Canadian Journal of Plant Science

Genome-wide identification, phylogeny and expression profile of the sucrose transporter multigene family in tobacco

Journal:	<i>Canadian Journal of Plant Science</i>
Manuscript ID	CJPS-2018-0187.R2
Manuscript Type:	Article
Date Submitted by the Author:	21-Nov-2018
Complete List of Authors:	Wang, Shanshan Yang, Jun Xie, Xiaodong Li, Feng Wu, Mingzhu Lin, Fucheng Wang, Zhong
Keywords:	Genetics, Genome
Is the invited manuscript for consideration in a Special Issue?:	Not applicable (regular submission)

SCHOLARONE™
Manuscripts

**Genome-wide identification, phylogeny and expression profile of the sucrose
transporter multigene family in tobacco**

Shanshan Wang^{1, 2}, Jun Yang¹, Xiaodong Xie¹, Feng Li¹, Mingzhu Wu¹, Fucheng Lin³, Zhong Wang^{1*}

¹China Tobacco Gene Research Center, Zhengzhou Tobacco Research Institute of CNTC, Zhengzhou 450001

²Xiangyang Cigarette Factory, Hubei China Tobacco, Xiangyang, Hubei 441000, China

³Zhejiang University, 866 Yu-Hang-Tang Road, Hangzhou 310058, China

* Corresponding author: Zhong Wang, e-mail address: zhongwang1123@126.com, phone number: +86 0371 67672069

Abstract

The transportation and distribution of sucrose in plants is mediated by the sucrose transporters (SUTs), which also participate in various plant developmental and resistance processes. However, no such study of the tobacco *SUT* family has been reported yet. In the present study, 11, 5, and 4 *SUT* genes were identified from the *Nicotiana. tabacum*, *N. sylvestris*, and *N. tomentosiformis* genome, respectively. The exon/intron structures of the tobacco *SUTs* genes were highly conserved among the three tobacco species. There were gene loss, duplication, and chromosome exchange occurred to the NtSUT family during the formation of the allotetraploid common tobacco. Expression profiling analysis revealed that the expression patterns of the *NtSUT* genes in common tobacco were more close to those in the *N. sylvestris* plants. The *NtSUT2s* and *NtSUT4* genes were ubiquitously expressed in various tobacco tissues, while the *NtSUT1s* gene was highly expressed in the maturing leaves, indicating their functional conservation and differentiation. The transcriptions of the *NtSUT2t*, *NtSUT3s*, *NtSUT4* and *NtSUT5s* genes in tobacco plants were dramatically induced under Pi starvation, drought, and salinity stresses, but their highest expression levels occurred in different tissues, suggesting the multiple roles of NtSUTs in plant resistance to various abiotic stresses. This study provides useful information for the further functional characterization of the *SUT* genes in tobacco.

Keywords: sucrose transporter, gene family, tobacco, expression analysis, abiotic stresses

Introduction

Sucrose is the main product of photosynthesis, which provides carbon and energy for plant growth and development, as well as for the synthesis of complex carbohydrates and storage reserves. Sucrose can also act as a signal molecule in the regulation of flowering (Chincinska et al. 2008; King and Ben-Tal 2001; Sivitz et al. 2007), tuberization (Chincinska et al. 2008), seed development (Iraqi and Tremblay 2001; Weschke et al. 2000), differentiation of vascular tissues (Uggla et al. 2001), and the expression levels of various genes in plants (Ciereszko et al. 2001; Sivitz et al. 2008; Vaughn et al. 2002). Besides, sucrose has been shown to be involved in plant resistance to various abiotic stresses. For instance, phosphate (Pi) starvation induced the biosynthesis of sucrose in plant leaves, and promoted the sucrose uploading into the phloem (Karthikeyan et al. 2007). Moreover, plants accumulated more sucrose under drought or salinity stress to maintain the normal functions of cells (Bethke et al. 2009; Hoffmann 2010; Pommerrenig et al. 2007).

Sucrose is the dominant form of carbohydrates for the long-distance transport in higher plants. Sucrose transporters (SUTs) are well known protein complex to control the sucrose distribution in plants via regulating the phloem uploading from the source tissues and the unloading into the sink tissues. SUTs were also proved to be involved in regulating many plant developmental processes, such as pollen germination (Sivitz et al. 2008), flowering (Sivitz et al. 2007), tuberization (Chincinska et al. 2008), restraining plant growth (Srivastava et al. 2009), fruit size reduction (Hackel et al. 2006), seed development (Weschke et al. 2000), biomass partitioning (Payyavula et

al. 2011), and ethylene biosynthesis (Chincinska et al. 2013). Moreover, SUTs were found to promote the sucrose transport in plants under Pi starvation (Hammond and White 2008; Lei et al. 2011), salinity (Gong et al. 2013; Noiraud et al. 2000), and drought stresses (Gong et al. 2015).

