CROP BREEDING AND APPLIED BIOTECHNOLOGY

ARTICLE

Role of sweet potato *GST* genes in abiotic stress tolerance revealed by genomic and transcriptomic analyses

Crop Breeding and Applied Biotechnology
22(1): e36852212, 2022
Brazilian Society of Plant Breeding.
Printed in Brazil
http://dx.doi.org/10.1590/1984-70332022v22n1a02

Deka Reine Judesse Soviguidi[1,2], Yi Liu[1,2], Rui Pan[1], Salah Fatouh Abou-Elwafa[3], Li-Ping Rao[1,2], Sefasi Abel[4], Wen-Ying Zhang[1*] and Xin-Sun Yang[2]

Abstract: *Glutathione S-transferases (GSTs) are proteins synthesized in plants and responsible for their tolerance to environmental stresses. However, little information is available on the GST gene family of sweet potato, a globally important crop. The genetic evolution of GSTs in sweet potato remains unclear. The present study investigated the GST gene family in sweet potato by transcriptomic and comparative genomic analyses. A total of 51 GSTs were identified. Gene expression analysis showed differential expression patterns of the GSTs between two investigated varieties. Some GST expression levels were either up- or downregulated under oxidative, salinity and drought stresses. The results of the investigation provided new insights on the GST gene family in sweet potato, which may further the understanding of the roles of these genes in regulating abiotic stresses.*

Keywords: *Gene expression, drought, salinity, oxidative stress, sweet potato*

***Corresponding author:**
E-mail: wyzhang@yangtzeu.edu.cn
ORCID: 0000-0002-9926-4115

Received: 05 February 2021
Accepted: 14 January 2022
Published: 20 February 2022

[1] Ministry of Education/Hubei Collaborative Innovation Center for Grain Industry, Yangtze University, Engineering Research Centre of Ecology and Agricultural Use of Wetland, Jingzhou 43.4025, China
[2] Hubei Academy of Agricultural Sciences, Institute of Food Corps/Hubei Sweet potato Engineering and Technology Research Centre, Hubei Key Laboratory of Food Crops Germplasm and Genetic Improvement, Wuhan 43.0064, China
[3] Assiut University, College of Agriculture, Agronomy Department, Assiut 71.526, Egypt
[4] Lilongwe University of Agriculture and Natural Resources, Faculty of Agriculture, Horticulture Department, P.O. Box 219, Lilongwe, Malawi

INTRODUCTION

Glutathione S-transferases (*GSTs*) represent a group of proteins found in various plant species. They mainly function as cytotoxic compounds helping to reduce damages caused by environmental stresses (Estévez and Hernández 2020). Moreover, *GST* genes can act as carriers of secondary metabolites such as anthocyanins and flavonoids, transporting them to vacuoles for sequestration (Wei et al. 2019). Previously, the *GST* gene family has been investigated in the genome of different plants and 85, 49, and 52 copies, respectively, were observed in *Capsicum annuum* (Islam et al. 2019), melon (Wang et al. 2020) and apple (Fang et al. 2020). Glutathione S-transferases could be found in three different subcellular localizations including microsomes, mitochondria and the cytoplasm in which they are most prominent (Hu et al. 2018). The *GST* gene family can be grouped into eight subfamilies, namely: eukaryotic translation elongation factor 1 gamma (EF1Bγ), Tau (U), Zeta (Z), dehydroascorbate reductase (DHAR), Lambda (L), tetrachlorohydroquinone dehalogenase (TCHQD), Phi (F) and Theta (T) (Liu et al. 2019). Tau and Phi are the most important subfamilies, which are involved in the transport of various secondary metabolites and cell detoxification. The upregulation of these subfamilies reportedly increased plant tolerance to various stresses (He et al. 2016). For example, *TaGSTF62* improved

the tolerance of wheat to salinity stress (Wang et al. 2019), whereas *CmaGSTU14* was shown to enhance cold stress tolerance in pumpkin (Kayum et al. 2018). Other studies also demonstrated the involvement of *IbGSTF4* and *FaGST73* in anthocyanin accumulation in Arabidopsis (Kou et al. 2019) and strawberry (Lin et al. 2020), respectively.

Sweet potato (*Ipomoea batatas*), a globally important crop, is affected by various environmental constraints, particularly in a number of developing countries (Zhang et al. 2020). To date, the *GST* gene family of this crop has not been extensively investigated. However, a recent study based on transcriptomic analysis identified 42 *GST* genes in sweet potato, which were classified into eight different subfamilies and exhibited varying expression patterns across different varieties and tissues in response to oxidative and metal stresses (Ding et al. 2017). Nevertheless, the genetic evolution of *GST* genes in sweet potato and their functions under abiotic stress remain unclear. Therefore, a more in-depth analysis of the genome of this crop would be important to better understand the characteristics and transcript accumulation of sweet potato *GST* genes in stressful environments. In the current study, the *GST* gene family in the sweet potato genome was investigated using not only transcriptomic technology but also comparative genomic analyses. With a view to providing new insights on *GST* genes of sweet potato, a number of differentially expressed transcripts were identified and the expression level of 10 *GST* genes under salinity, drought and oxidative stress conditions determined by RT-qPCR.

