Journal of Food and Nutrition Research, 2023, Vol. 11, No. 1, 63-83
Available online at http://pubs.sciepub.com/jfnr/11/1/7
Published by Science and Education Publishing
DOI:10.12691/jfnr-11-1-7

Quinoa (*Chenopodium quinoa* Willd.) Genome-Wide Analysis of Glutathione S-transferase (*CqGSTs*) and Transcription Factor *bHLH* Family Study on Their Salt Tolerance Function

Kaiyuan Cui[1,2], Zhijun Qiang[1,2], Rongzhen Wang[1,2], Pengcheng Ding[1,2], Aixia Ren[1,2], Linghong Li[1,2], Hafeez Noor[1,2], Xiangyun Wu[3], Min Sun[1,2,*], Zhiqiang Gao[1,2]

[1]College of Agriculture, Shanxi Agricultural University, Taigu, 030801, Shanxi, China
[2]Key Laboratory of Functional Agriculture in the Loess Plateau,
Ministry of Agriculture and Rural Affairs, Taigu, 030801, Shanxi, China
[3]Shanxi Jiaqi Agricultural Technology Co., Ltd.
*Corresponding author: sunmin@sxau.edu.cn

Received November 21, 2022; Revised January 02, 2023; Accepted January 10, 2023

Abstract Since the industrial Revolution, as the global economy has boomed and the agricultural population has expanded. Excessive use of chemical fertilizer and unreasonable farming methods make soil salinization more and more serious. *Chenopodium quinoa*, it has unique nutritional value and strong stress resistance and adaptability, under the background of soil salinization, quinoa has been widely studied as a halophyte model. With the release of high-quality genome of quinoa, more and more salt-tolerant genes of quinoa have been cloned gradually. Bioinformatics and expression analysis of *GST* gene in quinoa in this study, 114 *CqGST* genes were identified from the whole genome of quinoa by bioinformatics methods. The phylogenetic tree showed that 114 *CqGST* genes were divided into seven subgroups: GSTU (68 members), GSTF (23 members), GST members), GSTZ (6 members), GSTT (5 members), DHAR (4 members) and TCHQD (2 members). Gene structure and Motif analysis showed high similarity among members of each subgroup. Phylogenetic analysis of these genes suggested that tandem and fragment replication events played a key role in the expansion of the *CqGSTs* gene family, and the *CqGST* genes may have undergone strong purification selection during the evolution process. Analysis of salt-treated transcriptome from the roots of salt-tolerant and salt-sensitive quinoa cultivars showed that Salt treatment induced changes in the expression levels of *CqGSTs* genes, and eight *CqGST* genes (*CqDHAR2*, *CqDHAR3*, *CqGSTU22*, *CqGSTU44*, *CqGSTU60*, *CqGSTU63*, *CqGSTU67*, *CqGSTU68*) were steadily up-regulated in both cultivars. *RT-qPCR* results showed that these selected *CqGST* genes were not only induced by salt stress, but also by drought stress.

Keywords: *Chenopodium quinoa willd, glutathione -s transferase, agrobacterium tumefaciens, salt stress*

Cite This Article: Kaiyuan Cui, Zhijun Qiang, Rongzhen Wang, Pengcheng Ding, Aixia Ren, Linghong Li, Hafeez Noor, Xiangyun Wu, Min Sun, and Zhiqiang Gao, "Quinoa (*Chenopodium quinoa* Willd.) Genome-Wide Analysis of Glutathione S-transferase (*CqGSTs*) and Transcription Factor *bHLH* Family Study on Their Salt Tolerance Function." *Journal of Food and Nutrition Research*, vol. 11, no. 1 (2023): 63-83. doi: 10.12691/jfnr-11-1-7.

1. Introduction

Quinoa (*Chenopodium quinoa* Willd), a heterologous tetraploid (2n=4x=36), is an annual crop of Amaranthus family, with a genome size of about 1.5 *Gbp* [1]. Quinoa grains are nutritionally rich and balanced, rich in protein, starch, VB1, folic acid, minerals (Ca, Zn, Fe) and other nutrients [2]. In addition, because quinoa grows at high altitude (> 3500 meters above sea level) all year round, it is subjected to drought, low temperature, salt and other abiotic stresses during its growth and development, which not only limits the planting area, but also affects the growth, development and yield of quinoa [3]. With the publication of high-quality reference genomes [4]. Quinoa has now become an important plant material for studying the mechanism of salt tolerance in plants. Basic helix-loop-helix (*bHLH*) family is the second largest gene family in plants, and is involved in many biological processes such as plant growth, development, metabolism, and resistance to abiotic stress [5]. However, this gene family has not been identified in Quinoa. The systematic identification of quinoa *bHLH* gene family members in quinoa genome, and the analysis of their chromosome location, gene structure, evolutionary relationship and

expression characteristics are of great significance to elucidate the response of quinoa *bHLH* transcription factor family to salt stress, which is of great significance to enrich the mechanism of salt tolerance of quinoa and breed new varieties of quinoa tolerant to salt stress. The *bHLH* transcription factor family contains two highly conserved domains. The alkaline domain is located at the amino terminus of *bHLH* and is responsible for binding to specific DNA sequences and recognizing E-box elements [6]. The *bHLH* region is located at the carboxyl terminal and is connected by a ring structure of two alpha-helices and hydrophobic amino acids, containing 40 amino acids [7]. Due to the flexibility of helices, the *bHLH* domain facilitates protein interactions and regulates the expression of target genes through the formation of homodimer or heterodimer complexes [8]. According to the conserved and structural characteristics of the gene sequences, the Arabidopsis family is divided into 21 subfamilies [9]. It has been reported that the *bHLH* transcription factor family is mainly involved in abiotic stress in plants, such as in response to drought, low temperature, salt, and abscisic acid hormones [10,11]. In Arabidopsis thaliana, *AtbHLH006*, *AtbHLH17*, *AtbHLH32*, *AtbHLH92*, *AtbHLH122*, *AtbHLH128* and *AtbHLH130* directly or indirectly participate in abscisic acid signaling pathway to improve drought resistance of Arabidopsis thaliana [12]. Overexpression of bHLH transcription factors MYC type *ICE1*, *ICE2* and CBF enhanced Arabidopsis tolerance to low temperature stress [13]. In wheat, TabHLH1 can regulate abscisic acid mediated stress tolerance pathway, thereby improving plant adaptability to drought and salt stress [14]. *TabHLH39* gene is involved in regulating gene expression in response to stress, thereby improving salt tolerance of overexpressed wheat plants [15]. In rice, *OsbHLH148* and *OsbHLH006* (*RERJ1*) respond to drought stress through jasmonic acid signaling pathway [16, 17]. In the process of drought and abscisic acid induction, the expression of PebHLH35 gene was increased in populus euphratica, which actively participated in the regulation of drought stress, thus improving the tolerance of populus euphratica [18]. And has been found to have different degrees of functional differentiation. Although the whole genome of Quinoa was sequenced by Jarvis DE et al in 2017 [19]. An annual dicotyledonous plant that originated in the Andes and has been cultivated for about 5,000 to 7,000 years, it was the main food crop of the local Inca people [20]. Quinoa seeds are rich in nutrients and are considered an international food of the United Nations Agriculture Organization (FAO) praised as the only "whole nutrition food" that can meet the basic nutrition of human body with all kinds of single plants [21]. Therefore, the general attention of the people, the output shows multiple growth. As a halophytes [22]. Quinoa is drought - tolerant, salt - alkali - tolerant and barren - resistant. With the publication of a high-quality reference genome for Quinoa [23]. The Plants have become one of the important materials for studying the mechanism of plant stress resistance. Quinoa is mainly distributed in Peru, Bolivia, Ecuador and Chile, and is native to South America the Andes Mountains region of Asia is the traditional food crops of the Inca indigenous people, which has been about 5000~7000 years ago Years of planting history [24,25]. Quinoa is grown widely in South America, from 2° north latitude in Colombia to southern Chile. It can be cultivated from sea level to 4,000 m above sea level at latitude 40° [26,27]. Because quinoa is also exposed to extreme environmental conditions it still has a high yield [28]. The discovery of special nutritional properties of quinoa, health care based on quinoa as a raw material Products and food are more and more consumers love, relying solely on the country of origin production in short supply. Quinoa leaves are broad with serrated edges, smooth or palmately divided margins, and generally have short velvet Hairy, annual dicotyledonous plants, single leaves alternate, depending on the variety of plant height, roots, stems, leaves, flowers, fruit And other plant characteristics, with the characteristics of genetic diversity. Quinoa has a strong resistance and adaptability, can be in the dry It grows in extremely harsh environment such as drought, salinity and frost, and is suitable for growing in arid, salinized and other marginal areas [29,30]. Quinoa was native to high altitudes of several thousand meters. In order to adapt to the harsh environment, quinoa has a net root system like distribution, very developed, this structure helps to resist drought, strong wind and barren environment [31,32]. The smaller leaf area and the abundant vesicle structure on the leaf surface of Quinoa are beneficial to reduce transpiration and thus resist drying arid environment [33,34]. Physiological, quinoa mainly through the accumulation of proline, betaine, soluble sugar and other organic as well Inorganic permeates form tissue elasticity to achieve osmotic regulation. Quinoa also relies on its vesicle structure to accumulate salt tolerance Organic osmotic regulators such as proline, betaine, polyamine and dehydrating protein can remove reactive oxygen species in vivo [35]. Overaccumulation of osmotic regulating substances, Na+ efflux and K+ retention achieve osmotic balance of inorganic ions. Dehydrin and the accumulation of soluble sugars explained the frost resistance of quinoa from the physiological level [36]. In this study, bioinformatic methods were used to identify the quinoa *bHLH* transcription factor family, and their bioinformatic characteristics such as distribution, gene structure, evolutionary differentiation, tissue expression, and salt-stress-induced expression specificity were systematically analyzed, laying a foundation for clarifying the differentiation process and biological function of quinoa *bHLH* transcription factor family. Phylogenetic analysis of these genes suggested that tandem and fragment replication events played a key role in the expansion of the *CqGSTs* gene family, and the *CqGST* genes may have undergone strong purification selection during the evolution process.

2. Materials and Methods

The experiment was conducted in the laboratory of Wheat Research Station, College of Agronomy, Shanxi Agricultural University during 2021-2022.

2.1. Plant Materials and Treatment

Quinoa variety Faro was the experimental material in this chapter. Quinoa seeds were placed in sterilized camps

culture in soil. The indoor culture temperature was 23°C and photoperiod was 16 h light / 8 h darkness. In order to study *CqGST* gene expression patterns under different stress, one-month old Quinoa seedlings with the same growth rate were selected, and 60 mL were used, respectively. The treatment was treated with 300 mmol/L NaCl and 20% PEG6000 solution for 3, 6, 12, 24 and 48 h. Acquisition and processing the above-ground and underground parts of the seedlings were immediately frozen in liquid nitrogen and stored in an ultra-low temperature refrigerator at −80℃ for subsequent analysis, each treatment was repeated three times.

Figure 1. The in florescence of quinoa Wheat Research Station, College of Agronomy, Shanxi Agricultural University

2.2. Primer Design and Verification

According to the quinoa genome database GST gene sequence information, use the Primer - BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) online website design *qRT - PCR* primers used to detect the base due to the expression in different treatments and different treatment times, the primer sequence is shown in Table 1.

Table 1. Primer sequences used for this chapter

Quote by name	Primer sequence
CqDHAR2-F	TGAGAAAACCAAAGCAGCCA
CqDHAR2-R	TTTGGTGACAAGCATTGGAGC
CqDHAR3-F	GGTAGCCCTTGGCCACTTTA
CqDHAR3-R	ACCTCACCAGCACAAGATCA
CqGSTU22-F	GGCTAAAGCTCTTCCTGACCA
CqGSTU22-R	AGAGACACTAGACTAGAAACAGTCG
CqGSTU44-F	ACAAGCTGGTTTTACACATTGGA
CqGSTU44-R	ACTTCATCGGAATCCGGCAA
CqGSTU60-F	GAGGCAGCAAAGAGGGAACT
CqGSTU60-R	GCCACATGTTTCAAGGGCG
CqGSTU63-F	AGCTTCGTCATGAAACTCTCG
CqGSTU63-R	ACTCCCACAATACAAGAGAAAATCG
CqGSTU67-F	GGCTAAAGCTCTTCCTGACCA
CqGSTU67-R	TAAATGGAGTCTCAACTCGCTTG
CqGSTU68-F	TGCAAAGGGAAAGTGTGGCT
CqGSTU68-R	TGGTCAAGTTTCTTCTCCTTACCA
CqGAPDH-F	TTGGGCATTGACTTGGTTATCG
CqGAPDH-R	AGCGTTACTGATGATGGTGTCT
CqActin-F	GGTATTGGAACGGTGCCAGT
CqActin-R	GGACTCGTGGTGCATCTCAA

2.2.1. Solution Preparation

The 50×TAE electrophoresis buffer (1L): Weigh 242g of Tris in an electronic balance and stir with appropriate amount of deionized water Mix and dissolve, add 57.1 mL glacial acetic acid, add 0.5 mol EDTA (pH 8.0) 100 mL, constant volume to 1L.

2.2.2. Identification and Naming of Quinoa GST Gene Family

From arabidopsis thaliana TAIR database download known arabidopsis Jsp (https://www.arabidopsis.org/index.). Mustard GST protein sequence, known rice downloaded from the National Rice Data Center (https://www.ricedata.cn/). The sequence of the quinoa genome was downloaded from Chenopodium DB (h ttps://www.cbrc.kaust.edu.sa/chenopodiumdb/). Based on known GST sequences in Arabidopsis thaliana and rice. The TBTools tool kit was used to compare the data in the quinoa genome database, and the e value was set as E-10. The use CDD NCBI database (https://www.ncbi.nlm.nih.gov/cdd/) and SMART (http://smart.embl-heidelberg.de) database was used to screen quinoa GST genes with conserved domains *gST-C* and *gST-N* NCBI database (https://www.ncbi.nlm.nih.gov/) to verify the accuracy of the results. According to quinoa G the relationship between ST gene and Arabidopsis GST gene and the position of quinoa *GST* gene on chromosome were investigated Name it.

2.2.3. Identification of Members of the *bHLH* Gene Family in Quinoa

The genome sequence and gene structure annotation of Quinoa used in the study were downloaded from the NCBI Assembly database (Accession number: PI614886, Biosample accession: SAMN04338310) [37]. Using the published protein sequences of the *bHLH* family members of Arabidopsis [38] and rice [19] as reference sequences, local *BLastPs* were constructed by Tbtools, and homologous alignment was performed in the protein sequence library of Quinoa, with E value < 1E-10. In addition, the hidden Markov (HMM) model of characteristic conserved domain *bHLH* (PF00010) was downloaded from the Pfam database, and the "hmm search" program of Hmmer 3.0 software was used to conduct homology search in the quinoa protein sequence library, and E value < 1E-5 was set. The results of the above two methods are combined to obtain non-redundant candidate protein sequences. Using *NCBI-CDD* database (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi) and SMART database (http://smart.embl-heidelberg.de/) on the above candidate eggs the white sequences were searched for conserved domains to determine that the candidate proteins contained the *bHLH* domain. Finally, by removing incomplete sequences and redundant protein sequences, high confidence members of the quinoa *bHLH* gene family were identified.