In plants, the SUTs are encoded by a multi-gene family. The SUT family has been characterized in many plant species, including rice (Aoki et al. 2003), grapevine (Afoufa-Bastien et al. 2010), populus (Payyavula et al. 2011), sweet orange (Zheng et al. 2014), cacao (Li et al. 2014), maize (Usha et al. 2015), and sugarcane (Zhang et al. 2016). The molecular characteristics, structures, and phylogeny of plant SUTs have been well studied and reviewed (Geiger 2011; Peng et al. 2014; Sauer 2007). SUTs contain 12 transmembrane spanning helices, and belong to the major facilitator super-family (MFS) (Marger and Saier 1993; Stolz et al. 1999). The *SUT* families were recently reported to evolve from two different ancestor genes (named AG1 and AG2), but could be classified into four phylogenetic groups (Peng et al. 2014). Plant *SUTs* in the AG1 and AG2 group harbored 5 and 13 introns, respectively, except for the intronless *SUTs* in moss (Peng et al. 2014). The expression patterns of different *SUT* members showed various tissue and/or developmental specificity. For instance, the *TcSUT1* gene was dominantly expressed in the pods of cacao (*Theobroma cacao* L.), while the *TcSUT3* and *TcSUT4* genes were highly expressed in both the pods and the bark with phloem. The transcripts of *TcSUT1* and *TcSUT4* genes were more abundant in the seed and seed coat from the pod enlargement stage to the ripening stage. The expression level of *TcSUT5* gene was sharply increased in the cacao seed

coat during the ripening stage (Li et al. 2014).

Tobacco is an important model plant for the researches in plant science, and is also of important agricultural and economic value worldwide. There were only three *SUTs* genes, which have been identified from the tobacco genome so far (Burkle et al. 1998; Lemoine et al. 1999; Okubo-Kurihara et al. 2011). In this study, we identified all the members of the tobacco *SUT* family, and characterized their molecular characteristics, structures, phylogeny, and expression profiles. Our findings will provide a foundation for further functional studies of SUTs in tobacco, and contribute to elucidate SUT roles in plant responses to various stresses.

Materials and methods

Plant materials and treatments

Seeds of the tobacco variety Honghua Dajinyuan (*N. tabacum* L.) were germinated and maintained in plastic pots placed in the green house with 28 °C for 16 h light / 23 °C for 8 h dark cycles. After two months, tobacco plantlets with 6-7 true leaves were either transferred to open fields for continuing growing or used for stress treatment. The plant roots, stems, leaves, buds, sepals, stamens, pistils, and capsules were collected at flowering stage. The leaf vein, blade tip, blade middle and blade root of the tenth leaves were collected, separately. Leaves, stems and roots from different developmental stages were also collected for analyzing gene temporal expression patterns.

For the salinity and Pi starvation treatment, seedlings at the six-leaf stage were transported to a plastic container containing half-strength MS nutrient liquid medium

with 300 mM NaCl or replacing KH_2PO_4 with K_2SO_4 apart. Drought stress was induced by withholding water and allowing the water in the soil to drain out. After treatment, tobacco tissues were harvested and immediately submerged in liquid nitrogen, stored at -80°C and used for total RNA extraction.

Identification and sequence analysis of tobacco SUTs

Putative tobacco *SUT* genes were screened from the China tobacco genome database v4.0 with a series of blastp and blastn using the AtSUT sequences. All the reliable *SUT* genes in the tobacco genome were finally identified via multiple sequence alignments, and conserved domain and motif analyses. The genomic DNA sequences and the coding sequences were subsequently verified by cloning and sequencing. The SMART web tool was used to identify the MFS domain (<http://smart.embl-heidelberg.de/>) (Letunic and Bork 2018). Motif analysis of putative SUT protein sequences was firstly conducted with the multiple EM (<http://meme-suite.org/>) for motif elicitation program (Bailey et al. 2009), and the InterPro database (<http://www.ebi.ac.uk/interpro/>) was then used to check the identified motifs and find their detailed signature matches (Hunter et al. 2009). The chromosomal location image for *NtSUTs* was obtained according to the chromosomal position information from the database with the MapInsect software. Multiple alignments of the *SUT* sequences were performed using the programs DNAMAN (version 6.0) and Clustal X (version 1.83) with default gap penalties (Larkin et al. 2007). The software MEGA 5.2.1 was used to construct a phylogenetic tree with the deduced SUT amino acid sequences using the maximum-likelihood algorithm (Sohpal

et al. 2010). The exon/intron structures were mapped with an online tool GSDS (<http://gsds.cbi.pku.edu.cn/>) (Hu et al. 2015).

cDNA sample preparation and gene transcription level analysis

Total RNA was isolated using a Superpure Plantpoly RNA Kit (Gene Answer, Beijing, China). DNase I (Gene Answer) was applied to remove genomic DNA contaminants during the extraction process. First cDNA strand was synthesized from approximately 1 µg of total RNA using Reverse Transcriptase M-MLV (Takara, Biomedical Technology, Beijing, China) with OligodT primers according to manufacturer's instructions. We then evaluated the cDNA concentration with the NanoDrop 2000 (Thermo Fisher Scientific, Wilmington, DE, USA), and diluted it to 40 ng/µl as the templates for the following real-time quantitative polymerase chain reaction (qRT-PCR) analyses.

qRT-PCR analyses were conducted using the iCycler iQ thermo cycler (Bio-Rad, Hercules, CA, USA), monitored by a SYBR Green fluorescence (Bio-Rad). The PCR program was as follows: 95 °C for 5 min, 40 cycles of 95 °C for 30 s, 60 °C for 30 s, signal acquisition, and then a final melting curve of 65-95 °C. Tobacco *L25* was used as internal control (Schmidt and Delaney 2010). Primer sequences were shown in Table S1. The $2^{-\Delta Ct}$ and $2^{-\Delta\Delta Ct}$ methods were used to calculate the relative expression level for the *NtSUT* genes (Livak and Schmittgen 2001). Values in the figures were mean $\pm$ SD of three independent biological replicates.

