Genome-Wide Identification and Evolution Analysis of Trehalose-6-Phosphate Synthase Gene Family in *Nelumbo nucifera*

Qijiang Jin[1], Xin Hu[1], Xin Li[2], Bei Wang[1], Yanjie Wang[1], Hongwei Jiang[2], Neil Mattson[3] and Yingchun Xu[1*]

[1] College of Horticulture, Nanjing Agricultural University, Nanjing, China, [2] Leisure Agricultural Section, Institute of Agricultural Science of Taihu Lake District, Suzhou, China, [3] Horticulture Section, School of Integrative Plant Science, Cornell University, Ithaca, NY, USA

Trehalose-6-phosphate synthase (TPS) plays a key role in plant carbohydrate metabolism and the perception of carbohydrate availability. In the present work, the publicly available *Nelumbo nucifera* (lotus) genome sequence database was analyzed which led to identification of nine lotus *TPS* genes (*NnTPS*). It was found that at least two introns are included in the coding sequences of *NnTPS* genes. When the motif compositions were analyzed we found that *NnTPS* generally shared the similar motifs, implying that they have similar functions. The d_N/d_S ratios were always less than 1 for different domains and regions outside domains, suggesting purifying selection on the lotus *TPS* gene family. The regions outside TPS domain evolved relatively faster than NnTPS domains. A phylogenetic tree was constructed using all predicted coding sequences of lotus *TPS* genes, together with those from *Arabidopsis*, poplar, soybean, and rice. The result indicated that those *TPS* genes could be clearly divided into two main subfamilies (I-II), where each subfamily could be further divided into 2 (I) and 5 (II) subgroups. Analyses of divergence and adaptive evolution show that purifying selection may have been the main force driving evolution of plant *TPS* genes. Some of the critical sites that contributed to divergence may have been under positive selection. Transcriptome data analysis revealed that most *NnTPS* genes were predominantly expressed in sink tissues. Expression pattern of *NnTPS* genes under copper and submergence stress indicated that NNU_014679 and NNU_022788 might play important roles in lotus energy metabolism and participate in stress response. Our results can facilitate further functional studies of *TPS* genes in lotus.

Keywords: trehalose-6-phosphate synthase, *Nelumbo nucifera*, protein properties, molecular evolution, expression pattern

INTRODUCTION

Trehalose, known as a non-reducing disaccharide of glucose, has been found in wide range of diverse organisms, for example, bacteria, fungi, invertebrates, and plants (Lyu et al., 2013). In higher plants, trehalose accumulates at quite low concentrations, with the exception of resurrection plants (Paul et al., 2008). It was found that in trehalose-rich organisms, trehalose functions both

as a carbon source and as an osmoprotectant such in certain microorganisms and insects (Elbein et al., 2003; Zang et al., 2011). There is crucial evidence that trehalose helps prevent heat, osmotic, nutrient, dehydration stress, and toxic chemicals from damaging organisms (Chary et al., 2008). Trehalose synthesis in plants only occurs via the TPS/TPP pathway (Avonce et al., 2006; Jiang et al., 2010). In the first step of this pathway, trehalose-6-phosphate synthase (TPS) joins uridine diphosphoglucose with glucose 6-phosphate, thus forming trehalose 6-phosphate (T6P; Jiang et al., 2010). Then, T6P is dephosphorylated as catalyzed by trehalose-6-phosphate phosphatase (TPP) to produce the disaccharide trehalose. TPS expression levels in cotton increased under drought stress (Mu et al., 2016). When maize was under salt and temperature stress, *TPS* genes were upregulated (Jiang et al., 2010). Genes which controlled the production of enzymes of trehalose synthesis, including the *AtTPS1* gene from *Arabidopsis thaliana* (Blázquez et al., 1998; Jiang et al., 2010). These findings made it practical to attain transgenic plants altered by a *TPS* gene of plant origin. Most transgenic plants with overexpressed TPS and/or TPP genes exhibited great tolerance to abiotic stress (Garg et al., 2002; Jang et al., 2003; Karim et al., 2007; Ge et al., 2008).

Although numerous studies found that over-expressing *TPS* and *TPP* genes could improve abiotic stress tolerance (Holmström et al., 1996; Jang et al., 2003; Schluepmann et al., 2004; Jiang et al., 2010), observed stress tolerance of transgenic lines did not correlate with trehalose amount (van Dijken et al., 2004). Researchers have investigated function of plant trehalose biosynthetic enzymes by identifying mutations in *Arabidopsis TPS1* gene and seeing which were still effective in recovering homologous gene function for yeast mutants deficient in TPS functions (Blázquez et al., 1998; Müller et al., 2001; Vogel et al., 2001). Beyond these stress response, recent intriguing evidence has implicated that *TPS* genes are important modulators in plant development and inflorescence architecture (Satoh-Nagasawa et al., 2006).

The plant *TPS* gene family belongs to a small gene family with multiple copies. Members of this gene family showed widespread functional diversification (Lunn, 2007; Chary et al., 2008; Vandesteene et al., 2010; Yang et al., 2012). The *Arabidopsis TPS* gene family contains 11 members (*AtTPS1-11*; Shima et al., 2007); rice also contains 11 members (*OsTPS1-11*; Tuskan, 2006; Vandesteene et al., 2010), and poplar contains 12 members (*PtTPS1-12*; Vandesteene et al., 2010). The TPS family is divided into two distinct clades, class I and class II, and this dichotomy occurred early in the evolution of green plants (Jang et al., 2003; Shima et al., 2007; Suzuki et al., 2008; Ramon et al., 2009). Apart from polyploid and allopolyploid species, *Arabidopsis* is quite uncommon among all the angiosperms which have large number class I *TPS* genes (Delorge et al., 2015). Particularly, distinct characteristics of class I and class II *TPS* genes are displayed in gene expression patterns, copy number, physiological functions, and enzyme activity.

Sacred lotus (*Nelumbo nucifera*, Gaertn.) is a basal eudicot which is of significant cultural, agricultural, medicinal, as well as religious importance. TPS plays important role in plant response to environmental stimuli, and thus the study of TPS

would be very important for lotus breeding and the research of stress-resistance mechanism in lotus. In this study, we investigated the lotus *TPS* genes from whole genome-wide studies and their evolutionary relationship with other *TPS* genes. We explored TPS gene family members' expression patterns in various tissues and in response to various stresses. Having conducted a thorough analysis of gene sequences, molecular evolution, gene structures, and gene expression patterns, we offer an effective framework for further functional characterization of lotus *TPS* gene families.

METHODS

Identification of Lotus *TPS* Genes

The lotus genome DNA data (Ming et al., 2013) was downloaded from COGE (v2, id16885, https://www.genomevolution.org/coge/) and gene prediction was performed with hidden Markov model (HMM) program (Eddy, 1998). Hidden Markov models of TPP and TPS domain were obtained from PFAM (http://pfam.xfam.org/). Motif scanning in TPS was done by PROSITE scan (Hulo et al., 2006). Gene structures of *TPS* genes were analyzed on the Gene Structure Display Server 2.0 (GSDS; http://gsds.cbi.pku.edu.cn/).

Multiple Sequence Alignment and Phylogenetic Analysis

Multiple sequence alignments in full-length protein sequences were executed by using MUSCLE 3.52, and this was followed by manual refinement and comparisons (Edgar, 2004; Cao et al., 2014). Phylogenetic analysis was performed with MEGA 6.0 using a maximum likelihood method with the JTT (Jones, Taylor, and Thornton) amino acid substitution model. Bootstrap support values were estimated using 1000 pseudo-replicates.

Protein Properties and Conserved Motif Analysis

Protein properties of *TPS* genes, e.g., the molecular weight (MW), isoelectric point (*p*I), and grand average of hydropathicity (GRAVY) were calculated using ProParam (http://web.expasy.org/protparam). Analysis for conserved motifs in TPS proteins was carried out using MEME (http://meme.sdsc.edu/meme/cgi-bin/meme.cgo) with the default parameters.

Functional Divergence Analysis

To analyze the functional divergence between groups and predict the amino acid residues resulting in functional difference, program DIVERGE (Gu and Velden, 2002) was applied to calculate the coefficient of type I functional divergence θ_I as well as that of type II functional divergence θ_{II}. The likelihood test is approximated to a χ^2 distribution with 1 degree of freedom. After duplication, a significant θ_I means site-specific changed selective constraints (Xun, 2003; Gu, 2006; Yang et al., 2009). After duplication, a significant θ_{II} indicates amino acid physiochemical property's obvious shifts (Xun, 2003; Gu, 2006; Yang et al., 2009).

