

# Identification and Expression Analysis of DUF4228 Domain Containing (DDP) Genes in Potato Under Abiotic and Phytohormone Stress

yasir sharif (✉ [yasirsharif3336@gmail.com](mailto:yasirsharif3336@gmail.com))

Fujian Agriculture and Forestry University

**Madiha Zaynab**

Shenzhen University

**Khalid Ali Khan**

King Khalid University

**Shuangfei Li**

Shenzhen University

---

## Research Article

**Keywords:** Gene family,miRNA,Expression,Duplication,Phylogeny

**Posted Date:** September 29th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-902964/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

DUF4228 (Domain of unknown function 4228) proteins are widely distributed in plants and performed a vital role in abiotic stress response. In potato (*Solanum tuberosum*), the study of the DUF4228 gene family is sparse. To understand the role of DUF4228 in potato a comprehensive genome-wide analysis was carried out in potato genome. Further, the *StDUF4228* genes relative expression was also evaluated in various plant tissues. In the present study identified 31 *StDUF4228* genes were clustered into six groups. The promoter analysis of *StDUF4228* revealed various cis-acting regulatory components response to abiotic and phytohormones stresses. Under selection pressure, gene duplication also showed various positive and purifying selections in *StDUF4228* genes. Expression profiling indicated that *StDUF4228* genes were broadly expressed in various tissues and organs in potato. Moreover, the *StDUF4228* genes profiles expression was characterized through RNA-sequence analysis data under salt and heat treatments. In potato, *StDUF4228-4*, *StDUF4228-14*, *StDUF4228-21*, and *StDUF4228-30* represented a high expression under salt and heat stress. Furthermore, 23 genes under IAA treatment and 18 under ABA treatment, showed that IAA and ABA played a vital role in stresses. The RT-qPCR and expression profiling exhibited high expression in root tissues. Moreover multiple miRNA target *StDUF4228* genes in potato. These expression results provide primary reference data and functional analysis data for other commercial crops.

## Introduction

Plants often experience various environmental stresses, including biotic (diseases, insect pests, weeds, etc.) and abiotic stress factors such as high and low temperature, drought, and salinity (Zaynab et al., 2018; Zaynab et al., 2021b). These factors attenuate plant development and growth, badly affecting plant productivity and yield quality (Zaynab et al., 2018). Therefore, it is important to understand plants' stress damage and stress response mechanisms to improve their resistance level (Chen et al., 2012; Zaynab et al., 2020). Abiotic stress factors are considered chief limiting factors affecting plant yield and growth, causing worldwide significant economic and agricultural loss. The estimated annual loss is about 70% in yield due to inadequate chemical and physical environment.

The DUFs contained various functions in plants and current studies revealed that some important DUFs played vital roles in different plants against abiotic stress. The rice plant genes *DUF1644* and *OsSIDP366* significantly regulated in salinity and drought stress condition through the transgenic plant, in over-expressing *OsSIDP366* indicated high tolerance against enhanced salinity and drought conditions (Guo et al., 2016). In transgenic *A. thaliana* and *O. sativa* *OsSGL* confers improved drought tolerance and various genes response against stress was designated which significantly improved in transgenic rice (Cui et al., 2016). Moreover, in *O. sativa* other members of *DUF* genes were also identified and concerned with abiotic stress factors, for example, *OsDUF810* (Li et al., 2018), *OsDSR2(DUF966)* (Luo et al., 2014), and *SIDP361 (DUF1644)* (Li et al., 2016). Xin et al. (2007) result in *A. thaliana* showed that *DUF231* gene family member *ESK1 (AT3g55990)* is identified as a novel negative regulator of cold adaptation (Xin et al., 2007). Further analysis in *A. thaliana* species explained that suppressing the *ATRDU2* and *ATRDU1* (both are ubiquitin ligases of RING-DUF1117 E3) gene expression reduced the ABA-mediated drought stress (Kim et al., 2012). The *MsDUF (Medicago sativa DUF)* gene overexpression in the tobacco plants resulted in increased MDA (malondialde) content and significantly reduced the soluble sugar and chlorophyll contents. *MsDUF* gene expression was significantly lowered under different treatments, including gibberellic acid (GA), abscisic acid (ABA), PEG6000, and NaCl, indicating the negative regulatory function of *MsDUF* in *Medicago sativa* as a stress resistance (Wang et al., 2018). The expression of the *CiDUF4228-3 (Caragana intermedia DUF4228-3)* gene positively up-regulated under drought, dehydration, and low temperature showed its association in stress conditions (Leng et al., 2021). Domain-containing protein (DUF4228) is exclusively found in the genomic data of plants.

In March 2019, the Pfam database retrieved a total of 2882 DUF4228 (Pfam accession: 14009) family genes from 80 different species. Thus, only one gene (DUF4228) has been illustrated from plants. Stress-responsive novel gene *MsDUF* belonging to *DUF4228* family obtained from *Medicago sativa* played a negative regulatory role in osmotic stress and seed vigor in plants (Wang et al., 2018). Current studies illustrated that the *DUF4228* genes had been involved in cadmium tolerance and abiotic stress indication. Analysis of the *AtDUF4228* gene expression under the various co-expression networks of the *DUF4228* gene family and stress treatments in various plant species indicated the synergistic role of *DUF4228* genes in plant immune response (Didelon et al., 2020). These obtaining results proposed that *DUF4228* gene family may also associate with abiotic stress response in soybean.

Potato is an important economic crop and widely used as a food throughout the world (Zaynab et al., 2021a). However, as with other plants, potato yield is also vulnerable to biotic and abiotic stresses (Dahal et al., 2019; Sattar et al., 2021). Until now, very little is known about DUF4228 in the potato genome. Therefore, the current study involved a genome-wide screening of DUF4228 in the potato genome. In addition, several bioinformatics analyses were also carried out to explore the basic and advanced features of DUF4228, including gene structure, chromosome localization, phylogenetic relations, conserved protein domains, regulatory networks, expression profiles in different tissues, and relative expression through real-time qPCR. The outcomes of this study will provide the basis for further functional analysis of potato DUF4228 and can also contribute to a better understanding of their molecular mechanisms.