Results

Genome-wide identification and chromosomal localization of tobacco *SUT* genes

11, 5 and 4 *SUT* genes were identified from the *N. tabacum*, *N. sylvestris* and *N. tomentosiformis* genome, respectively. The genomic DNA sequences, coding sequences, and amino acid sequences of these tobacco *SUT* genes were listed in the Table S2. Putative polypeptides of the tobacco SUTs harbored 366 to 605 amino acids, with molecular weights between 39.79 and 65.13 kDa (predicted 6.13-9.43 in isoelectric points). The tobacco *SUTs* genes could be divided into five groups according to their sequence identities (Table 2), which were further designated as SUT1-5 based on their sequence similarities with the previously published SUT sequences. Sequence identities showed that the *NtSUT* gene was either close to the *Nsyt* gene or close to the *Ntom* gene, so we designated them as *NtSUTs* or *NtSUTt*, separately (Table 3). It's worth noting that there was no gene encoding SUT1 in the *N. tomentosiformis* genome, and no *NtSUT3t* gene in the *N. tabacum* genome. However, there were two genes encoding the *NtSUT4s* (*Ntab0332220/NtSUT4sa*, *Ntab0245750/NtSUT4sb*) and *NtSUT5s* (*Ntab0067760/NtSUT5sa*, *Ntab0158640/NtSUT5sb*) (Table 3). So it could be inferred that there were gene duplications or losses occurring to the SUT family in the *N. tabacum* and *N. tomentosiformis* genomes. As shown in Fig. 1, 10 *NtSUTs* were distributed in seven chromosomes, and the *NtSUT1t* gene was distributed in the scaffold of *Ntab_scaffold_339*. *NtSUT4sa*, *NtSUT4sb*, *NtSUT5sa*, and *NtSUT5sb* were located on four different chromosomes, suggesting chromosome exchange might have occurred after the duplication event during the formation of the tetraploid genome.

Structural and phylogenetic analysis of tobacco SUTs

Domain analysis showed that all tobacco SUTs had two highly conserved MFS domains (Fig. 2), except NtomSUT4. Multiple sequence alignments revealed that there was one highly conserved histidine residue located on the first extracellular loop, and one (R/K)X₂/3(R/K) motif in both the second and eighth loop (Fig. 2). The tobacco SUT4 and SUT5 had the conserved LX₂LL motif in the N termini, while the SUT2 contained 15-20% more amino acids, leading to the formation of an enlarged cytoplasmic loop (about 60 amino acids) in the N termini. In summary, all the highly conserved motifs found in the tobacco SUTs supported the fact that all the tobacco *SUT* genes identified in this study could encode functional SUT proteins.

Phylogenetic analysis showed that plant SUTs fell well into four clades (Fig. 3), named as SUT I -IV. The *SUT* genes in the SUT I group contained 2 to 5 introns, while most genes in the SUT IV group had 4 introns. However, the *SUT* genes in the SUT II and SUT III groups possessed much more introns. Most genes in the SUT II group had 13 introns, while the intron numbers of the *SUT* genes in the SUT III group varied from 9 to 13. The tobacco *SUT1* and *SUT3* belonged to the SUT I clade (Fig. 3), while the tobacco *SUT2*, *SUT4*&*5* belonged to the SUT II and SUT IV clade, respectively. There were 13 introns in the tobacco *SUT2* genes, 5 in the *SUT3* genes, and 4 in the other tobacco *SUT* genes. The gene structures of tobacco *SUT* genes were highly conserved among the allotetraploid tobacco and its two diploid progenitors (Fig. S1).

Expression patterns of tobacco *SUT* genes

To further clarify the roles of each NtSUT, we detected the relative expression

levels of the *NtSUTs* genes in different tobacco tissues by qRT-PCR. As shown in Fig. 4A, the transcript of all the *NtSUTs* genes was most abundant in the roots, which actually was at least more than 2-folds higher than those in any other tissues. Pistils harbored the lowest transcription level of the *NtSUTs* genes, indicating a weak sucrose demand in this tissue. The expression level of the *NtSUT2s* gene was dominant in the roots, stems, buds, sepals, stamen and pistils, whereas in the capsule and seeds the dominant gene was *NtSUT4s*. In the leaves, the *NtSUT1* and *NtSUT4* showed higher expression levels than the other *NtSUT* genes. The *NtSUT4s* gene was mainly expressed in the leaf vein, while the two *NtSUT1* genes were highly expressed in the leaf vein, blade tip, blade middle, and blade root (Fig. 4B). The transcripts of *NtSUT4*, *NtSUT5s* and *NtSUT2s* genes in the roots were highly up-regulated at the bud and flowering stage, but sharply down-regulated at the topping stage (Fig. 4C). However, the transcriptions of these *NtSUTs* gene were highly induced in the leaves at the topping stage (Fig. 4D). Moreover, the expression levels of the two *NtSUT1* genes were dramatically up-regulated in the maturing leaves, indicating that these two genes might play important roles in sucrose transportation during leaves maturation. The transcript of the *NtSUT3s* and *NtSUT2t* genes could hardly be detected in all the tissues examined (Fig. 4).

We further detected the expression patterns of the *SUT* genes in the two diploid parental tobacco species. In the *N. sylvestris* plants, the *NsylSUT2* gene was ubiquitously highly expressed in the leaves and roots, while the *NsylSUT1* and *NsylSUT4* genes were specifically highly expressed in the roots (Fig. 5A). The

expression levels of other *NsylSUT* genes were quite low in the tissues examined. In the *N. tomentosiformis* plants, the dominantly expressed gene in the leaves was the *NtomSUT5* gene (Fig. 5B), but the *NtomSUT2*, *NtomSUT4* and *NtomSUT5* genes showed high expression levels in the roots.