2.2.4. Analysis of Physicochemical Properties of Proteins of Quinoa bHLH Gene Family Members

Based on the gene structure annotation file, the sequence length, chromosome localization, encoding

protein length and intron number of the *bHLH* gene of Quinoa were sorted out. Using Expasy online tools (https://web.expasy.org/compute_pi/) to calculate the quinoa *bHLH* proteins isoelectric point (pI) and molecular weight (Mw). Use of Cell PLoc 2.0 online tools (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/) to predict the subcellular localization of the members of the family. The hydrophilicity of quinoa *bHLH* protein coefficient calculated by ProtScale online tools (https://web.expasy.org/protscale/).

2.2.5. Gene Expression Analysis of Quinoa *bHLH* Gene Family Members

Expression data of CqbHLH gene in the root, stem and leaf of Quinoa (Accession number: SRP278144) were obtained from NCBI-SRA database and used for expression pattern analysis. Expression data of CqbHLH gene in salt-tolerant and salt-sensitive cultivars of quinoa (Accession number: SRP247883) were obtained from NCBI-SRA database and used for salt-tolerant expression pattern analysis. Tbtools was used to draw the heat map of expression patterns, and the expression level Log2 (FPKM) of the gene family members in different tissues was standardized. Euclidean-style distance method was used to cluster genes with similar expression patterns. The database and website used in this experiment are as follows:Chenopodium DB (https://w) Ww.cbrc.kaust.edu.sa/chenopodiumdb/) Arabidopsis thaliana database (https://www.arabidopsis.org/); National Rice Data Center (https://www.ricedata.cn/); NCBI (https://www.ncbi.nlm.nih.gov/). NCBI CDD (https://www.ncbi.nlm.nih.gov/cdd/); SMART (http://smart.embl-heidelb) erg.de).

2.2.6. Health Information Analysis and Website

The online website and software involved in the bioinformatics analysis of this experiment are as follows: Evolview (https://evolge nius.info//evolview-v2/); MEME (http://memesuite.org/tools/meme).*ExPASy*(https://web.ex pasy.org/); CELLO (http://cello.life.nctu.edu.tw/); Plant CARE(http://biOinformatics.PSB.Ugent.Be/webtools/plan tcare/HTML /); MEGA6.0; TB Tools; Excel 2016.

2.2.7. Analysis of Family Physicochemical Properties of Quinoa GST Genes

Use ExPASy online website (http://www.Expasy.ch/tools/pi_tool.html) for analysis and identification Amino acid number, molecular weight, isoelectric point, instability index, fat index and hydrophilicity of quinoa GST protein Isophysical and chemical properties. Use the CELLO (http://cello.life.nctu.edu.tw/) online prediction quinoa GST gene subcellular localization.

2.2.8. Analysis and Classification of Quinoa GST Gene Family Phylogenetic Tree

The MEGA6.0 software was used to compare the full-length sequences of GSTs in quinoa, Arabidopsis and rice the phylogenetic evolutionary tree was constructed by likelihood method, and the bootstrap value was set to 1000. With Evolview (https://evol genius.info//evolview-v2/). Visual analysis of phylogenetic trees on online website.

2.2.9. Phylogenetic Analysis, Gene Structure and Conserved Motifs of Quinoa *GSTs*

The gene structure information of Quinoa GSTs was extracted from the gene annotation information of Quinoa reference genome. Will identify. The protein sequence of quinoa GSTs was submitted to MEME (http://meme-suite.org/tools/meme). Row motif search and identification, the upper limit of the conserved domain is set to 10, and the conserved domain is allowed to repeat. The phylogeny, gene structure and conserved motifs of Quinoa were analyzed using TBtools.

2.2.10. Chromosome Distribution, Gene Replication, Molecular Evolution and Synchronization Analysis of *CqGST* Gene

The distribution of Quinoa GST gene on chromosomes was mapped using TBtools. Using the McScanx method. The replication modes of quinoa GST gene were divided into tandem replication and segmenial replication. Ka/Ks are calculated using TBtools toolkit. To understand collinearity between homologous GST genes in Quinoa and other species, the TBtools toolkit was used to analyze the data Visualization is performed.

2.2.11. Statistical Analysis

Microsoft Excel 2016 was used to sort out the original data, Origin 8.5 and Graphpad were adopted. Analysis of variance and significance test were performed by prism 8 software.

3. Results

3.1. Member Identification and Bioinformatics Analysis of CqbHLH Transcription Factor Family in Quinoa

A total of 250 *CqbHLH* transcription factor family members were identified from the whole genome of Quinoa using bioinformatics methods and subsequent screening verification, which were successively named *CqbHLH1-CqbHLH250* according to chromosome sequence. 237 *CqbHLH* transcription factors were randomly distributed on all 18 chromosomes. There are 6 to 27 transcription factors per chromosome, of which 13 transcription factors (*CqbHLH238-CqbHLH250*) are not anchored to the chromosome (Table 2). The amino acid length of the CqbHLH transcription factor family was between 83(*CqbHLH243*) and 962(*CqbHLH100*). The molecular weight ranged from the minimum 8.90kD (*CqbHLH243*) to the maximum104.51kD (*CqbHLH100*). Isoelectric points ranged from 4.49 (*CqbHLH112*) to 11.84 (*CqbHLH107*), of which 173 were below 7 (Table 2). Therefore, the length span of quinoa bHLH transcription factor family genes was large and correlated with molecular weight, and 157 transcription factors were located in the nucleus. The total average of hydrophilicity (*GRAVY*) of all *CqbHLH* transcription factors is positive, indicating that all *CqbHLH* transcription factors are likely to be soluble proteins, which is in line with functional positioning as transcription factors. By comparing the sequences of the *bHLH* gene family members of

Arabidopsis thaliana and the bHLH transcription factor family members of Quinoa, it can be seen that most of the bHLH transcription factors of quinoa have high homology with that of Arabidopsis thaliana (Table 2), which provides a reference for studying the biological functions of the *bHLH* transcription factor family of quinoa. In order to further study the evolutionary relationship of the *CqbHLH* transcription factor family, the phylogenetic tree of the family members was established using NJ method. The results showed that, as shown in Figure 2, *CqbHLH* transcription factor members could be divided into 21 branches, named A~U, and the number of branch members ranged from 2 (A and D) to 36(Q).

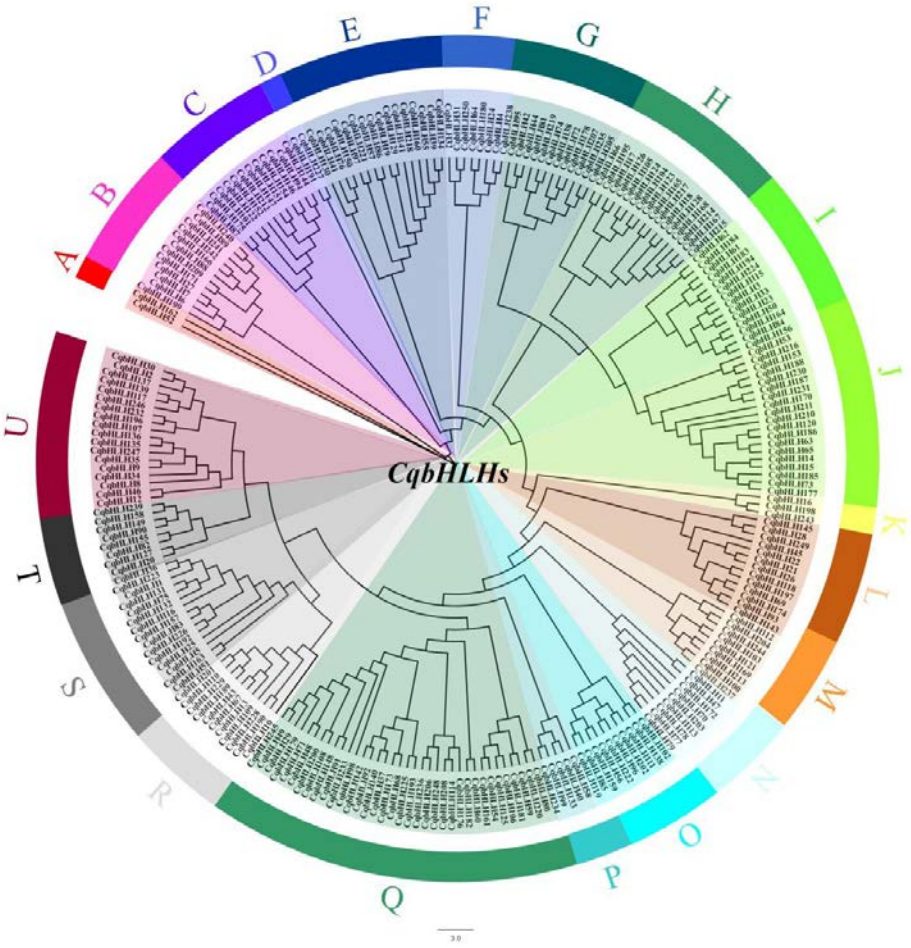

Figure 2. The unrooted phylogenetic tree of *bHLH* transcription factor family in quinoa

Table 2. The information of *CqbHLH* gene family in quinoa

Gene Name	Sequence ID	Number of Amino Acid	Genomic Location	Cellular Location	Molecular Weight (kD)	Isoelectric point pI	Arabidopsis homologous	E-value	Grand Average of Hydropathicity
CqbHLH1	AUR62021459-RA	276	Chr01:18329999-18334332	Nuclear	30.31	6.79	AT4G29100.1	7.83E-73	-0.739
CqbHLH2	AUR62031273-RA	209	Chr01:24171968-24172741	Nuclear	23.29	7.11	AT5G67060.1	3.04E-48	-0.738
CqbHLH3	AUR62014614-RA	244	Chr01:28148298-28154552	Nuclear	27.42	6.26	AT5G54680.1	2.04E-55	-0.439
CqbHLH4	AUR62014640-RA	141	Chr01:28520166-28520797	Nuclear	16.02	4.58	AT4G21330.1	1.32E-24	-0.351
CqbHLH5	AUR62014656-RA	470	Chr01:28728695-28733139	Extracellular	51.08	5.04	AT1G61660.1	1.11E-55	-0.676
CqbHLH6	AUR62014828-RA	158	Chr01:31260626-31262736	Extracellular	17.30	4.93	AT4G20970.1	1.40E-28	-0.479
CqbHLH7	AUR62014829-RA	179	Chr01:31277966-31280017	Extracellular	20.09	9.25	AT4G20970.1	5.44E-30	-0.376
CqbHLH8	AUR62038707-RA	349	Chr01:61672150-61673781	Extracellular	38.88	4.8	AT4G33880.1	8.24E-58	-0.667
CqbHLH9	AUR62038708-RA	182	Chr01:61677141-61679352	Nuclear	20.21	4.92	AT4G33880.1	1.82E-28	-0.886
CqbHLH10	AUR62040948-RA	203	Chr01:85943481-85944873	Extracellular	23.14	9.68	AT3G20640.1	2.08E-21	-0.546

Gene Name	Sequence ID	Number of Amino Acid	Genomic Location	Cellular Location	Molecular Weight (kD)	Isoelectric point pI	Arabidopsis homologous	E-value	Grand Average of Hydropathicity
CqbHLH11	AUR62009472-RA	310	Chr01:112000955-112007417	Extracellular	34.30	5.48	AT1G69010.1	1.30E-50	-0.754
CqbHLH12	AUR62023232-RA	319	Chr01:115484720-115486330	Extracellular	35.16	5.89	AT5G37800.1	1.33E-54	-0.719
CqbHLH13	AUR62004153-RA	323	Chr01:115999083-116004866	Nuclear	35.18	6.51	AT4G00050.1	1.34E-53	-0.741
CqbHLH14	AUR62004246-RA	286	Chr01:116894629-116897096	Nuclear	31.52	5.41	AT4G37850.1	2.27E-38	-0.406
CqbHLH15	AUR62004248-RA	278	Chr01:116962892-116965972	Nuclear	31.30	5.7	AT4G37850.1	4.42E-37	-0.552
CqbHLH16	AUR62004352-RA	830	Chr01:117857259-117866878	Extracellular	93.18	4.85	AT3G59970.3	5.09E-156	-0.248
CqbHLH17	AUR62004505-RA	571	Chr01:119528192-119530195	Nuclear	63.62	6.4	AT1G01260.1	0	-0.625
CqbHLH18	AUR62004545-RA	457	Chr01:120169303-120170676	Nuclear	50.64	6.32	AT4G00870.1	4.00E-69	-0.362
CqbHLH19	AUR62004564-RA	256	Chr01:120438699-120440991	Nuclear	28.91	7.19	AT4G01460.1	2.34E-37	-0.76
CqbHLH20	AUR62004612-RA	206	Chr01:121007276-121017934	Nuclear	22.40	8.53	AT4G09180.1	3.22E-65	-0.78
CqbHLH21	AUR62036595-RA	257	Chr01:124823875-124826160	Nuclear	28.89	5.73	AT2G40200.1	2.01E-30	-0.411
CqbHLH22	AUR62036594-RA	130	Chr01:124890305-124890697	Nuclear	14.41	5.38	AT1G68810.1	2.49E-19	0.024
CqbHLH23	AUR62012588-RA	187	Chr02:3592321-3593819	Nuclear	20.94	5.46	AT5G54680.1	1.51E-29	-0.334
CqbHLH24	AUR62012560-RA	140	Chr02:3892353-3892978	Nuclear	15.82	4.57	AT4G21330.1	1.14E-23	-0.262
CqbHLH25	AUR62012538-RA	353	Chr02:4121596-4124980	Extracellular	38.53	4.69	AT1G61660.1	4.89E-55	-0.646
CqbHLH26	AUR62031178-RA	257	Chr02:7454940-7457226	Nuclear	28.72	6.55	AT1G68810.1	1.93E-31	-0.332
CqbHLH27	AUR62017023-RA	171	Chr02:21105107-21107729	Extracellular	19.24	9.12	AT4G20970.1	1.50E-32	-0.458
CqbHLH28	AUR62029052-RA	531	Chr02:37006988-37012343	Nuclear	59.85	8.92	AT5G56960.1	1.41E-66	-0.475
CqbHLH29	AUR62027348-RA	543	Chr02:38318324-38328299	Extracellular	60.08	9.18	AT4G35790.1	3.25E-53	-0.262
CqbHLH30	AUR62027424-RA	141	Chr02:41096875-41097300	Nuclear	15.96	10.21	AT5G67060.1	7.43E-49	-0.59
CqbHLH31	AUR62027480-RA	787	Chr02:42649556-42666238	Extracellular	86.24	6.84	AT2G37690.1	0	-0.142
CqbHLH32	AUR62027481-RA	333	Chr02:42718823-42731109	Extracellular	36.34	8.55	AT5G57150.4	1.69E-43	-0.063
CqbHLH33	AUR62034025-RA	124	Chr02:44842669-44843505	Extracellular	13.85	8.71	AT5G54680.1	2.05E-11	-0.052
CqbHLH34	AUR62034089-RA	348	Chr02:47841944-47843587	Extracellular	38.88	4.8	AT4G33880.1	3.30E-58	-0.76
CqbHLH35	AUR62034090-RA	180	Chr02:47846958-47848942	Extracellular	20.37	5.24	AT2G14760.3	9.07E-24	-0.827
CqbHLH36	AUR62010948-RA	186	Chr02:54729172-54733181	Nuclear	20.28	7.67	AT4G29100.1	9.86E-73	-0.534
CqbHLH37	AUR62037403-RA	407	Chr03:19606481-19610661	Extracellular	43.82	6.31	AT1G10120.1	2.35E-79	-0.719
CqbHLH38	AUR62035700-RA	84	Chr03:21792605-21793053	Nuclear	9.84	10.07	AT5G65640.1	5.04E-28	-0.771
CqbHLH39	AUR62027824-RA	106	Chr03:69820795-69823571	Nuclear	11.95	9.1	AT1G22490.1	5.51E-22	-0.791
CqbHLH40	AUR62027851-RA	466	Chr03:70202336-70204330	Extracellular	50.93	5.02	AT5G50010.1	1.25E-19	-0.427
CqbHLH41	AUR62014519-RA	405	Chr03:74942399-74944322	Nuclear	44.20	6.34	AT2G31210.1	2.35E-74	-0.61
CqbHLH42	AUR62012336-RA	162	Chr03:77703910-77705575	Extracellular	17.95	4.98	AT2G40435.1	3.59E-59	-0.366
CqbHLH43	AUR62012272-RA	396	Chr03:78395388-78402093	Extracellular	44.13	8.95	AT3G09230.1	9.55E-14	-0.772