## Materials And Methods

### Identification of DUF4228 genes in the potato genome

In the present study, *Solanum tuberosum* genome data were obtained from Phytozome (<https://phytozome.jgi.doe.gov/>) and *Arabidopsis* genome data were retrieved from TAIR (<http://www.arabidopsis.org/>). To find the protein sequence of *StDUF4228* Pfam no. (PF14009) and HMMER3.0 software was used. The *StDUF4228* sequence was obtained through the BLASp program by using the Arabidopsis protein sequence as queries. Moreover, the *StDUF4228* protein sequence was next authenticated on the base of conserved domains through SMART (<http://smart.emblheidelberg.de>). The Physicochemical properties, such as molecular weight (kDa) and isoelectric point (pI) of each protein, were determined using ExPASy (<http://web.expasy.org/protparam/>).

## Multiple sequence alignment and phylogenetic analysis

Multiple sequence alignment and phylogenetic analysis of all identified protein sequences were aligned using MUSCLE with 16 iterations. Next, the associated sequences have been used to make a phylogenetic tree with a maximum probability approach by 1000 bootstrap values.

## Gene structure and motif analysis

Gene Structure Display Server was used for the identification of *StDUF4228* gene family exon-intron characteristics. Protein motifs were identified through Multiple Expectation Maximization for Motif Elicitation (MEME) software (<http://meme-suite.org/>). The following parameters were used for the motif analysis: the maximum motif number identified is 10, and other obtaining parameters are non-remittance values. The graphical display is depending on the incredible possible gene viewer part in TBtools software.

## Cis-elements Analysis and Chromosomal locations

To study the cis-element in the 31 *StDUF4228* gene promoter sequence, the start codon's upstream genomic sequences (2,000bp) for all *StDUF4228* genes have been submitted online to PlantCARE site (<http://bioinformatics.psb.ugent.be/webtools/plntcare/html/>). Regarding information to each known *StDUF4228* was restored as of PGSC website.

## Synteny and Selective Pressure Analysis

For genome conservation visualization, the relative synteny analysis was performed by the Circoletto tool ([tools.bat.infospire.org/circoletto/](http://tools.bat.infospire.org/circoletto/)). The duplicated genes coding sequence was aligned with MEGA7 through the Muscle/ codon method. The nonsynonymous and synonymous substitution rates (number of nonsynonymous substitution per nonsynonymous site: Ka, number of synonymous substitutions per synonymous site: Ks) were computed through KaKs-Calculator 2.0 software with the MYN process. Further, the time divergence ( $t = Ks/2r$ ) was computed with rate exchange ( $r = 2.6 \times 10^{-9}$ ) [31].

## Identification of miRNA

The coding sequence (CDS) of *StDUF4228* was used to identify possible target miRNAs in the psRNATarget database (Available online: <http://plantgrn.noble.org/psRNATarget/>), with default parameters.

## Expression analysis of StDUF4228 Genes

For expression analysis of *StDUF4228* genes, FPKM (fragments per kilobase million) were observed in different tissues. The obtained data were computed regarding more expressive tissues for example, roots, stems, and leaves. The observed FPKM values were utilized to construct the heatmap through TBtools [32]. Moreover, to identify the genes expression, we used RT-qPCR analysis in potato.

## Plant Material and Sample Collection

Potato tubers used as experimental material, obtained from NARC (National Agriculture Research Center) Islamabad, planted under control conditions in the glasshouse at NIGAB (National Institute for Genomics and Advanced Biotechnology), NARC Islamabad. After twenty-five days post-germination, young leaves, stems, and roots were gathered in replicates. For RNA extraction analysis, tissue samples were stored in liquid nitrogen.

## RNA extraction, cDNA synthesis, and qRT-PCR analysis

Using PureLink RNA Mini Kit (Invitrogen USA), total RNAs were extracted from each sample through the manufacturer's protocol. The extracted RNAs samples were quantified through a BioSpec-nano spectrophotometer. The total RNA 1 $\mu$ g sample was reverse transcribed through Reverse Transcriptase (M-MLV) with oligodT primer. qRT-PCR was performed through SYBER GREEN qPCR Master MIX (Thermo-Fisher Scientific, USA) in StepOnePlus RT-PCR, Applied Biosystem. qRT-PCR analysis profile was used 95°C for ten minutes, 94°C followed by 40 cycles for 30 seconds, and 58°C. During this analysis, elongation factor 1-alpha was used as a housekeeping gene. The further procedure was carried out through three biological replicates and analyzed using the 2- $\Delta\Delta$ Ct method [33]. The standard errors of replicates were graphically represented.

## Results

### Identification of StDUF4228 Genes

To identify *S.tuberosum* *DUF4228* genes, through their Pfam number (PF14009) and HMMER, investigate the *DUF4228* protein sequence using the *S.tuberosum* protein database. Moreover, the local algorithm BLASTP was used through *StDUF4228* genes for queries. Each gene candidate's conserved domain was predicted through the SMART database. A total of 31 *StDUF4228* genes were identified. For basic gene features analysis for *S.tuberosum* *StDUF4228*

proteins, we have studied gene location, gene number, molecular weight (MW), and isoelectric point (PI) (Table 1). The identified *StDUF4228* proteins length ranged from 104 to 318 amino acids, with isoelectric points varying from 5.49 to 10.25 and molecular weight ranging from 11513.26kD to 35578.61kD.

Table 1  
List of identified *StDUF4228* and their features.