Site-Specific Selection Assessment and Testing

The value of d_N/d_S ratio (or ω, the ratio of non-synonymous to synonymous distances) for each group was analyzed with the program codeml from PAML version 4 (Yang, 2007). A likelihood ratio test (LRT) was used to detect variation in ω among sites between M0 vs. M3 and M7 vs. M8, respectively. The χ^2 statistics with degrees of freedom equal to the differences between parameters is twice the log likelihood difference between the two models in the LRT (3 for M0/M3 tests and 2 for M7/M8 tests; Mondragon-Palomino and Gaut, 2005). If the LRT was statistically significant and the evolution of genes below groups, then sites are assumed to be under positive selection pressure through Bayes methods (Yang et al., 2005).

In silico Expression Analysis

To identify expression pattern of *TPS* genes, expression level in different tissues and abiotic stress treatments, transcriptome datasets of lotus was analyzed. The data sets obtained from GenBank include: rhizome apical meristem (SRS413495), rhizome elongation zone (SRS413496), rhizome internode (SRS414455, SRS414456, SRS414457, SRS414458, SRS414459), leaf (SRS413488, SRS413489, SRS413490, SRS413491, SRS413492), petiole (SRS413498, SRS413499, SRS413500, SRS413501, SRS413502), root (SRS412272, SRS412273, SRS412278, SRS412279, SRS412280). The library preparation methods and statistics of sequencing data were detailed in Kim et al. (2013). To take advantage of our in house transcriptome data, we analyzed the expression pattern of *TPS* genes in response to submerged and 5 mm Copper-treated

TABLE 1 | List of the *NnTPS* genes identified in this study.

Gene accession number	Scaffold ID	Predicted position	CDS (bp)	Deduced polypeptide				Predicted subcellular localization
				Length (aa)	MW (kDa)	pI	GRAVY	
NNU_020889	Scaffold_3	36472545-36476497	2598	865	97.58	5.86	−0.195	Cytoplasmic
NNU_016707	Scaffold_5	12364436-12368019	2580	859	97.05	6.02	−0.143	Plasma membrane
NNU_024672	Scaffold_8	6741636-6749305	2535	845	95.37	5.72	−0.167	Cytoplasmic
NNU_014679	Scaffold_20	449424-453667	2604	867	97.60	6.10	−0.183	Cytoplasmic
NNU_020115	Scaffold_17	1846465-1854756	2616	871	97.98	5.64	−0.183	Cytoplasmic
NNU_004429	Scaffold_4	32965362-32971524	2586	861	97.13	6.33	−0.092	Plasma membrane
NNU_000253	Scaffold_4	13382224-13386331	2649	883	100.39	5.76	−0.187	Cytoplasmic
NNU_016432	Scaffold_5	37054048-37058407	2814	937	106.30	5.94	−0.155	Plasma membrane
NNU_022788	Scaffold_9	4294971-4298958	2547	848	96.35	5.60	−0.275	Cytoplasmic

FIGURE 1 | Structure and phylogenetic analysis of NnTPS proteins. The unrooted phylogenetic tree resulting from the full-length amino acid alignment of all the NnTPS proteins is shown on the left side of the figure. Exon-intron structures of the identified *NnTPS* genes are shown on the right side. The graphic representation of the optimized gene model displayed using GSDS.

conditions. Three-month-old lotus seedlings were used for submergence (24 h) and 5 mm Copper-treatment (24 h). To achieve submergence, seedlings were transferred to a plastic tank and water was added into the tank to a water depth of 20 cm from top of lotus. The sample without copper or submergence treatment was the control (Con). After the treatments, total RNA was isolated using the CTAB-LiCl method. For mRNA-seq, three sets of total RNA were used for cDNA library construction and sequencing (without replicates) at the Beijing Genomics Institute (BGI, Shenzhen, China), following the manufacture's protocols.

All SRA format data were converted into FASTQ using the SRA toolkit. All raw reads were first filtered with NGS QC Toolkit (v2.3.3) to remove adapter sequences, reads containing poly-N, and low-quality sequences ($Q < 20$). An index of the lotus genome was built using Bowtie 2 and paired-end or single-end clean reads were aligned to the reference genome using TopHat. The expression levels (fragments per kb per million mapped read, FPKM) from the representative transcript were determined using cufflinks program (cufflinks v2.2.142) with default settings. We used a threshold of FPKM $\geq$ 1 to define a gene as "expressed."

Real-Time Quantitative PCR

Two month-old lotus seedlings were used for submergence treatment and copper treatment. Seedlings were completely submerged or treated with 5 mm copper for 24 h (as described above) and then rhizomes were sampled. The sample without copper or submergence treatment was the control (Con). Meanwhile, different tissues of control seedlings were sampled. Tissues were homogenized with mortar and pestle in liquid nitrogen. Total RNA was isolated using CTAB-LiCl method. About 4 μg of total RNA was reverse-transcribed using an oligo(dT) primer and SuperScript Reverse Transcriptase

(Invitrogen, USA). Real-time quantitative PCR was conducted using SYBR green (TaKaRa Biotechnology) on Mastercycler ep realplex real-time PCR system (Eppendorf, Hamburg, Germany) with a final volume of 20 μl per reaction. The specific primers were listed in **Table S1**. The relative transcript abundance was normalized using lotus *Actin* gene. Gene expression in various tissues were represented by *NnTPS/Actin*. Fold changes of genes under copper stress and submergence treatment were

FIGURE 2 | Pairwise sequence identity of full-length TPS proteins. Pairwise sequence identities between TPS domain, TPP domain, full length protein sequence, and sequence outside domain were calculated. The boxplot shows the median (black line), interquartile range (box), and maximum and minimum scores (whiskers) for each data set.
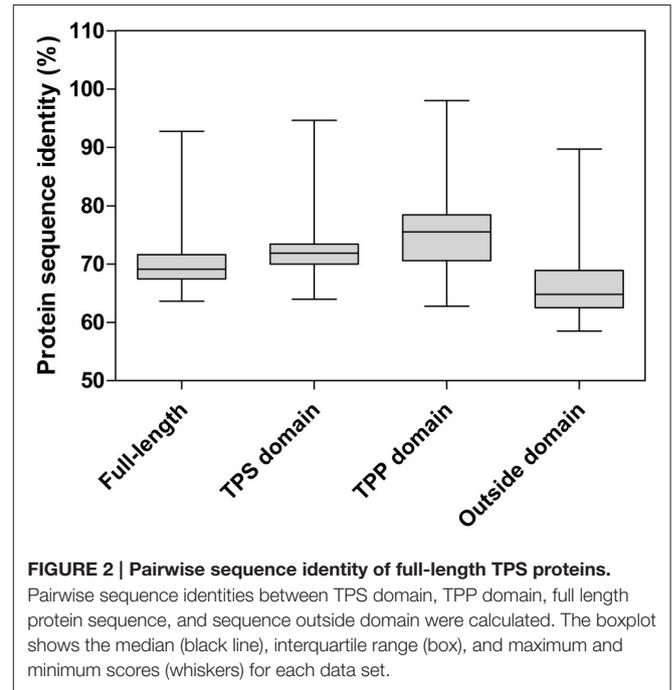
TABLE 2 | Functional divergence between groups of the *TPS* gene family in plants.

	$\Theta_I \pm SE$[a]	Type I		$Q_k \geq 0.80$[b]	$Q_k \geq 0.95$[c]	Type II	
		LRT	P			$\Theta_{II} \pm SE$	P
II-1/II-2	0.217 ± 0.101	4.607	0.032	1	0	-0.014 ± 0.072	0.846
II-1/II-3	0.352 ± 0.118	8.919	0.003	0	0	0.014 ± 0.08	0.861
II-1/II-4	0.444 ± 0.107	17.262	0.000	4	1	-0.019 ± 0.096	0.843
II-1/II-5	0.525 ± 0.127	16.947	0.000	6	0	-0.043 ± 0.107	0.688
II-1/I-1	0.57 ± 0.123	21.432	0.000	14	0	0.477 ± 0.066	0
II-2/II-3	0.143 ± 0.083	2.998	0.085	0	0	-0.002 ± 0.087	0.982
II-2/II-4	0.191 ± 0.075	6.425	0.011	1	0	-0.043 ± 0.102	0.673
II-2/II-5	0.266 ± 0.07	14.453	0.000	3	1	-0.04 ± 0.113	0.723
II-2/I-1	0.386 ± 0.092	17.500	0.000	6	0	0.425 ± 0.075	0
II-3/II-4	0.153 ± 0.075	4.162	0.041	1	0	-0.046 ± 0.106	0.664
II-3/II-5	0.281 ± 0.079	12.654	0.000	2	0	-0.066 ± 0.116	0.569
II-3/I-1	0.434 ± 0.089	24.043	0.000	6	1	0.487 ± 0.074	0
II-4/II-5	0.081 ± 0.047	2.979	0.085	0	0	-0.101 ± 0.125	0.419
II-4/I-1	0.304 ± 0.089	11.711	0.001	2	0	0.408 ± 0.088	0
II-5/I-1	0.415 ± 0.093	19.781	0.000	4	1	0.432 ± 0.093	0

[a] The coefficient of functional divergence between two subgroups and its standard error.
[b] The number of critical amino acid residues with posterior probability (Q_k) >0.80.
[c] The number of critical amino acid residues with posterior probability (Q_k) >0.95.

values relative to control samples after normalization to *Actin* transcript levels. RT-qPCR data presented were the means ± SE of at least four independent experiments. Differences among treatments were analyzed by on-way ANOVA, taking $P < 0.05$ as significant according to multiple comparisons or *t*-test.