MATERIAL AND METHODS

Identification of *GST* genes in sweet potato

The *Arabidopsis thaliana* GST protein data were retrieved from the Arabidopsis Information Resource (TAIR release 10 http://www.arabidopsis.org) and used as query to perform a BLASTp search against the *Ipomoea triloba* genome database NSP323.v3 (http://sweetpotato.uga.edu/), at a cut-off value of $\leq e^{-20}$. The NCBI-CDD and Pfam databases (El-Gebali et al. 2019) were further employed to examine the presence of conserved domains in the identified candidate proteins, and 55 proteins containing GST domains were obtained. In order to confirm the expression of these 55 *GSTs* in *I. batatas*, the genes were compared to those in the transcriptome database, which was constructed based on Illumina High-Seq 2500 sequencing technology (Illumina, Inc.; San Diego, CA, US). The transcriptomic analysis material consisted of leaves of two sweet potato varieties, namely Fushu 7-6 (FS7-6) and EC16 collected from 65- and 85-day-old plants (NCBI SRA, accession number PRJNA592001) (Liu et al. 2021). In the two varieties, 51 *GSTs* were found. The four not expressed mRNA were manually removed and the remaining 51 were finally confirmed as members of the *GST* gene family in *I. batatas*. The FPKM (fragments per kilobase of exon model per million reads mapped) values of these genes were used to generate a heatmap showing the average relative expression of these genes using R packages (https://bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html). The ExPASy Prot-Param tool (https://web.expasy.org/protparam/) was used to determine the molecular weight (Mw), isoelectric points (pI) and polypeptide length of the genes. The tool CELLO v2.5 (http://cello.life.nctu.edu.tw/) served for subcellular localization prediction.

Sequence analysis of sweet potato *GST* genes

Based on comparison with Arabidopsis *GST* proteins, the identified sweet potato *GST*s were subgrouped with ClustalX2 (Larkin et al. 2007). The Neighbor–Joining method of MEGA7 (Kumar et al. 2016) was used to generate the phylogenetic tree. The online program GSDS (http://gsds.cbi.pku.edu.cn) analyzed the structure of introns and exons of the *GST* genes, while MEME (http://memesuite.org/index.html) served as motif identification tool. The promoter regions in the *GST* genes were identified by the database PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html).

Plant material and abiotic stress treatments

The sweet potato varieties FS7-6 and EC16 were used in this study. To identify the *GST* genes involved in abiotic stress tolerance, the transcript abundance of these genes in the transcriptome database was analyzed. Among the known 51 *GSTs*, 10 gene transcription levels were relatively high in both varieties and developmental stages. The expression levels of these 10 *GST* genes were analyzed using real-time PCR under optimum and stressed conditions. Young sweet potato plants were grown in Hoagland hydroponic solution (1%) in a growth chamber at 26±2 ºC under 16 h light/ 8 h dark. Seven days after acclimatization, the plants were treated with 200 mM NaCl, 30% PEG6000 solution and 5% H_2O_2 for salinity, drought and oxidative stresses, respectively, while unstressed plants grown in 1% Hoagland solution served

as control. The test consisted of three replicates. Leaves were collected at 0, 6, 12 and 24 h after stress induction.

Quantitative real-time PCR analysis

Leaf samples of FS7-6 and EC16 were collected and ground in liquid nitrogen. A TransZol Up Kit was used for total RNA extraction and DNase treatment. A TransScript® All-in-One First-Strand cDNA Synthesis SuperMix for PCR served as cDNA synthesis kit and TransStart® Tip Green qPCR SuperMix was used for qRT-PCR analysis. All these kits were purchased from TransGen Biotech (China). The normalized values of the *GST* gene expression level were assessed using the internal control β-actin gene. The qRT-PCR thermal cycling profile consisted of: 94 °C for 30s, 40 cycles of 94 °C for 5s, 56 °C for 15s and 72 °C for 10s. The experiment was run in independent biological triplicates and the relative expression was calculated by the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen 2001). The primer sequences of the selected genes are listed in Additional file 1: Table S1.

RESULTS AND DISCUSSION

Characterization and sequence analysis of *GST* genes

In a BLASTp search against the *I. triloba* genome database, 55 putative genes containing conserved GST domains were obtained. These genes were compared with those of the transcriptome database to confirm their expression in *I. batatas*. Fifty-one *GSTs* were found in both varieties and finally confirmed as *GST* members in *I. batatas*. Among these proteins, *itb11g03260* was the largest with 389 amino acids (aa) and *itb12g06130* the smallest (110 aa). The isoelectric point (pI) of the *GSTs* varied widely, from 4.85 (*itb06g23080*) to 9.19 (*itb12g06130*) and the molecular weight (MW) from 12.91 to 43.69 kilodalton (kDa). Moreover, 44 *GSTs* were localized in the cytoplasm, one in the outer membrane and the remaining six in the periplasmic region (Table 1). Previously, a study identified 42 *IbGST* members and classified them into eight subfamilies in sweet potato (Ding et al. 2017). No similarity was observed between the genes identified in the two studies, except for *IbGSTU12* in the previous paper, named *itb01g20610* in this study. The sequences of the reported motifs were also different. These results could be explained by the fact that our study was based on the *I. triloba GST* genes present in the varieties FS7-6 and EC16, while Ding et al. (2017) characterized the *GSTs* specific to *I. batatas*. Phylogenetic analysis allocated the 51 *GST* members to five different subfamilies: Tau, Lambda, Theta, TCHQD and EF1Bγ, according to their similarity to the Arabidopsis *GST* proteins (Figure 1). However, no *GST* genes were assigned to the Phi and Zeta classes. A possible explanation is that each *GST* subfamily followed distinct evolutionary paths and the number of subfamilies varied greatly among plants. The Tau class was the largest, with 40 members. Three proteins were assigned to the Theta, four to the Lambda, two to the TCHQD and two to the EF1Bγ subfamily. The sequence information of sweet potato and Arabidopsis GST proteins is listed in Additional file 2: Table S2 and Additional file 2: Table S3, respectively. Usually, the subfamilies Phi and Tau are the most represented in many plant species, however, in our study Tau (40 *GST* members) and Lambda (4 *GST* members) were the most dominant, probably because of the absence of Phi subfamily members in the identified *GST* genes. This is consistent with Islam et al., who discovered that Tau (57 *GST* members) and Lambda (7 *GST* members) were the most prominent *GST* subfamilies in the tomato genome (Islam et al. 2017). Moreover, the subfamilies EF1Bγ and TCHQD were less represented in the analyzed sweet potato varieties, suggesting few duplication events of these genes during speciation or that duplicated copies were lost during evolution (Abou-Elwafa et al. 2011).