Gene Name	Sequence ID	Number of Amino Acid	Genomic Location	Cellular Location	Molecular Weight (kD)	Isoelectric point pI	Arabidopsis homologous	E-value	Grand Average of Hydropathicity
CqbHLH44	AUR62022980-RA	162	Chr04:1820784-1822436	Extracellular	17.94	4.98	AT2G40435.1	1.63E-59	-0.345
CqbHLH45	AUR62022910-RA	255	Chr04:2497297-2500553	Nuclear	29.06	9.08	AT2G40200.1	1.54E-45	-0.494
CqbHLH46	AUR62023185-RA	319	Chr04:7116819-7118412	Extracellular	35.10	6.06	AT5G37800.1	2.90E-54	-0.726
CqbHLH47	AUR62023150-RA	318	Chr04:7511299-7518302	Extracellular	35.26	5.5	AT1G69010.1	1.50E-53	-0.781
CqbHLH48	AUR62031607-RA	148	Chr04:39037121-39041660	Nuclear	16.38	6.51	AT1G59640.2	9.26E-36	-1.058
CqbHLH49	AUR62026945-RA	141	Chr04:43346566-43347840	Nuclear	15.37	9.3	AT1G59640.2	1.13E-30	-0.303
CqbHLH50	AUR62015636-RA	239	Chr05:3732235-3735278	Nuclear	26.55	7.61	AT5G54680.1	1.31E-105	-0.696
CqbHLH51	AUR62041153-RA	394	Chr05:5055075-5058071	Nuclear	41.99	5.38	AT3G20640.1	1.06E-64	-0.604
CqbHLH52	AUR62036174-RA	393	Chr05:11435332-11439658	Extracellular	44.13	5.64	AT5G53900.2	5.37E-158	-0.646
CqbHLH53	AUR62013988-RA	389	Chr05:30869823-30872911	Extracellular	43.36	5.4	AT1G60060.1	1.87E-127	-0.571
CqbHLH54	AUR62004950-RA	285	Chr05:67954966-67963038	Nuclear	30.02	5.92	AT4G02590.1	1.47E-79	-0.38
CqbHLH55	AUR62005105-RA	307	Chr05:70775476-70781608	Nuclear	34.68	6.02	AT1G72210.1	1.31E-68	-0.69
CqbHLH56	AUR62006825-RA	619	Chr05:75708868-75712350	Extracellular	67.33	6.37	AT1G09530.1	4.27E-37	-0.658
CqbHLH57	AUR62019337-RA	338	Chr05:79951489-79954715	Nuclear	37.34	5.01	AT5G53210.1	1.68E-98	-0.294
CqbHLH58	AUR62019260-RA	376	Chr05:80960304-80963038	Nuclear	42.04	6.07	AT2G42280.1	3.60E-62	-0.872
CqbHLH59	AUR62028866-RA	372	Chr06:3401420-3407507	Nuclear	39.94	5.11	AT2G24260.1	5.22E-68	-0.478
CqbHLH60	AUR62028856-RA	211	Chr06:3512960-3516427	Nuclear	23.68	8.78	AT1G68920.1	2.61E-33	-0.747
CqbHLH61	AUR62032867-RA	284	Chr06:6247053-6249642	Extracellular	32.23	5.42	AT4G37850.1	5.95E-32	-0.313
CqbHLH62	AUR62032868-RA	211	Chr06:6257256-6258391	Nuclear	23.85	8.88	AT4G37850.1	1.53E-35	-0.452
CqbHLH63	AUR62032870-RA	194	Chr06:6294705-6295377	Nuclear	21.45	8.98	AT4G37850.1	2.55E-47	-0.321
CqbHLH64	AUR62014049-RA	441	Chr06:7530704-7533192	Extracellular	48.97	5.81	AT1G10610.1	1.98E-62	-0.397
CqbHLH65	AUR62014013-RA	333	Chr06:8026773-8029425	Extracellular	37.21	5.05	AT4G37850.1	1.52E-48	-0.454
CqbHLH66	AUR62003365-RA	259	Chr06:8605318-8611639	Nuclear	28.97	5.75	AT3G19860.2	8.35E-82	-0.853
CqbHLH67	AUR62003052-RA	343	Chr06:12903814-12914751	Nuclear	37.95	5.76	AT1G69010.1	1.72E-94	-0.881
CqbHLH68	AUR62003039-RA	567	Chr06:13141068-13150299	Nuclear	61.54	5.2	AT1G68920.1	5.62E-114	-0.695
CqbHLH69	AUR62002923-RA	302	Chr06:15813766-15815843	Nuclear	33.03	5.85	AT1G68810.1	9.71E-102	-0.634
CqbHLH70	AUR62002849-RA	164	Chr06:18643884-18644466	Extracellular	18.06	8.93	AT5G50915.1	1.38E-11	-0.748
CqbHLH71	AUR62026495-RA	280	Chr06:66087485-66091048	Nuclear	30.37	7.7	AT1G59640.2	1.80E-65	-0.628
CqbHLH72	AUR62026493-RA	313	Chr06:66178472-66180921	Extracellular	33.81	4.75	AT3G26744.1	7.16E-26	-0.457
CqbHLH73	AUR62026381-RA	201	Chr06:73106200-73110076	Extracellular	22.74	5.32	AT2G22770.1	8.61E-19	-1.015
CqbHLH74	AUR62038620-RA	478	Chr07:14175496-14178956	Nuclear	52.99	5.55	AT4G09820.1	9.23E-56	-0.514
CqbHLH75	AUR62043180-RA	252	Chr07:15296711-15299529	Nuclear	27.49	6.07	AT1G69010.1	1.22E-39	-0.433
CqbHLH76	AUR62032672-RA	353	Chr07:25984126-25986127	Nuclear	39.22	6.47	AT2G31210.1	2.01E-71	-0.474

Gene Name	Sequence ID	Number of Amino Acid	Genomic Location	Cellular Location	Molecular Weight (kD)	Isoelectric point pI	Arabidopsis homologous	E-value	Grand Average of Hydropathicity
CqbHLH77	AUR62034672-RA	311	Chr07:28935564-28938833	Nuclear	34.05	5.86	AT1G69010.1	1.04E-32	-0.458
CqbHLH78	AUR62034673-RA	403	Chr07:28947818-28955089	Nuclear	43.68	5.86	AT1G69010.1	3.94E-35	-0.572
CqbHLH79	AUR62034677-RA	251	Chr07:29068554-29070900	Nuclear	27.90	5.68	AT1G69010.1	8.27E-28	-0.529
CqbHLH80	AUR62036801-RA	312	Chr07:52860622-52865295	Extracellular	34.81	4.86	AT5G64980.1	3.88E-25	-0.446
CqbHLH81	AUR62036822-RA	360	Chr07:53825937-53829358	Nuclear	40.19	5.23	AT4G09820.1	2.81E-44	-0.512
CqbHLH82	AUR62038167-RA	191	Chr07:55101317-55109310	Nuclear	21.36	6.66	AT2G43140.2	1.17E-54	-0.621
CqbHLH83	AUR62002393-RA	148	Chr07:63418406-63425928	Nuclear	16.78	9.27	AT2G31730.1	1.26E-30	-0.139
CqbHLH84	AUR62002110-RA	244	Chr07:67080347-67083914	Nuclear	27.79	5.59	AT4G14410.1	1.00E-58	-0.88
CqbHLH85	AUR62001913-RA	358	Chr07:69105078-69108192	Extracellular	38.82	5.66	AT4G36930.1	5.99E-53	-0.564
CqbHLH86	AUR62006151-RA	352	Chr07:72124882-72127807	Nuclear	38.35	4.89	AT5G53210.1	8.56E-101	-0.244
CqbHLH87	AUR62016228-RA	245	Chr07:75582592-75583329	Nuclear	27.09	11.49	AT3G17100.1	9.66E-47	-0.663
CqbHLH88	AUR62016073-RA	168	Chr07:77158330-77159433	Extracellular	19.12	8.38	AT5G51790.1	4.09E-22	-0.663
CqbHLH89	AUR62016072-RA	197	Chr07:77176470-77177799	Extracellular	22.02	4.85	AT1G12540.1	1.39E-35	-0.572
CqbHLH90	AUR62025963-RA	358	Chr07:79196982-79202501	Extracellular	39.66	9.01	AT2G42280.1	3.67E-55	-0.785
CqbHLH91	AUR62025981-RA	281	Chr07:79552009-79557215	Nuclear	31.43	5.68	AT3G57800.1	3.77E-48	-0.675
CqbHLH92	AUR62018515-RA	375	Chr07:85310416-85312916	Nuclear	41.58	6.1	AT3G07340.1	3.25E-75	-0.77
CqbHLH93	AUR62018364-RA	197	Chr07:87631344-87634341	Nuclear	22.08	6.85	AT2G41130.1	1.85E-37	-0.518
CqbHLH94	AUR62018341-RA	421	Chr07:88049542-88055844	Extracellular	47.97	5.59	AT3G56970.1	2.11E-45	-0.318
CqbHLH95	AUR62006335-RA	152	Chr07:89627497-89631179	Extracellular	16.53	5.13	AT2G40435.1	7.27E-56	-0.261
CqbHLH96	AUR62031810-RA	305	Chr07:91968257-91974953	Extracellular	33.01	5.23	AT2G20180.2	1.10E-43	-0.905
CqbHLH97	AUR62031845-RA	184	Chr07:92459227-92461578	Nuclear	21.14	9.05	AT3G06120.1	1.02E-57	-0.316
CqbHLH98	AUR62001169-RA	541	Chr07:96839783-96842563	Nuclear	59.03	7.06	AT3G07340.1	2.23E-97	-0.757
CqbHLH99	AUR62001482-RA	241	Chr07:101073151-101076296	Nuclear	27.08	5.5	AT3G47640.1	3.28E-48	-0.724
CqbHLH100	AUR62001552-RA	962	Chr07:101914788-101921824	Nuclear	104.51	6.17	AT2G27230.1	4.62E-116	-0.161
CqbHLH101	AUR62011773-RA	186	Chr08:970531-973108	Nuclear	20.78	9.44	AT1G71200.1	9.67E-28	-0.413
CqbHLH102	AUR62011824-RA	277	Chr08:1398410-1402947	Extracellular	31.00	9.14	AT4G02090.1	3.95E-24	-0.524
CqbHLH103	AUR62011893-RA	705	Chr08:2250226-2257348	Nuclear	78.22	6.11	AT2G31280.3	6.97E-152	-0.393
CqbHLH104	AUR62011907-RA	510	Chr08:2532315-2533847	Nuclear	55.89	6.01	AT4G16430.1	5.08E-145	-0.376
CqbHLH105	AUR62011908-RA	502	Chr08:2536290-2537798	Nuclear	55.63	6.24	AT4G16430.1	2.22E-145	-0.402
CqbHLH106	AUR62011948-RA	251	Chr08:2895972-2899594	Nuclear	27.24	9.07	AT4G02590.1	8.57E-113	-0.469
CqbHLH107	AUR62028696-RA	160	Chr08:4443729-4445373	Extracellular	18.59	11.84	AT2G43060.1	4.95E-18	-0.648
CqbHLH108	AUR62021777-RA	314	Chr08:7863144-7868070	Nuclear	33.34	6.02	AT5G62610.1	5.88E-73	-0.521
CqbHLH109	AUR62021718-RA	92	Chr08:8643519-8644695	Nuclear	10.52	9.61	AT3G47710.1	1.05E-43	-0.714