Differential expression patterns of the *NtSUT* genes in response to abiotic stresses

To investigate the possible roles of the *NtSUTs* in plant responses to abiotic stresses, we measured their transcript levels under Pi starvation, drought, and salinity stresses (Fig. 6 & Table S4). When tobacco seedlings suffered Pi starvation, the transcripts of the *NtSUT* genes were greatly induced at 48h after treatment, and this induction mainly occurred in the stems (Fig. 6A-C). The expression levels of the *NtSUT2t*, *NtSUT4*, and *NtSUT5s* genes increased by more than 20 folds in the stems under Pi starvation, whereas in the roots and leaves their expression levels just showed slight increase. The expression patterns of the *NtSUT* genes varied among different plant tissues under drought treatment. As shown in Fig. 6D, the expression of most *NtSUT* genes in the roots seemed unaffected by the drought stress, except the *NtSUT2t* gene, which was up-regulated by more than five folds at 2d and 8d. In contrast, the expression of most *NtSUT* genes in the stems was induced by the drought treatment (Fig. 6E). For instance, the expression levels of the *NtSUT3s*, *NtSUT1s*, *NtSUT2t*, *NtSUT4* and *NtSUT5s* genes in the stems were up-regulated at 2d, but slightly decreased at 4d, and then increased again at 8d. Moreover, the expression

levels of most *NtSUTs* genes were repressed in the leaves at 2d, and then gradually increased at 4d and 8d (Fig. 6F). In addition, the transcript levels of *NtSUT3s*, *NtSUT4* and *NtSUT5s* were highly up-regulated in both the stems and leaves at 8d, indicating an important role of these four genes in plant response to drought stress.

The expression levels of the *NtSUT* genes showed drastic changes under salinity stress. In the roots, the transcript levels of the *NtSUT3s* and *NtSUT4* genes dramatically increased by about 120 folds at 12h under salinity stress, and then sharply decreased to the control level at 24h (Fig. 6G). The expression level of the *NtSUT2t* gene also greatly increased (about 50 folds) in the roots at 12h and 48h. There was no significant changes occurred to the expression levels of all the *NtSUT* genes in the stems, except that of the *NtSUT4s* gene, which was clearly up-regulated at 6h (Fig. 6H). Surprisingly, the expression patterns of the *NtSUTs* genes in the leaves could be divided into two groups under salinity stress. In detail, the expression levels of the *NtSUT2s*, *NtSUT5t* and *NtSUT1* genes were suppressed along with the salinity treatment, but those of the *NtSUT3s*, *NtSUT5s* and *NtSUT4* genes were up-regulated at the same time (Fig. 6I).

Discussion

Sucrose is important for plant growth and development, and the *SUTs* determine the distribution of sucrose in plants. In the present study, we identified 11, 5, 4 *SUT* genes from the *N. tabacum*, *N. sylvestris*, and *N. tomentosiformis* genome, respectively. The molecular structures, evolutionary relationships, and expression profiles of tobacco *SUTs* gene were fully characterized to make a good foundation for

the functional study of the tobacco sucrose transporters.

Evolutionary of tobacco SUT genes

Plant *SUT* genes showed a great variation in the intron numbers, ranging from 0 to 14. Moreover, the descendant plant *SUTs* originated from the progenitor AG1 usually comprised less introns than those derived from the ancestor AG2. From the analyses of exon-intron structure within the Solanaceae species, the 64bp- and 44bp-exon ubiquitously existed in all the Solanaceae *SUT* genes except *LeSUT1*. *LeSUT1* had a larger exon, which might originate from the fusion of the 64bp- and 44bp-exon. These results suggested that Solanaceae *SUTs* gene structures might have formed before the divergence of the Solanaceae species.

In the phylogenetic analysis, plant SUTs were divided into four distinct groups, which were considered to be evolved from two ancient ancestors AG1 and AG2. This was consistent with the previous studies (Doidy et al. 2012; Payyavula et al. 2011). In tobacco, only the *NtSUT2* gene was classed to the ancestor AG2. The remaining four genes shared the same ancestor AG1, supporting the hypothesis that AG1 was preferentially expanded in the eudicots (Peng et al. 2014). In the AG1 group, the four tobacco *SUT* genes were clearly divided into two subgroups. *NtSUT1* and *NtSUT5* clustered together to the SUT1 subfamily, while *NtSUT3* and *NtSUT4* were the members of the SUT4 subfamily. However, these two subgroups of tobacco SUTs were more close to those of *Solanum lycopersicum* and *Solanum tuberosum* in the phylogenetic tree. And this situation also occurred to the *NtSUT2* gene indicating that

the generation of these three subgroups occurred before the divergence of Solanaceae.