Sequence composition of sweet potato *GST* genes

The phylogenetic classification of the 51 identified *GST* genes matched the classification of motifs, domains and structures of these genes (Figure 2A). Online software MEME identified 10 different motifs for all *GST* genes. These motifs were almost identical within each subfamily, but different among subfamilies (Figure 2B). Motif 3 was the most widely represented in the *GSTs*. Moreover, motifs 7 and 8 are present in the Lambda, while motif 9 is unique in the Theta subfamily. Motif 4 belongs to Tau and TCHQD, while other motifs (1, 2, 5, 6 and 10) were present in the Tau subfamily. Sequence information for each motif is provided in Additional file 3: Table S4. The GST domains were highly conserved in each subfamily (Figure 2C). Within the subfamilies, they shared a similar structure, although the domains were different among subfamilies. For instance, most members of subfamily Tau comprise GST_C_Tau and GST_N_*Tau,* except *itb15g08160*, *itb12g06130* and *itb08g15850*, and most members of subfamily Theta comprise GST_C_Theta and

Table 1. Characteristics of the identified 51 GST genes of sweet potato

Order	Locus ID	CDS length (bp)	Protein length (aa)	MW (KDa)	pI	Predicted subcellular localization
1	itb01g15260	1131	376	41.18	8.78	Periplasmic
2	itb01g20560	663	220	25.69	5.51	Cytoplasmic
3	itb01g20590	669	222	25.49	5.01	Cytoplasmic
4	itb01g20600	648	215	24.98	5.12	Cytoplasmic
5	itb01g20610	669	222	25.81	5.59	Cytoplasmic
6	itb01g20620	684	227	26.05	5.33	Cytoplasmic
7	itb01g20640	666	221	25.36	6.71	Cytoplasmic
8	itb01g20650	669	222	25.64	6.18	Cytoplasmic
9	itb01g35330	981	326	36.51	8.7	Periplasmic
10	itb02g09490	663	220	25.48	5.33	Cytoplasmic
11	itb02g16280	702	233	26.87	5.6	Cytoplasmic
12	itb02g20780	675	224	25.91	5.89	Cytoplasmic
13	itb04g03890	666	221	25.68	5.4	Cytoplasmic
14	itb06g23080	669	222	25.36	4.85	Cytoplasmic
15	itb06g25110	744	247	28.37	5.26	Cytoplasmic
16	itb06g25120	714	237	26.99	5.42	Cytoplasmic
17	itb07g03710	1089	362	40.26	8.73	Outer membrane
18	itb08g03140	966	321	36.17	8.75	Periplasmic
19	itb08g06500	396	131	14.89	5.91	Cytoplasmic
20	itb08g15840	693	230	25.37	5.48	Cytoplasmic
21	itb08g15850	681	226	25.32	5.42	Cytoplasmic
22	itb09g30700	723	240	27.49	8.99	Cytoplasmic
23	itb10g18320	669	222	25.67	5.19	Cytoplasmic
24	itb11g02640	705	234	26.85	5.15	Cytoplasmic
25	itb11g02650	705	234	26.99	5.55	Cytoplasmic
26	itb11g03220	660	219	25.33	5.83	Cytoplasmic
27	itb11g03240	669	222	25.87	6.16	Cytoplasmic
28	itb11g03250	669	222	25.86	5.55	Cytoplasmic
29	itb11g03260	1170	389	43.69	8.83	Cytoplasmic
30	itb11g07110	711	236	26.93	5.29	Periplasmic
31	itb11g07120	804	267	30.3	6.84	Periplasmic
32	itb12g01870	675	224	25.29	5.3	Cytoplasmic
33	itb12g06130	333	110	12.91	9.19	Cytoplasmic
34	itb12g10540	672	223	25.25	6.4	Cytoplasmic
35	itb13g02460	657	218	25.1	6.01	Cytoplasmic
36	itb13g03920	723	240	27.19	8.87	Cytoplasmic
37	itb13g03940	720	239	27.03	5.77	Cytoplasmic
38	itb13g24640	684	227	25.94	5.81	Cytoplasmic
39	itb14g03990	678	225	24.78	5.97	Periplasmic
40	itb14g04060	687	228	25.61	5.34	Cytoplasmic
41	itb14g04110	684	227	25.23	5	Cytoplasmic
42	itb15g08110	672	223	25.19	5.56	Cytoplasmic
43	itb15g08120	660	219	25.16	6.55	Cytoplasmic
44	itb15g08130	672	223	25.25	5.92	Cytoplasmic
45	itb15g08150	678	225	25.3	5.91	Cytoplasmic
46	itb15g08160	573	190	21.51	6.54	Cytoplasmic
47	itb15g08170	672	223	24.92	6.92	Cytoplasmic
48	itb15g08180	675	224	25.25	6.02	Cytoplasmic
49	itb15g08190	672	223	25.31	5.54	Cytoplasmic
50	itb15g08200	672	223	25.18	5.65	Cytoplasmic
51	itb15g08210	759	252	28.38	5.67	Cytoplasmic

