

Analysis of GSH1-GSH3 Genes by Comprehensive Use of Bioinformatics Method in Tomato

Yulong Pang

School of Ocean and Agricultural Engineering, Yantai Institute, China Agricultural University,
Yantai, China

dh15169639317@163.com

Abstract. GSH is a ubiquitous gene family in animals and plants, which plays a key regulatory role in plant growth, development and stress response. In this study, three GSH genes are identified in the Phytozome database, and bioinformatics analysis of physicochemical properties, structural domains, promoter cis-acting elements and phosphorylation sites is performed. The results show that the numbers of introns and exons of tomato GSH1-GSH3 are quite different. Promoters cis-acting elements are analyzed in GSH gene promoters, which account for a large proportion of light-responsive, hormone-responsive and protein synthesis elements. Transcription factors based on different tissues group data analysis reveal that the expression of GSH genes is different in different tissue development processes. The highest expression sites of GSH1-GSH3 are in 1-cm fruit, GSH2 currant tomato fruit and roots five days after veraison. According to physicochemical properties, transmembrane structure and signal peptide, the study finds that GSH proteins are all hydrophilic proteins and do not have a transmembrane structure and signal peptides. Based on the data analysis of protein phosphorylation sites and glycosylation sites, it is found that the number of the two sites differ significantly between GSH proteins. By means of prediction of secondary and tertiary structures in proteins, it finds that the three GSH proteins have four secondary structures: α -helix, extended chain, β -turn and random coil, but the proportions are different, and the tertiary structures of the three are quite different. Based on the analysis of protein phylogenetic tree, it is believed that tomato and potato are closely related. According to the protein interaction network, it points out that there is a strong link between GSH1 and GSH3. Through the analysis of tomato GSH family system, this study is expected to provide a theoretical basis for further research on the biological function of GSH gene.

Keywords: Bioinformatics; Tomato; GSH.

1. Introduction

Tomato is the second most important plant in the world and one of the most widely grown fruit or vegetable crops in China, enjoying a high status in vegetable production. GSH (glutathione) is the most abundant small molecule thiol compound in plant organisms, with the main non-protein reducing sulfur and bioactive peptide with strong antioxidant properties, and is involved in protein synthesis, sulfur transfer storage and other processes in the growth and development of tomato [2-3]. Related studies have shown that GSH plays an important role in plant resistance to heavy metal and other adversity stress.

GSH genes have been identified in many plants, such as wheat, *arabidopsis thaliana*, potato, and cucumber, where 23, 2, 6, and 3 GSH family members are identified respectively. Han et al. found that SO₂ is able to up-regulate GSH gene expression in cereals by exposing cereal seedlings under drought stress to a certain concentration of SO₂ gas environment. Kudelko et al. discovered that changes in GSH concentration affected growth hormone production by analyzing the effect of reduced GSH concentration on embryogenesis potential in *arabidopsis thaliana* in vitro cultures, which in turn affects embryogenesis in explants. Zhao et al. argued that exogenous GSH significantly reduces the symptoms of toxicity and induces the expression of GSH genes in plants by treating leaves of *hydrocharis dubia* poisoned by Zn²⁺ with different GSH concentration solutions. Qiu et al. used CO₂ laser treatment on drought-stressed wheat seedlings and measured the GSH

content in the plants, proving that a certain time of CO₂ laser treatment could effectively promote the expression of GSH genes in wheat seedlings under drought stress.

In this study, the tomato GSH gene is analyzed for the first time by bioinformatics methods, and phylogenetic tree, structural domains, promoter cis-acting elements and phosphorylation sites of the gene members are studied to provide a theoretical basis for further exploration of the response mechanism of tomato under adversity stress and further study of the functions of the tomato GSH gene.

2. Materials and Methodology

2.1 Determination of Research Object

Using the Phytozome v13 database (<https://phytozome-next.jgi.doe.gov/>), we searched for "GSH", downloaded the tomato GSH genomic data obtained from the search, and named GSH1 (Solyc08g068800), GSH2 (Solyc08g081010), and GSH3 (Solyc01g098610) in order .

2.2 Analysis of GSH1- GSH3 Gene Structure and Cis-Acting Elements

The gene structure is mapped by GSDS v2.0 (<http://gsds.gao-lab.org/>) . Based on the data information from the tomato GSH genome, the promoter sequence of the GSH gene (the sequence 2000 bp upstream of the start codon) is extracted. The promoter cis-acting elements within the gene region are predicted by PlantCARE(<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) online and TBtools (<https://github.com/CJ-Chen/TBtools/releases>) to visualize the data [11-12].

2.3 Specific Expression of GSH1- GSH3 Genes Analysis

According to the data of the tomato genome, by using the ePlants module on website BAR (<http://bar.utoronto.ca/>) and entering the GSH gene ID number, a graph of the coloring tomatoes based on the dynamics of gene expression can be obtained, which can be used to determine the relative expression of the selected genes in different organs of tomatoes and in different developmental stages of the organs. At the same time, the data is exported to draw a heat map of gene tissue expression and we use MATLAB (<https://www.mathworks.com/>) to draw a heat map of gene expression for visual analysis of the data [13-14].

2.4 Amino Acid Composition and Physicochemical Properties of GSH1-GSH3 Proteins Analysis

Expasy (<https://web.expasy.org/protparam/>) is used to predict the physicochemical property indicators and amino acid composition of tomato GSH1- GSH3 proteins , while CELLO (<http://cello.life.nctu.edu.tw/>) is used to determine the protein distribution on the subcellular structure. The imageGP (<http://www.ehbio.com/ImageGP/>) [16-17] helps to visualize the data.