RESULTS AND DISCUSSION

Genome-Wide Identification of Lotus *TPS* Gene Family Members

Accumulating evidence showed that *TPS* genes play a vital role in modulating plants response to abiotic stresses, i.e., *OsTPS1* could improve rice-seed tolerance to low temperature, salt, and drought (Jiang et al., 2010; Vandesteene et al., 2010; Li et al., 2011; Zang et al., 2011). However, identity and function of lotus *TPS* gene family was previously undescribed. The lotus genome was sequenced recently, which makes it possible to carry out genome-wide study on *TPS* genes in lotus. Plant TPS proteins contain a TPS as well as a trehalose-6-phopshate phosphatase (TPP) domain (Yang et al., 2012). Thus, HMM profile of two conserved TPS domains (TPS and TPP domain) was used to search local lotus protein data with HMMER program (Sonnhammer et al., 1998; Ross et al., 2007). Nine lotus TPSs (NnTPSs) gene family members were identified in

lotus genome (**Table 1**). Taking advantage of these identified genes, we further conducted a homology search from genome and amino acid sequences of lotus. However, the search did not yield new potentially TPS encoding genes for lotus beyond the nine initially identified. The identified NnTPSs contained a similar number of amino acid from 845 (NNU_024672) to 937 (NNU_016432). The isoelectric point (PI) ranged from 5.60 (NNU_022788) to 6.33 (NNU_004429). The protein weight ranged from 95.37 (NNU_024672) to 106.21 (NNU_016432) kDa. Nine NnTPS scattered in 7 scaffolds. Scaffold 4 and 5 had 2 *NnTPS* genes, respectively. Previous studies have identified 11 TPS in *Arabidopsis*, 11 in rice, and 12 in poplar (Yang et al., 2012), suggesting that the number of genes remains stable in this family and that usually duplicated copies of *TPS* genes are not retained after a whole-genome duplication event (Hao et al., 2014).

Gene Structure Analysis of TPS in Lotus

To assess the evolutionary relationship of nine *NnTPS* genes, a small evolutionary tree was constructed using the nine *NnTPS* genes to exam *NnTPS* gene structures which are typical imprints of evolution within some gene families. The length of gene introns was larger and genetic structures were more complicated in NNU_024672, NNU_020115, NNU_000253, and NNU_016432 than the other *NnTPS* genes (**Figure 1**). All

FIGURE 3 | Schematic diagram of amino acid motifs of TPS protein. Motif analysis was performed using MEME 4.11.2 as described in the methods. The different-colored boxes represent different motifs and their position in each TPS sequence. The sequences of key motifs that located in TPS and TPP domain were indicated.
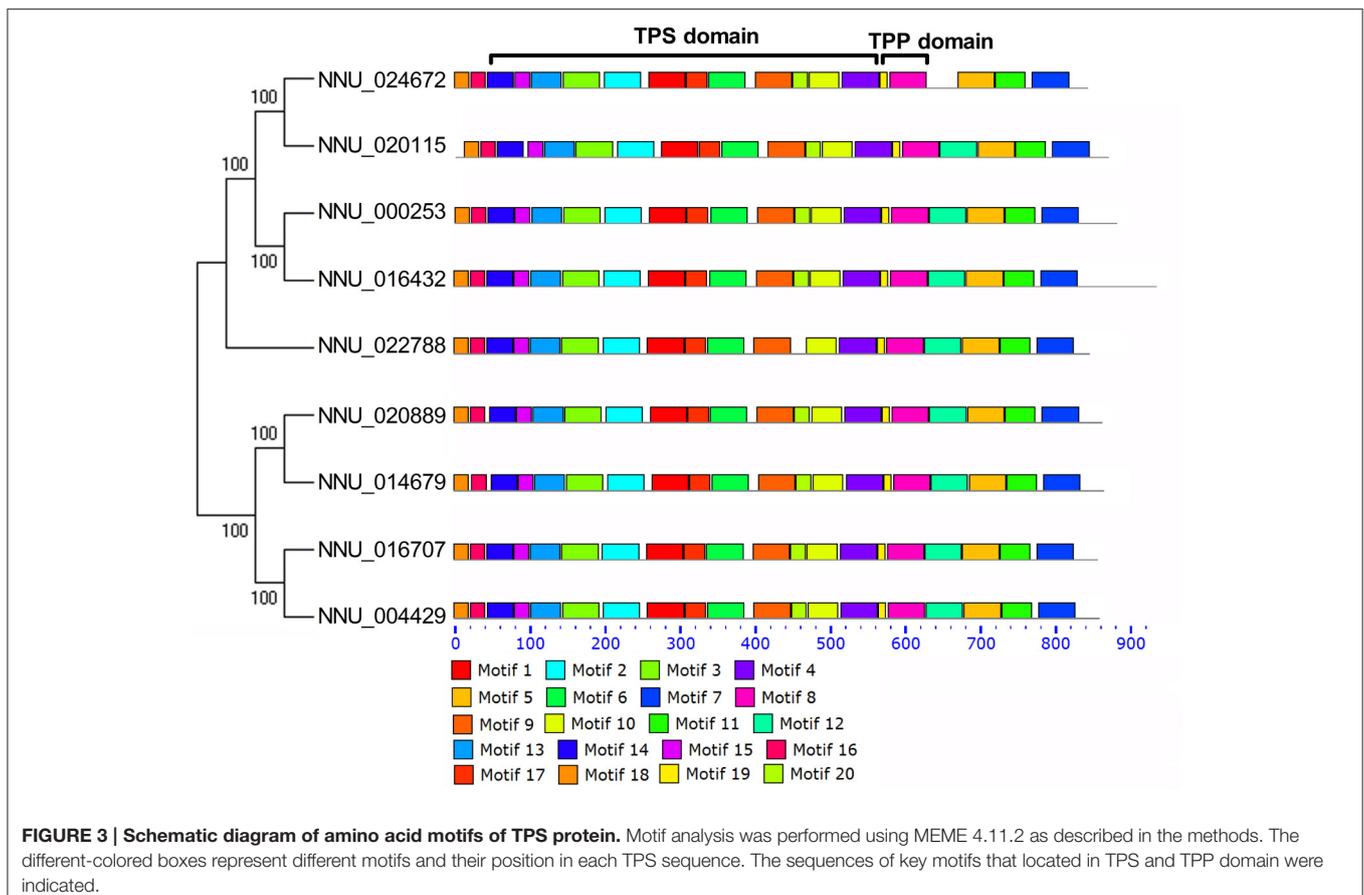
NnTPS genes contain at least one intron in their CDSs: most of them have three introns, one (NNU_004429) has four introns and one (NNU_022788) has five introns. According to the previous reports (Nuruzzaman et al., 2010; Hu et al., 2016), the rate of intron gain is slower than that of intron loss after segmental duplication in rice. Thus, it can be concluded that NNU_022788 might be the most ancestral with the other *NnTPS* genes derived from it. Compared with similar CDS length (2535–2814 bp) among 9 *NnTPS* genes, their gene length is even more variable (3583–8291 bp).

Conserved Sequences in TPS Proteins

A homology analysis of nine full-length protein sequences indicated that nucleotide identity among them ranged from 63.66% (between NNU_016707 and NNU_022788) to 92.78% (between NNU_024672 and NNU_020115; **Figure 2**; **Figure S1**). Reflecting what was found in the protein sequences, a high similarity (ranging from 65.14 to 90.26%) was shared among nine *NnTPSs* full-length sequences. Several residues in the catalytic center were highly conserved in all *NnTPS* genes (**Figure S1**) suggested that corresponding genes encoding active enzymes. Nevertheless, relatively high divergence was observed in some regions of amino acid sequences outside of the domain. It is likely that these un-conserved regions might contribute to functional distinction.

Subcellular localization of NnTPS isoforms was predicted using web subcellular localization predictor server (http://cello.life.nctu.edu.tw/; **Table 1**). Translated proteins of NNU_016707, NNU_004429, and NNU_016432 were predicted as being located on plasma membrane while most of NnTPSs were located in cytoplasm.

