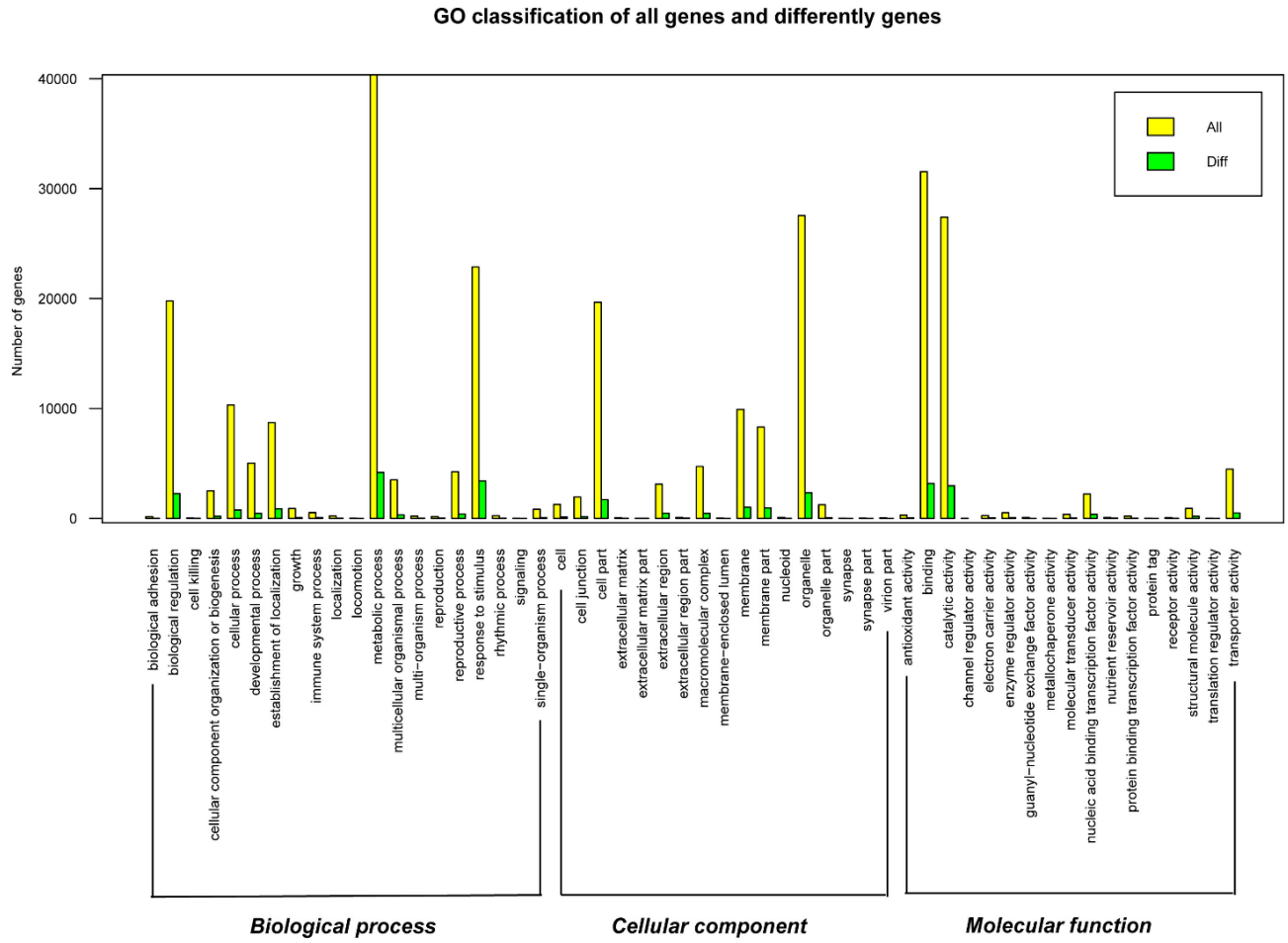

Fig. 3. Gene Ontology (GO) classification of 65,535 annotated unigenes and 1,530 annotated differentially expressed unigenes (DEGs).