2.5 Analysis of Signal Peptide, Transmembrane Structure and Hydrophobicity of GSH1-GSH3 Proteins

Based on the tomato GSH protein information, SignalP (<http://www.cbs.dtu.dk/services/SignalP/>), DTU (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>), and Expasy (<https://web.expasy.org/protparam/>) are taken to predict GSH protein signal peptide and its composition, transmembrane structure composition and hydrophobicity respectively [18-19].

2.6 Prediction of GSH1-GSH3 Protein Phosphorylation Sites and Glycosylation Sites

With online platforms NetPhos-3.1 (<https://services.healthtech.dtu.dk/service.php?NetPhos-3.1>), YinOYang-1.2(<http://www.cbs.dtu.dk/services/YinOYang-1.2/>), DictyOGlyc-1.1 (<https://services.healthtech.dtu.dk/service.php?DictyOGlyc-1.1>), NetNGlyc-1.0 (<https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0>), the study predicts GSH protein phosphorylation

sites, GSH protein O- β -glucose glycosylation sites, GSH protein O-glycosylation sites, and GSH protein N-glycosylation sites [20-22] .

2.7 Prediction of Secondary and Tertiary Structures of GSH1-GSH3 Proteins

The secondary structure of GSH proteins is predicted by SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html). Besides, SWISS-MODEL (<https://swissmodel.expasy.org/interactive>) is employed for homology modeling to predict the tertiary structure of GSH proteins [24-26].

2.8 Construction of Protein Phylogenetic Tree of Different Plants

Multiple sequence comparisons of GSH protein sequences in tomato, maize, sorghum, wheat, rice, arabidopsis thaliana, potato, and cucumber are performed by the software MEGA11 (<https://www.megasoftware.net/>), and then phylogenetic tree is constructed by the neighbor-joining method . Bootstrap is set to 500 replicates.

2.9 Protein Interaction Network Predictions

Based on the information of tomato GSH protein sequence, the sequence of GSH protein is entered online by the website STRING (<https://www.string-db.org/>) to predict the GSH protein interactions network .

3. Results and Analysis

3.1 GSH1-GSH3 Gene Structure and Cis-Acting Elements

The basic information of SlGSH gene family is given in Table 1. The results show that the longest length of SlGSH2 is 8.020Kb and the shortest is 3.218Kb. Both SlGSH1 and SlGSH2 are located on chromosome 8 and SlGSH3 is on chromosome 1. The gene length ranges from 3.218Kb to 8.020Kb, the transcribed sequence from 1.011Kb to 2.222Kb, and the CDS from 0.597Kb to 1.641Kb. The structural analysis of the GSH gene reveals that the SlGSH1 gene has 3 exons and 2 introns, the SlGSH2 gene has 8 exons and 7 introns, and the SlGSH3 gene has 12 exons and 11 introns. The analysis of cis-acting elements indicates that there are five categories. The first one is the cis-acting elements involved in hormone response including TGACG-motif, CGTCA-motif, GARE-motif, AuxRR-core, TATC-box, P-box, ABRE. The second category is the cis-acting elements involved in response to adversity stress including TC-rich repeats, MBS, ARE, LTR. The third category consists of cis-acting elements related to light response, including TCCC-motif, TCT-motif, chs-CMA1a, GATA-motif, ATCT-motif, Box 4, I-box, TCT-motif, AT1-motif, Gap-box, ACE, G-Box, chs-CMA2a, GT1-motif, G-box. The fourth category is for cis-acting elements directly with regard to plant growth and development, including CAT-box. Finally, the fifth category is for cis-acting elements related to genes involved in transcription and protein binding processes, including TATA-box, AT-rich element, O2-site, CAAT-box, and G-box. O2-site, CAAT-box, Unnamed__1, CCAAT-box, and MBSI.

Table 1. Basic information of GSH

Gene	Gene ID	Gene Length (Kb)	Transcription sequence length (Kb)	CDS sequence length (Kb)
<i>SlGSH1</i>	Solyc08g068800	3.218	1.011	0.597
<i>SlGSH2</i>	Solyc08g081010	8.020	2.178	1.572
<i>SlGSH3</i>	Solyc01g098610	5.434	2.222	1.641