The identification of motifs for all NnTPS proteins were performed by software MEME (Bailey and Elkan, 1994) with default settings. We obtained sequences of 20 motifs and the distribution of these motifs in TPS proteins (**Figure 3**; **Table S2**). These motifs were conserved in all *NnTPS* genes except NNU_024672 which lacked motif12. According to **Figure 3**, motifs 1, 2, 3, 4, 6, 9, 10, 13, 14, 15, 17, and 20 were located in the TPS domain and motifs 8 and 19 were located in TPP domain. These features in conserved motifs of NnTPS were also observed in other plants (Mu et al., 2016).

To explore whether selection pressure had affected *NnTPS* genes, we calculated ratio (ω) of the synonymous substitution rate (d_S) vs. the non-synonymous substitution rate (d_N) using the program codeml from PAML (version 4). Generally, d_N/d_S ratio >1 indicates positive selection, a ratio <1 indicates negative or purifying selection and a ratio = 1 indicates neutral evolution (Wang et al., 2005; Song et al., 2015). Results in **Figure 5** showed that d_N/d_S ratios were always less than 1 for different domains and regions outside domain, suggesting purifying selection on lotus *TPS* gene family. The d_N/d_S ratios outside NnTPS domain were found to be much higher than the ratios inside the NnTPS domains (**Figure 4**). These results showed that NnTPS domains evolved faster than regions outside TPS domain. The relaxed purifying or positive selection in the regions outside TPS domain may lead to the above results.

Relationships of Lotus TPS Family with Other Plant TPSs

A phylogenetic tree was used to reveal homologous relationships and evolutionary roots of TPS from *Arabidopsis*, rice, poplar (Yang et al., 2012), and soybean (Yang et al., 2012). The amino acid sequence alignment of TPS in these plants was conducted

TABLE 3 | Detection of positive selection under site-specific model.

Group	N[a]	d_N/d_S(ω) under M0[b]	2Δ/M3 vs. M0[c]	2Δ/M8 vs. M7[c]	M8 estimates[d]	Positively selected sites[e]
I-1	7	0.140	620.331**	123.550**	$p_1 = 0.117$, $\omega = 1.833$ β ($p = 1.000$, $q = 10.380$)	31
I-2	3	0.425	72.153**	16.600**	$p_1 = 0.031$, $\omega = 19.426$ β ($p = 0.428$, $q = 0.560$)	–
II-1	5	0.095	145.763**	1.774	$p_1 = 0.018$, $\omega = 1.328$ β ($p = 0.333$, $q = 2.962$)	–
II-2	8	0.079	610.610**	32.373**	$p_1 = 0.019$, $\omega = 217.11$ β ($p = 0.345$, $q = 3.406$)	4
II-3	9	0.104	609.630**	0.008	$p_1 = 0.00001$, $\omega = 3.705$ β ($p = 0.389$, $q = 2.518$)	–
II-4	13	0.149	981.750**	7.786**	$p_1 = 0.015$, $\omega = 1.263$ β ($p = 0.511$, $q = 2.624$)	–
II-5	10	0.164	852.843**	0.621	$p_1 = 0.002$, $\omega = 4.772$ β ($p = 0.532$, $q = 2.021$)	–

[a]*Number of the sequences in the group.*
[b]*d_N/d_S (ω) is the average ratio over sites under a codon model with one-ratio.*
[c]*Double asterisks represent $P < 0.01$ for chi-square (χ^2) test, while single asterisk represents $P < 0.05$.*
[d]*ω is estimated under model M8; p1 is the inferred proportion of positively selected sites; p and q are the parameters of the beta distribution.*
[e]*The amino acid, which is listed in positively selected position, had a posterior probability > 0.95 under M8.*

FIGURE 4 | The d_N/d_S ratios for different regions of lotus TPS proteins.
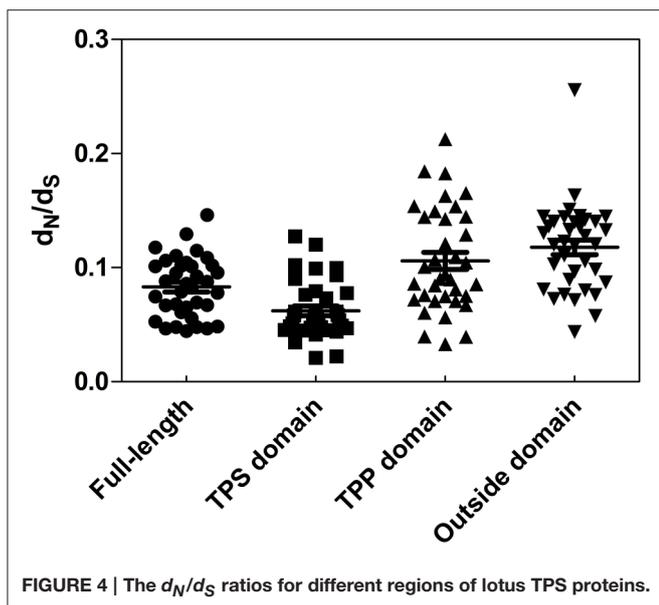
for phylogenetic tree construction using MEGA 5.1 (**Figure 5**). According to the phylogenetic tree and previous division of *Arabidopsis* and rice *TPS* genes, these *TPS* genes are divided into two subfamilies (I and II; Vandesteene et al., 2010; Zang et al., 2011). *Arabidopsis*, rice, soybean, and poplar contain four (*AtTPS1*, *2*, *3*, and *4*), one (*OsTPS1*), four (*GmTPS1*, *2*, *3*, and *4*), and one (*PtTPS1*) subfamily I TPS, respectively. However, no lotus TPS subfamily I members was found. It should be noted that the absence of subfamily I TPS in lotus might be due to incomplete genome sequencing coverage (86.5%). However, we cannot exclude the possibility that these genes have been fatally lost in evolution, as the number of *TPS* genes in subfamily I is far less than *TPS* genes in subfamily II in previous reported species (Zang et al., 2011). In order to describe paralogous and orthologous relations among this family, the subfamily II *TPS* genes were further divided into five group (II-1, 2, 3, 4, and 5) with high bootstrap support, suggesting that genes in each subgroup might share a similar origin. In most cases, subgroups are duplicates of at least one lineage. For example, subgroup II-2 and II-3 were duplicated after the separation between dicot and monocots. We identified nodes that lead to dicots (lotus, *Arabidopsis*, poplar, and soybeans)-specific and monocots (rice)-specific subfamilies. Group II-2 only contained lotus, *Arabidopsis*, and poplar Class II *TPS* genes, indicating that genes might have been lost from the rice genome. Subfamily I-1 only contained three *Arabidopsis* Class I TPS. Moreover, most proteins in lotus TPS family were contained in paralogous pairs, including 4 pairs of NnTPS accounting for 89% NnTPSs. This result indicated that most lotus *TPS* genes expanded in a species-specific manner.

Functional divergence of the duplicate genes has been recognized as an important source of evolutionary novelty (Yang et al., 2012). To evaluate potential functional divergence, type I and type II functional divergence between groups of the TPS family were estimated by posterior analysis (Gu and