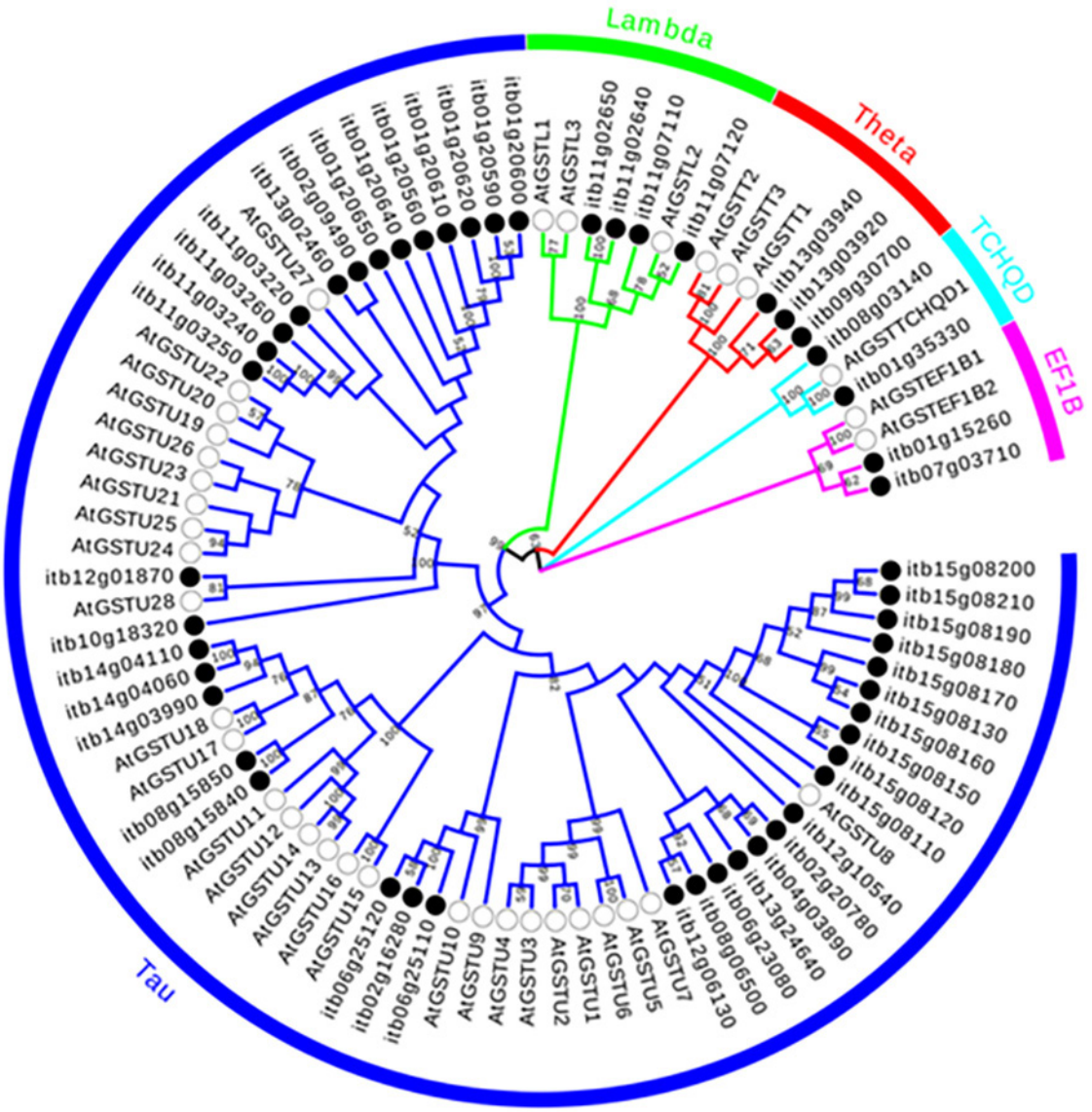

Figure 1. Phylogenetic tree representing relationships among sweet potato and Arabidopsis *GSTs*. The colors indicate distinct subfamilies of the *GST* gene family. GST proteins in Arabidopsis with the prefix 'At' indicate 'AtGST'. MEGA 7, Neighbor-Joining method and 1000 replicates for the bootstrap test were used to generate the tree. Percentage bootstrap scores of >50% were displayed.

GST_N_Theta, except *itb13g03940*. In general, the analysis of conserved domains showed 37 *GST* genes with a highly conserved N-terminal domain. It has been reported that the residue in N-terminal domain regulates the catalytic function of *GST* genes (Nishida et al. 1998). The gene structure analysis revealed great variation among the genes (Figure 2D). Most *GST* genes in subfamily Tau contained two exons. The Theta and TCHQD *GSTs* have seven and six exons, respectively, while the Lambda and *EF1Bγ* contain approximately 10. Moreover, variations in intron numbers among subfamilies were observed, whereas their positions within subfamilies were conserved.

Expression profiling of sweet potato *GST* genes

To explore the role of the *GST* genes, the expression of all 51 members was analyzed based on RNA-sequencing data derived from the varieties EC16 and FS7-6, sampled at two developmental stages (from 65- and 85-day-old plants). The