Gene	Transcript Name	Gene Start (bp)	Gene End (bp)	Chromosome	No. of Amino Acids	Molecular Weight	pl
<i>StDUF4228-1</i>	PGSC0003DMT400078616	41686489	41687034	ST4.03ch02	181	20315.41	7.71
<i>StDUF4228-2</i>	PGSC0003DMT400007187	24430059	24430517	ST4.03ch02	152	16679.11	7.7
<i>StDUF4228-3</i>	PGSC0003DMT400026013	44717174	44717662	ST4.03ch02	162	17551	8.99
<i>StDUF4228-4</i>	PGSC0003DMT400078646	41801189	41801677	ST4.03ch02	162	18206.31	9.11
<i>StDUF4228-5</i>	PGSC0003DMT400055171	42653187	42657421	ST4.03ch02	177	19112.88	9.63
<i>StDUF4228-6</i>	PGSC0003DMT400073981	6655855	6656454	ST4.03ch07	199	22851.28	9.57
<i>StDUF4228-7</i>	PGSC0003DMT400073980	6642898	6643476	ST4.03ch07	192	22108.42	9.67
<i>StDUF4228-8</i>	PGSC0003DMT400030549	7320224	7320538	ST4.03ch01	104	11513.26	5.49
<i>StDUF4228-9</i>	PGSC0003DMT400064780	85542220	85542687	ST4.03ch01	155	17204.21	8.29
<i>StDUF4228-10</i>	PGSC0003DMT400023326	62565576	62566169	ST4.03ch01	197	22088.64	9.33
<i>StDUF4228-11</i>	PGSC0003DMT400078351	4062424	4063872	ST4.03ch05	236	26549.79	9.71
<i>StDUF4228-12</i>	PGSC0003DMT400060331	51430374	51430974	ST4.03ch05	157	17572.32	6.71
<i>StDUF4228-13</i>	PGSC0003DMT400040233	7307326	7308049	ST4.03ch05	165	18458.08	10.11
<i>StDUF4228-14</i>	PGSC0003DMT400078350	4068529	4070256	ST4.03ch05	237	26292.31	9.58
<i>StDUF4228-15</i>	PGSC0003DMT400078471	4083644	4086424	ST4.03ch05	243	27232.75	9.78
<i>StDUF4228-16</i>	PGSC0003DMT400074187	6111438	6112394	ST4.03ch12	318	35578.61	10.25
<i>StDUF4228-17</i>	PGSC0003DMT400052957	2913780	2914277	ST4.03ch06	165	18169.07	9.72
<i>StDUF4228-18</i>	PGSC0003DMT400083024	52156117	52156781	ST4.03ch06	161	18024.8	9.76
<i>StDUF4228-19</i>	PGSC0003DMT400096537	43862131	43863054	ST4.03ch09	307	35416.36	9.83
<i>StDUF4228-20</i>	PGSC0003DMT400022906	2745385	2745966	ST4.03ch09	193	21851.33	8.46
<i>StDUF4228-21</i>	PGSC0003DMT400063216	53980754	53981900	ST4.03ch03	187	20649.6	10.04
<i>StDUF4228-22</i>	PGSC0003DMT400043061	5389486	5390001	ST4.03ch03	171	19379.48	8.49
<i>StDUF4228-23</i>	PGSC0003DMT400036819	57055529	57056170	ST4.03ch03	213	24075.1	8.69
<i>StDUF4228-24</i>	PGSC0003DMT400011206	1562505	1563080	ST4.03ch03	191	21509.43	9.81
<i>StDUF4228-25</i>	PGSC0003DMT400059007	9596151	9596588	ST4.03ch03	145	16139.72	9.48
<i>StDUF4228-26</i>	PGSC0003DMT400014973	5221989	5224980	ST4.03ch08	201	22562.72	8.65
<i>StDUF4228-27</i>	PGSC0003DMT400031886	56747122	56747694	ST4.03ch08	190	21361.3	8.52
<i>StDUF4228-28</i>	PGSC0003DMT400015362	1998775	1999281	ST4.03ch04	168	18344.09	6.36
<i>StDUF4228-29</i>	PGSC0003DMT400043595	49123526	49124140	ST4.03ch04	204	22937.24	8.46
<i>StDUF4228-30</i>	PGSC0003DMT400025724	72069470	72070060	ST4.03ch04	196	21949.27	9.85
<i>StDUF4228-31</i>	PGSC0003DMT400029320	108614	109869	ST4.03ch10	214	23894.93	5.79

Table 1  
List of identified of miRNA Target StDUF4228

miRNA_Acc.	Target_Gene	Expectation	UPE\$	miRNA_start	miRNA_end	Target_start	Target_end	miRNA_aligned_fragment	a
stu-miR8001b-3p	StDUF4228-1	3.5	-1	1	24	390	413	GGAUUUUCAUACUAAUUCUAGAA	:: ::
stu-miR827-5p	StDUF4228-24	3.5	-1	1	21	182	202	UUUGUUGAUGGUCAUCUAUUC	:
stu-miR167a-5p	StDUF4228-30	4	-1	1	21	140	160	UGAAGCUGCCAGCAUGAUCUA	::
stu-miR167b-5p	StDUF4228-30	4	-1	1	21	140	160	UGAAGCUGCCAGCAUGAUCUA	::
stu-miR167c-5p	StDUF4228-30	4	-1	1	21	140	160	UGAAGCUGCCAGCAUGAUCUA	::
stu-miR167d-5p	StDUF4228-30	4	-1	1	21	140	160	UGAAGCUGCCAGCAUGAUCUA	::
stu-miR5303e	StDUF4228-16	4	-1	1	24	791	814	UUUUGGAGAAUCUGACACGGGUGU	:: ::
stu-miR7990a	StDUF4228-19	4	-1	1	24	213	236	UUCAAAUGAUCGUAACUUGGCCU	:: ::
stu-miR8033-3p	StDUF4228-10	4	-1	1	22	73	94	UCAAUUCUGCAGCUUUAGGAGU	:: ::
stu-miR8035	StDUF4228-20	4	-1	1	22	444	465	UCCAUCUCAAUAUCACUUUCU	:: ::
stu-miR8038a-3p	StDUF4228-20	4	-1	1	21	338	358	GUUCAACUUGCUCACUUGGAG	:: ::
stu-miR8038b-3p	StDUF4228-20	4	-1	1	21	338	358	GUUCAACUUGCUCACUUGGAG	:: ::
stu-miR156d-3p	StDUF4228-13	4.5	-1	1	22	251	272	GCUCUCUAUGCUUCUGUCAUCA	:: ::
stu-miR172e-5p	StDUF4228-1	4.5	-1	1	21	426	446	GCAACAUCAUCAAGAUUCACA	:: ::
stu-miR1886g-3p	StDUF4228-26	4.5	-1	1	22	17	38	UUUCAUUAUGAUUUCUUCUCAU	::
stu-miR1886i-5p	StDUF4228-12	4.5	-1	1	24	4	27	AUGAGAUGAAAUJAGCGUUUGGAU	:: ::
stu-miR319a-5p	StDUF4228-28	4.5	-1	1	21	447	467	AGAGCUUUCUUCGGUCCACAC	:: ::
stu-miR482a-5p	StDUF4228-9	4.5	-1	1	21	186	206	GGAAUUGGUGGAUUGGAAAGC	:: ::
stu-miR8011a-5p	StDUF4228-10	4.5	-1	1	24	182	205	UUGUGUGAGGUUUCUUUUUGUUUC	:: ::
stu-miR8015-5p	StDUF4228-19	4.5	-1	1	24	769	792	UAUUGGAUUAUGAAAUGAAACUU	:: ::
stu-miR8020	StDUF4228-26	4.5	-1	1	24	7	29	AAUUUCAUUGAGUAUGUUGUUGUU	:: ::
stu-miR8037	StDUF4228-12	4.5	-1	1	22	233	255	AUAAUUUGGAGGAAUAG-GAACC	:: ::
stu-miR8040-3p	StDUF4228-14	4.5	-1	1	21	385	405	CUUAUAAUUGUAAUUAUGAUC	:: ::
stu-miR8047	StDUF4228-3	4.5	-1	1	21	134	154	CCAUUUUUCGAAAUAGACC	:: ::
stu-miR827-5p	StDUF4228-23	4.5	-1	1	21	427	447	UUUGUUGAUGGUCAUCUAUUC	:: ::
stu-miR156a	StDUF4228-7	5	-1	1	21	133	153	UUGACAGAAGAUAGAGAGCAC	:: ::