FIGURE 5 | Phylogenetic relationships of the TPS from lotus, Arabidopsis (AtTPS), poplar (PtTPS), soybean (GmTPS), and rice (OsTPS). Red circles indicate the most recent common ancestral *TPS* genes among dicots and rice.
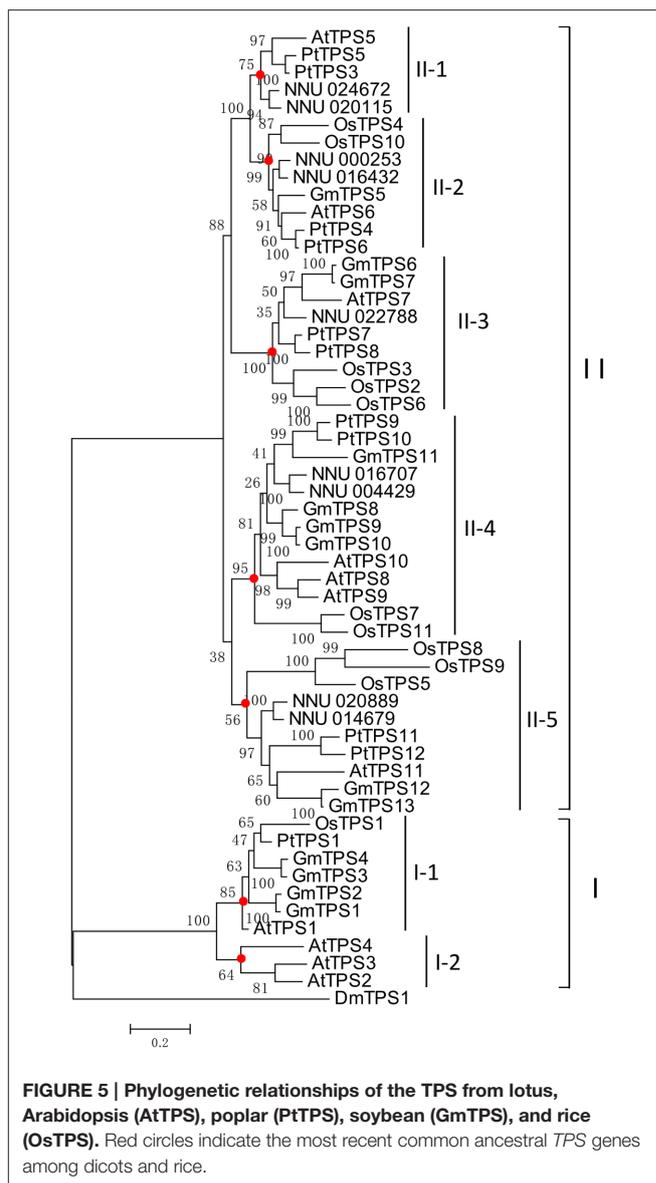
Velden, 2002). For this analysis, the collected plant TPS proteins were used and estimation was based on multiple alignments of proteins for any two groups (**Table 2**). Previous work proposed several possible evolutionary fates of duplicate genes (Hughes, 1994; Force et al., 1999). Our results showed that most type I coefficients (θ_I) of functional divergence were significantly greater than zero ($P < 0.01$), while few of the type II coefficients (θ_{II}) was statistically greater than zero. These results suggested that type I functional divergence was the dominant pattern for evolution of TPS family in these plants, and significantly site-specific altered selective constraints should contribute to most of the TPS members which would, after the diversification, cause a group-specific functional evolution.

Further, the calculation of site-specific profiles on the basis of posterior analysis for all group pairs with functional divergence would predict the critical amino acid residues, which are

accountable for the functional divergence. Among aligned sites, most of them had low posterior probabilities. $Q_k \geq 0.80$ and $Q_k \geq 0.95$ were utilized empirically as cutoffs to determine type I functional divergence-related residues existing between groups, so as to greatly reduce false positive (Wang et al., 2011). For most of pairs of groups, at least one site had the posterior probability higher than 0.8, and four pairs of groups had one site with posterior probability higher than 0.95 (**Figure 6**). A cut-off value $Q(k) \geq 0.95$ was used for estimated key amino acid residues. The number of amino acid residues that presumably contributed to alter functional constraints was small between TPS groups. Three amino acid residues (669, 704, and 753) in all comparisons were identified as being most important for the functional divergence. These amino acids are all localized in the TPP domain.

To examine if selective pressure between the seven groups of *TPSs* genes would change, the likelihood ratio tests of positive selection were used via the ML methods as well as codon substitution models (Yang, 1998; Yang and Bielawski, 2000; Yang et al., 2012). Two tests were performed during the analysis process. First, models M0 and M3 were compared in order to assess whether d_N/d_S ratio variations existed among codon positions below each group. On the whole, the

maximum likelihood estimates of d_N/d_S values as for groups under model M0 were close to zero, with three exceptions of *Arabidopsis TPS* genes in group I-2. This implied that the principal force which acting on the evolution of TPS family was purifying selection (**Table 3**). It is obvious that although the majority of protein residues are heavily dependent on constant purifying selection, some sites are also being impacted by positive selection. Positive selection is an important adaption mechanism; thus, sites under positive selection pressure may have sped up functional divergence of TPS; in this case it would allow the plants to adjust to its environment. The results we get are in accordance with this conclusion. As the statistical differences of log likelihood between model M0 and M3 were significant for all groups, it suggests that the overall selective constraint levels varied across the TPS family group lineages. Secondly, the LRT which is applied to compare data fit to models M7 vs. M8 was used to determine whether positive selection promoted divergence of TPS family below groups or not. Two groups (group I-1 and II-2) of the 7 groups
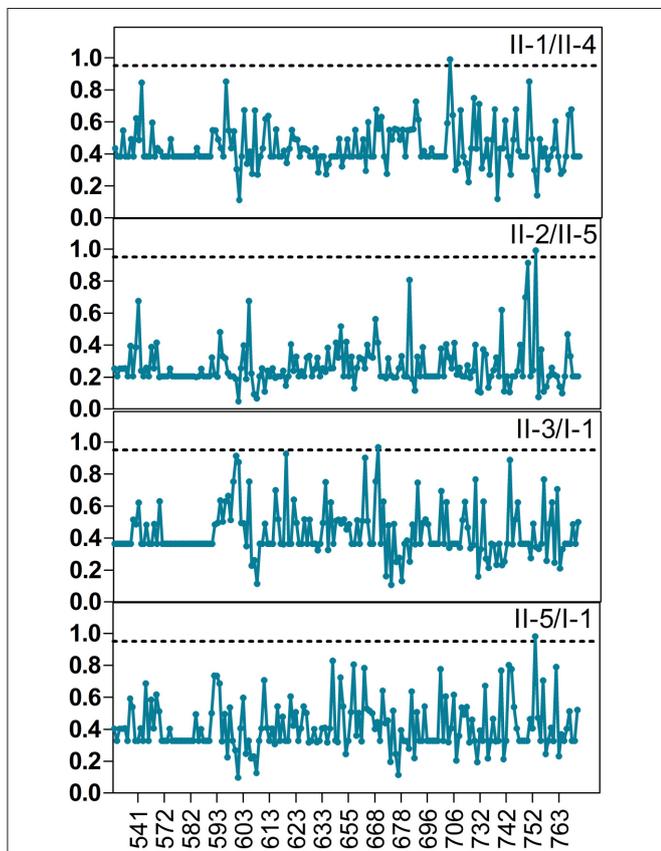
FIGURE 6 | Type I functional divergence among the plant *TPS* gene members. Posterior probability (PP) profiles of the site-specific type I functional divergence. The line indicates cutoff = 0.95.
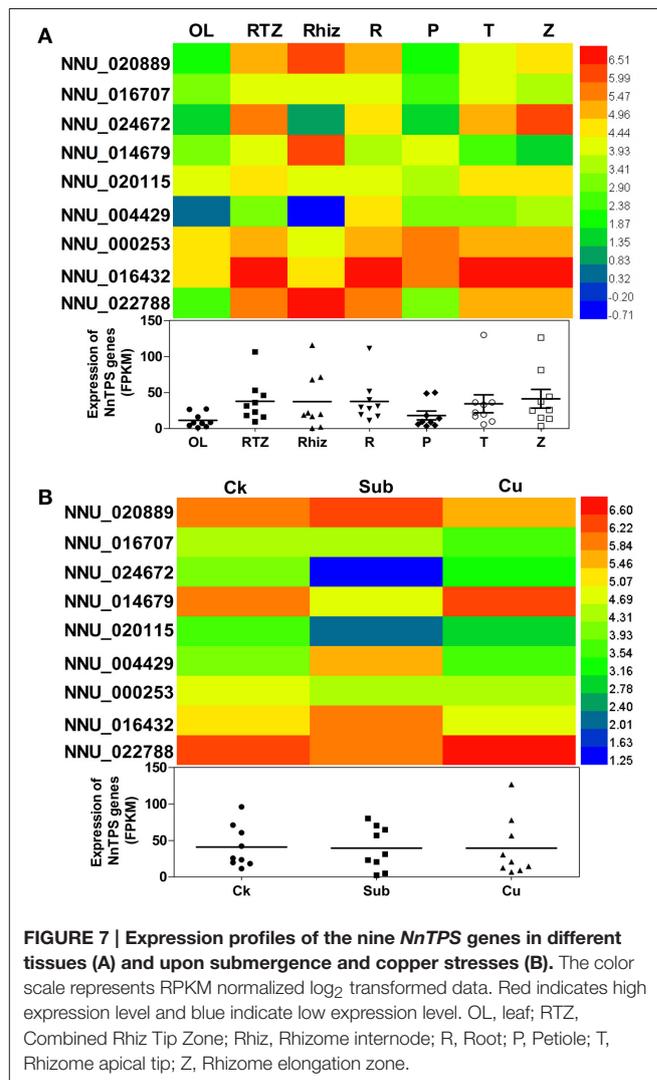
FIGURE 7 | Expression profiles of the nine *NnTPS* genes in different tissues (A) and upon submergence and copper stresses (B). The color scale represents RPKM normalized $\log_2$ transformed data. Red indicates high expression level and blue indicate low expression level. OL, leaf; RTZ, Combined Rhiz Tip Zone; Rhiz, Rhizome internode; R, Root; P, Petiole; T, Rhizome apical tip; Z, Rhizome elongation zone.

analyzed, were believed to have undergone positive selection, because they satisfied that (1) an estimate of $\omega > 1$ under M8, (2) sites identified to be under positive selection, and (3) a significant LRT. This result suggests that positive selection contributed to the evolution of genes in these groups. When the LRT suggested positive selective action had occurred, positively selected sites are identified under model M8 using Bayesian method (Nielsen and Yang, 1998; Yang and Bielawski, 2000). We found 31 and 4 positively selected sites in subgroup I-1 and II-2, respectively.