Gene Name	Sequence ID	Number of Amino Acid	Genomic Location	Cellular Location	Molecular Weight (kD)	Isoelectric point pI	Arabidopsis homologous	E-value	Grand Average of Hydropathicity
CqbHLH110	AUR62021715-RA	87	Chr08:8713054-8713739	Nuclear	10.08	6.09	AT1G26945.1	2.03E-19	-0.385
CqbHLH111	AUR62021604-RA	362	Chr08:10429556-10432292	Extracellular	39.37	9.08	AT4G00050.1	2.08E-41	-0.504
CqbHLH112	AUR62032125-RA	719	Chr08:12273831-12281493	Extracellular	80.87	4.49	AT2G27230.1	8.14E-45	-0.357
CqbHLH113	AUR62032143-RA	516	Chr08:12759378-12767674	Nuclear	55.86	6.15	AT5G08130.5	1.07E-75	-0.565
CqbHLH114	AUR62029874-RA	208	Chr08:35247070-35248492	Nuclear	23.16	8.82	AT1G73830.1	5.53E-56	-0.875
CqbHLH115	AUR62042938-RA	182	Chr08:40354830-40356238	Extracellular	20.34	9.18	AT5G54680.1	3.31E-31	-0.322
CqbHLH116	AUR62007612-RA	274	Chr09:1019568-1024817	Extracellular	29.54	7.67	AT1G27660.1	4.21E-20	-0.454
CqbHLH117	AUR62003602-RA	154	Chr09:4056201-4058715	Nuclear	17.52	11.27	AT4G00120.1	9.90E-46	-0.479
CqbHLH118	AUR62003654-RA	249	Chr09:4801701-4804740	Nuclear	27.01	9.53	AT1G68810.1	2.57E-35	-0.36
CqbHLH119	AUR62003896-RA	384	Chr09:7458256-7461950	Extracellular	41.46	5.13	AT5G67110.1	8.97E-16	-0.526
CqbHLH120	AUR62004073-RA	163	Chr09:9759472-9761127	Nuclear	17.89	5.46	AT4G37850.1	7.34E-25	-0.235
CqbHLH121	AUR62024734-RA	705	Chr09:10554663-10561121	Nuclear	78.10	5.79	AT2G31280.3	4.21E-152	-0.365
CqbHLH122	AUR62024616-RA	187	Chr09:12005222-12007608	Extracellular	21.08	9.26	AT1G71200.1	8.72E-27	-0.46
CqbHLH123	AUR62032220-RA	510	Chr09:13097728-13099260	Nuclear	55.80	5.92	AT4G16430.1	5.78E-147	-0.395
CqbHLH124	AUR62032221-RA	502	Chr09:13100968-13102476	Nuclear	55.58	6.14	AT4G16430.1	7.07E-147	-0.411
CqbHLH125	AUR62032262-RA	242	Chr09:13522871-13526432	Nuclear	26.36	8.74	AT4G02590.1	8.40E-113	-0.441
CqbHLH126	AUR62023449-RA	582	Chr10:2408456-2410453	Nuclear	64.81	6.57	AT1G01260.1	0	-0.595
CqbHLH127	AUR62022652-RA	282	Chr10:5176778-5183813	Nuclear	29.96	6.54	AT4G09180.1	2.55E-78	-0.689
CqbHLH128	AUR62022659-RA	457	Chr10:5375814-5377187	Nuclear	50.54	6.19	AT4G00870.1	3.27E-69	-0.319
CqbHLH129	AUR62022678-RA	256	Chr10:5613683-5615383	Nuclear	28.69	9.6	AT4G01460.1	1.45E-72	-0.561
CqbHLH130	AUR62038015-RA	405	Chr10:6774889-6776779	Nuclear	44.09	6.46	AT2G31210.1	8.30E-74	-0.608
CqbHLH131	AUR62038049-RA	193	Chr10:7048708-7051490	Nuclear	21.23	5.35	AT1G27660.1	4.07E-41	-0.514
CqbHLH132	AUR62034506-RA	245	Chr10:7292290-7295092	Nuclear	26.95	5.81	AT1G27660.1	1.73E-50	-0.527
CqbHLH133	AUR62019010-RA	488	Chr10:10262810-10265467	Extracellular	53.84	5.26	AT5G50010.1	3.24E-20	-0.283
CqbHLH134	AUR62018984-RA	120	Chr10:10644877-10647830	Nuclear	13.50	7.75	AT1G22490.1	8.47E-22	-0.739
CqbHLH135	AUR62038801-RA	306	Chr10:34854781-34856099	Extracellular	34.43	5.28	AT4G33880.1	2.36E-51	-0.565
CqbHLH136	AUR62038798-RA	346	Chr10:35285606-35287046	Extracellular	38.92	5.13	AT4G33880.1	8.92E-50	-0.556
CqbHLH137	AUR62043211-RA	138	Chr10:39691439-39691855	Nuclear	15.55	10.49	AT3G50330.1	6.36E-48	-0.378
CqbHLH138	AUR62035463-RA	210	Chr10:41390858-41393904	Nuclear	23.27	8.42	AT4G36930.1	1.40E-38	-0.707
CqbHLH139	AUR62035415-RA	262	Chr10:43532201-43555002	Nuclear	29.95	8.33	AT5G65640.1	3.58E-76	-0.579
CqbHLH140	AUR62034991-RA	409	Chr10:54445665-54449798	Extracellular	44.17	6.27	AT1G10120.1	3.92E-81	-0.737
CqbHLH141	AUR62020815-RA	468	Chr11:126562-128329	Nuclear	51.39	5.6	AT3G24140.1	3.13E-123	-0.567
CqbHLH142	AUR62020904-RA	426	Chr11:1195234-1203761	Nuclear	47.63	5.4	AT3G07340.1	7.30E-67	-0.817

Gene Name	Sequence ID	Number of Amino Acid	Genomic Location	Cellular Location	Molecular Weight (kD)	Isoelectric point pI	Arabidopsis homologous	E-value	Grand Average of Hydropathicity
CqbHLH143	AUR62030298-RA	182	Chr11:3550377-3553398	Nuclear	20.37	8.73	AT2G41130.1	9.05E-36	-0.484
CqbHLH144	AUR62030326-RA	271	Chr11:4277683-4279104	Nuclear	30.75	5.68	AT3G56970.1	1.28E-48	-0.421
CqbHLH145	AUR62030329-RA	151	Chr11:4383637-4387400	Extracellular	17.31	4.82	AT3G56980.1	5.72E-11	-0.346
CqbHLH146	AUR62030330-RA	404	Chr11:4393724-4401706	Extracellular	45.99	5.6	AT3G56970.1	2.88E-42	-0.642
CqbHLH147	AUR62030331-RA	138	Chr11:4409276-4410650	Extracellular	15.92	6.5	AT3G56970.1	2.28E-12	-0.636
CqbHLH148	AUR62024373-RA	339	Chr11:7996765-8001917	Nuclear	37.32	5.38	AT3G57800.1	9.96E-56	-0.606
CqbHLH149	AUR62024365-RA	271	Chr11:8119809-8123354	Nuclear	29.72	9.08	AT2G42280.1	1.51E-58	-0.706
CqbHLH150	AUR62035295-RA	330	Chr11:13385739-13388321	Nuclear	36.10	4.99	AT2G31210.1	3.00E-55	-0.264
CqbHLH151	AUR62031987-RA	138	Chr11:29758063-29759504	Extracellular	15.90	6.28	AT3G56970.1	6.32E-12	-0.651
CqbHLH152	AUR62031986-RA	275	Chr11:29773137-29775369	Nuclear	31.26	5.83	AT3G56970.1	5.85E-47	-0.516
CqbHLH153	AUR62031985-RA	472	Chr11:29782837-29807428	Extracellular	51.89	4.82	AT3G56970.1	8.49E-13	-0.831
CqbHLH154	AUR62031984-RA	274	Chr11:29818045-29821357	Extracellular	31.63	6.12	AT3G56970.1	1.52E-43	-0.451
CqbHLH155	AUR62038082-RA	258	Chr11:58104915-58113722	Nuclear	28.45	7.78	AT1G05805.1	3.87E-62	-0.727
CqbHLH156	AUR62028228-RA	244	Chr11:67530295-67533797	Nuclear	27.83	5.59	AT4G14410.1	2.80E-58	-0.876
CqbHLH157	AUR62023796-RA	135	Chr11:71646590-71654129	Nuclear	15.01	8.55	AT2G31730.1	3.63E-54	-0.364
CqbHLH158	AUR62000088-RA	297	Chr12:1284942-1287137	Nuclear	32.93	7.8	AT2G42280.1	2.23E-61	-0.78
CqbHLH159	AUR62000483-RA	778	Chr12:5336725-5348877	Extracellular	84.68	6.54	AT1G09530.1	7.67E-52	-0.481
CqbHLH160	AUR62000871-RA	211	Chr12:10419006-10421055	Nuclear	23.49	5.98	AT1G72210.1	2.08E-63	-0.512
CqbHLH161	AUR62001007-RA	299	Chr12:12750252-12757715	Nuclear	31.23	6.45	AT4G02590.1	2.25E-85	-0.312
CqbHLH162	AUR62019456-RA	390	Chr12:49297755-49302310	Extracellular	43.80	5.7	AT5G53900.2	4.18E-156	-0.642
CqbHLH163	AUR62029280-RA	265	Chr12:52433330-52436356	Nuclear	28.49	6.84	AT3G20640.1	5.67E-66	-0.624
CqbHLH164	AUR62029235-RA	239	Chr12:53198025-53201320	Nuclear	26.57	7.61	AT5G54680.1	2.82E-106	-0.741
CqbHLH165	AUR62007294-RA	654	Chr13:2650505-2652469	Nuclear	71.40	5.68	AT1G32640.1	9.26E-175	-0.606
CqbHLH166	AUR62007178-RA	172	Chr13:3921862-3922546	Extracellular	19.68	5.07	AT1G12540.1	1.86E-16	-0.615
CqbHLH167	AUR62007060-RA	328	Chr13:5583358-5584344	Extracellular	36.29	7.79	AT4G00870.1	1.17E-29	-0.28
CqbHLH168	AUR62007059-RA	353	Chr13:5587781-5588842	Extracellular	39.54	5.6	AT5G46760.1	3.07E-37	-0.097
CqbHLH169	AUR62007032-RA	681	Chr13:5852485-5856769	Extracellular	74.82	5.25	AT1G64625.1	2.17E-93	-0.261
CqbHLH170	AUR62010677-RA	310	Chr13:9264026-9265820	Extracellular	34.72	5.02	AT4G37850.1	3.69E-46	-0.445
CqbHLH171	AUR62010555-RA	255	Chr13:11313385-11315474	Extracellular	29.00	6.12	AT5G51790.1	9.98E-27	-0.645
CqbHLH172	AUR62007865-RA	412	Chr14:3910916-3921726	Nuclear	45.41	5.98	AT1G69010.1	1.05E-67	-0.664
CqbHLH173	AUR62007846-RA	529	Chr14:4203634-4209016	Nuclear	57.00	5.63	AT1G68920.1	4.10E-114	-0.66
CqbHLH174	AUR62035285-RA	344	Chr14:9063449-9065311	Nuclear	37.73	6.09	AT1G68810.1	4.95E-104	-0.613
CqbHLH175	AUR62037297-RA	164	Chr14:10107378-10111778	Nuclear	18.23	11.06	AT3G17100.1	2.86E-36	-0.466

Gene Name	Sequence ID	Number of Amino Acid	Genomic Location	Cellular Location	Molecular Weight (kD)	Isoelectric point pI	Arabidopsis homologous	E-value	Grand Average of Hydropathicity
CqbHLH176	AUR62037320-RA	260	Chr14:10754482-10756288	Nuclear	29.05	6.09	AT1G25330.1	6.31E-55	-0.546
CqbHLH177	AUR62026248-RA	247	Chr14:23404668-23405578	Nuclear	27.31	8.92	AT2G22750.2	6.77E-33	-0.457
CqbHLH178	AUR62026299-RA	464	Chr14:25645862-25654440	Nuclear	50.46	6.03	AT3G26744.1	9.81E-74	-0.155
CqbHLH179	AUR62026301-RA	351	Chr14:25693939-25707915	Nuclear	38.87	8.69	AT1G59640.2	2.35E-70	-0.521
CqbHLH180	AUR62005387-RA	431	Chr14:50695436-50697838	Extracellular	48.03	5.87	AT1G10610.1	6.32E-64	-0.41
CqbHLH181	AUR62005693-RA	316	Chr14:55587380-55593208	Nuclear	33.72	5.24	AT2G24260.1	2.14E-70	-0.776
CqbHLH182	AUR62005699-RA	301	Chr14:55724056-55726776	Extracellular	33.83	6.12	AT1G68920.1	1.73E-31	-0.697
CqbHLH183	AUR62005873-RA	268	Chr14:58421747-58423566	Nuclear	30.36	5.41	AT4G37850.1	1.70E-30	-0.312
CqbHLH184	AUR62005874-RA	313	Chr14:58429015-58449887	Extracellular	35.55	5.89	AT2G22760.1	5.36E-37	-0.419
CqbHLH185	AUR62005875-RA	278	Chr14:58458662-58459798	Nuclear	31.09	8.84	AT4G37850.1	3.32E-47	-0.323
CqbHLH186	AUR62005877-RA	321	Chr14:58500042-58502159	Extracellular	35.34	4.85	AT4G37850.1	4.83E-43	-0.458
CqbHLH187	AUR62027063-RA	329	Chr15:920400-922688	Nuclear	36.19	8.1	AT4G37850.1	5.41E-48	-0.233
CqbHLH188	AUR62027064-RA	319	Chr15:932794-935356	Nuclear	36.19	6.09	AT4G37850.1	6.41E-53	-0.428
CqbHLH189	AUR62010034-RA	353	Chr15:4509834-4511502	Nuclear	39.98	5.08	AT2G16910.1	3.23E-72	-0.694
CqbHLH190	AUR62010097-RA	280	Chr15:5286250-5289703	Nuclear	30.61	5.9	AT3G19860.2	1.14E-74	-0.893
CqbHLH191	AUR62010242-RA	255	Chr15:6848438-6852765	Nuclear	28.52	6.99	AT1G49770.1	3.25E-36	-0.371
CqbHLH192	AUR62010325-RA	242	Chr15:7657350-7659950	Nuclear	26.06	8.38	AT3G19500.1	3.67E-40	-0.519
CqbHLH193	AUR62011307-RA	459	Chr15:9288006-9290834	Extracellular	50.64	5.72	AT4G34530.1	5.16E-62	-0.515
CqbHLH194	AUR62011472-RA	203	Chr15:10926407-10928777	Nuclear	23.08	8.74	AT4G37850.1	4.19E-31	-0.414
CqbHLH195	AUR62029468-RA	323	Chr15:15184574-15189250	Nuclear	35.96	5.74	AT3G19860.2	1.85E-80	-0.99
CqbHLH196	AUR62031976-RA	306	Chr15:18482340-18491734	Nuclear	33.61	5.12	AT3G21330.1	3.45E-57	-0.77
CqbHLH197	AUR62017962-RA	249	Chr15:21162653-21164684	Nuclear	27.01	9.53	AT1G68810.1	2.57E-35	-0.36
CqbHLH198	AUR62014973-RA	244	Chr15:55874186-55885747	Nuclear	27.45	9.37	AT2G28160.1	1.31E-79	-0.487
CqbHLH199	AUR62008225-RA	206	Chr16:1827538-1835992	Extracellular	23.09	5.8	AT4G20970.1	1.44E-13	-0.666
CqbHLH200	AUR62008318-RA	242	Chr16:3057643-3061591	Nuclear	25.64	5.74	AT5G62610.1	1.41E-72	-0.634
CqbHLH201	AUR62008383-RA	87	Chr16:3897832-3898446	Nuclear	10.18	7.75	AT1G26945.1	2.73E-19	-0.603
CqbHLH202	AUR62008486-RA	352	Chr16:5191581-5194392	Extracellular	37.73	6.27	AT3G62090.2	8.82E-11	-0.413
CqbHLH203	AUR62039664-RA	566	Chr16:6746155-6753952	Nuclear	61.73	6.46	AT5G08130.5	7.07E-89	-0.616
CqbHLH204	AUR62039677-RA	754	Chr16:7091310-7106955	Extracellular	85.33	4.5	AT2G27230.1	4.49E-44	-0.363
CqbHLH205	AUR62034764-RA	354	Chr16:9898480-9901986	Nuclear	38.32	8.97	AT3G26744.1	1.24E-115	-0.489
CqbHLH206	AUR62034777-RA	339	Chr16:10198339-10200944	Extracellular	37.41	6.85	AT5G50915.1	3.09E-57	-0.627
CqbHLH207	AUR62041687-RA	282	Chr16:12722481-12725629	Nuclear	30.68	8.32	AT3G26744.1	7.47E-109	-0.569
CqbHLH208	AUR62025839-RA	196	Chr16:21556896-21560515	Nuclear	21.11	9.41	AT1G73830.1	4.23E-44	-0.912