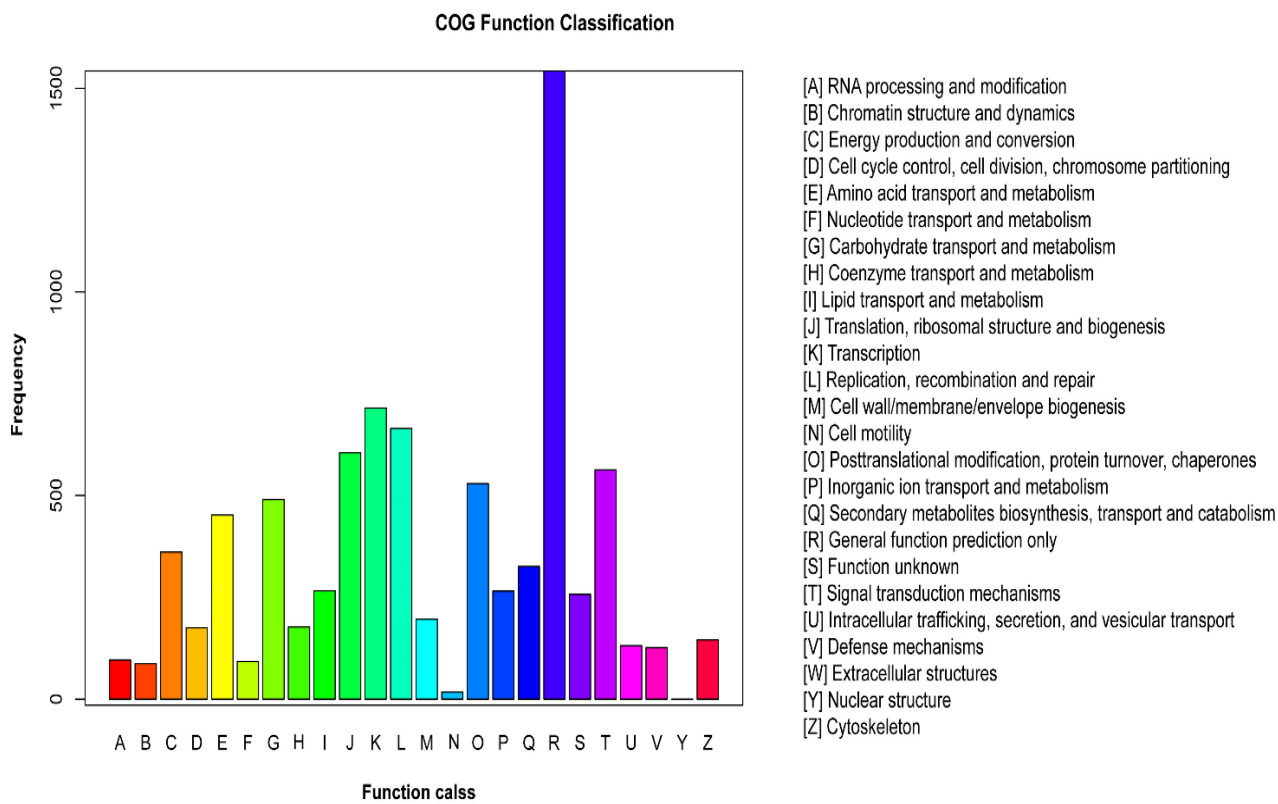

Fig. 4. COG classification of all unigenes.