miRNA_Acc.	Target_Gene	Expectation	UPE\$	miRNA_start	miRNA_end	Target_start	Target_end	miRNA_aligned_fragment	a
stu-miR156b	StDUF4228-7	5	-1	1	21	133	153	UUGACAGAAGAUAGAGAGCAC	:: ::
stu-miR156c	StDUF4228-7	5	-1	1	21	133	153	UUGACAGAAGAUAGAGAGCAC	:: ::
stu-miR156d-5p	StDUF4228-7	5	-1	1	21	133	153	UUGACAGAAGAUAGAGAGCAC	:: ::
stu-miR167b-3p	StDUF4228-13	5	-1	1	20	430	449	GAUCAUGUGGCAGCAUCACC	:: ::
stu-miR172e-5p	StDUF4228-31	5	-1	1	21	300	320	GCAACAUCAUCAAGAUUCACA	::
stu-miR482a-5p	StDUF4228-4	5	-1	1	21	333	353	GGAAUUGGUGGAUUGGAAAGC	:: ::
stu-miR482d-5p	StDUF4228-24	5	-1	1	21	323	343	CGUGAGUGGUGGGUAAGUA	:: ::
stu-miR5303j	StDUF4228-9	5	-1	1	24	351	374	AAUAUUUUUGAAGAGUCUGAGCAA	:: ::
stu-miR6149-3p	StDUF4228-18	5	-1	1	22	284	305	UGAUUCAGGUUUUGUAUGCAAAC	:: ::
stu-miR7981-3p	StDUF4228-19	5	-1	1	24	799	822	AUAGGACUUUAGUUUAGUUAAAGGU	:: ::
stu-miR7985	StDUF4228-16	5	-1	1	24	542	565	CGGGCUUGCCUAGAACGGGUUACC	:: ::
stu-miR7987	StDUF4228-22	5	-1	1	24	44	68	ACACUAGUUGAUCUU-AUUGAUGAC	:: ::
stu-miR7990b	StDUF4228-28	5	-1	1	24	24	47	GAAUUUUCAAUAUGAUCGUAACUUU	:: ::
stu-miR7991a	StDUF4228-3	5	-1	1	24	330	353	AGGAGGUCGGAAUUUUUAAUGAAU	:: ::
stu-miR7991b	StDUF4228-3	5	-1	1	24	330	353	AGGAGGUCGGAAUUUUUAAUGAAU	:: ::
stu-miR7991c	StDUF4228-3	5	-1	1	24	330	353	AGGAGGUCGGAAUUUUUAAUGAAU	:: ::
stu-miR7996a	StDUF4228-29	5	-1	1	24	452	475	AUGUGGUACAUUGAAAUUUGAAA	:: ::
stu-miR7996b	StDUF4228-29	5	-1	1	24	452	475	AUGUGGUACAUUGAAAUUUGAAA	:: ::
stu-miR7996c	StDUF4228-29	5	-1	1	24	452	475	AUGUGGUACAUUGAAAUUUGAAA	:: ::
stu-miR7999-5p	StDUF4228-6	5	-1	1	24	536	559	CUGGGUCACUUCUACGGGUCCUUC	:: ::
stu-miR8001b-5p	StDUF4228-2	5	-1	1	24	11	34	AUGGGGAUUAGUAUGAAAUUUGC	::
stu-miR8002-5p	StDUF4228-20	5	-1	1	24	367	390	UUUUUCGUGAUAAUAAUGGAAUCA	:: ::
stu-miR8007b-5p	StDUF4228-6	5	-1	1	24	569	592	AUGUGACACUUUUUGAAUUUCGAG	:: ::
stu-miR8014-3p	StDUF4228-22	5	-1	1	24	242	265	AUGAAUACAAUGUUUGGAUAAAUU	:: ::
stu-miR8019-3p	StDUF4228-4	5	-1	1	24	94	117	AAAAGAAUGACCUGGUUUGACUUG	:: ::
stu-miR8023	StDUF4228-19	5	-1	1	24	792	815	UUUGGCACAAUUUCAUUGGCAACC	:: ::
stu-miR8032d-3p	StDUF4228-21	5	-1	1	22	407	427	AGUGUGAGUUGGUGCGAUUAGG	:: ::
stu-miR8033-3p	StDUF4228-2	5	-1	1	22	250	271	UCAAUUCUGCAGCUUUAGGAGU	:: ::

miRNA_Acc.	Target_Gene	Expectation	UPE\$	miRNA_start	miRNA_end	Target_start	Target_end	miRNA_aligned_fragment	a
stu-miR8040-3p	StDUF4228-31	5	-1	1	21	607	627	CUUAUAAUUGUAAUUUAUGAUC	:
stu-miR8041a-5p	StDUF4228-18	5	-1	1	20	45	64	GUGCUUUGCUAUUUUCAUUG	::
stu-miR8041b-5p	StDUF4228-18	5	-1	1	20	45	64	GUGCUUUGCUAUUUUCAUUG	::
stu-miR8048-5p	StDUF4228-25	5	-1	1	21	373	393	CUCAUUAGCAUCUCCAUCUUG	:
stu-miR827-5p	StDUF4228-16	5	-1	1	21	850	870	UUUGUUGAUGGUCAUCUAUUC	:

Table 2  
Gene duplication and selection pressure of StDUF4228. Ka: no. of nonsynonymous substitutions per nonsynonymous site, Ks: no. of synonymous substitutions per synonymous site, MYA: million years ago.