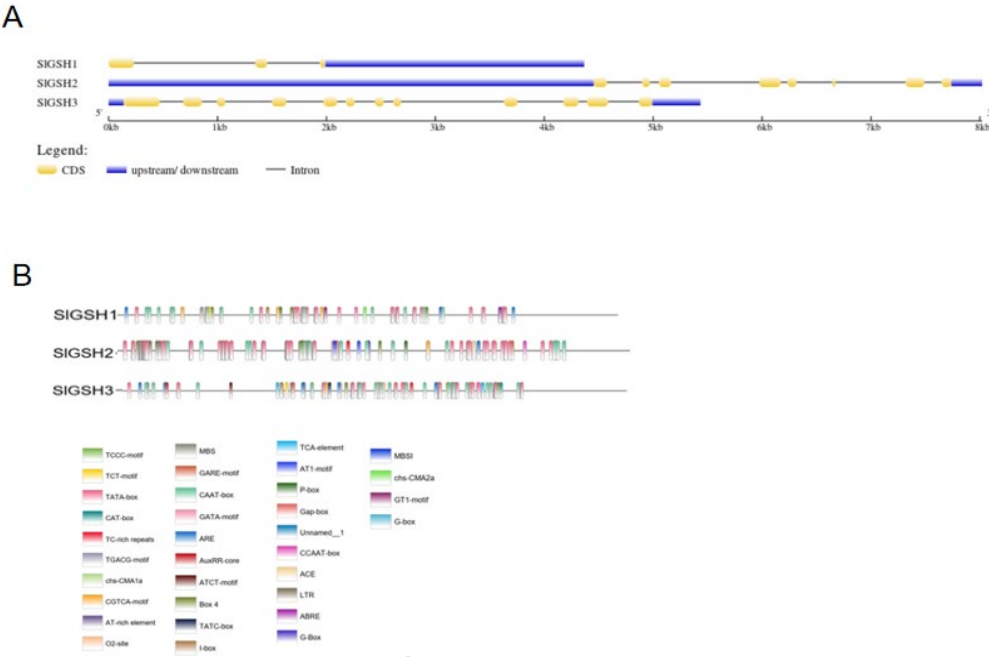


Figure 1. Analysis of Structure of GSH Gene (A) and Cis-regulatory Elements (B)

3.2 Specific Expression of GSH1-GSH3 Genes

From gene expression heat map analysis, in terms of individual gene expression, GSH1 expression is highest in 1-cm fruits and lowest in common tomatoes ten-day veraison and currant tomatoes five-day veraison. GSH2 expression is highest in currant tomatoes in five-day veraison and lowest in unopened buds, and GSH3 expression is highest in roots and lowest in ten-day veraison In terms of overall gene expression, GSH2 is significantly more expressed in each site than the other two genes, with GSH2 being the most highly expressed in the five-day currant tomatoes, followed by higher expression in the ten-day veraison. GSH1 is significantly less expressed in each site than the other two genes, with GSH1 being the most highly expressed in the 1-cm fruits, followed by higher expression in the blooming buds.

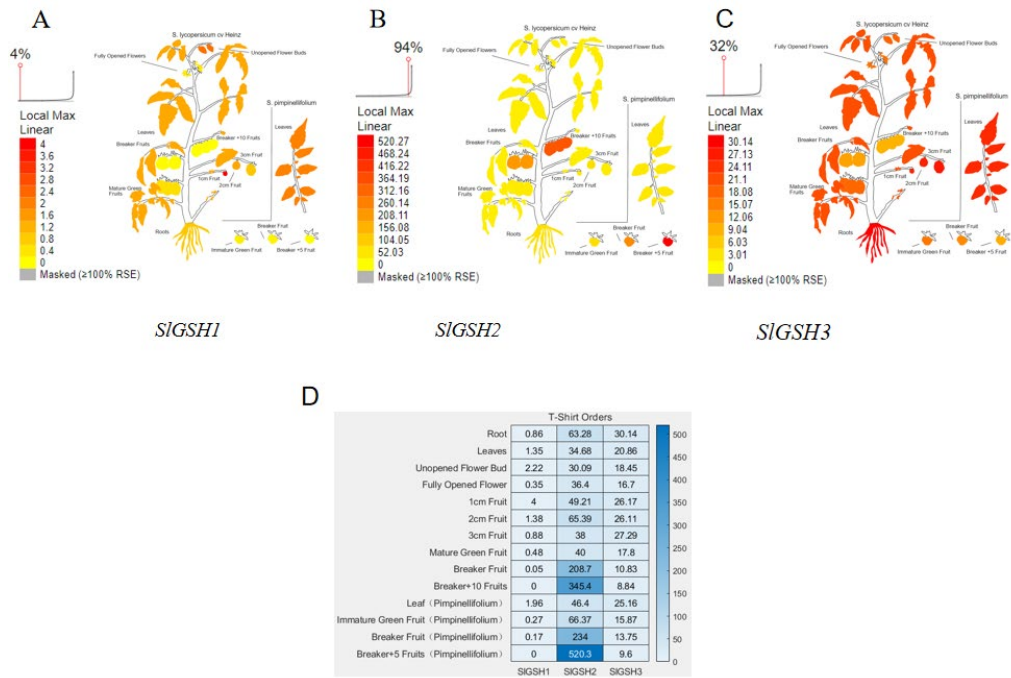


Figure 2. Heat Map of Specific Expression of SlGSHs Genes

3.3 Amino Acid Composition and Physicochemical Properties of GSH1-GSH3 Proteins

Analysis of the amino acid composition of the GSH1-GSH3 proteins reveals that the proteins have the highest number of leucine, followed by serine, glutamic acid and glycine and the least amount of tryptophan. The three proteins do not contain pyrrolysine and selenocysteine. The analysis of the subcellular localization heat map indicates that GSH1 is located in mitochondria, GSH2 in chloroplasts and cytoplasm, and GSH3 in cytoplasm. To understand the physicochemical properties of amino acids of GSH1-GSH3 proteins, analysis from isoelectric point and hydrophilicity index implies that the total number of amino acids of GSH1-GSH3 proteins ranged from 3128 to 8618, with the total number of amino acids from 198 to 546, and the relative molecular mass from 22220.36 to 61163.73. The isoelectric points of GSH1-GSH3 are 9.44, 6.26 and 5.85. The positively charged residues are 23, 64 and 61 and negatively charged residues are 14, 68, 71 respectively, where GSH2 is a stable protein and GSH1 and GSH3 are unstable proteins. The hydrophilicity index is between -0.304 and -0.285, so GSH1-GSH3 are hydrophilic proteins.