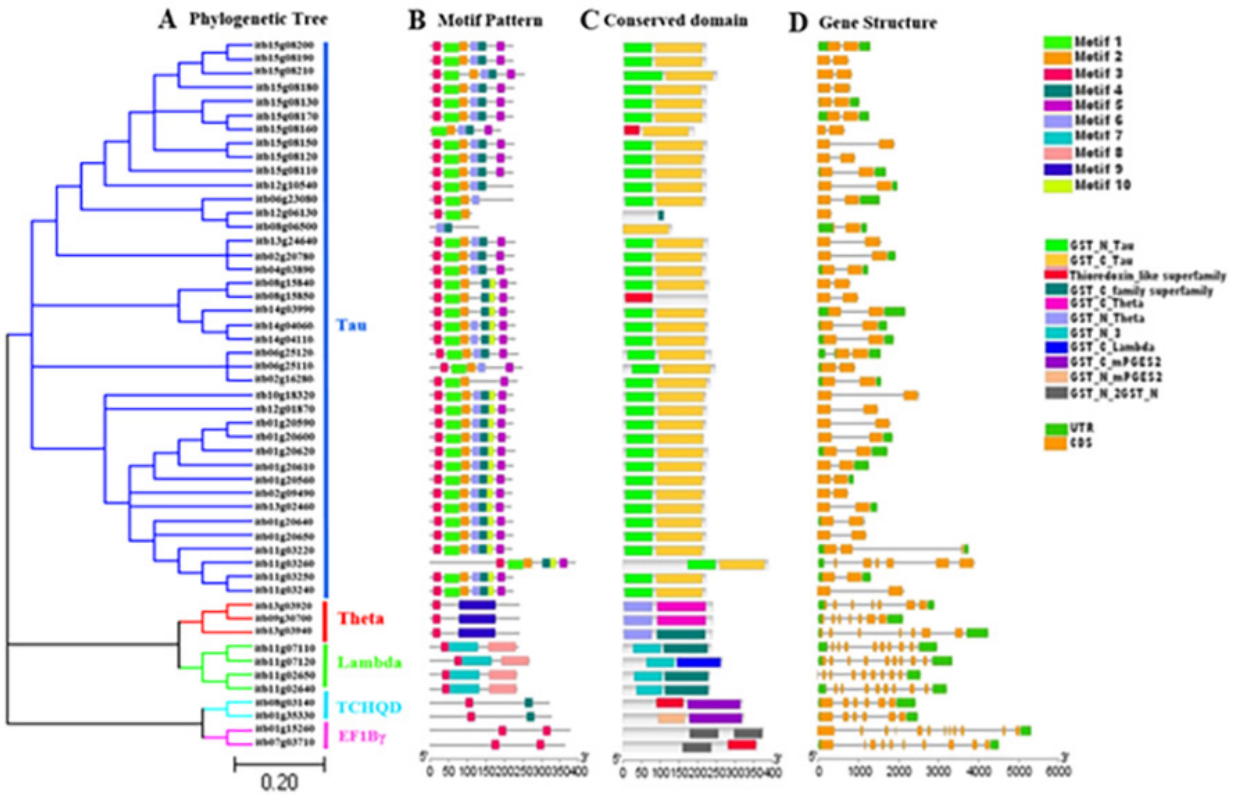

Figure 2. Sequence analysis of sweet potato *GST* genes. (**A**) Phylogenetic tree generated using MEGA 7 software, (**B**) motif pattern of *GST* genes, (**C**) conserved domain, **(D)** *GST* gene structure. The GST proteins in the four plots (**A, B, C, D**) are in the same order.

51 *GSTs* had different expression patterns for the varieties and developmental stages (Figure 3A). Among them, 29 *GST* genes exhibited constitutive expression (FPKM>1), whereas 22 were expressed with a FPKM value >0 in all analyzed samples. Some *GSTs* were highly expressed in all samples, while the expression levels of others were extremely low. For example, transcript abundance was highest in *itb15g08190* and *itb14g03990*, and lowest in all analyzed samples of *itb01g20560* and *itb15g08120*. The expression levels of *itb15g08190* and *itb14g03990* genes were higher in variety EC16 than FS7-6. Moreover, the expression levels of *itb13g03940* and *itb11g03220* were higher at the sampling of the first developmental stage compared to the second. It can be speculated that the expression of the *GST* genes is influenced by the genetic constitution and developmental stages of each sweet potato variety.

Transcript accumulation of sweet potato *GST* genes under abiotic stress

To determine the functions of *GSTs* under abiotic stress, transcriptome profiling data of all *GST* genes from two other sweet potato varieties (Lizixiang and ND98) were retrieved from the NCBI SRA database (accession number SRP092215) (Zhang et al. 2017). These varieties were chosen because no *GST* transcript levels of EC16 and FS7-6 under abiotic stress conditions were available. Transcriptomic data of 0, 12 and 48 hours after salt stress treatment were extracted from the database. The expression levels of most *GSTs* were low in Lizixiang and ND98, except for *itb12g10540* and *itb01g20600*, which were upregulated at 12 and 48 h after salinity treatment in both varieties. Some other genes, such as *itb15g08200*, *itb15g08210*, *itb13g02460*, *itb14g03990* and *itb15g08150,* were also found to be downregulated, indicating the non-responsiveness to high salinity conditions of these *GSTs* in Lizixiang and ND98 (Figure 3B).

To further confirm the abiotic stress-responsiveness of *GST* genes in the varieties EC16 and FS7-6, the 10 most highly expressed genes in both varieties and developmental stages, i.e., *itb01g15260*, *itb09g30700*, *itb11g02640*, *itb11g03220*, *itb11g07110*, *itb11g07120*, *itb13g03940*, *itb14g03990*, *itb14g04110* and *itb15g08190,* were selected from the 51 *GST*