Seq_1	Seq_2	Ka	Ks	Ka_Ks	Time(MYA)
StDUF4228-1	StDUF4228-22	0.168109464	0.989504106	0.169892639	190.2892512
StDUF4228-2	StDUF4228-25	0.245129022	1.045515748	0.234457513	201.0607207
StDUF4228-6	StDUF4228-7	0.051053921	0.181191713	0.281767419	34.84456026
StDUF4228-11	StDUF4228-15	0.119124683	0.439450788	0.271076276	84.50976689
StDUF4228-12	StDUF4228-29	0.236721002	1.014966048	0.233230464	195.1857786
StDUF4228-17	StDUF4228-24	0.561923199	1.302641069	0.431372242	250.5078979
StDUF4228-18	StDUF4228-21	0.196926033	0.716174658	0.274969284	137.7258957
StDUF4228-19	StDUF4228-23	0.878722096	2.352538655	0.37352079	452.4112798

## Phylogenetic analysis of StDUF4228

The evolutionary relationships of 31 *StDUF4228* were investigated by constructing a phylogenetic tree using MEGA7.0 software, using the neighbor-joining method. The multiple sequence alignment demonstrated critical conserved residues in *DUF4228* (Fig. 1). The phylogenetic tree was divided into Six groups: Group 1, Group 2, Group 3, Group 4, Group 5, and Group 6. Of these, Group 4 was the largest. In *StDUF4228*, all genes were distributed into all six groups (Fig. 2).

## Gene structure organization and motif analysis

For phylogenetic reconstruction supported by performing structure analysis of exon-intron with the comparison of genomic and coding DNA sequences. The exon-intron distribution, length, and number were diverse in all genes represented in Fig. 2a, while the *StDUF4228-9* gene was identified as the longest sequence and *StDUF4228-5* was identified as the smallest sequence among all *StDUF4228* genes (Figure.3B). To identify the architecture of StDUF4228 proteins in potato, here used StDUF4228 amino acid sequences and submitted them for online motif analysis by using MEME software. After the analysis 10 conserved motifs were observed in potato (Fig. 3A). All *StDUF4228* genes had motif 1. In the present study motif, 2 was present on all the genes except *StDUF4228-8*. Motif 1 has 29 amino acids while motif 6 and 10 had 41, motif 2 has 21 amino acids, and motif 9 has 39 amino acids.

## Cis-Element Analysis and Chromosomal distribution

For Cis-Regulatory elements analysis, we have used the 2000bp promoter sequence region of the 10 *StDUF4228* gene represented in Fig. 02. The *Cis*-element analysis by PlantCARE predicted that most *Cis*-acting sites were from three groups: phytohormones-responsive, growth- and development-related, and stress-responsive (Lescot et al., 2002). For instance, the MYB binding site and light responses were the main element considered a growth- and development-responsive factor. In contrast, the anaerobic induction and defense and stress responsive elements were enriched in their promoters among stress-response elements. For phytohormones, the GBRE, ABRE, and MeJRE response factors were observed highly enriched. As a result, it indicated that the expression of the *StDUF4228* gene was carried out through different *Cis*-Regulatory elements (Fig. 4). The *StDUF4228* chromosomal distribution illustrated that all genes are unequally distributed on entire chromosomes. The higher number of three genes was present on Chr2 and Chr4 while Chr3 had two genes while Chr1 and Chr1 had only one gene (Table.1).

## Gene Duplications of *DUF4228* Genes

To study the molecular evolution rate for all duplicating genes were computed through Ka/Ks value estimation. The Ka/Ks > 1 was measured as a positive selection effect. Ka/Ks < 1 was considered purifying selection value, and Ka/Ks = 1 was considered as neutral selection value between all duplicating gene pairs (Yang and Bielawski, 2000). Our results show that most of the *DUF4228* duplicated genes endured purifying selection pressure during the duplication process, implying that the function of duplicated *DUF4228* genes might not change significantly in the succeeding evolutionary process. In addition, the deviation time between pairs of duplicated genes was also estimated. The cosmic mass of *DUF4228* genes showed a Ks value > 0.52, whereas the significant time deviation can be greater than 100 MYA (million years ago). Interestingly, the value of Ks for duplicated genes (*StDUF4228-19/StDUF4228-23*) was 2.35, whereas the significant time duplication may be 452.41 MYA (million years ago) (Table 02).

### Analysis of miRNA Targeting *StDUF4228* Genes

In the recent past, several studies have revealed that miRNA-mediated regulation accompanies the stress responses in plants. Thus, for a deep understanding of miRNA-mediated post-transcriptional regulation of *StDUF4228*, we predicted 53 miRNAs targeting 25 genes. Some of the miRNA-targeted sites are presented in Table (3), whereas the detailed information of all miRNAs targeted genes/sites is presented in Table S6. Our results showed that four members of the *stu-miR167a-5p*, *stu-miR167b-5p*, *stu-miR167c-5p*, and *stu-miR167d-5p* targeted *StDUF4228-30*. For members *stu-miR8035*, *stu-miR8038a-3p*, *stu-miR8038b-3p* and *stu-miR8002-5p* targeted *StDUF4228-20*. Three members *stu-miR7996a*, *stu-miR7996b* and *stu-miR7996c* targeted one gene (*StDUF4228-29*) (Table.3). One member *stu-miR827-5p* targeted three genes *StDUF4228-24*, *StDUF4228-23*, and *StDUF4228-16*. Furthermore one member *stu-miR8033-3p* targeted gene *StDUF4228-10*, and *StDUF4228-2*. Two genes *StDUF4228-1* and *StDUF4228-31* targeted by one member *stu-miR172e-5p*. one member *stu-miR482a-5p* targeted genes *StDUF4228-4* and *StDUF4228-4*.