Gene Name	Sequence ID	Number of Amino Acid	Genomic Location	Cellular Location	Molecular Weight (kD)	Isoelectric point pI	Arabidopsis homologous	E-value	Grand Average of Hydropathicity
CqbHLH209	AUR62017336-RA	266	Chr16:66830740-66832518	Nuclear	30.40	6.41	AT4G25410.1	3.16E-27	-0.64
CqbHLH210	AUR62017206-RA	292	Chr16:68524854-68527010	Extracellular	32.90	5.35	AT4G37850.1	1.26E-41	-0.409
CqbHLH211	AUR62017204-RA	289	Chr16:68549573-68551812	Nuclear	32.38	5.65	AT4G37850.1	2.05E-44	-0.413
CqbHLH212	AUR62017199-RA	283	Chr16:68620325-68621907	Extracellular	32.31	7.16	AT3G56970.1	1.31E-36	-0.455
CqbHLH213	AUR62019896-RA	665	Chr16:72627022-72631075	Nuclear	72.92	5.1	AT1G64625.1	4.13E-99	-0.291
CqbHLH214	AUR62019922-RA	352	Chr16:72860313-72861371	Extracellular	39.88	6.61	AT1G32640.1	9.95E-34	-0.229
CqbHLH215	AUR62019923-RA	394	Chr16:72888587-72889838	Extracellular	44.23	7.12	AT4G00870.1	3.61E-41	-0.333
CqbHLH216	AUR62018826-RA	234	Chr16:74621689-74626406	Extracellular	25.70	4.84	AT1G47530.1	1.46E-48	0.065
CqbHLH217	AUR62018713-RA	631	Chr16:75941569-75943546	Nuclear	68.67	5.74	AT1G32640.1	3.70E-175	-0.581
CqbHLH218	AUR62018701-RA	299	Chr16:76133234-76135409	Nuclear	33.57	8.29	AT1G22490.1	2.01E-69	-0.328
CqbHLH219	AUR62043163-RA	549	Chr17:11636780-11646281	Nuclear	60.81	5.19	AT4G09820.1	2.14E-117	-0.486
CqbHLH220	AUR62039903-RA	313	Chr17:28819813-28823393	Extracellular	34.91	4.77	AT5G64980.1	3.44E-26	-0.424
CqbHLH221	AUR62008866-RA	161	Chr17:67853667-67856879	Nuclear	18.58	9.48	AT3G06120.1	5.67E-43	-0.243
CqbHLH222	AUR62008898-RA	383	Chr17:68299633-68306807	Extracellular	41.71	5.86	AT2G20180.2	1.94E-42	-0.543
CqbHLH223	AUR62016510-RA	105	Chr17:69736075-69738594	Nuclear	11.34	9.57	AT4G29100.1	1.88E-36	-0.556
CqbHLH224	AUR62016579-RA	394	Chr17:72705224-72711265	Extracellular	44.54	5.71	AT2G22770.1	9.76E-12	-0.331
CqbHLH225	AUR62016750-RA	417	Chr17:74437946-74440934	Extracellular	46.08	5.62	AT4G34530.1	8.53E-58	-0.428
CqbHLH226	AUR62039063-RA	244	Chr17:76394139-76396907	Nuclear	26.23	8.38	AT3G19500.1	4.20E-40	-0.54
CqbHLH227	AUR62013224-RA	171	Chr17:77612637-77614595	Nuclear	19.01	9.58	AT1G49770.1	4.38E-29	-0.127
CqbHLH228	AUR62013093-RA	280	Chr17:78859944-78863214	Nuclear	30.47	5.73	AT3G19860.2	5.22E-75	-0.869
CqbHLH229	AUR62013022-RA	364	Chr17:79661324-79662891	Extracellular	41.48	5.23	AT2G16910.1	6.46E-69	-0.63
CqbHLH230	AUR62030905-RA	155	Chr17:82569496-82570282	Nuclear	17.51	5.42	AT4G37850.1	3.44E-32	-0.052
CqbHLH231	AUR62030904-RA	364	Chr17:82590039-82591662	Extracellular	40.74	6.19	AT4G37850.1	4.69E-50	-0.542
CqbHLH232	AUR62022271-RA	322	Chr18:19401657-19405092	Nuclear	36.00	5.48	AT3G21330.1	3.67E-63	-0.811
CqbHLH233	AUR62022154-RA	390	Chr18:21412961-21414963	Nuclear	41.99	6.31	AT3G07340.1	4.25E-37	-0.792
CqbHLH234	AUR62009884-RA	240	Chr18:25582977-25586751	Nuclear	26.89	5.39	AT3G47640.1	3.04E-46	-0.683
CqbHLH235	AUR62009524-RA	329	Chr18:29562680-29566543	Nuclear	35.42	6.09	AT3G26744.1	2.14E-95	-0.483
CqbHLH236	AUR62009509-RA	342	Chr18:29870193-29877684	Extracellular	37.66	6.15	AT5G50915.1	2.88E-57	-0.611
CqbHLH237	AUR62020231-RA	907	Chr18:32215050-32222447	Nuclear	98.16	6.09	AT2G27230.1	2.48E-120	-0.228
CqbHLH238	AUR62044052-RA	141	Chr00:23256986-23257611	Nuclear	16.02	4.58	AT4G21330.1	1.32E-24	-0.351
CqbHLH239	AUR62043582-RA	222	Chr00:27685189-27687093	Nuclear	24.95	6.11	AT2G42280.1	1.69E-58	-0.857
CqbHLH240	AUR62021280-RA	150	Chr00:36438326-36440203	Extracellular	16.63	4.66	AT1G12540.1	1.63E-12	-0.728
CqbHLH241	AUR62021281-RA	127	Chr00:36445607-36446218	Nuclear	14.27	4.75	AT4G25410.1	2.74E-19	-0.57

Gene Name	Sequence ID	Number of Amino Acid	Genomic Location	Cellular Location	Molecular Weight (kD)	Isoelectric point pI	Arabidopsis homologous	E-value	Grand Average of Hydropathicity
CqbHLH242	AUR62021282-RA	198	Chr00:36468477-36471168	Extracellular	22.60	8.85	AT5G51790.1	8.02E-21	-0.638
CqbHLH243	AUR62029160-RA	83	Chr00:40716687-40718076	Nuclear	8.90	7.81	AT2G28160.1	4.99E-14	-0.117
CqbHLH244	AUR62044349-RA	705	Chr00:43294219-43304209	Extracellular	79.92	4.57	AT2G27230.1	3.34E-44	-0.402
CqbHLH245	AUR62042871-RA	311	Chr00:69188097-69191660	Extracellular	34.94	6.47	AT4G35790.1	1.07E-53	-0.239
CqbHLH246	AUR62043707-RA	158	Chr00:83604600-83605073	Nuclear	18.55	6.98	AT4G00120.1	8.83E-45	-0.302
CqbHLH247	AUR62042990-RA	258	Chr00:98418369-98419998	Nuclear	28.89	4.6	AT4G33880.1	1.88E-57	-0.62
CqbHLH248	AUR62044545-RA	339	Chr00:166171863-166174471	Extracellular	37.44	6.02	AT5G50915.1	1.91E-56	-0.618
CqbHLH249	AUR62033947-RA	532	Chr00:184685685-184692352	Extracellular	60.04	8.9	AT5G56960.1	5.61E-66	-0.504
CqbHLH250	AUR62033996-RA	249	Chr00:186161265-186168312	Nuclear	28.72	6.12	AT5G57150.4	6.59E-60	-0.552

3.3. Identification and Naming of Quinoa GST Gene Family

The 114 members of Quinoa GST gene family were identified by MEGA6.0. The maximum likelihood method was used to construct phylogenetic tree for 228 protein sequences from Quinoa, Arabidopsis and rice (Figure 3). The results showed that 114 members of Quinoa GST gene family were divided into seven groups, including GSTZ, GSTT, TCHQD. DHAR, GSTF, GSTL, GSTU. Among them, GSTU has 68 family members, accounting for 59.6%, which is quinoa. The largest branch of the GST gene family. GSTF, GSTL, GSTZ, GSTT, DHAR and TCHQD each contain 2. Three, six, six, five, four, two members of the family. They were named C for their classification and distribution on chromosomes, *qGSTU1~68*, *CqGSTF1~23*, *CqGSTL1~6*, *CqGSTZ1~Z6*, *CqGSTT1~T5*, *CqDHAR1~4*.

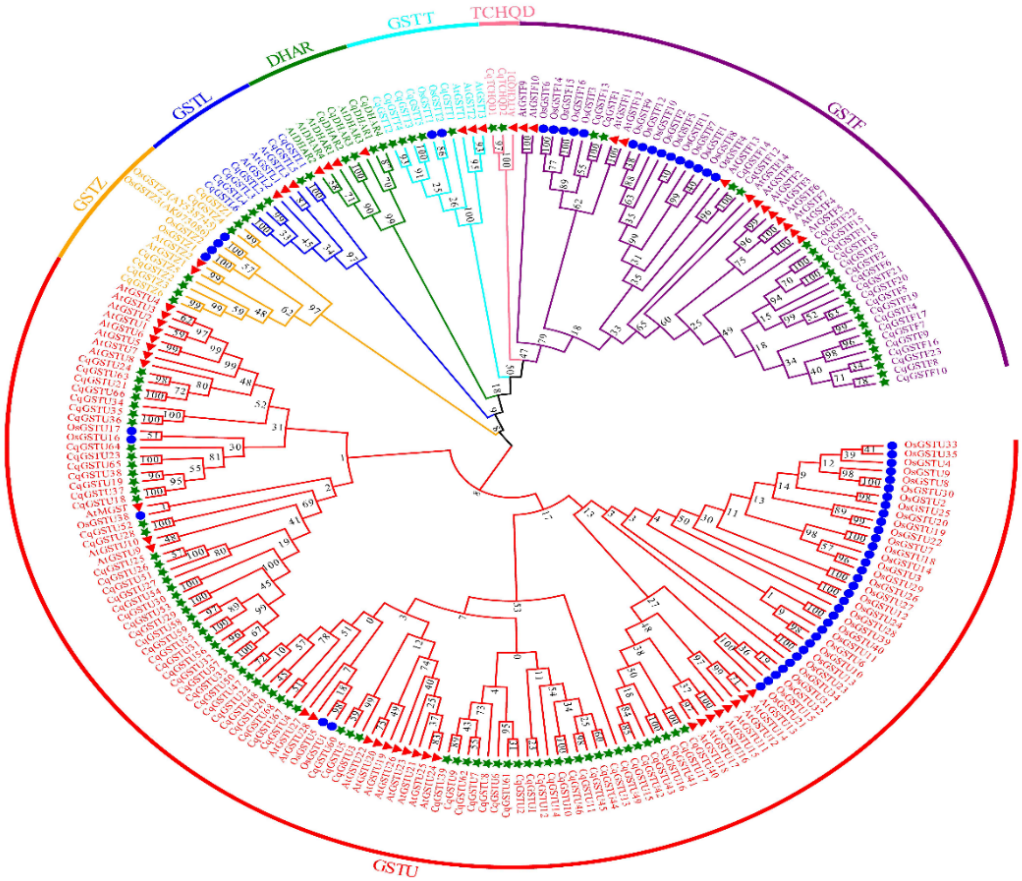

Figure 3. Phylogenetic analyses of GST proteins. The 7 subgroups were represented in different colors. The green stars represent *CqGSTs*, the red triangles represent *AtGSTs* and the blue circles represent *OsGSTs*

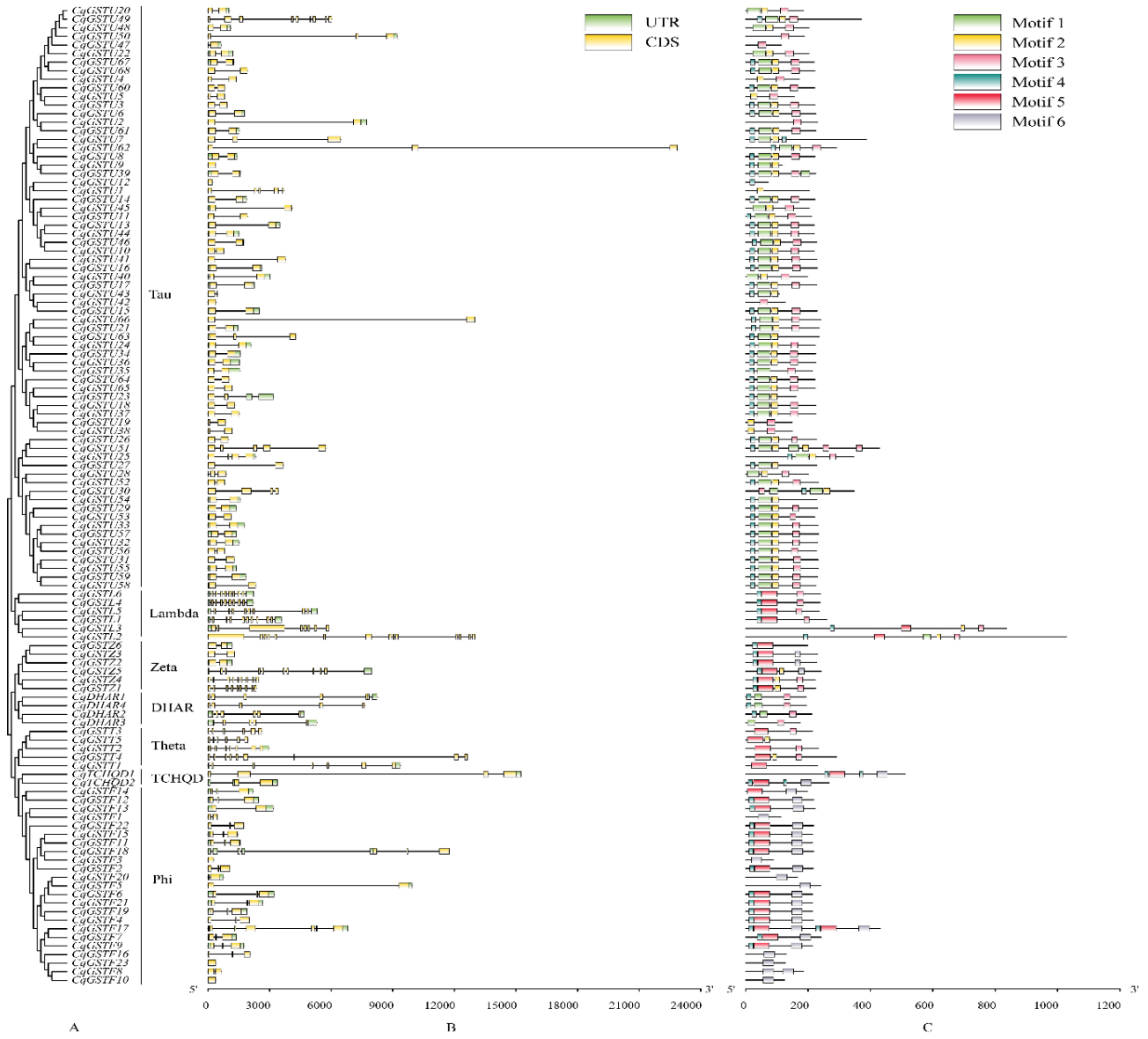

Figure 4. The phylogenetic tree, gene structure and Motif of *CqGSTs* were analyzed (Note: A, evolutionary tree; B, genetic structure; C, conservative motif)