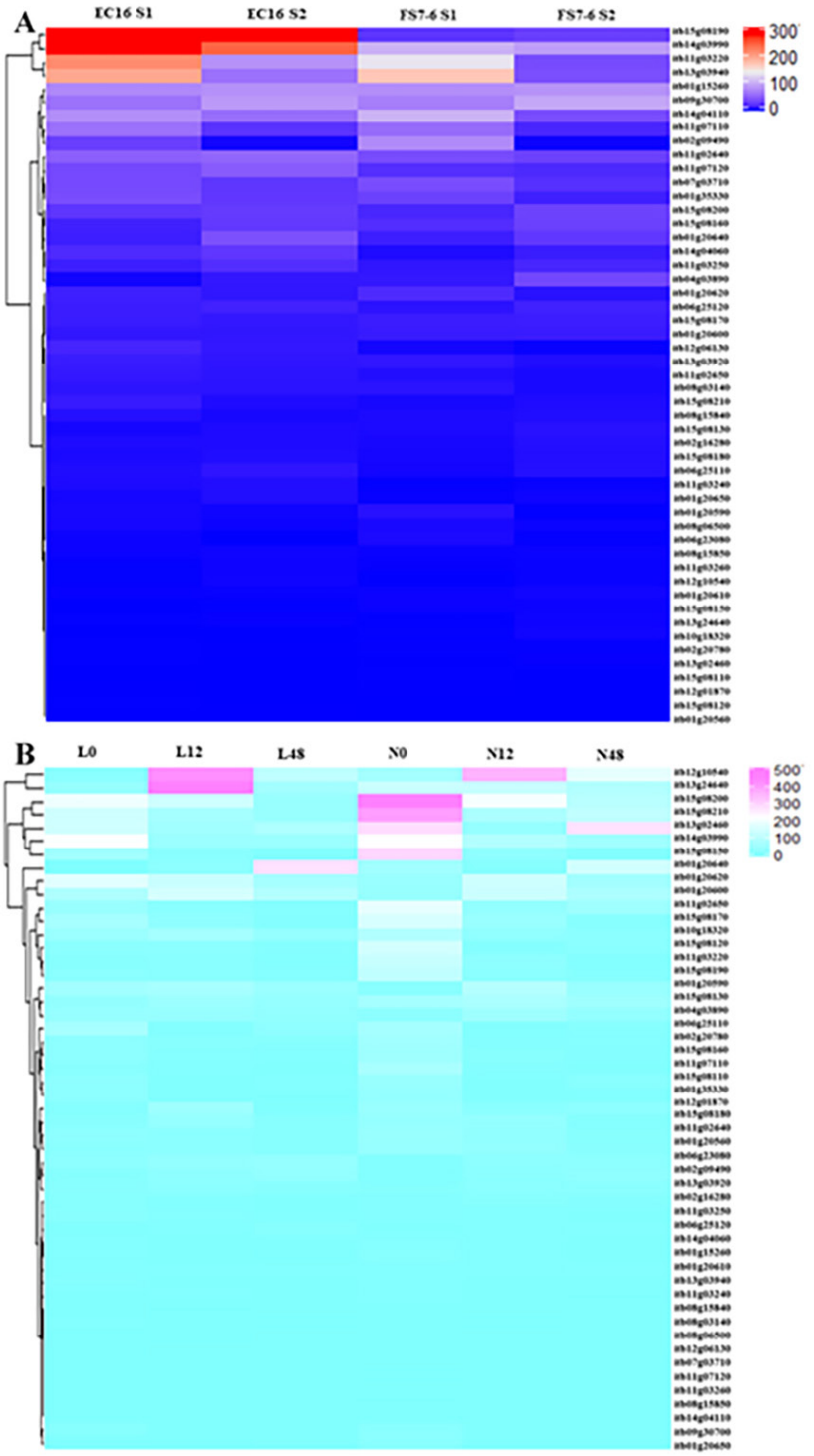

Figure 3. Transcript analysis of the 51 GSTs. Longitudinal direction indicates the varieties EC16 and FS7-6 and developmental stages S1 (stage 1: 65 days) and S2 (stage 2: 85 days). Horizontal direction indicates the 51 sweet potato GST genes. High expression levels are shown in red and low levels in blue **(A)**. GST gene expression levels of sweet potato varieties Lizixiang (L) and ND98 (N) in response to salinity stress. Treatment at 0, 12 and 48 h. L0; Lizixiang at 0 h, L12; Lizixiang at 12 h, L48; Lizixiang at 48 h, N0; ND98 at 0 h, N12; ND98 at 12 h and N48; ND98 at 48 h. The colors represent the gene transcript levels (FPKM value) **(B)**.