Gene Expression Analysis

TPS genes are known to be involved in development and stress responses. In this study, expression pattern of *TPS* genes in different lotus tissues has been studied using transcriptome data set of lotus (available at NCBI; Kim et al., 2013) as well as our in house transcriptome dataset of lotus under submergence and copper stress. *NnTPS* genes showed differential expression level in various lotus tissues (**Figure 7A**). No tissue-specific expressed *NnTPS* genes were found. Among them, NNU_016432 showed the highest expression level while
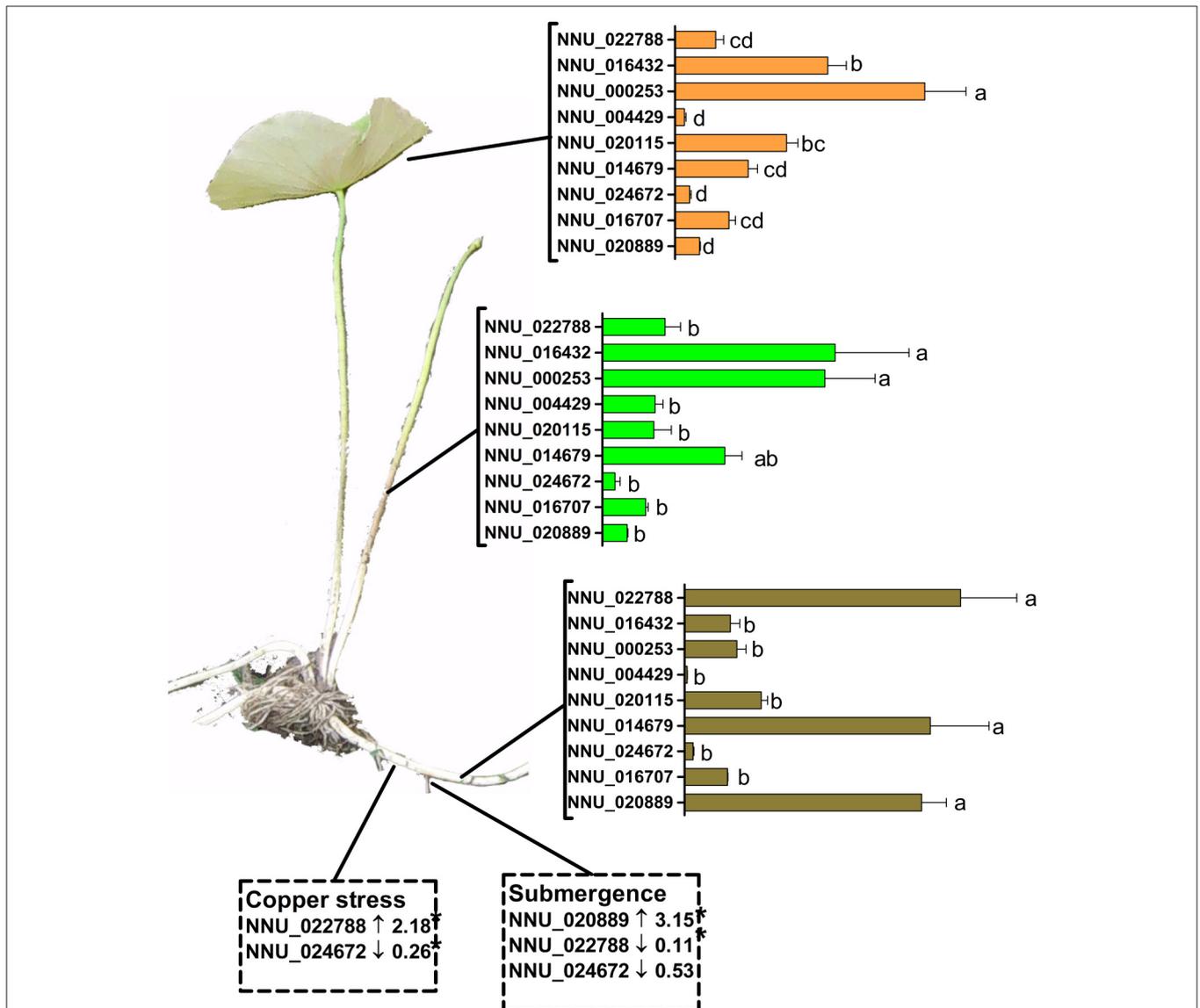
FIGURE 8 | Validation of expression of selected *NnTPS* genes in various tissues and in rhizome upon exogenous stimuli. Tissues including leaf, petiole, and rhizome were sampled from 2 month old lotus seedlings. Meanwhile, lotus seedlings were treated with submergence stress or copper stress for 24 h and then rhizome was sampled. The sample without 5 mm copper treatment or submergence treatment was the control (Con). Gene expression was analyzed by real-time RT-PCR. The relative transcript abundance was normalized using lotus actin gene. Values are means ± SE of at least four independent experiments. Bars with different letters are significantly different at $P < 0.05$ according to Turkey's multiple range test. Fold changes of genes under copper stress and submergence treatment were values relative to control samples. Asterisks indicate significant differences compared with control sample at $P < 0.05$ levels according to *t*-test.

NNU_004429 showed the lowest transcript level in most examined tissues, in comparisons with other *NnTPS* genes. Genes from the same group frequently showed similar expression pattern within specific tissues. For example, both NNU_000253 and NNU_016432 had relatively low expression in leaf and rhizome internode, and high expression in root, rhizome apical tip, and rhizome elongation zone. Based on FPKM calculations, the total transcript abundance of *NnTPS* genes were obviously low in leaf and petiole and high in root and rhizome. These results indicated that *NnTPS* genes might primarily have functions in sink tissues.

In response to submergence and copper stress, the total transcript abundance of NnTPS did not show apparent difference (**Figure 7B**). Two *NnTPS* genes (NNU_0044229 and NNU_016432) were highly induced by submergence treatment. In contrast, most *NnTPS* genes (i.e., NNU_024672, NNU_014679, and NNU_022788) were down-regulated. The abundance of NNU_024672 decreased over 80%. Submergence leads to lower concentration of oxygen in a plant as a result of a lower diffusion rate of molecular oxygen in water as opposed to air, which would ultimately influence energy metabolism. The reaction of NNU_024672 to low oxygen demonstrates possible involvement in energy metabolism in the seedlings and that highly expressed *NnTPS* genes could play vital roles during this complicated biological process. Some plant metabolic processes contain copper as an essential micronutrient, functioning as a constituent of enzymes and other proteins, and in photosynthetic reactions (Maksymiec, 1998; Hall, 2002; Yruela, 2009; Martins et al., 2014). Excessive copper can lead to oxidative stress and be phytotoxic (Mithöfer et al., 2004). It can inhibit growth, influence plant metabolism and affect the function of several enzymes (Martins and Mourato, 2006; Mourato et al., 2009). Under copper stress, the transcript abundance of most *NnTPS* genes was not obviously changed, except two *NnTPS* genes (NNU_014679 and NNU_022788). It is interesting that abundance of two genes were apparently decreased in response to submergence while increased upon copper stress. We speculate that under copper stress more energy was needed to strengthen antioxidant system, while more energy was needed to be preserved under submergence as photosynthesis is inhibited. The contrasting expression pattern under submergence and copper stress indicated that NNU_014679 and NNU_022788 might play important roles in lotus energy metabolism and participate in stress response. Similar to previous work, TPS has been reported to play an important role in the regulation of carbohydrate metabolism and especially the perception of carbohydrate availability (Müller et al., 2001).

In silico analysis showed that *NnTPS* genes exhibited differential abundance in various tissues and some of them are obviously regulated by environmental stimuli. Some of these obviously changed genes were then validated by RT-qPCR. As shown in **Figure 8**, RT-qPCR results of genes under various tissues and stresses were similar in magnitude to those obtained by deep sequencing. NNU_020889, NNU_014679, and NNU_022788 were the most abundant genes in rhizome. It is interesting, that while under submergence, NNU_020889

was significantly induced while NNU_000788 was significantly down-regulated. Moreover, NNU_024672, which has the minimum transcript level were further down-regulated upon submergence.