3.4. Phylogenetic Tree Analysis and Classification of Quinoa GST Gene Family

In general, the position of exons and introns can provide the most important information for the evolutionary relationship of species. The 114 gene structures of all identified quinoa *GST* gene families were analyzed, and the results were shown in (Figure 4). The number of exons contained in *CqGST* gene ranged from 1 to 14, among which *CqGSTL2* contained 14 exons, indicating that *CqGSTL2* contained 14 exons. The gene with the highest number of exons. Among the 68 *CqGSTU* members, except *CqGSTU49*, there are 9 exons. All the other Tau members contained 1-5 exons, of which 55 members contained 2 exons, accounting for 80.9% of the Tau group. Among the 6 *CqGSTL* members, *CqGSTL2* contains 14 exons and CqGSTL1 contains 9 exons all other members contain eight exons. The number of exons in 6 *CqGSTZ* members was significantly different, including 3 members have 2 exons and 3 have 9 exons. *CqDHAR2* was included in 5 *CqDHAR* members. It has 6 exons, and all the others have 5 exons. The five members of the *CqGSTT* class contain 6~9 exons. The two *CQTCHQDS* contain two and four exons respectively. 23 *CqGSTF* class members, except *CqGSTF17*. Except for 6 exons, the others all contained 1~3 exons, among which the most contained 3 exons, accounting for *CqGSTF* 60.9% of the class members. Genes in the same class often have similar structures, for example, 68 *CqGSTU* members *CqGSTU49* contains 9 exons, and all other members contain 1~5 exons, including 55 exons. There are 2 exons, accounting for 80.9% of the Tau class. Similar to the gene structure, the conserved motif of quinoa *GST* protein was in the same class is usually composed of similar motifs. Similar gene structures and conserved motifs in each class are advanced Step proves the credibility of classification.

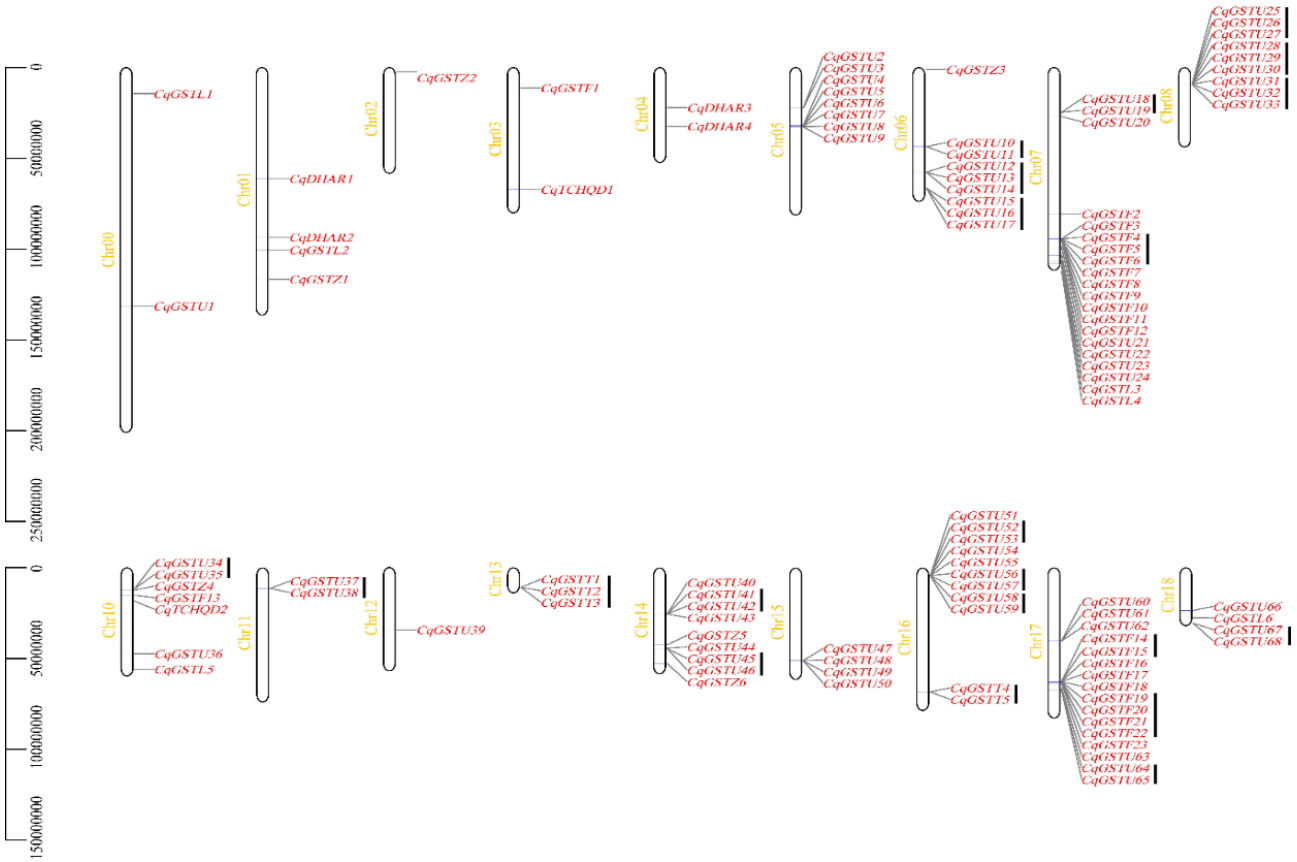

Figure 5. Distribution of *CqGST* gene on chromosome (Note: Black lines represent gene clusters)

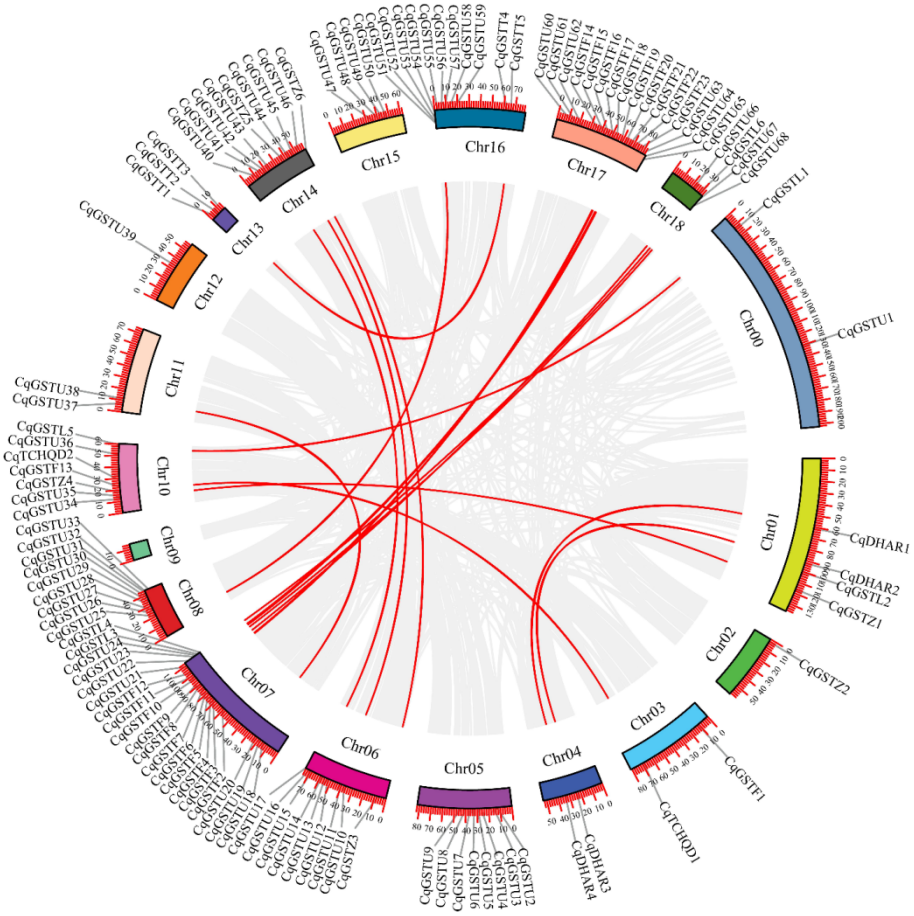

Figure 6. Schematic representations for the chromosomal distribution and inter chromosomal relationships of *CqGST* **genes (**Note: The chromosomes are shown in different colors. All the colinear blocks in the quinoa genome are indicated by gray lines, and the *CqGST* gene pairs of fragment replication events are represented by red lines)

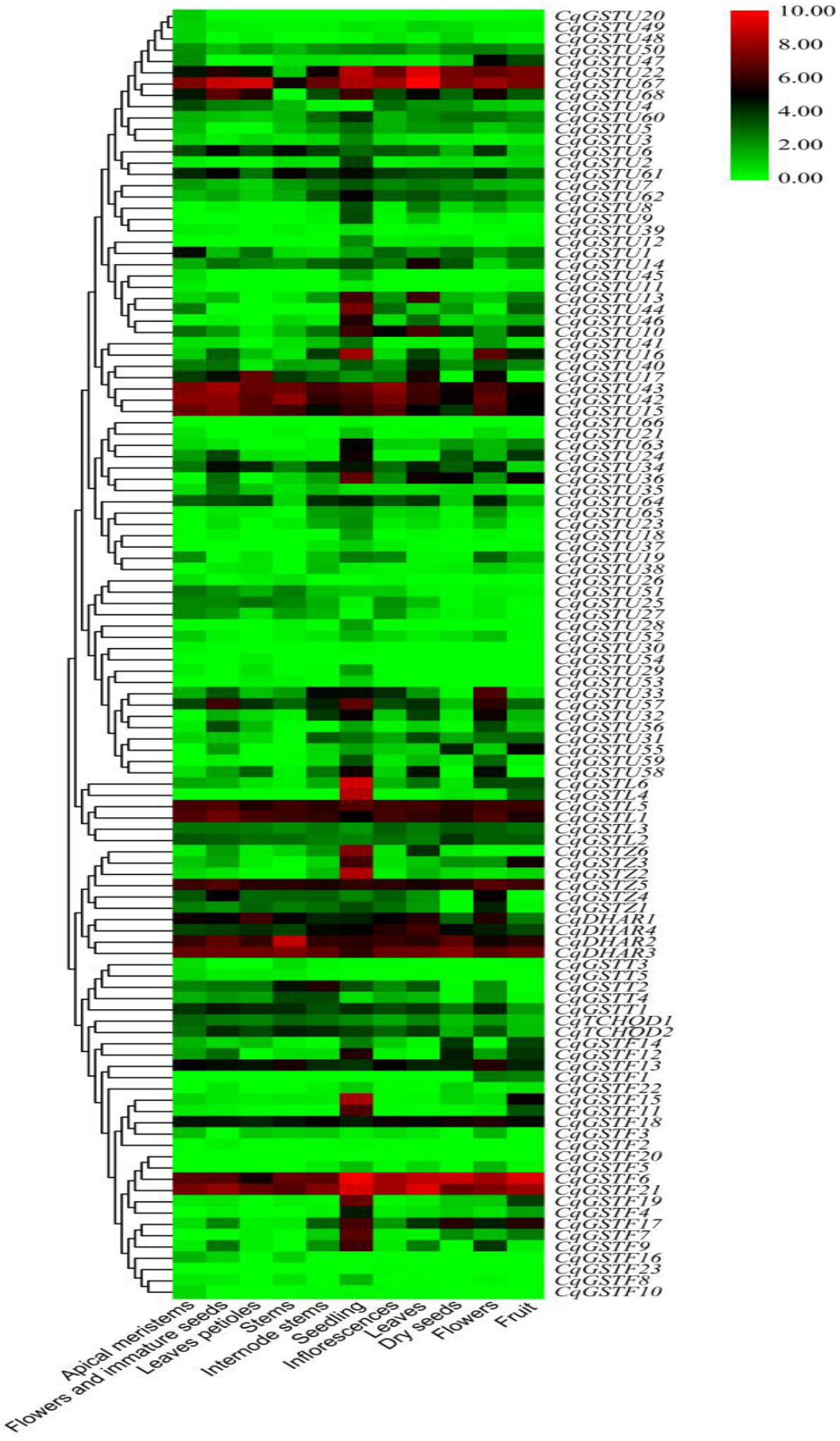

Figure 7. Analysis of GST gene expression patterns in different tissues and organs of quinoa