Comparison of the *SUT* expression patterns among three tobacco species

The *N. tabacum* is believed to be developed from the interspecific hybridization of two diploid ancestors *N. sylvestris* and *N. tomentosiformis* (Gazdova et al. 1995; Kenton et al. 1993). Polyploidization occurred to the hybridized genome, which gave rise to the allotetraploid common tobacco. After duplication, the fate of the homeologous genes might be conservation or redundancy, non-functionalization, subfunctionalization, and neofunctionalization (Hahn 2009; Rastogi and Liberles 2005). Analysis of gene expression patterns can provide an insightful indication of gene functions. In general, the *NtSUTs* originated from the *N. sylvestris* seemed more active in the common tobacco, and the expression patterns of the *SUT* genes in the common tobacco were more close to those in the *N. sylvestris* plants. For instance, the transcripts of *NtSUT2s*, *NtSUT4s*, and *NtSUT5s* genes could be detected in the tobacco leaves or roots, but only *NtSUT4t* gene was found to be expressed in the leaves and roots (Fig. 4A, B). Similarly, the *NsylSUT2* gene showed high expression levels in both the leaves and roots, while the *NsylSUT4* and *NsylSUT1* genes were mainly expressed in the roots (Fig. 5A, B). Although the *Ntom5* gene was dominantly expressed in the *N. tomentosiformis* leaves (Fig. 5A), the expression of *NtSUT5t* gene could hardly be detected in the leaves (Fig. 4B).

Multifunction analysis of tobacco *NtSUT* genes

It's worth pointing that the expression levels of the *NtSUT4*, *NtSUT5s* and *NtSUT2s* genes greatly increased in the roots when the inflorescence tissues emerged at the bud and flowering stages, but sharply decreased when the inflorescence tissues were removed at the topping stage (Fig. 4C). It has been reported that improving the root development in the rice at the flowering stage could be a key factor to obtain higher grain yield and good agronomic performance (He et al. 2014). Thus it could be inferred that the high expression levels of the *NtSUTs* genes in the roots at the bud and flowering stages might be determined by the development of the inflorescence tissues.

When plant suffered drought and salinity stresses, the roots were the major tissues, which sensed and responded to these stresses. So in theory, more sucrose should be transported to the roots, and the expression levels of the *NtSUTs* genes should be up-regulated in the roots under drought and salinity stresses. But the fact was that the expression levels of the *NtSUTs* genes in the roots under drought treatment were similar to (or even lower than) those in the control roots (Fig. 6D). At the same time, the expression levels of the *NtSUTs* genes were highly up-regulated in the stems under drought treatment. Therefore, it could be concluded that the *NtSUTs* should be multifunctional under abiotic stresses other than sucrose transport. The expression levels of the *NtSUT4* genes were drastically up-regulated under the abiotic stresses, while those of the *NtSUT2s* gene were hardly affected, indicating that there were functional redundancy and differentiation among *NtSUTs*. Moreover, the expression levels of *NtSUT3s* and *NtSUT2t* could hardly be detected in all the tissues from various developmental stages. However, the transcripts of these two genes was

obviously induced by abiotic stresses (Fig. 6), further indicating there were highly functional differentiation among *NtSUTs*.

In this study, we identified 11, 5, 4 genes encoding putative sucrose transporter from the *N. tabacum*, *N. sylvestris*, and *N. tomentosiformis* genome, respectively. The chromosome location, exon/intron structure, and phylogeny of the tobacco *SUTs* were systematically characterized. Expression profiling analysis revealed that the expression patterns of the *NtSUT* genes in common tobacco were more close to those in the *N. sylvestris* plants. The expression of the *NtSUTs* gene showed tissue-specific and development-dependent, and the *NtSUT2t*, *NtSUT3s*, *NtSUT4* and *NtSUT5s* were greatly induced by Pi starvation, drought and salinity stress, indicating their potential roles in stress tolerance. This study provides new clues for functional characterization of sucrose transporters during the developmental process, and in response to abiotic stresses of plants.

Acknowledgements

This work was supported by the Project of Regulation of the Transcription Factor NtMYB27 in the Biosynthesis of Flavonoids (902017CA0220) and the Guizhou tobacco company project (201501).

Conflicts of interest

We declare that we have no conflict of interest.

References

Afoufa-Bastien, D., Medici, A., Jeauffre, J., Coutos-Thevenot, P., Lemoine, R., Atanassova, R. and Laloi, M. 2010. The *Vitis vinifera* sugar transporter gene family: phylogenetic overview and

macroarray expression profiling. *BMC Plant Biol* 10:245.

Aoki, N., Hirose, T., Scofield, G. N., Whitfield, P. R. and Furbank, R. T. 2003. The sucrose transporter gene family in rice. *Plant Cell Physiol* 44(3):223-32.

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W. and Noble, W. S. 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 37(Web Server issue):W202-8.

Bethke, P. C., Sabba, R. and Bussan, A. J. 2009. Tuber Water and Pressure Potentials Decrease and Sucrose Contents Increase in Response to Moderate Drought and Heat Stress. *American Journal of Potato Research* 86 (6):519-532.

Burkle, L., Hibberd, J. M., Quick, W. P., Kuhn, C., Hirner, B. and Frommer, W. B. 1998. The H⁺-sucrose cotransporter NtSUT1 is essential for sugar export from tobacco leaves. *Plant Physiol* 118(1):59-68.

Chincinska, I., Gier, K., Krugel, U., Liesche, J., He, H., Grimm, B., Harren, F. J., Cristescu, S. M. and Kuhn, C. 2013. Photoperiodic regulation of the sucrose transporter StSUT4 affects the expression of circadian-regulated genes and ethylene production. *Front Plant Sci* 4:26.

Chincinska, I. A., Liesche, J., Krugel, U., Michalska, J., Geigenberger, P., Grimm, B. and Kuhn, C. 2008. Sucrose transporter StSUT4 from potato affects flowering, tuberization, and shade avoidance response. *Plant Physiology* 146(2):515-28.