CONCLUSION

Collectively, our results characterize the structure, evolutionary, and some functional features of the lotus *TPS* gene family. Based on systematic analysis of lotus genomes, we identified 9 *TPS* genes. Gene structure analysis showed that NNU_022788 might be the ancestral member, from which other genes were derived. Regions outside the TPS domain in nine *NnTPS* genes evolved faster than regions within the TPS domains. Based on phylogenetic tree constructed using *TPS* genes from lotus, *Arabidopsis*, polar, soybean, and rice, *TPS* genes could be classified into two main subfamilies (I-II). Lotus *TPS* genes showed species-specific expanded manner. Three amino acid residues (669, 704, and 753) in TPP domain were identified as the most important residues for the functional divergence of these *TPS* genes. Moreover, 2 groups (group I-1 and II-2) were believed to have undergone positive selection based on likelihood ratio test. Expression pattern analysis showed that *NnTPS* genes might mainly have function in sink tissues and two *NnTPS* genes (including NNU_014679 and NNU_022788) might play important roles in lotus energy metabolism and stress responses.

AUTHOR CONTRIBUTIONS

QJ and XH conceived the study; QJ, BW, and XL did data analysis and drafted the manuscript; NM, YW, HJ, and YX revised the manuscript. All authors read and approved the final manuscript.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016. 01445

Figure S1 | Amino acid sequence alignment of three lotus TPS proteins. Strictly conserved sequence is in white on black background; similar amino acids are in black on green background. Residues involved in the catalytic center are placed in boxes.

Table S1 | List of primers used in this study.

Table S2 | Detail information of identified motifs.

REFERENCES

Avonce, N., Mendoza-Vargas, A., Morett, E., and Iturriaga, G. (2006). Insights on the evolution of trehalose biosynthesis. *BMC Evol. Biol.* 6:109. doi: 10.1186/1471-2148-6-109

Bailey, T. L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2, 28–36.

Blázquez, M. A., Santos, E., Flores, C. L., Salinas, J., and Gancedo, C. (1998). Isolation and molecular characterization of the *Arabidopsis TPS1* gene, encoding trehalose-6-phosphate synthase. *Plant J.* 13, 685–689. doi: 10.1046/j.1365-313X.1998.00063.x

Cao, J., Han, X., Zhang, T. C., Yang, Y. P., Huang, J. L., and Hu, X. Y. (2014). Genome-wide and molecular evolution analysis of the subtilase gene family in *Vitis vinifera. BMC Genomics* 15:1116. doi: 10.1186/1471-2164-15-1116

Chary, S. N., Hicks, G. R., Choi, Y. G., Carter, D., and Raikhel, N. V. (2008). Trehalose-6-phosphate synthase/phosphatase regulates cell shape and plant architecture in Arabidopsis. *Plant Physiol.* 146, 97–107. doi: 10.1104/pp.107.107441

Delorge, I., Figueroa, C. M., Feil, R., Lunn, J. E., and Van Dijck, P. (2015). Trehalose-6-phosphate synthase 1 is not the only active TPS in *Arabidopsis thaliana. Biochem. J.* 466, 283–290. doi: 10.1042/BJ20141322

Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics* 14, 755–763. doi: 10.1093/bioinformatics/14.9.755

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Elbein, A. D., Pan, Y. T., Pastuszak, I., and Carroll, D. (2003). New insights on trehalose: a multifunctional molecule. *Glycobiology* 13, 17R–27R. doi: 10.1093/glycob/cwg047

Force, A., Lynch, M., Pickett, E. B., Amores, A., Yan, Y. L., Postlethwait, J., et al. (1999). Preservation of duplicate gene by complementary, degenerative mutations. *Genetics* 151, 1531–1545.

Garg, A. K., Kim, J. K., Owens, T. G., Ranwala, A. P., Choi, Y. D., Kochian, L. V., et al. (2002). Trehalose accumulation in rice plants confers high tolerance levels to different abiotic stresses. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15898–15903. doi: 10.1073/pnas.252637799

Ge, L. F., Chao, D. Y., Shi, M., Zhu, M. Z., Gao, J. P., and Lin, H. X. (2008). Overexpression of the trehalose-6-phosphate phosphatase gene OsTPP1 confers stress tolerance in rice and results in the activation of stress responsive genes. *Planta* 228, 191–201. doi: 10.1007/s00425-008-0729-x

Gu, X. (2006). A simple statistical method for estimating type-II (cluster-specific) functional divergence of protein sequences. *Mol. Biol. Evol.* 23, 1937–1945. doi: 10.1093/molbev/msl056

Gu, X., and Velden, K. V. (2002). DIVERGE: phylogeny-based analysis for functional-structural divergence of a protein family. *Bioinformatics* 18, 500–501. doi: 10.1093/bioinformatics/18.3.500

Hall, J. L. (2002). Cellular mechanisms for heavy metal detoxification and tolerance. *J. Exp. Bot.* 53, 1–11. doi: 10.1093/jexbot/53.366.1

Hao, Y. W., Wang, X. Y., Li, X., Bassa, C., Mila, I., Audran, C., et al. (2014). Genome-wide identification, phylogenetic analysis, expression profiling, and proteinprotein interaction properties of TOPLESS gene family members in tomato. *J. Exp. Bot.* 65, 1013–1023. doi: 10.1093/jxb/ert440

Holmström, K., Mäntylä, E., Welin, B., Mandal, A., Palva, E. T., Tunnela, O. E., et al. (1996). Drought tolerance in tobacco. *Nature* 379, 683–684. doi: 10.1038/379683a0

Hu, W., Yang, H. B., Yan, Y., Wei, Y. X., Tie, W. W., Ding, Z. H., et al. (2016). Genome-wide characterization and analysis of bZIP transcription factor gene family related to abiotic stress in cassava. *Sci. Rep.* 6:22783. doi: 10.1038/srep22783

Hughes, A. L. (1994). The evolution of functionally novel proteins after gene duplication. *Proc. Biol. Sci.* 256, 119–124. doi: 10.1098/rspb.1994.0058

Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De, C. E., Langendijk-Genevaux, P. S., et al. (2006). The PROSITE database. *Nucleic Acids Res.* 34, D227–D230. doi: 10.1093/nar/gkj063

Jang, I. C., Oh, S. J., Seo, J. S., Choi, W. B., Sang, I. S., Kim, C. H., et al. (2003). Expression of a bifunctional fusion of the Escherichia coli genes for trehalose-6-phosphate synthase and trehalose-6-phosphate phosphatase in transgenic rice

plants increases trehalose accumulation and abiotic stress tolerance without stunting growth. *Plant Physiol.* 131, 516–524. doi: 10.1104/pp.007237

Jiang, W., Fu, F. L., Zhang, S. Z., Wu, L., and Li, W. C. (2010). Cloning and characterization of functional trehalose-6-phosphate synthase gene in maize. *J. Plant Biol.* 53, 134–141. doi: 10.1007/s12374-010-9098-7

Karim, S., Aronsson, H., Ericson, H., Pirhonen, M., Leyman, B., Welin, B., et al. (2007). Improved drought tolerance without undesired side effects in transgenic plants producing trehalose. *Plant Mol. Biol.* 64, 371–386. doi: 10.1007/s11103-007-9159-6

Kim, M.-J., Nelson, W., Soderlund, C. A., and Gang, D. R. (2013). Next-generation sequencing-based transcriptional profiling of sacred lotus "China Antique." *Trop. Plant Biol.* 6, 161–179. doi: 10.1007/s12042-013-9130-4

Li, H. W., Zang, B. S., Deng, X. W., and Wang, X. P. (2011). Overexpression of the trehalose-6-phosphate synthase gene *ostps1*, enhances abiotic stress tolerance in rice. *Planta* 234, 1007–1018. doi: 10.1007/s00425-011-1458-0

Lunn, J. E. (2007). Gene families and evolution of trehalose metabolism in plants. *Funct. Plant Biol.* 34, 550–563. doi: 10.1071/FP06315

Lyu, J. I., Min, S. R., Lee, J. H., Lim, Y. H., Kim, J. K., Bae, C. H., et al. (2013). Overexpression of a trehalose-6-phosphate synthase/phosphatase fusion gene enhances tolerance and photosynthesis during drought and salt stress without growth aberrations in tomato. *Plant Cell Tiss Org.* 112, 257–262. doi: 10.1007/s11240-012-0225-7

Maksymiec, W. (1998). Effect of copper on cellular processes in higher plants. *Photosynthetica* 34, 321–342. doi: 10.1023/A:1006818815528