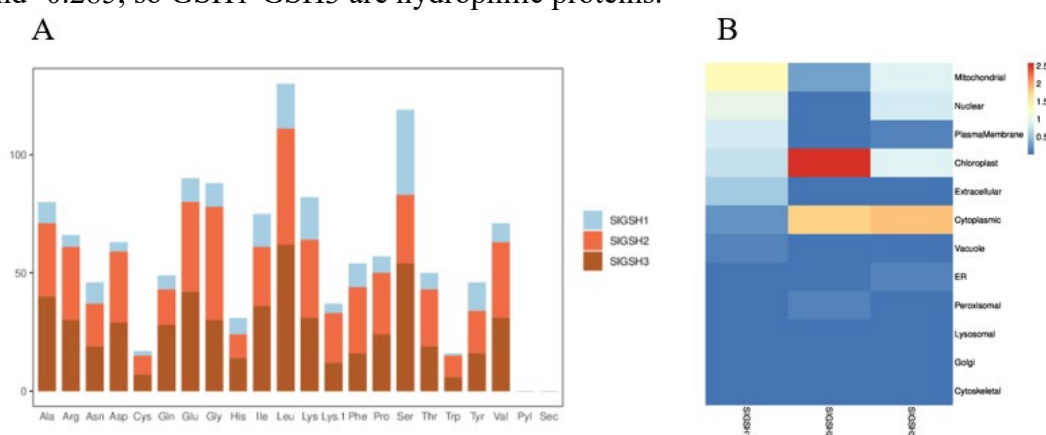


Figure 3. Amino Acid Composition Diagram and Subcellular Localization Heat Map

Table 2. Physicochemical Properties of Amino Acids

Gene Name	Total number of atoms	Number of amino acids	Relative molecular mass (Da)	Isoelectric point	Number of positively charged residues	Number of negatively charged residues	Instability index	Aliphatic index	Hydrophilicity index
SIGSH1	3128	198	22220.36	9.44	23	14	51.43	81.26	-0.285
SIGSH2	8266	523	59074.84	6.26	64	68	38.84	78.85	-0.304
SIGSH3	8618	546	61163.73	5.85	61	71	48.67	93.79	-0.297

3.4 Signal Peptide, Transmembrane Structure and Hydrophobicity of GSH1-GSH3 Proteins

Using the website DTU (<http://www.cbs.dtu.dk/services/SignalP/>) to predict the signal peptide and composition of GSH proteins, this paper finds that the probability of GSH1-GSH3 proteins being signal peptides are 0.2%, 0.07%, and 0.12% respectively, which are very small probability events, so they are not signal peptides and have no signal peptide sites. There is no localization in the plasma membrane in subcellular localization, so according to Figure 4B, GSH1-GSH3 proteins are not transmembrane proteins without transmembrane region. By analyzing the hydrophobicity, this paper finds that the maximum value of GSH1 is 1.800, the minimum value is -3.189, and the overall value is -0.285. The maximum value of GSH2 is 2.022, the minimum value is -2.356, and the overall value is -0.304. The maximum value of GSH3 is 2.722, the minimum value is -2.733, and the overall value is -0.297. The results of the hydrophobicity analysis are consistent with the results of the physicochemical property analysis.

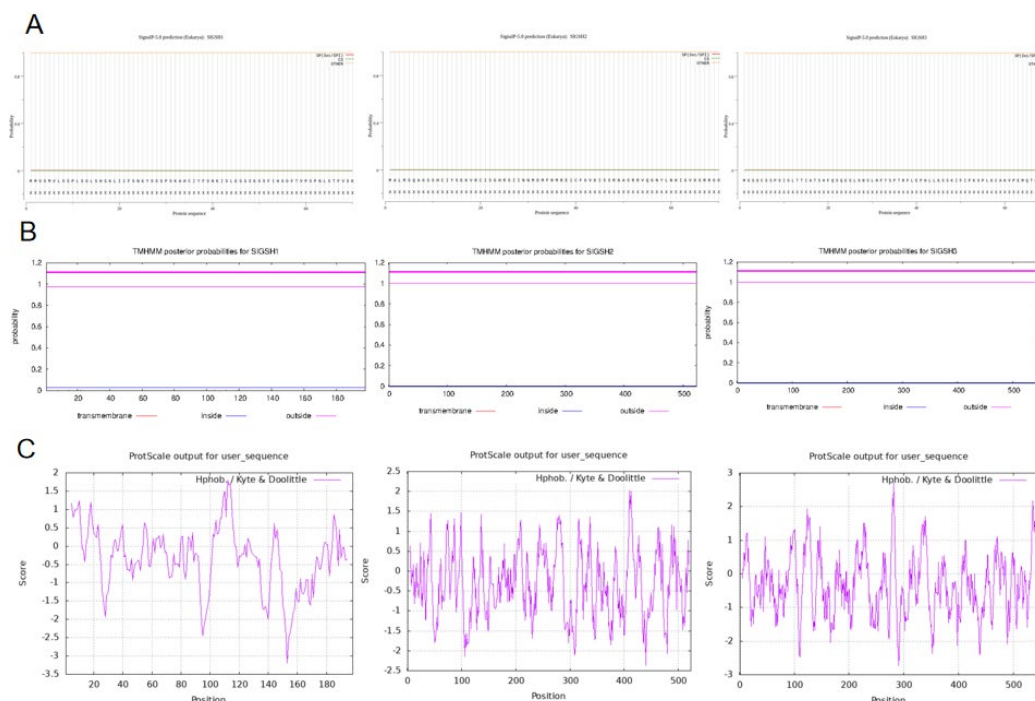


Figure 4. Signal Peptide Judgment (A), Transmembrane Region Prediction (B), and Hydrophobicity Analysis (C) of SIGSH