Ciereszko, I., Johansson, H. and Kleczkowski, L. A. 2001. Sucrose and light regulation of a cold-inducible UDP-glucose pyrophosphorylase gene via a hexokinase-independent and abscisic acid-insensitive pathway in *Arabidopsis*. *Biochem J* 354(Pt 1):67-72.

Doidy, J., van Tuinen, D., Lamotte, O., Corneillat, M., Alcaraz, G. and Wipf, D. 2012. The *Medicago truncatula* sucrose transporter family: characterization and implication of key members in carbon partitioning towards arbuscular mycorrhizal fungi. *Molecular Plant* 5(6):1346-58.

Gazdova, B., Siroky, J., Fajkus, J., Brzobohaty, B., Kenton, A., Parokonny, A., Heslop-Harrison, J. S., Palme, K. and Bezdek, M. 1995. Characterization of a new family of tobacco highly repetitive DNA, GRS, specific for the *Nicotiana tomentosiformis* genomic component. *Chromosome Res* 3(4):245-54.

Geiger, D. 2011. Plant sucrose transporters from a biophysical point of view. *Mol Plant* 4(3):395-406.

Gong, X., Liu, M., Zhang, L., Ruan, Y., Ding, R., Ji, Y., Zhang, N., Zhang, S., Farmer, J. and Wang, C. 2015. *Arabidopsis* AtSUC2 and AtSUC4, encoding sucrose transporters, are required for abiotic stress tolerance in an ABA-dependent pathway. *Physiol Plant* 153(1):119-36.

Gong, X. L., M. L., Wang, C., Zhang, L. J. and Liu, W. 2013. Sucrose Transporter Gene AtSUC4 Regulates Sucrose Distribution and Metabolism in Response to Salt Stress in *Arabidopsis thaliana*. *Advanced Materials Research* 731(12):217-221.

Hackel, A., Schauer, N., Carrari, F., Fernie, A. R., Grimm, B. and Kuhn, C. 2006. Sucrose transporter LeSUT1 and LeSUT2 inhibition affects tomato fruit development in different ways. *The Plant Journal* 45(2):180-92.

Hahn, M. W. 2009. Distinguishing among evolutionary models for the maintenance of gene duplicates. *J Hered* 100(5):605-17.

Hammond, J. P. and White, P. J. 2008. Sucrose transport in the phloem: integrating root responses to phosphorus starvation. *J Exp Bot* 59(1):93-109.

He, H. B., Yang, R., Chen, L., Fan, H., Wang, X., Wang, S. Y., Cheng, H. W. and Ma, F. Y. 2014. Rice root system spatial distribution characteristics at flowering stage and grain yield under plastic mulching drip irrigation (PMDI). *Journal of Animal & Plant Sciences* 24(1):290-301.

- Hoffmann, C. M. 2010. Sucrose Accumulation in Sugar Beet Under Drought Stress. *Journal of Agronomy & Crop Science* 196(4):243-252.
- Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J. and Gao, G. 2015. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31(8):1296-7.
- Hunter, S., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L. and others. 2009. InterPro: the integrative protein signature database. *Nucleic Acids Res* 37(Database issue):D211-5.
- Iraqi, D. and Tremblay, F. M. 2001. Analysis of carbohydrate metabolism enzymes and cellular contents of sugars and proteins during spruce somatic embryogenesis suggests a regulatory role of exogenous sucrose in embryo development. *J Exp Bot* 52(365):2301-11.
- Karthikeyan, A. S., Varadarajan, D. K., Jain, A., Held, M. A., Carpita, N. C. and Raghothama, K. G. 2007. Phosphate starvation responses are mediated by sugar signaling in Arabidopsis. *Planta* 225(4):907-18.
- Kenton, A., Parokony, A. S., Gleba, Y. Y. and Bennett, M. D. 1993. Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. *Mol Gen Genet* 240(2):159-69.
- King, R. W. and Ben-Tal, Y. 2001. A florigenic effect of sucrose in *Fuchsia hybrida* is blocked by gibberellin-induced assimilate competition. *Plant Physiol* 125(1):488-96.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R. and others. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947-8.
- Lei, M., Liu, Y., Zhang, B., Zhao, Y., Wang, X., Zhou, Y., Raghothama, K. G. and Liu, D. 2011. Genetic and genomic evidence that sucrose is a global regulator of plant responses to phosphate starvation in Arabidopsis. *Plant Physiol* 156(3):1116-30.
- Lemoine, R., Burkle, L., Barker, L., Sakr, S., Kuhn, C., Regnacq, M., Gaillard, C., Delrot, S. and Frommer, W. B. 1999. Identification of a pollen-specific sucrose transporter-like protein NtSUT3 from tobacco. *FEBS Letters* 454(3):325-30.
- Letunic, I. and Bork, P. 2018. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res* 46(D1):D493-D496.
- Li, F., Wu, B., Qin, X., Yan, L., Hao, C., Tan, L. and Lai, J. 2014. Molecular cloning and expression analysis of the sucrose transporter gene family from *Theobroma cacao* L. *Gene* 546(2):336-41.
- Livak, K. J. and Schmittgen, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻($\Delta\Delta C(T)$) Method. *Methods* 25(4):402-8.
- Marger, M. D. and Saier, M. H., Jr. 1993. A major superfamily of transmembrane facilitators that catalyze uniport, symport and antiport. *Trends in Biochemical Sciences* 18(1):13-20.
- Noiraud, N., Delrot, S. and Lemoine, R. 2000. The sucrose transporter of celery. Identification and expression during salt stress. *Plant Physiol* 122(4):1447-55.
- Okubo-Kurihara, E., Higaki, T., Kurihara, Y., Kutsuna, N., Yamaguchi, J. and Hasezawa, S. 2011. Sucrose transporter NtSUT4 from tobacco BY-2 involved in plant cell shape during miniprotoplast culture. *Journal of Plant Research* 124(3):395-403.
- Payyavula, R. S., Tay, K. H., Tsai, C. J. and Harding, S. A. 2011. The sucrose transporter family in *Populus*: the importance of a tonoplast PtaSUT4 to biomass and carbon partitioning. *The Plant Journal* 65(5):757-70.
- Peng, D., Gu, X., Xue, L. J., Leebens-Mack, J. H. and Tsai, C. J. 2014. Bayesian phylogeny of sucrose transporters: ancient origins, differential expansion and convergent evolution in monocots and dicots.