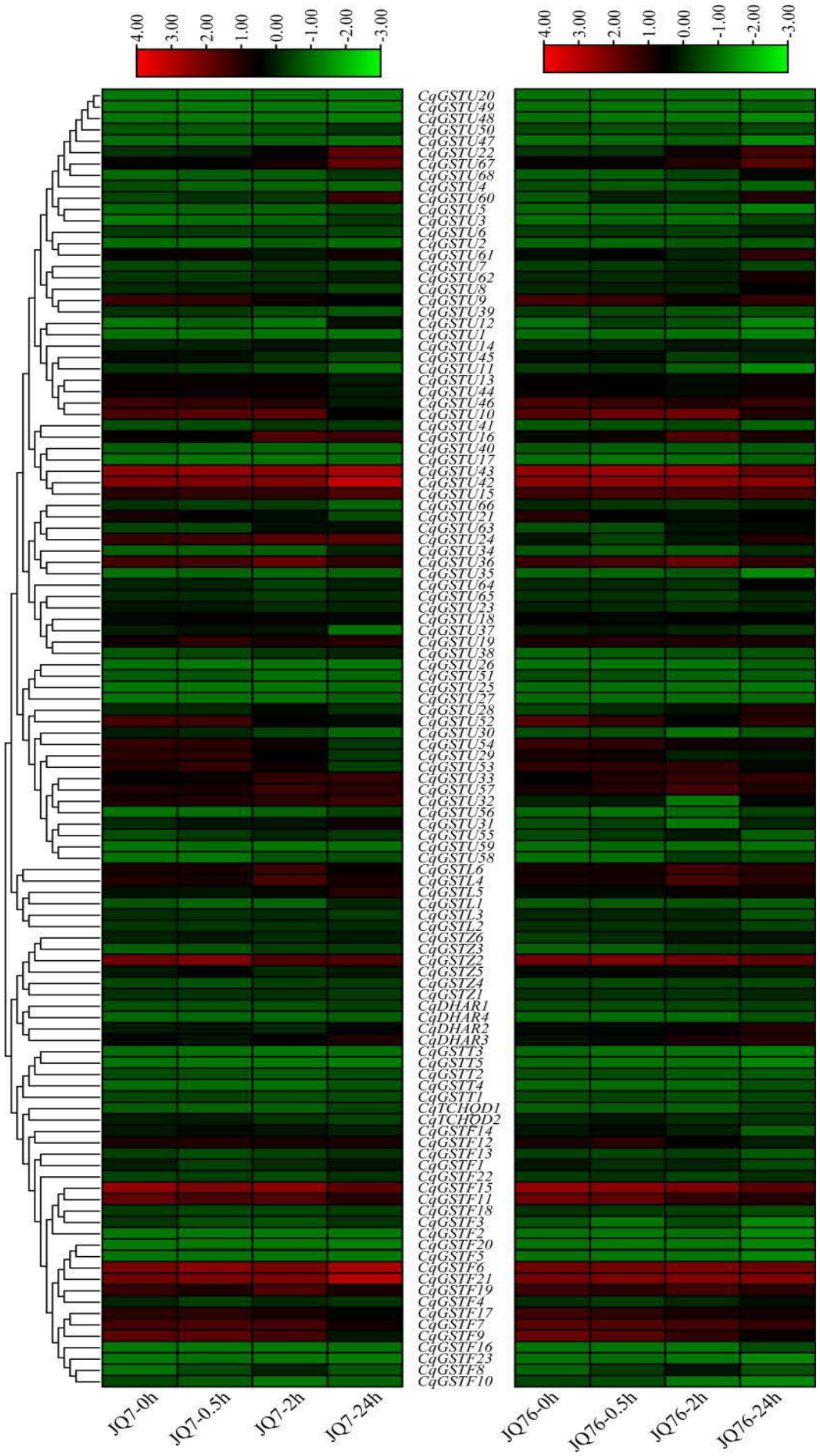

Figure 8. Expression patterns of GST gene in roots of two different varieties of Quinoa under salt

3.5. Chromosome Distribution, Gene Replication, Molecular Evolution and Synchronization Analysis of *CqGST* Gene

Studies on chromosome distribution of Quinoa showed in (Figure 5) that except for no *CqGST* gene on chr09, all the other chromosomes were identical there was *CqGST* gene distribution, and chr07 contains 20 *CqGST* genes, which was one of *CqGST* gene distribution. Hot spots. *CqGSTL1* and *CqGSTU1* cannot locate specific chromosomes due to the quality of the genome. Overall, there was no correlation between the number of GST genes on chromosomes and chromosome length.

The presence of two or more homologous genes in the 200 kb chromosome region was defined as tandem duplication. As shown in (Figure 6), there are 21 tandem repeats including 51 genes on 10 chromosomes of Quinoa. Accounting for 44.7%, referred to as tandem repeat gene cluster. Each of the 10 chromosomes contains 1-4 bases Chr16 has 4 GST gene clusters, which is the chromosome with the most GST gene clusters. The number chromosomes contain only one set of gene clusters (GST clusters). Each GST cluster contains 2, 4 genes, most of them The GST cluster has only two genes. Chr17 contains three gene clusters, one of which has four genes, was most genes GST cluster. In addition, a total of 51 genes formed 30 tandem duplicate gene pairs. The repeat event, *MCScanX's* method was used to obtain 27 duplicate gene pairs consisting of 51 genes. These results suggest that some *CqGST* genes may be produced by gene replication, while tandem and Fragment replication events may play a key role in the amplification of *CqGST*.

3.6. Analysis of Expression Pattern of *CqGST* Gene

Transcription analysis by RNA sequencing (RNA-SEQ) is used to identify the tissue type-specific genes of interest mature tool. In order to explore the expression pattern of GST gene in Quinoa, different tissues of Quinoa or RNA-seq data from various developmental stages of organs were systematically analyzed. These samples were collected from the apical fractions. Biological tissue and other 11 tissues and organs (Figure 7). In class U, *CqGSTU22, CqGSTU67, CqGSTU43, CqGSTU42, CqGSTU15*; in class L, *CqGSTL1* and *CqGSTL5*; in class Z, *CqGSTZ5; DHAR* class *CqDHAR2* and *CqDHAR3* were strongly expressed in various varieties, tissues and organs (Figure 8). And the *CqGST, U66, CqGSTF20 and CqGSTF23* were not expressed in the 11 tissues or organs tested. It may be related to the temporal and spatial expression patterns of these genes. In addition, there were multiple genes in each type in the seedlings Medium and high expression, such as *CqGSTL4, CqGSTL6.*

4. Discussions

4.1. Identification of Quinoa *bHLH* Transcription Factor Family Members

After the release of quinoa genome sequencing data, many gene families have been identified and verified at the whole genome level, including GRAS, ZIP [38,39]. The second largest transcription factor family in plants, *bHLH* transcription factor family is involved in many pathways of plant growth and metabolism [40]. However, no such detailed study has been done on the quinoa *bHLH* family, and we identified 250 *CqbHLH* genes in quinoa. Based on the phylogenetic tree analysis, we divided these CqbHLHTFs into 21 clades, with 19 clades, the number of genes ranging from 3 (K) to 36 (D), and 2 clades A and D containing 2 genes. This is the same number of branches that have been reported for tomatoes [41], soybeans [42], and white pears [44]. However, these plants *SlbHLHs* (159) and *MdbHLHs* (188) have fewer *bHLHs* members than quinoa. This study was based on the published high-quality genome of Quinoa 114 members of Quinoa's GST gene family were identified, divided into seven major types, and like most other plants, *GSTUs* and *GSTFs* are the two largest, with 68 members and 23 members, respectively, while the other five divisions All have fewer than 7 members. The number of GST genes in Quinoa was much higher than that in Arabidopsis thaliana (66 GST genes) [45], Rice (59 GST genes) [46], poplar (81 GST genes) [47], tomato (90 GST genes) [48], soybean (101 GST genes) [49], potato (42 GST genes) [50] and other plants were advanced in the genome In the process of mutation, its size depends mainly on tandem and fragment replication, which is involved in the evolution of gene families. The number of *CqGST* genes in Quinoa has important implications, often leading to chromosome rearrangement and genomic instability amount is relatively high, which may be caused by species-specific fragment replication [51,52]. In the course of evolution, gene replication plays an important role in the function of new genes and the expansion of gene families it plays a role [53] and is associated with chromosome and morphological variation [54]. To better understand the *CqGST* gene family *MCScanX's* method was used to analyze the expansion mechanism of *CqGST* gene family. The analysis was obtained in total thirty tandem gene pairs formed by 51 genes and 27 fragment-replicating gene pairs formed by another 51 genes. These results suggest that both tandem and fragment replication events are major drivers of *CqGST* evolution. In addition, the synchro between Quinoa and its ancestral species C. allidicaule, C.suecicum, Arabidopsis and millet was constructed System comparison chart. It has been shown that quinoa is composed of C. allidicaule (A genomic diploid) and C.suecicum (B genome diploid) was hybridized about 3.3-6.3 million years ago [55]. There were several identical collinear gene pairs between C. allidicaule and C. Suggested that these genes had existed before quadrupling of quinoa while other genes may have evolved over time after the two species interbred or interbred to form quinoa In the process of gene replication and other forms. In addition, quinoa and Arabidopsis possible functional differences between *CqGSTs* under salt stress were analyzed using *RNA-seq* data from a public database The results showed that the expression level of *CqGSTs* gene was induced by salt stress, and part of *Cq* gene was induced by salt stress The expression of GST gene in both salt-tolerant and salt-sensitive quinoa cultivars was up-regulated with the extension of salt treatment time It is suggested that these genes play an important role in the response of quinoa to salt stress.

Some studies have reported differences in the expression of GST genes in plants under abiotic stress, but only A few reports have confirmed their role in abiotic stress tolerance [56,57]. Currently, it's used to study salinity most plants are CAM (crassulacean acid metabolism) plants, and this plant's genome is Unknown, quinoa has been sequenced as an ideal model for studying salt tolerance mechanisms in plants [58]. Therefore, this research Bioinformatics was used to analyze the transcriptase of two different Quinoa varieties under salt stress, and *qRT*-PC was used after verification, R selected eight *CqGST* genes in DHAR and GSTU subclasses. The expression of genes depends on the interaction between cis-acting and trans-acting elements can be analyzed by analyzing the cis-acting elements upstream of the promoter to predict the function of genes [59]. After the target genes were identified, the cis-elements in the promoter regions of these genes were identified and analyzed A piece. The results showed that these eight genes, in addition to having a large number of hormone-induced response elements and cis-regulatory elements, A large number of cis-acting elements related to plant stress response were also included, which was consistent with the experimental results of *qRT-PCR* above under salt and drought stress.

4.2. Replication Analysis of Quinoa *bHLH* Transcription Factor Family Members

In the process of evolution, plants produce multiple members of a gene family mainly through extensive genome replication and diversification. WGD/ segment replication and tandem replication are the main causes of plant gene family expansion [60]. According to gene replication analysis, the main driving force of CqbHLH family expansion was genome replication and fragment replication, which was the same as that of apple [61]. For example, using *MCScanX* (incomplete), 78% of the *bHLH* genes in quinoa were classified as WGD/ fragment copies, despite which 10% of the *bHLH* family members came by tandem. In addition, chromosome localization showed that *CqbHLH* genes were randomly distributed on 19 chromosomes of Quinoa, and there was evidence of fragment replication and tandem replication, which further indicated that these two replication modes played an important role in the expansion and evolution of the *CqbHLH* gene family in Quinoa. In addition, the Ka/Ks ratio suggests that *CQBHLHS* evolved mainly through purification selection, which may be due to selection to maintain important biological functions during their evolutionary history in order to adapt to the current environment and under the pressure of purification evolution. Exons and introns play key roles in plant evolution, and a large number of introns are lost over time in the early stages of species expansion [62]. Thus number of introns increased through evolution, which may be a necessary mechanism for generating new gene functions. Intron-free members of the *bHLH* gene family in plants may have originated in prokaryotes and then replicated widely in plants. In this study, the structure of the *bHLHTFs* exon intron gene of quinoa showed that only 6.0% (15 genes) of *CqbHLH TFs* contained no intron and only one exon, which was basically consistent with the number of rice [63]. Therefore, we speculate that the *bHLH* gene family may be involved in many biological processes and molecular functions due to its evolutionary diversity and low conserved pattern. In addition, most *bHLH* proteins belonging to the same clave have similar motifs. Therefore, we speculate that members of the same branch may have similar functions. Numerous studies have shown that the *bHLH* transcription factor family plays an important regulatory role in plant abiotic stress response. Therefore, transcriptome data were used in this study to explore the expression patterns of *CqbHLH* gene family members in different tissues and the response of *CqbHLH* family members to salt stress. The *bHLH* transcription factor family was found to have significant tissue expression differences in Quinoa, suggesting that they play an important role in different development and physiological functions of quinoa. In addition, quinoa has been studied as a model for understanding plant salt tolerance, and *bHLH* genes identified in various plant species have been shown to play an important role in salt stress response. In addition, some *CqbHLH* transcription factors, such as *CqXXX*, were strongly expressed in Quinoa leaves according to the analysis results of the salt stress transcript data. After reviewing the literature, it was found that the *CqXXX*, gene did have salt tolerance function in other species. We speculated that these *CqbHLHTFs* might play an important role in regulating the formation of salt vesicles in Quinoa leaves and channling salt. This is consistent with previous reports. By comparing the expression differences of *bHLHTFs* between the two varieties before and after salt stress, we found that the expression levels of *genes CqbHLH167, CqbHLH180, CqbHLH115 and CqbHLH215* in the sprouts of salt-tolerant varieties were significantly higher than those of sensitive varieties. To find the evidence of salt tolerance of the above differentially expressed genes in other species, this is the discussion, which cannot be limited to its own results.) Other genes, such as *CqbHLH232, CqbHLH169, CqbHLH71, CqbHLH97*, were significantly more expressed in the roots of salt-tolerant varieties than those of sensitive varieties. After stress treatment, its expression increased significantly. This suggests that these genes may be responsible for the differences in salt tolerance between the two varieties. In addition, in this study, some *bHLH* transcription factors, such as *CqbHLH194, CqbHLH231, CqbHLH89 and CqbHLH62*, were positively regulated by salt stress in both salt-tolerant varieties and salt-sensitive varieties of quinoa. These genes may not be the cause of salt stress resistance in salt-tolerant varieties and salt-sensitive varieties of Quinoa. The results of this study provide evidence for further identification of the functions of candidate genes in the process of plant salt stress resistance. In addition, some *CQBHLHS* were negatively responsive to salt stress, suggesting that they may be responding to other stresses or participating in other biological processes.

5. Conclusion

In this study, we identified 250 *CqbHLHTFs* from Quinoa and systematically analyzed their gene structure, conserved motifs, gene replication, evolutionary

relationships and tissue expression patterns. According to phylogenetic analysis, the *CqbHLH* family was divided into 21 groups. According to collinearity analysis, WGD and tandem replication may play a role in the evolution of the *CqbHLH* family. In addition, according to *RNA-seq* data, *CqbHLH* gene may play an important role in coping with abiotic stress, and different varieties of Quinoa have different expression patterns under salt stress. (It was important to point out which genes are possible salt tolerance genes in the conclusion.). The data collected in this study provided the basis for further research on the *bHLH* gene and salt tolerance stress and other abiotic stress in quinoa, which will help to improve the economic value of quinoa.