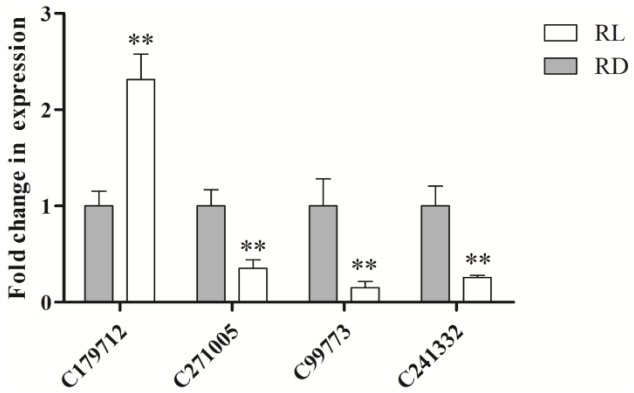

Fig. 5. The results of real-time quantitative PCR validation for candidate genes.

qRT-PCR validation for candidate genes: Although the precision rate for raw data from high-throughput sequencing was high, the procedure for constructing a database and subsequent assembly may affect the accuracy of the obtained transcript data. Thus, we validated the expression of some interesting genes mentioned in Table 5 using qRT-PCR analysis. The results showed in Fig. 5 were in agreement with that in the RNA-Seq, indicating that the results of sequencing and post-analysis were highly reliable.

Discussion

Lm type female strain of castor was widely used in the breeding of castor seeds. The willow-shaped functional leaves are typical features of this single female strain. However, the molecular mechanisms responsible for the formation of willow-shaped functional leaves have not yet been clarified. In the present study, 1,530 DEGs in this single female mutant strain were identified using RNA-Seq analysis, among which including 1,205 up-regulated genes and 325 down-regulated ones. In addition, several DEGs that have been related to leaf morphogenesis were validated using qRT-PCR analysis and discussed in this section. DEGs involved in male flower formation were also selected based on the speculation that formation of willow-shaped functional leaves and male flower were regulated by some common genes. Because willow-shaped functional leaves appear with the emergence of buds and shed until the fruit ripening.

Gibberellin (GA) is a plant hormone that regulates plant growth and affects various developmental processes, including leaf morphogenesis. Previous studies have indicated that GA accelerated cell growth and differentiation by promoting degradation of DELLA protein. Either exogenous GA or DELLA mutant could change compound leaves to simple leaves in tomato (Hay *et al.,* 2002; Jasinski *et al.,* 2008). GA suppresses the maturity of leaves to accelerate the formation of young leaves through the particular growth period of compound leaves (Shani *et al.,* 2009). In this work, we found that genes encoding gibberellin receptor GID1 and gibberellin-responsive protein had a 10.978 and 2.824-fold lower expression level in single female strain compared to normal amphiprotic strain, respectively, indicating that disorder of GA related signal transduction pathways may participate in the abnormal leaf formation.

The results of this study showed that the expression of ribulose bisphosphate carboxylase activase in mutants was downregulated by 3.91 times relative to normal strains. Ribulose bisphosphate carboxylase is a key enzyme in the Calvin cycle of plant photosynthetic carbon assimilation. Increased expression of ribulose bisphosphate leads to increased carbon assimilation efficiency during the flowering process, as well as increased photosynthetic product and carbon nitrogen ratio. Based on the carbon-nitrogen hypothesis, the plant can flower when the ratio of saccharides and nitrogen compounds (C/N) is high, while it cannot flower when the ratio is low, which is related to the disappearance of male flowers (Luo *et al.,* 2003). Conversely, increased photosynthesis leads to increased synthesis of signaling molecule sugars, which can regulate plant growth and development. This consequently affects the normal floral process, leading to the growth of willow-shaped functional leaves. However, the transcription data revealed that the expression level of *Ricinus communis* cultivar Hale chloroplast in mutant strains was 7 times of normal strains, indicating the enhancement of photosynthesis.