Front Plant Sci 5:615.

Pommerrenig, B., Papini-Terzi, F. S. and Sauer, N. 2007. Differential regulation of sorbitol and sucrose loading into the phloem of *Plantago major* in response to salt stress. *Plant Physiol* 144(2):1029-38.

Rastogi, S. and Liberles, D. A. 2005. Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evol Biol* 5:28.

Sauer, N. 2007. Molecular physiology of higher plant sucrose transporters. *FEBS Lett* 581(12):2309-17.

Schmidt, G. W. and Delaney, S. K. 2010. Stable internal reference genes for normalization of real-time RT-PCR in tobacco (*Nicotiana tabacum*) during development and abiotic stress. *Mol Genet Genomics* 283(3):233-41.

Sivitz, A. B., Reinders, A., Johnson, M. E., Krentz, A. D., Grof, C. P., Perroux, J. M. and Ward, J. M. 2007. Arabidopsis sucrose transporter AtSUC9. High-affinity transport activity, intragenic control of expression, and early flowering mutant phenotype. *Plant Physiology* 143(1):188-98.

Sivitz, A. B., Reinders, A. and Ward, J. M. 2008. Arabidopsis sucrose transporter AtSUC1 is important for pollen germination and sucrose-induced anthocyanin accumulation. *Plant Physiol* 147(1):92-100.

Sohpal, V. K., Dey, A. and Singh, A. 2010. MEGA biocentric software for sequence and phylogenetic analysis: a review. *Int J Bioinform Res Appl* 6(3):230-40.

Srivastava, A. C., Dasgupta, K., Ajieren, E., Costilla, G., McGarry, R. C. and Ayre, B. G. 2009. Arabidopsis plants harbouring a mutation in AtSUC2, encoding the predominant sucrose/proton symporter necessary for efficient phloem transport, are able to complete their life cycle and produce viable seed. *Annals of Botany* 104(6):1121-8.

Stolz, J., Ludwig, A., Stadler, R., Biesgen, C., Hagemann, K. and Sauer, N. 1999. Structural analysis of a plant sucrose carrier using monoclonal antibodies and bacteriophage lambda surface display. *FEBS Letters* 453(3):375-9.

Ugglä, C., Magel, E., Moritz, T. and Sundberg, B. 2001. Function and dynamics of auxin and carbohydrates during earlywood/latewood transition in scots pine. *Plant Physiol* 125(4):2029-39.

Usha, B., Bordoloi, D. and Parida, A. 2015. Diverse expression of sucrose transporter gene family in *Zea mays*. *Journal of Genetics* 94(1):151-4.

Vaughn, M. W., Harrington, G. N. and Bush, D. R. 2002. Sucrose-mediated transcriptional regulation of sucrose symporter activity in the phloem. *Proc Natl Acad Sci USA* 99(16):10876-80.

Weschke, W., Panitz, R., Sauer, N., Wang, Q., Neubohn, B., Weber, H. and Wobus, U. 2000. Sucrose transport into barley seeds: molecular characterization of two transporters and implications for seed development and starch accumulation. *The Plant Journal* 21(5):455-67.

Zhang, Q., Hu, W., Zhu, F., Wang, L., Yu, Q., Ming, R. and Zhang, J. 2016. Structure, phylogeny, allelic haplotypes and expression of sucrose transporter gene families in *Saccharum*. *BMC Genomics* 17:88.

Zheng, Q. M., Tang, Z., Xu, Q. D. and eng, X. X. 2014. Isolation, phylogenetic relationship and expression profiling of sugar transporter genes in sweet orange (*Citrus sinensis*). *Plant Cell* 119(3):609-624.

Table and Figure legends

Table 1 Characteristics of the tobacco *SUT* genes

Table 2 Identity matrix of NtSUT amino acid sequences and their coding sequences

Table 3 The SUT orthologs among *N. tabacum*, *N. sylvestris*, and *N. tomentosiformis*

Fig. 1. Location of the *NtSUT* genes on *N. tabacum* chromosomes. Chromosome sizes (Mb) and numbers are indicated at the end of each scaffold. The paralogous genes were linked with solid lines.

Fig. 2. Alignment of the predicted amino acid sequences of the tobacco SUT proteins. The red box indicated the conserved histidine residue. The black box represented the LX2LL motif. The two yellow boxes indicated the (R/K)X2/3(R/K) motifs. The black lines indicated the MFS domains.

Fig. 3. Phylogenetic analysis and exon-intron structure of plant *SUT* genes. The genes used in the analysis were listed in the Table S3.

Fig. 4. Expression patterns of the *NtSUT* genes in common tobacco. The transcription levels of the *NtSUT* genes were detected from different tissues at flowering stage (a), different parts of tobacco leaves at flowering stage (B), the roots of different developmental stages (C), and the leaves of different developmental stages (D). Values are mean $\pm$ SD of three independent biological replicates.