3.5 GSH1-GSH3 Protein Phosphorylation Sites and Glycosylation Sites

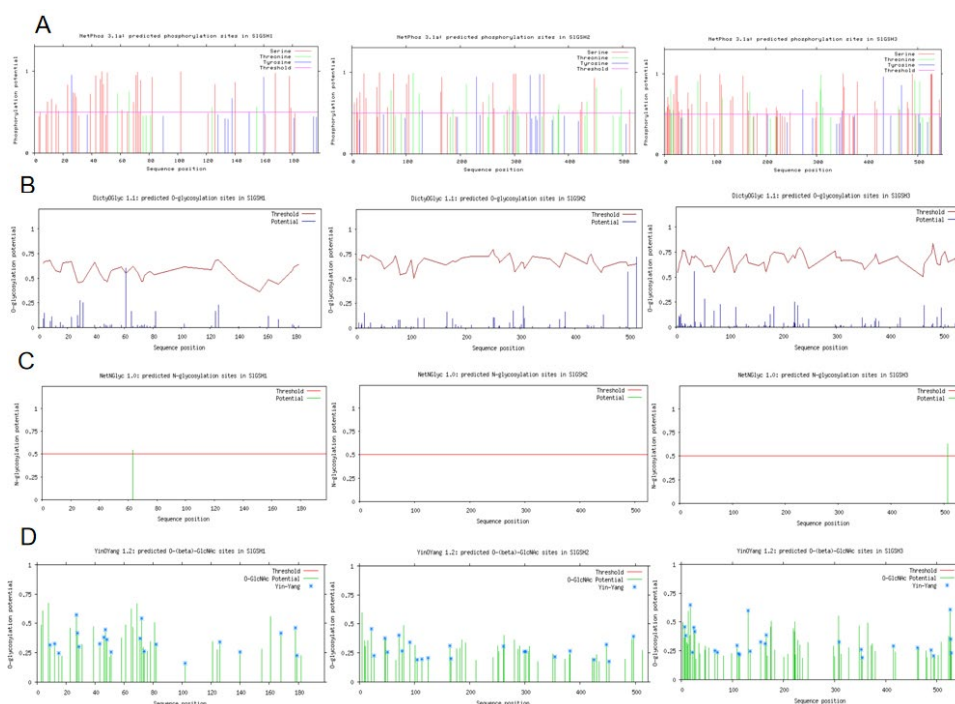


Figure 5. Phosphorylation Site Prediction (A), O-Glycosylation Site Prediction (B), N-Glycosylation Site Prediction (C), O-β-Glucose Glycosylation Site Prediction (D)

To study the phosphorylation sites and glycosylation sites of GSH1-GSH3 proteins, analysis of Figure 5A reveals that the phosphorylation threshold is 0.5, and there are 27, 25 and 44 phosphorylation sites for serine, 4, 11 and 11 phosphorylation sites for threonine and 3, 6 and 4 phosphorylation sites for tyrosine of GSH1-GSH3 proteins respectively. Analysis of Figure 5B implies that GSH1 and GSH2 both have only one O-glycosylation site, located at positions 61 and

513 respectively, and GSH3 has no O-glycosylation site. Analysis of Figure 5C reveals that the N-glycosylation threshold is 0.5, GSH2 has no N-glycosylation site, and GSH1 and GSH3 each has one N-glycosylation site, located at position 63 and 507 respectively. Analysis of Figure 5D indicates that GSH1-GSH3 proteins have 43, 53, and 73 O- β -glucose glycosylation sites and 21, 21, and 26 Yin-Yang sites respectively.

3.6 Secondary and Tertiary Structures of GSH1-GSH3 Proteins

In order to study the secondary structure of GSH proteins, analysis of Figure 7A reveals that the four secondary structures, including α -helix, extended chain, β -turn, and random coil, account for 15.66%, 32.83%, 10.10%, and 41.41% in GSH1, while it is 41.30%, 15.11%, 7.07%, and 36.52% in GSH2, and 41.94%, 15.93%, 6.04%, and 36.08% in GSH3. The selected models for the tertiary structure homology modeling of GSH1-GSH3 proteins are 2he3.1, 6gmo.1, and 5oes.1 respectively. Analysis of the Ramachandran plot Figure 7D reveals that green and light green are allowed regions and white is not allowed region, and most amino acids are located in green and light green regions, so homology modeling is reasonable.