Mainly, *StDUF4228-3*, *StDUF4228-7*, *StDUF4228-19*, *StDUF4228-20* and *StDUF4228-30* were prophesied to be targeted by a greater number of miRNAs. The expression levels of these miRNAs and their targeted genes require validation in additional research to govern their biological roles in the rapeseed genome.

## Comparative Synteny Analysis of Identified *DUF4228* Protein Sequences

The analysis of comparative synteny among *Arabidopsis thaliana* (*Arabidopsis*), *Solanum lycopersicum* (tomato), and *Solanum tuberosum* (potato) reported a remarkable association in terms of gene expression, gene evolution, gene function, gene duplication, and triplication. It was observed that *Solanum lycopersicum* (tomato) Solyc02g091180.1 gene sequence demonstrated synteny with the *Solanum tuberosum* (potato) *StDUF4228-1* gene sequences. Correspondingly, *Solanum tuberosum* (potato) gene *StDUF4228-31* showed synteny with tomato Solyc10g005020.3. Potato *StDUF4228-29* presented synteny with tomato Solyc04g057830.1 (Fig. 5).

## Tissue-specific analysis expression of *StDUF4228* genes

To find out the *StDUF4228* genes expression analyses in different tissues, including root, stem, and leaf, a widely accessible RNA database sequence of potato was investigated. Our consequences showed that most *StDUF4228* genes illustrated relatively large transcriptional abundance in the three tissues described above (Fig. 6). Furthermore, some members were showed just in 1 tissue but not showed these members in other tissues. The *StDUF4228-3*, *StDUF4228-4*, and *StDUF4228-30* genes were expressed in all tissues (root, stem, and leaf). Additionally, 8 *StDUF4228* genes were expressed in leaves. Maximum numbers of *StDUF4228* genes were expressed, which were 15 in the root. In stem 14 *StDUF4228* genes were expressed. Besides, the *StDUF4228-20* genes were illustrated high expression in root tissues, while in leaf tissue *StDUF4228-30* gene showed high expression. Gene *StDUF4228-3* expression was observed maximum in the stem. The profiling data expression of all potato *StDUF4228* genes was studied to establish the heat-map illustrated in Fig. 6. Moreover, our consequences represented that genes have been differentially expressed in three organs and highly expressed in plant stem and root tissues.

## *StDUF4228* Genes Expression Patterns in Response to Heat, Salt, and Phytohormones

For the expression pattern structure of *StDUF4228* genes in potato at the transcriptional level during heat stress, we deduced that few *StDUF4228* genes might be intricate with heat stress in potato. Under heat stress treatment, the *StDUF4228-4* and *StDUF4228-21* are extraordinarily enhanced in potatoes (Fig. 7), while the expression of *StDUF4228-4* showed more than the *StDUF4228-21* genes. Total 16 *StDUF4228* genes were expressed under heat treatment response. Under salt stress treatment, the *StDUF4228-28* and *StDUF4228-21* are extraordinarily enhanced in potato (Fig. 7), while the expression of *StDUF4228-28* showed more than the *StDUF4228-21* genes. Total 23 *StDUF4228* genes were expressed under salt treatment response. Indole acetic acid and abscisic acid have been chosen to investigate transcriptional responses of *StDUF4228* to hormone treatments. For the expression patterns analysis of *StDUF4228* genes, leaf tissue was treated with abscisic acid. The 18 *StDUF4228* genes were expressed treated with abscisic acid. Out of all, *StDUF4228-21* and *StDUF4228-30* gene models show higher expression when treated with abscisic acid. Likely, to investigate the expression patterns of *StDUF4228* genes, leaf tissue was treated with IAA. The 23 genes were expressed treated with IAA. Out of all, *StDUF4228-21* and *StDUF4228-30* gene models show higher expression when treated with IAA, while *StDUF4228-30* showed more expression than *StDUF4228-21* (Fig. 8). Furthermore, to investigate the expression patterns of *StDUF4228* genes, leaf tissue was treated with GA3. The 23 genes were expressed treated with GA3. Out of all, *StDUF4228-28* and *StDUF4228-30* gene models show higher expression when treated with GA3, while *StDUF4228-30* showed more expression than *StDUF4228-28*.

## mRNA e xpression of *StDUF4228* Genes

In potato, the *StDUF4228* genes transcriptional profile expressions were validated in three tissue types (stem, root, and leaf) qRT-PCR analysis were performed (Fig. 9). The expression analysis validation was carried through RT-qPCR analysis by using *StDUF4228-3*, *StDUF4228-4*, and *StDUF4228-30* genes in stem, root, and leaf. The comparative expression pattern illustrated that the *StDUF4228-4* gene had a higher expression pattern in the root while relative expression of *StDUF4228-30* was higher in leaf than root and stem (Fig. 9).

## Discussion

Potatoes are an economically important crop owing to food importance world widely. During the development and growth process potato faced various stresses (Quandahor et al., 2021; Zaynab et al., 2021c). The DUF4228 proteins belong to the DUF superfamily. The current study illustrated that the DUF4228 family proteins played a vital role in abiotic stress response in plants (Leng et al., 2021)(Qi et al., 2019). Sequence analysis of plant genomes has revealed the presence of DUF4228 genes in *Medicago truncatula* (Albornos et al., 2017). In *Arabidopsis thaliana* three DUF4228 family members for example, *AT1G21010*, *AT1G28190*, and *AT1G10530*, are important in drought stress (Yang et al., 2020). The *CiDUF4228* (*Caragana intermedia DUF4228*) genes member Expression of the *Caragana intermedia DUF4228-3* (*CiDUF4228-3*) genes were significantly up-regulated in low temperature, drought, and dehydration indicating its response in stress condition (Na et al., 2016). However, little information is available for the potato MAPK gene members. In this article, a total of 10 *StDUF4228* genes were observed from the genome sequence of potatoes.