Martins, L. L., and Mourato, M. P. (2006). Effect of excess copper on tomato plants: growth parameters, enzyme activities, chlorophyll, and mineral content. *J. Plant Nutr.* 29, 2179–2198. doi: 10.1080/01904160600972845

Martins, L. L., Mourato, M. P., Baptista, S., Reis, R., Carvalheiro, F., Almeida, A. M., et al. (2014). Response to oxidative stress induced by cadmium and copper in tobacco plants (*Nicotiana tabacum*) engineered with the trehalose-6-phosphate synthase gene (AtTPS1). *Acta Physiol. Plant.* 36, 755–765. doi: 10.1007/s11738-013-1453-0

Ming, R., Vanburen, R., Liu, Y., Yang, M., Han, Y., Li, L. T., et al. (2013). Genome of the long-living sacred lotus (*Nelumbo nucifera*, Gaertn.). *Genome Biol.* 14:R41. doi: 10.1186/gb-2013-14-5-r41

Mithöfer, A., Schulze, B., and Boland, W. (2004). Biotic and heavy metal stress response in plants: evidence for common signals. *FEBS Lett.* 566, 1–5. doi: 10.1016/j.febslet.2004.04.011

Mondragon-Palomino, M., and Gaut, B. S. (2005). Gene conversion and the evolution of three leucine-rich repeat gene families in *Arabidopsis thaliana. Mol. Biol. Evol.* 22, 2444–2456. doi: 10.1093/molbev/msi241

Mourato, M. P., Martins, L. L., and Campos-Andrada, M. P. (2009). Physiological responses of *Lupinus luteus* to different copper concentrations. *Biol. Plant.* 53, 105–111. doi: 10.1007/s10535-009-0014-2

Müller, J., Aeschbacher, R. A., Wingler, A., Boller, T., and Wiemken, A. (2001). Trehalose and trehalase in *Arabidopsis. Plant Physiol.* 125, 1086–1093. doi: 10.1104/pp.125.2.1086

Mu, M., Lu, X.-K., Wang, J.-J., Wang, D.-L., Yin, Z.-J., Wang, S., et al. (2016). Genome-wide identification and analysis of the stress-resistance function of the TPS (Trehalose-6-Phosphate Synthase) gene family in cotton. *BMC Genet.* 17:54. doi: 10.1186/s12863-016-0360-y

Nielsen, R., and Yang, Z. (1998). Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 148, 929–936.

Nuruzzaman, M., Manimekalai, R., Sharoni, A. M., Satoh, K., Kondoh, H., Ooka, H., et al. (2010). Genome-wide analysis of NAC transcription factor family in rice. *Gene* 465, 30–44. doi: 10.1016/j.gene.2010.06.008

Paul, M. J., Primavesi, L. F., Jhurreea, D., and Zhang, Y. (2008). Trehalose metabolism and signaling. *Ann. Rev. Plant Biol.* 59, 417–441. doi: 10.1146/annurev.arplant.59.032607.092945

Ramon, M., De Smet, I., Vandesteene, L., Naudts, M., Laymen, B., Dijck, P. V., et al. (2009). Extensive expression regulation and lack of heterologous enzymatic activity of the Class II trehalose metabolism proteins from *Arabidopsis thaliana. Plant Cell Environ.* 32, 1015–1032. doi: 10.1111/j.1365-3040.2009.01985.x

Ross, C. A., Liu, Y., and Shen, Q. X. J. (2007). The WRKY gene family in rice (*Oryza sativa*). *J. Integr. Plant Biol.* 49, 827–842. doi: 10.1111/j.1744-7909.2007.00504.x

Satoh-Nagasawa, N. S., Nagasawa, N., Malcomber, S., Sakai, H., and Jackson, D. (2006). A trehalose metabolic enzyme controls inflorescence architecture in maize. *Nature* 441, 227–230. doi: 10.1038/nature04725

Schluepmann, H., Van, D. A., Aghdasi, M., Wobbes, B., Paul, M., and Smeekens, S. (2004). Trehalose mediated growth inhibition of Arabidopsis seedlings is due to trehalose-6-phosphate accumulation. *Plant Physiol.* 135, 879–890. doi: 10.1104/pp.104.039503

Shima, S., Matsui, H., Tahara, S., and Imai, R. (2007). Biochemical characterization of rice trehalose-6-phosphate phosphatases supports distinctive functions of these plant enzymes. *FEBS J.* 274, 1192–1201. doi: 10.1111/j.1742-4658.2007.05658.x

Song, J., Gao, Z. H., Huo, X. M., Sun, H. L., Xu, Y. S., Shi, T., et al. (2015). Genome-wide identification of the auxin response factor (ARF) gene family and expression analysis of its role associated with pistil development in Japanese apricot (*Prunus mume* Sieb. et Zucc). *Acta Physiol Plant.* 37:145. doi: 10.1007/s11738-015-1882-z

Sonnhammer, E., Eddy, S., Birney, E., Bateman, A., and Durbin, R. (1998). Pfam: multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Res.* 26, 322–325. doi: 10.1093/nar/26.1.320

Suzuki, N., Bajad, S., Shuman, J., Shulaev, V., and Mittler, R. (2008). The transcriptional co-activator MBF1c is a key regulator of thermotolerance in *Arabidopsis thaliana*. *J. Biol. Chem.* 283, 9269–9275. doi: 10.1074/jbc.M709187200

Tuskan, G. A. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & amp; amp; Gray). *Science* 313, 1596–1604. doi: 10.1126/science.1128691

Vandesteene, L., Ramon, M., Le Roy, K., Van Dijck, P., and Rolland, F. (2010). A single active trehalose-6-P synthase (TPS) and a family of putative regulatory TPS-like proteins in Arabidopsis. *Mol. Plant 3,* 406–419. doi: 10.1093/mp/ssp114

van Dijken, A. J., Schluepmann, H., and Smeekens, S. C. (2004). Arabidopsis trehalose-6-phosphate synthase 1 is essential for normal vegetative growth and transition to flowering. *Plant Physiol.* 135, 969–977. doi: 10.1104/pp.104.039743

Vogel, G., Fiehn, O., Jean-Richard-dit-Bressel, L., Boller, T., Wiemken, A., Aeschbacher, R. A., et al. (2001). Trehalose metabolism in Arabidopsis: occurrence of trehalose and molecular cloning and characterization of trehalose-6-phosphate synthase homologues. *J. Exp. Bot.* 52, 1817–1826. doi: 10.1093/jexbot/52.362.1817

Wang, Q., Wang, M., Zhang, X., Hao, B., Kaushik, S. K., and Pan, Y. (2011). Wrky gene family evolution in *Arabidopsis thaliana*. *Genetica* 139, 973–983. doi: 10.1007/s10709-011-9599-4

Wang, X., Shi, X., Hao, B., Ge, S., and Luo, J. (2005). Duplication and DNA segmental loss in the rice genome: implications for diploidization. *New Phytol.* 165, 937–946. doi: 10.1111/j.1469-8137.2004.01293.x

Xun, G. (2003). Functional divergence in protein (family) sequence evolution. *Genetica* 118, 133–141. doi: 10.1007/978-94-010-0229-5_4

Yang, H. L., Liu, Y. J., Wang, C. L., and Zeng, Q. Y. (2012). Molecular evolution of trehalose-6-phosphate synthase (TPS) gene family in Populus, Arabidopsis and rice. *PLoS ONE* 7:e42438. doi: 10.1371/journal.pone.0042438

Yang, Z. (1998). Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol. Biol. Evol.* 15, 568–573. doi: 10.1093/oxfordjournals.molbev.a025957

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088

Yang, Z., and Bielawski, J. P. (2000). Statistical methods for detecting molecular adaptation. *Trends Ecol Evol.* 15, 496–503. doi: 10.1016/S0169-5347(00)01994-7

Yang, Z., Gao, Q., Sun, C., Li, W., Gu, S., and Xu, C. (2009). Molecular evolution and functional divergence of HAK potassium transporter gene family in rice (*Oryza sativa* L.). *J. Genet. Genomics* 36, 161–172. doi: 10.1016/S1673-8527(08)60103-4

Yang, Z., Wong, W. S., and Nielsen, R. (2005). Bayes empirical bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* 22, 1107–1118. doi: 10.1093/molbev/msi097

Yruela, I. (2009). Copper in plants: acquisition, transport and interactions. *Funct. Plant Biol.* 36, 409–430. doi: 10.1071/FP08288

Zang, B., Li, H., Li, W., Deng, X. W., and Wang, X. (2011). Analysis of trehalose-6-phosphatesynthase (TPS) gene family suggests the formation of TPS complexes in rice. *Plant Mol. Biol.* 76, 507–522. doi: 10.1007/s11103-011-9781-1