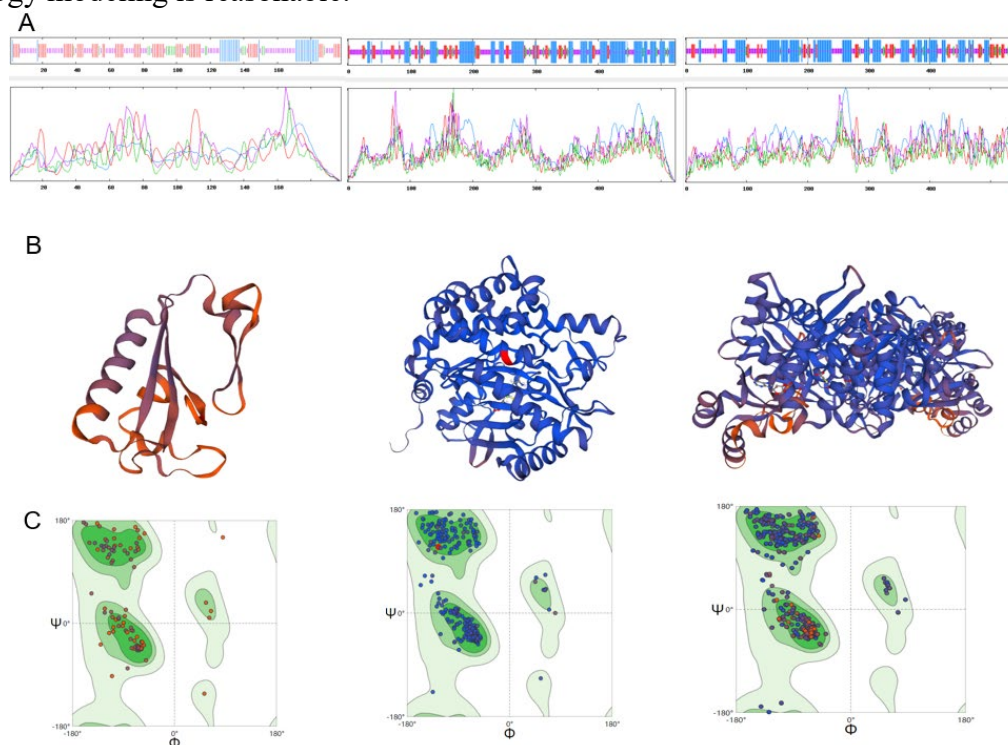


Figure 7. GSHs Protein Secondary Structure Prediction (A), Tertiary Structure Prediction (B), Ramachandran Plot (C)

3.7 Phylogenetic Tree of Different Plant Constituent Proteins

To investigate the evolutionary relationships of the GSH family, a study of tomato (*Solanum lycopersicum* ITAG4.0), maize (*Zea mays* PHJ40 v1.2), sorghum (*Sorghum bicolor* v3.1.1), wheat (*Triticum aestivum* v2.2), rice (*Oryza sativa* v7.0), arabidopsis thaliana (*Arabidopsis thaliana* TAIR10), potato (*Solanum tuberosum* v6.1), and cucumber (*Cucumis sativus* v1.0) are subjected to multiple sequence alignment and phylogenetic tree construction. GSH1 is found to be more closely related to GSH2, and GSH3 is more distantly related to GSH1 and GSH2. GSH1 is most closely related to potato GSH3-Soltu.DM.08G018060. GSH2 is most closely related to potato GSH1-Soltu.DM.08G027770. GSH3 is most closely related to potato GSH6- DM.04G008160. The results indicate that potato is most closely related to tomato in evolution among the above plants.

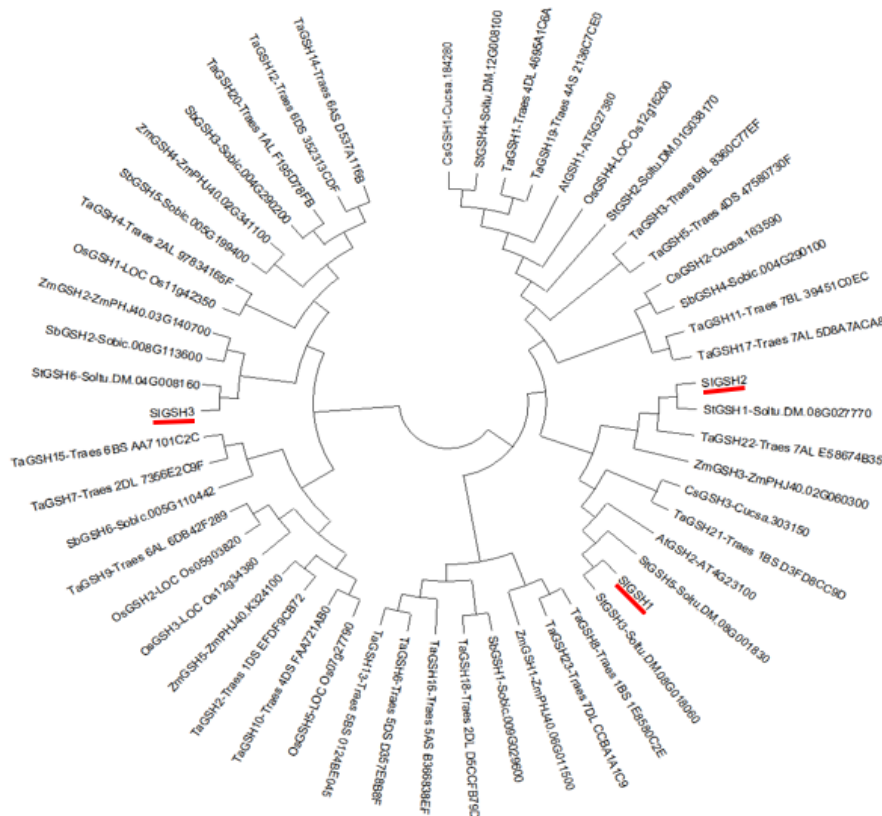


Figure 8. Protein Phylogenetic Tree

3.8 Protein Interaction Network

The protein interaction network is found to interact between any two of GSH1-GSH3. LAPA2, Solyc00g187050.2.1, Solyc08g081250.2.1, Solyc10g079720.1.1 all have interaction with GSH1-GSH3. Solyc05g051780.2.1, Solyc09g082060.2.1, Solyc12g008640.1.1, Solyc09g010560.1.1, 5-oxoprolinase have interacted relationship with GSH1 and GSH2, while they do not interact with GSH3.