Acknowledgements

This research was supported by the Research Program Sponsored by Ministerial and Provincial Co-Innovation Centre for Endemic Crops Production with High-quality and Efficiency in Loess Plateau, Taigu 030801, China (No. SBGJXTZX-15), the National Modern Agricultural Industrial Technology System (CARS-03-01-24), the Key Laboratory of Crop Ecology and Dryland Cultivation Physiology of Shanxi Province (201705D111007), the Key Innovation Team of the 1331 Project of Shanxi Province, the Scientific and Technological Innovation Project of colleges in Shanxi Province (2021L178), Science and technology innovation fund of Shanxi Agricultural University (2018YJ18).

References

[1] Ledent V, Vervoort M. The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis. Genome Res. 2001. 11(5):754-70.

[2] Wang Y M, Bai H, Cao Y. Research progress of *bHLH* transcription factor and its application in plantstress tolerance [J]. Anhui Agricultural Sciences, 2015. 43(21): 34-35, 50.

[3] Wang X J, Li K Z. Research progress of plant *bHLH* transcription factor involved in abiotic stress sig-naling pathway [J]. Life Science, 2015. 27(2): 208-216.

[4] Buck MJ, Atchley WR. Phylogenetic analysis of plant basic helix-loop-helix proteins. J Mol Evol. 2003. 56(6): 742-50.

[5] Chinnusamy V, Ohta M, Kanrar S, Lee BH, Hong X, Agarwal M, Zhu JK. ICE1: a regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis. Genes, 2003.15; 17(8):1043-54.

[6] Kiribuchi K, Sugimori M, Takeda M, Otani T, Okada K, Onodera H, Ugaki M, Tanaka Y, Tomiyama-Akimoto C, Yamaguchi T, Nishiyama M, Nojiri H, Yamane H. RERJ1, a jasmonic acid-responsive gene from rice, encodes a basic helix-loop-helix protein. Biochem Biophys Res Commun. 2004. 17; 325(3): 857-63.

[7] Le Hir R, Castelain M, Moritz T, Dinant S, Bellini C. *AtbHLH68* transcription factor contributes to the regulation of ABA homeostasis and drought stress tolerance in Arabidopsis thaliana. Physiol Plant, 2017.160(3):312-327.

[8] Fursova OV, Pogorelko GV, Tarasov VA. Identification of ICE2, a gene involved in cold acclimation which determines freezing tolerance in Arabidopsis thaliana. Gene, 2009. 15; 429(1-2): 98-103.

[9] Yang T, Yao S, Hao L, Zhao Y, Lu W, Xiao K. Wheat *bHLH*-type transcription factor gene TabHLH1 is crucial in mediating osmotic stresses tolerance through modulating largely the ABA-associated pathway. Plant Cell Rep, 2016. 35(11): 2309-2323.

[10] Zhai Y, Zhang L, Xia C, Fu S, Zhao G, Jia J, Kong X. The wheat transcription factor, TabHLH39, improves tolerance to multiple abiotic stressors in transgenic plants. Biochem Biophys Res Commun. 2016 May 13; 473(4): 1321-1327.

[11] Seo JS, Joo J, Kim MJ, Kim YK, Nahm BH, Song SI, Cheong JJ, Lee JS, Kim JK, Choi YD. OsbHLH148, a basic helix-loop-helix protein, interacts with OsJAZ proteins in a jasmonate signaling pathway leading to drought tolerance in rice. Plant J. 2011, 65(6): 907-21.

[12] Zhao L, Gong X, Gao J, Dong H, Zhang S, Tao S, Huang X. Transcriptomic and evolutionary analyses of white pear (Pyrus bretschneideri) β-amylase genes reveals their importance for cold and drought stress responses. Gene. 2019, 20; 689: 102-113.

[13] Dong Y, Wang C, Han X, Tang S, Liu S, Xia X, Yin W. A novel *bHLH* transcription factor PebHLH35 from Populus euphratica confers drought tolerance through regulating stomatal development, photosynthesis and growth in Arabidopsis. Biochem Biophys Res Commun. 2014, 18; 450(1): 453-8.

[14] Li X, Duan X, Jiang H, Sun Y, Tang Y, Yuan Z, Guo J, Liang W, Chen L, Yin J, Ma H, Wang J, Zhang D. Genome-wide analysis of basic/helix-loop-helix transcription factor family in rice and Arabidopsis. Plant Physiol. 2006, 141(4): 1167-84.

[15] Gao C, Sun J, Wang C, Dong Y, Xiao S, Wang X, Jiao Z. Genome-wide analysis of basic/helix-loop-helix gene family in peanut and assessment of its roles in pod development. PLoS One. 2017, 27; 12(7): e0181843.

[16] Sun H, Fan HJ, Ling HQ. Genome-wide identification and characterization of the *bHLH* gene family in tomato. BMC Genomics. 2015, 22; 16(1): 9.

[17] Mao K, Dong Q, Li C, Liu C, Ma F. Genome Wide Identification and Characterization of Apple *bHLH* Transcription Factors and Expression Analysis in Response to Drought and Salt Stress. Front Plant Sci. 2017 Apr 11; 8:480.

[18] Dong H, Chen Q, Dai Y, Hu W, Zhang S, Huang X. Genome-wide identification of *PbrbHLH* family genes, and expression analysis in response to drought and cold stresses in pear (*Pyrus bretschneideri).* BMC Plant Biol. 2021 Feb 9; 21(1): 86.

[19] Chen C, Chen H, Zhang Y. TBtools - an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020: 1-9.

[20] Pires N, Dolan L. Origin and diversification of basic-helix-loop-helix proteins in plants. Mol Biol Evol. 2010 Apr; 27(4):862-74.

[21] Carretero-Paulet L, Galstyan A, Roig-Villanova I, Robertson DL. Genome-wide classification and evolutionary analysis of the *bHLH* family of transcription factors in Arabidopsis, poplar, rice, moss, and algae. Plant Physiol, 2010.153(3): 1398-412.

[22] Gorlova O, Fedorov A, Logothetis C, Amos C, Gorlov I. Genes with a large intronic burden show greater evolutionary conservation on the protein level. BMC Evol Biol, 2014.16; 14(1): 50.

[23] Maher C, Stein L, Ware D. Evolution of Arabidopsis micro RNA families through duplication events. Genome Res. 2006 Apr; 16(4):510-9.

[24] Zhu X, Wang B, Wei X. Genome wide identification and expression pattern analysis of the GRAS family in quinoa. Funct Plant Biol. 2021 Aug; 48(9):948-962.

[25] Li F, Liu J, Guo X, Yin L, Zhang H, Wen R. Genome-wide survey, characterization, and expression analysis of *bZIP* transcription factors in *Chenopodium quinoa*. BMC Plant Biol. 2020 Sep 1; 20(1):405.

[26] Kavas M, Baloğlu MC, Atabay ES, Ziplar UT, Daşgan HY, Ünver T. Genome-wide characterization and expression analysis of common bean *bHLH* transcription factors in response to excess salt concentration. Mol Genet Genomics. 2016 Feb; 291(1): 129-43.

[27] Song, J., Gao, Z., Huo, X. 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 (2015).

[28] Roy SW, Penny D. A very high fraction of unique intron positions in the intron-rich diatom Thalassiosira pseudonana indicates widespread intron gain. Mol Biol Evol. 2007 Jul; 24(7):1447-57.

[29] Roy SW, Gilbert W. Complex early genes. Proc Natl Acad Sci U S A. 2005 Feb 8; 102(6): 1986-91.

[30] Gao C, Sun J, Wang C, Dong Y, Xiao S, Wang X, Jiao Z. Genome-wide analysis of basic/helix-loop-helix gene family in peanut and assessment of its roles in pod development. PLoS One. 2017 Jul 27; 12(7):e0181843.

[31] Sun W, Jin X, Ma Z, Chen H, Liu M. Basic helix-loop-helix (*bHLH*) gene family in Tartary buckwheat (*Fagopyrum tataricum*): Genome-wide identification, phylogeny, evolutionary expansion and expression analyses. Int J Biol Macromol. 2020 Jul 15; 155:1478-1490.

[32] Gu X. Statistical methods for testing functional divergence after gene duplication. Mol Biol Evol. 1999 Dec; 16(12): 1664-74.

[33] Bhargava A, Shukla S, Ohri D. *Chenopodium quinoa* - An Indian perspective [J]. Industrial Crops and Products, 2006, 23(1): 73-87.

[34] White P, Alvistur E, Diaz C, et al. Nutrient Content and Protein Quality of Quinua and Cafiihua Edible Seed Products of the Andes Mountains[J]. Agricultural and Food Chemistry, 1955, 3(6): 351-355.

[35] Jacobsen S E, Mujica A, Jensen C R. The resistance of quinoa (*Chenopodium quinoa Willd.)* to adverse abiotic factors [J]. Food Reviews International, 2003, 19(1/2): 99-109.

[36] Jarvis D E, Ho Y S, Lightfoot D J. The genome of *Chenopodium quinoa* [J]. Nature, 2017, 542(7641): 307-312.

[37] Zurita-Silva A, Fuentes F, Zamora P. breeding quinoa (*Chenopodium quinoa Willd*.): potential and perspectives [J]. Molecular Breeding, 2014, 34(1): 13-30.

[38] Fuentes F, bhargava A. Morphological analysis of quinoa germplasm grown under lowland desert conditions [J]. Journal of agronomy and cropscience, 2011, 197(2): 124-134.

[39] Gonzlez J A, Bruno M, Valoy M. Genotypic variation of gas exchange parameters and leaf stable carbon and nitrogen isotopes in ten quinoa cultivars grown under drought [J]. Journal of agronomy and crop science, 2011, 197(2): 81-93.

[40] Jacobsen S E. Adaptation of quinoa (*Chenopodium quinoa*) to Northern European agriculture: studies on developmental pattern [J]. Euphytica, 1997, 96(1): 41-48.

[41] Christensen S A, Pratt D B, Pratt C. Assessment of genetic diversity in the USDA and CIP-FAO international nursery collections of quinoa (*Chenopodium quinoa* Willd.) using microsatellite markers[J]. Plant Genetic Resources, 2007, 5(2): 82-95.

[42] Oyoo M E, Githiri S M, Ayiecho P O. Performance of some quinoa (Chenopodium quinoa Willd.) genotypes in Kenya[J]. South African Journal of Plant and Soil, 2010, 27(2): 187-190. 19-22.

[43] Razzaghi F, Ahmadi S H, Jacobsen S E. Effects of salinity and soil-drying on radiation use efficiency, water productivity and yield of quinoa (*Chenopodium quinoa* Willd.) [J]. Journal of agronomy and crop science, 2012, 198(3): 173-184.

[44] Bhargava A, Srivastava S. Quinoa: Botany, Production and Uses [M]. Wallingford, UK: CABI Publisher, 2013: 4-61.

[45] Maughan P J, Bonifacio A, Coleman C E. Quinoa (*Chenopodium quinoa*) [M]. Berlin: Springer, 2007. 147-158.

[46] Shabala L, Mackay A, Tian Y, et al. Oxidative stress protection and stomatal patterning as components of salinity tolerance mechanism in quinoa (*Chenopodium quinoa*)[J]. Physiological plantarum. 2013, 146(1): 26-38.

[47] Kiani P A, Roessner U, Jayasinghe N S. Epidermal bladder cells confer salinity stress tolerance in the halophyte quinoa and Atriplex species [J]. Plant cell and environment, 2017. 40(9): 1900-1915.

[48] Fairbanks D J, Waldrigues A, Ruas C F. Efficient characterization of biological diversity using field DNA extraction and random amplified polymorphic DNA markers [J]. Revista brasileira degenetica, 1993. 16: 11-22.

[49] Maughan P J, Smith S M, Rojas-Beltr N J A. Single nucleotide polymorphism identification, characterization, and linkage mapping in quinoa[J]. Plant genome, 2012. 5(3): 114-125.

[50] Zou C S, Chen A J, Xiao L H. A high-quality genome assembly of quinoa provides insights into the molecular basis of salt bladder-based salinity tolerance and the exceptional nutritional value [J]. Cell research, 2017. 27: 1327-1340.

[51] Frear D S, Swanson H R. Biosynthesis of S-(4-ethylamino-6-isopropylamino-2-s triazino) glutathione: Partial purification and properties of a glutathione S transferase from corn [J]. Phytochemistry, 1970, 9(10): 2123-2132.

[52] Mcgonigle B. A genomics approach to the comprehensive analysis of the glutathione s-transferase gene family in soybean and maize [J]. Plant physiology, 2000, 124(3): 1105-1120.

[53] Edwards R, Dixon D P, Walbot V. Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health [J]. Trends in Plant Science, 2000, 5(5): 193-198.

[54] Marrs, Kathleen A. The functions and regulation of glutathione s-transferases in plants [J]. Annu Rev Plant Physiol Plant Mol Biol, 1996. 47(1): 127-158.

[55] Droog F, Hooykaas P, Van D Z B J. 2, 4-Dichlorophenoxyacetic acid and related chlorinated compounds inhibit two auxin-regulated type-III tobacco glutathione S-transferases [J]. Plant physiology, 1995, 107(4): 1139-1146.

[56] Dixon D P, Cummins I, Cole D J. Glutathione-mediated detoxification systems in plants [J]. Current Opinion in Plant Biology, 1998. 1(3): 258.

[57] Droog F. Plant GlutathioneS-Transferases, a Tale of Theta and Tau [J]. Journal of Plant Growth Regulation, 1997. 16(2): 95-107.

[58] Board G P, Baker T R, Chelvanayagam G, et al. Zeta, a novel class of glutathione transferases in a range of species from plants to humans [J]. Biochemical Journal, 1998. 328(3): 929-935.

[59] Dixon D P, Lapthorn A, Edwards R. Plant glutathione transferases [J]. Genome Biology, 2002. 3(3): 1-10.

[60] Basantani M, Srivastava A. Plant glutathione transferases - A decade falls short. Canadian Journal of Botany, 2011, 85(5): 443-456.

[61] Lallement P A, Brouwer B, Keech O. The still mysterious roles of cysteine-containing glutathione transferases in plants [J]. Frontiers in Pharmacology, 2014. 5(192): 1-22.

[62] Lallement P, Meux E, Gualberto J. Structural and enzymatic insights into Lambda glutathione transferases from Populus trichocarpa, monomeric enzymes constituting an early divergent class specific to terrestrial plants [J]. Biochemical Journal, 2014. 462(1): 39-52.

[63] Liu Y J, Han X M, Ren L. Functional divergence of the glutathione S-transferase supergene family in Physcomitrella patens reveals complex patterns of large gene family evolution in land plants. [J]. Plant Physiology, 2013. 161(2): 773-786.