All identified *StDUF4228* genes were classified into six groups, e.g., Group 1, Group 2, Group 3, Group 4, Group 5, and Group 6. The comparative phylogenetic analysis revealed that the organization of St, Sl, and At proteins was relatively similar in groups 1, 2, 3, 4, 5, and 6., indicating that all *StDUF4228* genes in these groups may have descended from a common ancestor. The *StDUF4228* protein structural analysis will provide functional analysis significance. The evolutionary relic illustrated the arrangement of exon-intron that shaped the gene family evolution (Flagel and Wendel, 2009; Moore and Purugganan, 2005). This is associated with prior scientific findings, so as too few genes are subjected to be retained in plants, during evolution, some may not show introns and may show short introns (Mattick and Gagen, 2001). Without or with few introns, the gene expression level is lowered in plants (Mattick and Gagen, 2001). Furthermore, a complex gene structure may represent a quick response against endogenous and exogenous stimuli. (Jeffares et al., 2008). Structural gene analysis results reported that sequences of *StDUF4228* genes showed equal numbers and similar functional characters of introns-exons because they originated during duplication events in the course of evolution (Bari et al., 2018; Waqas et al., 2019). *Cis*-element studies provide vital foundations for further functional analysis of *StMAPK* genes. We observed that each *StDUF4228* promoters restrain more than one or one stress-responsive *cis*-elements, including light, anaerobic induction, low-temperature response elements, MeJA response element, SA response element (SARE), and ABA response element (ABRE). Furthermore, these mentioned *Cis*-elements play an important function through regulating stress-responsive genes in stress conditions (Wu et al., 2014). Thus, in *StDUF4228*, these significant *Cis*-acting sites suggest their response in different environmental stresses (Maruyama-Nakashita et al., 2007; Osakabe et al., 2014).

Over the past few years, abundant miRNAs have been recognized through genome-wide examination in rapeseed to participate in diverse environmental factors (Buhtz et al., 2008; Fu et al., 2019). The present study identified 53 miRNAs from different families targeting 25 *StDUF4228* genes (Fig. 6a; Table S6). Numerous genomic and proteomic studies have elucidated that plants' response to saline and other stresses comprises a broad spectrum of processes, such as protein biosynthesis, membrane trafficking, and signal transduction (Bej and Basak, 2014), and it has been well established that miRNAs and their targets influence directly on plant stress tolerance (Li, 2015; Villanueva et al., 2016). In this context, in *Arabidopsis*, maize, and cowpea an upregulation of miR156, of their corresponding targets such as SBPs/SPLs, TCP family transcription factor, ARFs, RNaseIII CAF protein, AGO1, and CBF during salinity stress has been documented (Ding et al., 2009; Liu et al., 2008). Similarly, it has been widely reported that miR156 has a significant role during abiotic stress conditions in many plants (Cui et al., 2014; Kohli et al., 2014). Similarly, a microarray experiment on cotton revealed that miR827 has a crucial role of salt stress responses (Covarrubias and Reyes, 2010). miR167 was previously reported to have key roles in abiotic stress responses (Khraiwesh et al., 2012). Shortly, these reports support our results and recommend that bna-miRNAs might play pivotal roles against several stresses by altering the transcript levels of *StDUF4228* genes in rapeseed.

The genome duplications, distribution of genes, and genome size are the chief feature of genetic diversity between land plants. Genetic duplication has long been observed in the above factors during gene families' evolutionary origins, complexity, and expression. Some duplication events were also discovered in *StDUF4228*, which play an important function in amplifying the *StDUF4228* genes. As gene duplication has been an important factor in diversification, expansion, and neofunctionalization of gene families (Lavin et al., 2005), similarly the *StDUF4228* gene mapping and distribution at a chromosomal level will provide the potato breeders through desired traits to develop novel varieties of potatoes.

The function of *DUF4228* genes has been suggested in different abiotic stress factors of plants (Wang et al., 2018). Therefore, the *DUF4228* gene function has not yet been observed in potato. Earlier studies reported a total of 25 *DUF4228* genes in *Arabidopsis*, 52 in *Populus trichocarpa*, 22 in *Aquilegia coerulea*, and 29 in *Medicago truncatula* (Yang et al., 2020). Therefore, *StDUF4228* expression profiling and their validation are helpful for a deep understanding of the potato genome. The recent study illustrated higher transcript abundances of *StDUF4228* in the roots; such indication was also identified in earlier studies (Yang et al., 2020) and supported our data where *StDUF4228-20*, *StDUF4228-21*, and *StDUF4228-28*, show a high expression in root tissues. Our RT-qPCR results show that *StDUF4228-3*, *StDUF4228-4*, and *StDUF4228-30* were specifically upregulated in stem, leaf and root tissues. This illustrates that potato *DUF4228* genes have a significant role in abiotic response against various stresses.

Thus, more studies on the *DUF4228* gene family revealed its significant role in abiotic stress response (Leng et al., 2021). In potato, *StDUF4228-21*, and *StDUF4228-4* are upregulated in response to heat and salt stress. Various signal transduction hormones may affect plants' biochemical and physiological processes (Fatima et al., 2021; Zaynab et al., 2017). Such as, ABA and IAA are considered essential immune responsible hormones in plants. A variety of studies reported that DUF4228 is responsible for stress control and is also involved in hormonal signaling and developmental signaling. Moreover, to find whether the *DUF4228* were expressed through hormonal signaling in potato, the leaves of the potato were treated by ABA, IAA and examined the gene

expression. After IAA and ABA treatment, genes were induced, showed different functions of different *StDUF4228* gene members in IAA- and ABA-induced immune responses. Total 18 *StDUF4228* genes under ABA and 23 *StDUF4228* genes under IAA and GA3 treatment, upregulation illustrated that IAA and ABA perform a vital role in this immune response. Phytohormones were identified to play a role in the *ATDUF4228* genes that support our results (Yang et al., 2020). The expression of genes and their clusters also mentioned a powerful correlation of gene clusters and gene expression in different tissues under different stress factors. This co-expression and co-occurrence show their putative function in the adaptation of plants under diverse environmental stress factors.

## Conclusion

Overall, a total of 31 *DUF4228* genes were identified in potatoes. The comparative studies of evolutionary analysis resulted in the presence of five major groups in the *DUF4228* family. The conserved functional and structural motifs were lying in each *StDUF4228*, with slight changes among members and groups. The presented results provide a deep understanding of major potato plant challenges under biotic stresses. In potato, *StDUF4228-21* and *StDUF4228-4* presented a higher expression against heat and stress response. Under IAA treatment 23 and ABA treatment, 18 genes expressions illustrated that IAA and ABA performed a vital function in the immune response. These consequences provoked an advanced understanding of *StDUF4228* gene family function and gave a fundamental study for further analysis of the *StDUF4228* genes on molecular processes in abiotic stress response in potato.