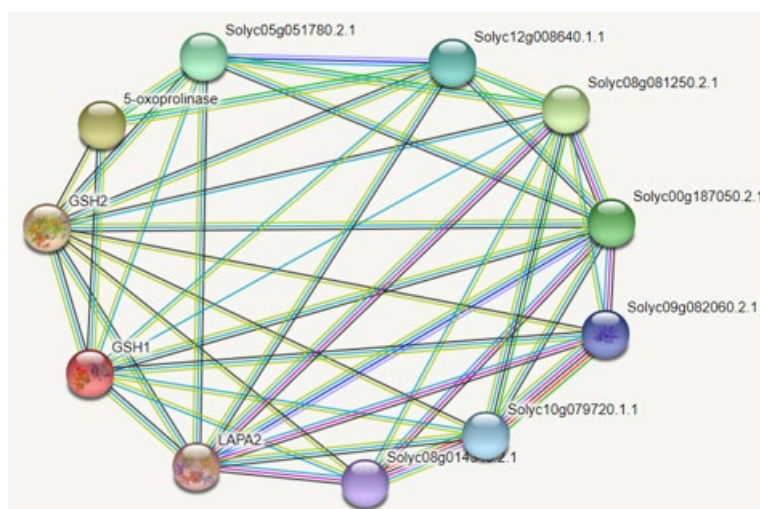


Figure 9. Protein Interaction Network

4. Discussion

GSH genes are widely presented in the plant body and play an important role in the response mechanism to salt stress and low temperature stress in plant seedlings [29-30]. So far, research topics on the role of GSH in plants involves its antioxidant capacity, growth performance and stress tolerance in plants and the role of exogenous GSH addition . Comprehensive studies on GSH genes in tomato and other plants are less available.

Exons and introns of genes can be regarded as important imprints to reveal the evolutionary relationships between members of gene families . Analysis of the structure of GSH genes reveals that the number of exons and introns differ greatly among GSH1-GSH3, and the gene structure is highly variable, with fewer introns in SlGSH1 presumed to be the more primitive type. First of all, SlGSH3 has 12 exons with multiple intron insertions in the gene sequence, which is a relatively new member of the GSH . Analysis of GSH gene cis-acting elements indicates that the light response element is the most prevalent, while the hormone response element, the protein synthesis and transcription element are more prevalent, which are involved in plant growth, adversity stress and the prevalent distribution of the light response element. Based on the role of GSH under lead stress, it is hypothesized that the protein synthesis element is closely linked to GSH regulation of heavy metal stress while the light response element affected plant photosynthesis . In addition, the high expression of GSH genes in fruits may be involved in redox homeostasis during fruit development and in response to adversity stress . The GSH protein is mainly localized in the cytoplasm, which is consistent with the research results of Edwards et al. that most GSH proteins are localized in the cytoplasm. Bioinformatics analysis reveals that GSH is a stable hydrophilic protein without transmembrane structural domains and signal peptides . Besides, due to the important role of polyphosphorylation sites in activating protein viability, the high number of phosphorylation sites of GSH proteins is hypothesized to regulate the activity and thus biological functions through phosphorylation modifications of threonine, tyrosine and serine, with multiple biological functions [38-39]. The tertiary structures of GSH1-GSH3 vary widely, inferring that the functions of the three proteins in tomato may be different . Finally, the protein phylogenetic tree reveals that the tomato GSH gene is highly homologous and evolutionarily related to potato, which is also a member of the Solanaceae family. It suggests that the biological functions performed by GSH proteins in tomato and potato may be the same . Solyc05g051780.2.1, Solyc09g082060.2.1, Solyc12g008640.1.1, Solyc09g010560.1.1, 5-oxoprolinase and GSH3 do not interact with GSH3, which may be due to its slightly different response mechanism in the GSH cascade pathway .

In this study, the tomato GSH gene is identified and analyzed with preliminary functional predictions, providing a theoretical reference to explore the role of tomato GSH genes in response to biotic and abiotic stress and in tomato growth and development. However, further studies on the specific functions played by GSH in tomato are still needed.

References

- [1] Zhang, Y.G. et al. (2016). Domestic Situation and Development Trend of Grafted Seedling of Tomato. Horticulture & Seed, 09, 5-9.
- [2] Chen, Y.Y. et al. (2020). Effects of foliar spraying oxidized glutathione on dry matter accumulation and yield of potato. Acta Agriculturae Boreali-Sinica, 48(04), 54-59.
- [3] Wang, K. H. et al. (2021). Review of the Role of GSH/GSSG in Plant Abiotic Stress Response. Acta Horticulturae Sinica, 48(04), 647-660.
- [4] Lu, L.D. et al. (2014). The Change of Glutathione (GSH) Content Induced by Cadmium and Zinc Stress in *Carthamus tinctorius* L. Journal of Anhui Agricultural Sciences, 42(26), 8960-8961+9005.
- [5] Han, Y. S. et al. (2020). Sulfur Dioxide Enhances Drought Tolerance of Foxtail Millet Seedlings by the Modulation of Glutathione Homeostasis. Chinese Journal of Biochemistry and Molecular Biology, 36(05), 559-565.