The expression of S-adenosyl methionine-dependent methyltransferase was downregulated by 2.31-fold in mutant strains. S-adenosyl methionine (SAM) participates in three major biochemical metabolism pathways, i.e. transmethylation, transsulfhydrylation and polyamine synthesis. S-adenosyl methionine synthetase utilizes methionine and ATP to synthesize SAM. The gene expression level of SAM plays a very important role in SAM content, and there is a close relationship between ethylene biosynthesis and the formation of cucumber male flowers (Li & Si, 2007). SAM is an important enzyme involved in the ethylene regulated process, and its excessive expression may result in disappearance of male flowers.

The results of investigations of different genetic backgrounds indicated that there was a close relationship between changes in peroxidase and male sterility (Zhao *et al.,* 2015). Changes in peroxidase, especially during the sensitive period, reflect changes in gene expression level and regulation activity of these enzymes. Moreover, these changes also regulate normal gene expression related to fertility, which can cause a series of changes in metabolic process that related to fertility, thus regulating the transition of plant fertility and ultimately leading to male sterility. Therefore, we inferred that peroxidase may play a role in the regulation of male sterility based on the twice times expression level of peroxidase in mutant castor strains of the normal strains. However, due to the significant differences in the expression of peroxidase in different organs, it was uncertain if peroxidase expression had some relevance and comparability between blades and male flowers. We speculate that peroxidase function in the early development stage of male flowers was not greater than in the blades because previous studies have indicated that catalase was involved in elimination and maintenance of the level of reactive oxygen and photorespiration, while the blade was the major organ involved in this function, which may be the cause of growth of willow-shaped functional leaves on the mutant castor strains.

AGAMOUS (AG) belongs to the C functional gene of ABC models of homeotic genes involved in the process of floral organ development, which expression in mutant castor strains being four times that of normal strains. Previous studies have reported that ABC genes regulated the transition from vegetative growth to reproductive growth, reactivity to photoperiod, floral organ formation, pollen development, and fertility. Studies have also indicated that it cannot generate stamens if the *AG* genes are mutated. The inaction of C function will change the stamen to corolla, carpel to calyx or other organ structures. The environmental light a plant is exposed to during the vegetative growth stage can also affect plant flowering (Bock *et al.*, 2006). The expression level of *Ricinus communis* phytochrome-interacting factor in the MADS-box protein family was 2.6 times that of normal strains, which regulated plant flowering.

The expression of a YABBY1 gene discovered based on the transcriptome data was down-regulated in the mutant castor strains. The YABBY family consists of transcription factors with C2C2 zinc finger domains and the YABBY structural domain (Bowman, 2000). YABBY transcription factor takes part in multiple biological processes, including morphogenesis of leaves and flowers, organ growth and activity control of top meristems. Although the inner links are still not clear, it can be inferred from previous studies that YABBY genes were closely related to organ vertical differentiation.

In future studies, key genes that may participate in the development of willow-shaped functional leaves and the disappearance of stamen should be further screened. Interfering with or knocking out these genes may help facilitate cultivation of good varieties of castor.

Acknowledgments

This study was supported by the Z.-S.S. Professor Program of Shenyang Agricultural University. We are grateful to LetPub (www. letpub.com) for their language embellishment.

References

Bock, K.W., D. Honys, J.M. Ward, S. Padmanaban, E.P. Nawrocki, K.D. Hirschi, D. Twell and H. Sze. 2006. Integrating membrane transport with male gametophyte development and function through transcriptomics. *Plant Physiol.*, 140(4): 1151-1168.

Bowman, J.L. 2000. Axial patterning in leaves and other lateral organs. *Curr. Opin. Genet. Dev.*, 10(4): 399-404.

Costa, V., C. Angelini, I.D. Feis and A. Ciccodicola. 2010. Uncovering the Complexity of Transcriptomes with RNA-Seq. *J. Biomed. Biotechnol.*, 2010: 853916.

Hay, A., H. Kaur, A. Phillips, P. Hedden, S. Hake and M. Tsiantis. 2002. The gibberellin pathway mediates KNOTTED1-type homeobox function in plants with different body plans. *Curr. Biol.*, 12(18): 1557-1565.