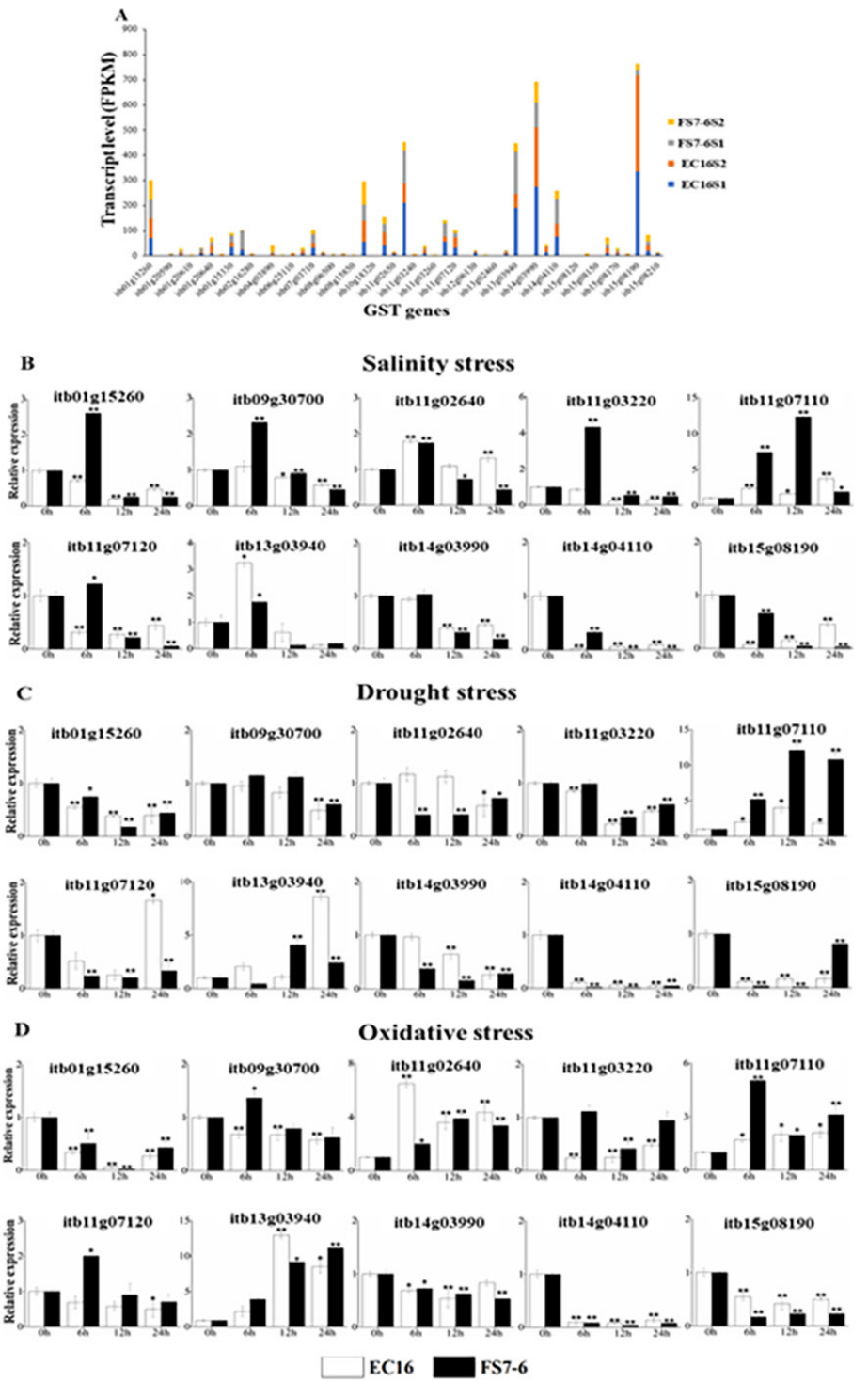

Figure 4. Transcript level of the *GSTs*. *itb01g15260*, *itb09g30700*, *itb11g02640*, *itb11g03220*, *itb11g07110*, *itb11g07120*, *itb13g03940*, *itb14g03990*, *itb14g04110* and *itb15g08190* had the highest expression levels in the varieties EC16 and FS7-6 and two developmental stages (65- and 85-day-old plants) **(A)**. Expression analysis of *GSTs* under salinity **(B)**, drought **(C)** and oxidative **(D)** stresses. Samples of each treatment were collected at 0, 6, 12 and 24 h after stress induction. Data were normalized to the β-actin gene. * and ** indicate significant differences compared to 0 h at $P < 0.05$ and $P < 0.01$ (t-test), respectively.

genes (Figure 4A). RT-qPCR experiments were then performed to identify the transcript accumulation of these genes in plants under drought (30% PEG6000), salinity (200 mM NaCl) and oxidative (5% H_2O_2) stresses. Samples in each treatment were collected at 0, 6, 12 and 24 h after stress induction. The *GST* genes were significantly up- and downregulated by the induced abiotic stresses. Under salinity, *itb11g07110* was significantly upregulated in both varieties throughout the stress period, compared to the control. In addition, *itb01g15260*, *itb09g30700*, *itb11g02640*, *itb11g03220*, *itb11g07120* and *itb13g03940* were upregulated only at 6h after treatment, especially in variety FS7-6. The other genes, namely *itb14g03990*, *itb14g04110* and *itb15g08190*, were mostly downregulated (Figure 4B). Under drought, apart from *itb11g07110* and *itb13g03940*, which were upregulated, the expression of *itb11g07120* was remarkable at 24h in variety EC16 (Figure 4C). Under oxidative stress, *itb11g07110*, *itb11g02640* and *itb13g03940* were the most responsive genes, whereas *itb09g30700* and *itb11g07120* were significantly upregulated only at 6h in FS7-6 (Figure 4D). Overall, *itb11g07110* responded positively to drought, salt and oxidative stresses, with a significantly higher expression level in FS7-6 compared to EC16. Moreover, the gene *itb11g02640* was induced under oxidative stress in both varieties, while *itb13g03940* was significantly expressed under drought and oxidative stress. In contrast, *itb14g03990*, *itb14g04110* and *itb15g08190* were downregulated in all treatments, indicating that these genes were insensitive to the tested stresses. The differences in stress-regulatory units present in the *GST* promoter could be the cause of variations observed in the expression pattern of these genes. Additionally, *itb15g08190* and *itb14g03990,* with the highest transcript abundance under unstressed conditions (Figure 3A), were significantly downregulated in all stress treatments. This confirms the non-responsiveness of the genes to abiotic stresses. Various studies have demonstrated the downregulation of *GST* genes under environmental stress (Islam et al. 2019, Wang et al. 2019). The *GSTs* of Lizixiang and ND98 were mostly downregulated under salt stress (Figure 3B), even the *itb11g07110* gene, which was highly expressed in EC16 and FS7-6 under the same stress condition. The ability of each variety to resist and thrive in a specific growing environment and certain developmental stages could influence the expression of genes responsible for defense against environmental stresses. In numerous studies, *GST* genes improved tolerance to abiotic factors. For example, *VvGSTF13* of grape increased Arabidopsis plant tolerance to drought and salinity stresses (Xu et al. 2018). Likewise, *GmGSTU63* enhanced drought tolerance in soybean (Hasan et al. 2020). These studies revealed the significant roles of GST proteins in modulating plant stress pathways. Further investigations may improve our understanding about roles and functions of *GST* genes in regulating the response of sweet potato to various abiotic stresses.

CONCLUSION

In this study, 51 *GST* genes were identified. The full-length genomic sequences of these 51 *GSTs* were characterized and allocated to five distinct subfamilies, based on their similarity with Arabidopsis *GSTs*. Different conserved motifs and domains were observed in the sweet potato *GST* genes. The up- and downregulation of the *GSTs* differed between the two analyzed varieties and abiotic stress treatments. Three genes were upregulated in a stress -specific manner; *itb11g07110* was upregulated in all treatments, *itb11g02640* was highly expressed under oxidative stress, while *itb13g03940* was induced in both drought and oxidative stress. This study provides new perspectives on the *GST* gene family in terms of its role in regulating abiotic stresses in sweet potato.

ACKNOWLEDGMENTS

This research project was supported by the National Key R&D Program of China (2018YFD1000700, 2018YFD1000705, 2019YFD1001300, 2019YFD1001305), Science and Technology Development Plan Project of Jingzhou City, Hubei Province, China (2018-37), Characteristic Discipline of Hubei Academy of Agricultural Sciences (2015TSXK06) and Science and Technology Innovation Center of Hubei Academy of Agricultural Sciences (2007-620-001-03). Additional Tables S1, S2, S3, S4 and Software and parameters used for RNA-seq data analysis can be required to the main author by e-mail (wyzhang@yangtzeu.edu.cn).

REFERENCES

Abou-Elwafa SF, Büttner B, Chia T, Schulze-Buxloh G, Hohmann U, Mutasa-Göttgens E, Jung C and Müller AE (2011) Conservation and divergence of autonomous pathway genes in the flowering regulatory network of *Beta vulgaris*. **Journal of Experimental Botany 62**: 3359-3374.

Ding N, Wang A, Zhang X, Wu Y, Wang R, Cui H, Huang R and Luo Y (2017) Identification and analysis of glutathione S-transferase gene family in sweet potato reveal divergent GST-mediated networks in aboveground and underground tissues in response to abiotic stresses. **BMC Plant Biology 17**: 225.