Fig. 5. Transcription levels of the *SUT* genes in diploid tobacco plants. The relative expression levels of the *SUT* genes in the leaves (A), roots (B) of the *N. sylvestris* and *N. tomentosiformis* plants. Values are mean $\pm$ SD of three independent biological replicates.

Fig. 6. Expression profiles of the *NtSUT* genes in response to drought, salt and Pi

starvation stresses. The relative expression levels of the *NtSUTs* genes in the roots, stems and leaves under Pi starvation (A, B, C), drought (D, E, F) and salt stresses (G, H, I), respectively. Values are mean $\pm$ SD of three independent biological replicates with $2^{-\Delta\Delta C_t}$ methods.

Supplementary material

For Review Only

Fig. 1

Fig. 2

Fig. 3

Fig. 4

Fig. 5

Fig. 6

Table1

	Gene ID	gDNA size (bp)	CDS size (bp)	Amino acid size	Molecular weight (kDa)	Isoelectric point
<i>N. tabacum</i>	Ntab0067760	13064	1506	501	54.27	8.89
	Ntab0158640	9789	1506	501	54.27	8.89
	Ntab0840010	7950	1506	501	54.33	9.07
	Ntab0332220	5813	1506	501	54.32	8.55
	Ntab0245750	5832	1506	501	54.31	8.55
	Ntab0472550	10302	1506	501	54.36	8.55
	Ntab0397480	6438	1818	605	65.13	6.65
	Ntab0049540	9562	1737	578	62.69	7.16
	Ntab0480720	5595	1533	510	54.11	9.43
	Ntab0592260	5264	1530	509	54.02	9.43
	Ntab0430110	7663	1566	521	55.96	8.96
<i>N. sylvestris</i>	Nsyl0255840	9490	1506	501	54.3	8.89
	Nsyl0469380	17022	1506	501	54.23	7.76
	Nsyl0087130	6435	1818	605	65.13	6.64
	Nsyl0323720	5602	1536	511	54.23	9.43
	Nsyl0151140	7773	1566	521	55.96	8.96
<i>N. tomentosiformis</i>	Ntom0251120	7900	1506	501	54.36	9.07
	Ntom0030920	2198	1101	366	39.79	6.13
	Ntom0237460	8863	1818	605	65.09	6.78
	Ntom0227400	2364	1212	403	43.65	8.68

Table 2

		Amino acid identity										
		Ntab0067	Ntab0158	Ntab0840	Ntab0332	Ntab0245	Ntab0472	Ntab0397	Ntab0049	Ntab0480	Ntab0592	Ntab0430
		760	640	010	220	750	550	480	540	720	260	110
Nucleotide identity	Ntab0067760		100	97.6	85.63	85.83	85.43	41.19	37.62	48.25	48.25	42.78
	Ntab0158640	99.87		97.6	85.63	85.83	85.43	41.19	37.46	48.25	48.25	42.78
	Ntab0840010	98.07	98.21		86.83	87.03	86.63	41.02	36.92	48.64	48.64	43.35
	Ntab0332220	84.29	84.29	85.09		99.8	98.6	39.87	36.47	48.25	48.25	43.73
	Ntab0245750	84.36	84.36	85.16	99.93		98.8	40.03	36.47	48.44	48.44	43.73
	Ntab0472550	84.36	84.03	85.16	98.14	98.21		40.03	36.3	48.44	48.44	43.54
	Ntab0397480	47.73	47.73	46.86	46.36	46.41	46.33		89.75	38.79	38.95	34.99
	Ntab0049540	41.72	41.72	41.35	39.89	39.89	39.53	90.99		35.35	35.19	32.05
	Ntab0480720	53.03	54.98	55.24	56.52	56.58	56.78	46.71	41.96		98.04	65.01
	Ntab0592260	54.72	53.45	54.85	55.81	55.88	56.07	47	42.18	96.87		65.71
Ntab0430110	51.92	51.8	52.43	52.45	52.52	51.74	45.31	41.15	69.85	69.92		

Table 3

Gene symbol	<i>N. tabacum</i>		<i>N. sylvestris</i>		<i>N. tomentosiformis</i>	
	Gene ID	Protein symbol	Gene ID	Protein symbol	Gene ID	Protein symbol
<i>SUT1</i>	Ntab0480720	NtSUT1s	Nsyl0323720	NsylSUT1	--	--
	Ntab0592260	NtSUT1t	--	--	NO	--
<i>SUT2</i>	Ntab0397480	NtSUT2s	Nsyl0087130	NsylSUT2	--	--
	Ntab0049540	NtSUT2t	--	--	Ntom0237460	NtomSUT2
<i>SUT3</i>	Ntab0430110	NtSUT3s	Nsyl0151140	NsylSUT3	--	--
	NO	--	--	--	Ntom0227400	NtomSUT3
<i>SUT4</i>	Ntab0332220	NtSUT4sa	Nsyl0469380	NsylSUT4	--	--
	Ntab0245750	NtSUT4sb				
	Ntab0472550	NtSUT4t	--	--	Ntom0030920	NtomSUT4
<i>SUT5</i>	Ntab0067760	NtSUT5sa	Nsyl0255840	NsylSUT5	--	--
	Ntab0158640	NtSUT5sb				
	Ntab0840010	NtSUT5t	--	--	Ntom0251120	NtomSUT5