Hu, Y., X.H. Yao, H.D. Ren, K.L. Wang and P. Lin. 2014. Sequencing of Transcriptome Relevant to Flowering and Analysis of Floral-Related Genes Expression in Camellia oleifera. *Scientia Silvae Sinicae*, 50(9): 36-43.

Jasinski S., A. Tattersall, P. Piazza, A. Hay, J.F. Martinez-Garcia, G. Schmitz, K. Theres, S. McCormick and M. Tsiantis. 2008. PROCERA encodes a DELLA protein that mediates control of dissected leaf form in tomato. *Plant J.*, 56(4): 603-612.

Li, H., Y. Dong, J. Yang, X. Liu, Y. Wang, N. Yao, L. Guan, N. Wang, J. Wu and X. Li. 2012. De novo transcriptome of safflower and the identification of putative genes for oleosin and the biosynthesis of flavonoids. *PLoS One*, 7(2): e30987.

Li, Z. and L.T. Si. 2007. Progress on Major Genes for Sex Determination in Cucumber. *Mol. Plant Breed.*, 5(6): 75-79.

Luo, R., Y. Huang, B. Cai, S. Chen and G. Su. 2003. Preliminary studies on relationship between flowering and element content in leaves of longan. *Guangxi Agr. Sci.*, 34(2): 9-11.

Mahan, M.J., J.M. Slauch and J.J. Mekalanos. 1993. Selection of bacterial virulence genes that are specifically induced in host tissues. *Science*, 259(5095): 686-688.

Qi, X., L. Zhang, Y. Han, X. Ren, J. Huang and H. Chen. 2015. De novotranscriptome sequencing and analysis of Coccinella septempunctata L. in non-diapause, diapause and diapause-terminated states to identify diapause-associated genes. *BMC Genomics*, 16: 1086.

Ruan, Y., P. Le Ber, H.H. Ng and E.T. Liu. 2004. Interrogating the transcriptome. *Trends Biotechnol.*, 22(1): 23-30.

Shani, E., Y. Burko, L. Benyaakov, Y. Berger, Z. Amsellem, A. Goldshmidt, E. Sharon and N. Ori. 2009. Stage-specific regulation of Solanum lycopersicum leaf maturation by class 1 KNOTTED1-LIKE HOMEOBOX proteins. *Plant Cell*, 21(10): 3078-3092.

Shen, H., J. Jiang, Z. Wang and S. Geng. 1994. Studies on the breeding and inheritance of male sterile lines of pepper (*Capsicum annuum* L.). *Acta Agri. Univ. Pekinensis*., 20(1): 25-30.

Wang, Z., M. Gerstein and M. Synder. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genetics.*, 10(1): 57-63.

Wilhelm, B.T. and J.R. Landry. 2009. RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing. *Methods (San Diego, Calif)*, 48(3): 249-257.

Zhao, L.H., Z. Fei, M.X. Guan, C. Liu, Z. Xin and X.M. Liu. 2015. Analysis of differentially expressed transcripts in early and middle development of Betula platyphylla female inflorescence. *J. Beijing Forest Univ.*, 37: 104-112.

Zhao, Y., G. Zhu, Z.B. He, C.G. Bao, X.F. Chen, T.T. Jiang, J.H. Li, F.L. Huang, T.T. Zhou, X.F. Chang and J.Q. Liu. 2016. Establish Castor Marked Female Lm Per Female Amphoteric Lineinflores Cence MSAP Reaction System. *Acta Agr. Boreali-Sinica*., 31(3): 114-120.

Zhou, H., X. Zhang, T.Y. Liu and F.X. Yu. 2012. Data Processing and Gene Discovery of High-throughput Transcriptome Sequencing. *Jiangxi Sci*., 30(5): 607-611.

(Received for publication 28 February 2018)