El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA and Smart A (2019) The Pfam protein families database in 2019. **Nucleic Acids Research 47**: D427-D432.

Estévez IH and Hernández MR (2020) Plant glutathione S-transferases: An overview. **Plant Gene 23**: 100233.

Fang X, An Y, Zheng J, Shangguan L and Wang L (2020) Genome-wide identification and comparative analysis of GST gene family in apple (*Malus domestica*) and their expressions under ALA treatment. **3 Biotech 10**: 1-16.

Hasan MS, Islam S, Hasan MN, Sajib SD, Ahmed S, Islam T and Ghosh A (2020) Genome-wide analysis and transcript profiling identify several abiotic and biotic stress-responsive Glutathione S-transferase genes in soybean. **Plant Gene 23**: 100239.

He G, Guan CN, Chen QX, Gou XJ, Liu W, Zeng QY and Lan T (2016) Genome-wide analysis of the Glutathione S-transferase gene family in *Capsella rubella*: Identification, expression, and biochemical functions. **Frontiers in Plant Science 7**: 1325.

Hu F, Ye K, Tu X-F, Lu Y-J, Thakur K, Jiang L and Wei Z-J (2018) Identification and expression profiles of twenty-six glutathione S-transferase genes from rice weevil, *Sitophilus oryzae* (Coleoptera: Curculionidae). **International journal of biological macromolecules 120**: 1063-1071.

Islam S, Rahman IA, Islam T and Ghosh A (2017) Genome-wide identification and expression analysis of glutathione S-transferase gene family in tomato: gaining an insight to their physiological and stress-specific roles. **PLoS One 12**: e0187504.

Islam S, Sajib SD, Jui ZS, Arabia S, Islam T and Ghosh A (2019) Genome-wide identification of glutathione S-transferase gene family in pepper, its classification, and expression profiling under different anatomical and environmental conditions. **Scientific Reports 9**: 9101.

Kayum A, Nath UK, Park J-I, Biswas MK, Choi EK, Song J-Y, Kim H-T and Nou I-S (2018) Genome-wide identification, characterization, and expression profiling of glutathione S-transferase (GST) family in pumpkin reveals likely role in cold-stress tolerance. **Genes 9**: 84.

Kou M, Liu Y-j, Li Z-y, Zhang Y-g, Tang W, Yan H, Wang X, Chen X-g, Su Z-x and Arisha MH (2019) A novel glutathione S-transferase gene from sweetpotato, IbGSTF4, is involved in anthocyanin sequestration. **Plant Physiology and Biochemistry 135**: 395-403.

Kumar S, Stecher G and Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. **Molecular Biology and Evolution 33**: 1870-1874.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A and Lopez R (2007) Clustal W and Clustal X version 2.0. **Bioinformatics 23**: 2947-2948.

Lin Y, Zhang L, Zhang J, Zhang Y, Wang Y, Chen Q, Luo Y, Zhang Y, Li M and Wang X (2020) Identification of Anthocyanins-related Glutathione S-transferase (GST) genes in the genome of cultivated strawberry (Fragaria× ananassa). **International Journal of Molecular Sciences 21**: 8708.

Liu Y, Jiang H, Zhao Y, Li X, Dai X, Zhuang J, Zhu M, Jiang X, Wang P and Gao L (2019) Three *Camellia sinensis* glutathione S-transferases are involved in the storage of anthocyanins, flavonols, and proanthocyanidins. **Planta 250**: 1163-1175.

Liu Y, Su W, Wang L, Lei J, Chai S, Zhang W and Yang X (2021) Integrated transcriptome, small RNA and degradome sequencing approaches proffer insights into chlorogenic acid biosynthesis in leafy sweet potato. **PLoS One 16**: e0245266.

Livak KJ and Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2 (-Delta Delta C (T)) method. **Methods 25**: 402-408.

Nishida M, Harada S, Noguchi S, Satow Y, Inoue H and Takahashi K (1998) Three-dimensional structure of *Escherichia coli* glutathione S-transferase complexed with glutathione sulfonate: catalytic roles of Cys10 and His106. **Journal of Molecular Biology 281**: 135-147.

Wang J, Zhang Z, Wu J, Han X, Wang-Pruski G and Zhang Z (2020) Genome-wide identification, characterization, and expression analysis related to autotoxicity of the GST gene family in *Cucumis melo* L. **Plant Physiology and Biochemistry 155**: 59-69.

Wang R, Ma J, Zhang Q, Wu C, Zhao H, Wu Y, Yang G and He G (2019) Genome-wide identification and expression profiling of glutathione transferase gene family under multiple stresses and hormone treatments in wheat (*Triticum aestivum* L.). **BMC genomics 20**: 1-15.

Wei L, Zhu Y, Liu R, Zhang A, Zhu M, Xu W, Lin A, Lu K and Li J (2019) Genome wide identification and comparative analysis of glutathione transferases (GST) family genes in *Brassica napus*. **Scientific Reports 9**: 9196.

Xu J, Zheng A-Q, Xing X-J, Chen L, Fu X-Y, Peng R-H, Tian Y-S and Yao Q-H (2018) Transgenic Arabidopsis plants expressing grape glutathione S-Transferase gene (VvGSTF13) show enhanced tolerance to abiotic stress. **Biochemistry (Moscow) 83**: 755-765.

Zhang H, Zhang Q, Zhai H, Li Y, Wang X, Liu Q and He S (2017) Transcript profile analysis reveals important roles of jasmonic acid signalling pathway in the response of sweet potato to salt stress. **Scientific Reports 7**: 1-12.

Zhang L, Yu Y, Shi T, Kou M, Sun J, Xu T, Li Q, Wu S, Cao Q and Hou W (2020) Genome-wide analysis of expression quantitative trait loci (eQTLs) reveals the regulatory architecture of gene expression variation in the storage roots of sweet potato. **Horticulture Research 7**: 1-12.