- [6] Kudelko K., & Gaj, M. D. (2019). Glutathione (GSH) induces embryogenic response in in vitro cultured explants of *Arabidopsis thaliana* via auxin-related mechanism. *Plant Growth Regulation*, 89(1), 25-36.
- [7] Zhao, J. et al. (2006). Alleviatory Effect of Exogenous Glutathione (GSH) on *Hydrocharis dubia* Toxicated by Zn²⁺. *Journal of Tropical and Subtropical Botany*, 03,213-217.
- [8] Qiu, Z.B. et al. (2008). Effects of CO₂ Laser on Glutathione-dependent Antioxidative System in Wheat Seedling under Drought Stress. *Journal of Wuhan Botanical Research*, 04,402-406.
- [9] Hosmani, P. S. (2019). An improved de novo assembly and annotation of the tomato reference genome using single-molecule sequencing, Hi-C proximity ligation and optical maps. *BioRxiv*, 767764.
- [10] Hu, B. et al. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*, 31(8), 1296-1297.
- [11] Lescot, M. et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic acids research*, 30(1), 325-327.
- [12] Chen, C. et al. (2018). TBtools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. *BioRxiv*, 289660.
- [13] Waese, J. et al. (2017). ePlant: visualizing and exploring multiple levels of data for hypothesis generation in plant biology. *The Plant Cell*, 29(8), 1806-1821.
- [14] Tomato Genome Consortium, X. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, 485(7400),635-641.
- [15] Gasteiger, E. et al. (2005). Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook*, 571-607.
- [16] Yu, C. S. et al. (2006). Prediction of protein subcellular localization. *Proteins: Structure, Function, and Bioinformatics*, 64(3), 643-651.
- [17] Zhang, R. et al. (2018). Survival associated alternative splicing events in diffuse large B-cell lymphoma. *American journal of translational research*, 10(8), 2636.
- [18] Armenteros, J. J. A. et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature biotechnology*, 37(4), 420-423.
- [19] Möller, S. et al. (2001). Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics*, 17(7), 646-653.
- [20] Blom, N. et al. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics*, 4(6), 1633- 1649.
- [21] Blom, N. et al. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of molecular biology*, 294(5), 1351-1362.
- [22] Gupta, R. et al. (1999). Scanning the available *Dictyostelium discoideum* proteome for O-linked GlcNAc glycosylation sites using neural networks. *Glycobiology*, 9(10), 1009- 1022.
- [23] Geourjon, C., & Deleage, G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics*, 11(6), 681-684.
- [24] Waterhouse, A. et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research*, 46(1), 296-303.
- [25] Benkert, P. et al. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*, 27(3), 343-350.
- [26] Studer, G. et al. (2020). QMEANDisCo-distance constraints applied on model quality estimation. *Bioinformatics*, 36(6), 1765-1771.
- [27] Tamura, K. et al. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, 38(7), 3022-3027.
- [28] Szklarczyk, D. et al. (2021). The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research*, 49(1), 605-612.
- [29] Zhou, Y. (2019). Salt-Tolerant Mechanism of GSH Alleviates Sate-induced Stress in Tomato Seedlings. Doctorial dissertation. Shihezi University, Shihezi.

- [30] Li, J. et al. (2021). Study on the Response Mechanism of the AsA-GSH Cycle in Cotton Seedling Under Low Temperature Stress. *Journal of Nuclear Agricultural Sciences*, 35(01), 221-228.
- [31] Sun, G.X. et al. (2018). Application of Glutathione in Animal and Plant Fields. *Journal of Anhui Agricultural Sciences*, 46(08), 42-45.
- [32] Long, M. et al. (2003). The origin of new genes: glimpses from the young and old. *Nature Reviews Genetics*, 4(11), 865-875.
- [33] Yang, L.Y. et al. Genome-wide identification and expression analysis of maize ACO family. *Journal of Natural Science of Hunan Normal University*, 04,1-12.
- [34] Zhao, L.Q. et al. (2021). Effects of reduced glutathione on the growth and photosynthesis of perennial ryegrass under lead stress. *Acta Prataculturae Sinica*, 30(09),97-104.
- [35] Zhang L. et al. (2016). Cloning and Prokaryotic Expression of TaGST from *Triticum aestivum* L. *Acta Agriculturae Boreali-Sinica*, 31(03), 38-43.
- [36] Kumar, S., &Trivedi, P. K. (2018). Glutathione S-transferases: role in combating abiotic stresses including arsenic detoxification in plants. *Frontiers in Plant Science*, 9, 751.
- [37] Wang, T. K., & Liu, F. (2022). Cloning and bioinformatics analysis of EPSPS gene in tomato. *Molecular Plant Breeding*, 04, 1-9.
- [38] Jin, Y.X. (2022). Cloning and bioinformatics analysis of the LmGST gene of *Phytophthora floxica*. *Molecular Plant Breeding*, 04, 1-27.
- [39] Jin, B.X. et al. (2021). Cloning, Bioinformatics and Expression Analysis of ERFb.2 Transcription Factor in Tomato. *Journal of Fruit Science*, 23,1-10.
- [40] Li, S.Q. et al. (2021). Bioinformatics Analysis of NADP-ME Gene Family in Sugar Beet. *Sugar Crops of China*, 43(04), 17-25.
- [41] Zhou, H.N. et al. (2021). Genome-Wide Identification and Bioinformatic Analysis of Mitogen Activated Protein Kinase Gene Family in Sweet Potato. *Journal of Shenyang Agricultural University*, 52(05),513-520.