## Declarations

## Author contributions

MZ designed the study. MZ, Y.S, conducted the experiments and analyzed the data. S.L KAK contributed to writing and editing the manuscript. All authors have read and agreed to the published version of the manuscript.

## Funding

This work was supported by Shenzhen Sustainable Development Science and Technology Project (KCXFZ20201221173404012, KCXFZ20201221173211033). The authors appreciate the support of the Research Center for Advanced Materials Science (RCAMS) at King Khalid University Abha, Saudi Arabia through a Grant KKU/RCAMS/

G002-21.

## Ethics declarations

## Conflict of interest

The authors declare that there are no conflict of interest.

## Research involving human participants and/or animals

This study was Not involving human participants and/or animals

## Informed consent

All authors informed consent.

## References

1. Albornos, L., Martín, I., Labrador, E., and Dopico, B. (2017). Three members of *Medicago truncatula* ST family are ubiquitous during development and modulated by nutritional status (MtST1) and dehydration (MtST2 and MtST3). *BMC plant biology* 17, 1–19.
2. Bari, A., Farooq, M., Hussain, A., Muhammad, T., Abbas, M., Mustafa, G., Karim, A., Ahmed, I., and Hussain, T. (2018). Genome-wide bioinformatics analysis of aquaporin gene family in maize (*Zea mays* L.). *Journal of Phylogenetics & Evolutionary Biology* 6, 197.
3. Bej S, Basak J (2014) MicroRNAs: the potential biomarkers in plant stress response. *American Journal of Plant Sciences* 2014
4. Buhtz, A., Springer, F., Chappell, L., Baulcombe, D. C., and Kehr, J. (2008). Identification and characterization of small RNAs from the phloem of *Brassica napus*. *The Plant Journal* 53, 739–749.
5. Chen, T., Zhu, H., Ke, D., Cai, K., Wang, C., Gou, H., Hong, Z., and Zhang, Z. (2012). A MAP kinase kinase interacts with SymRK and regulates nodule organogenesis in *Lotus japonicus*. *The Plant Cell* 24, 823–838.
6. Covarrubias, A. A., and Reyes, J. L. (2010). Post-transcriptional gene regulation of salinity and drought responses by plant microRNAs. *Plant, Cell & Environment* 33, 481–489.

7. Cui, L. G., Shan, J. X., Shi, M., Gao, J. P., and Lin, H. X. (2014). The miR156-SPL 9-DFR pathway coordinates the relationship between development and abiotic stress tolerance in plants. *The Plant Journal* 80, 1108–1117.
8. Cui Y, Wang M, Zhou H, Li M, Huang L, Yin X, Zhao G, Lin F, Xia X, Xu G (2016) OsSGL, a novel DUF1645 domain-containing protein, confers enhanced drought tolerance in transgenic rice and Arabidopsis. *Frontiers in plant science* 7, 2001
9. Dahal, K., Li, X.-Q., Tai, H., Creelman, A., and Bizimungu, B. (2019). Improving potato stress tolerance and tuber yield under a climate change scenario—a current overview. *Frontiers in plant science* 10, 563.
10. Didelon, M., Khafif, M., Godiard, L., Barbacci, A., and Raffaele, S. (2020). Patterns of Sequence and Expression Diversification Associate Members of the PADRE Gene Family With Response to Fungal Pathogens. *Frontiers in genetics* 11, 491.
11. Ding, D., Zhang, L., Wang, H., Liu, Z., Zhang, Z., and Zheng, Y. (2009). Differential expression of miRNAs in response to salt stress in maize roots. *Annals of botany* 103, 29–38.
12. Fatima, M., Ma, X., Zhou, P., Zaynab, M., and Ming, R. (2021). Auxin regulated metabolic changes underlying sepal retention and development after pollination in spinach. *BMC plant biology* 21, 1–15.
13. Flagel, L. E., and Wendel, J. F. (2009). Gene duplication and evolutionary novelty in plants. *New Phytologist* 183, 557–564.
14. Fu, Y., Mason, A. S., Zhang, Y., Lin, B., Xiao, M., Fu, D., and Yu, H. (2019). MicroRNA-mRNA expression profiles and their potential role in cadmium stress response in *Brassica napus*. *BMC plant biology* 19, 1–20.
15. Guo, C., Luo, C., Guo, L., Li, M., Guo, X., Zhang, Y., Wang, L., and Chen, L. (2016). OsSIDP366, a DUF1644 gene, positively regulates responses to drought and salt stresses in rice. *Journal of integrative plant biology* 58, 492–502.
16. Jeffares, D. C., Penkett, C. J., and Bähler, J. (2008). Rapidly regulated genes are intron poor. *Trends in genetics* 24, 375–378.
17. Khraiwesh, B., Zhu, J.-K., and Zhu, J. (2012). Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* 1819, 137–148.
18. Kim, S. J., Ryu, M. Y., and Kim, W. T. (2012). Suppression of Arabidopsis RING-DUF1117 E3 ubiquitin ligases, AtrDUF1 and AtrDUF2, reduces tolerance to ABA-mediated drought stress. *Biochemical and biophysical research communications* 420, 141–147.
19. Kohli, D., Joshi, G., Deokar, A. A., Bhardwaj, A. R., Agarwal, M., Katiyar-Agarwal, S., Srinivasan, R., and Jain, P. K. (2014). Identification and characterization of wilt and salt stress-responsive microRNAs in chickpea through high-throughput sequencing. *PLoS one* 9, e108851.
20. Lavin, M., Herendeen, P. S., and Wojciechowski, M. F. (2005). Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Systematic biology* 54, 575–594.
21. Leng Z-X, Liu Y, Chen Z-Y, Guo J, Chen J, Zhou Y-B, Chen M, Ma Y-Z, Xu Z-S, Cui X-Y (2021) Genome-Wide Analysis of the DUF4228 Family in Soybean and Functional Identification of GmDUF4228–70 in Response to Drought and Salt Stresses. *Frontiers in plant science* 12
22. Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P., and Rombauts, S. (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, 325–327.
23. Li, L.-H., Lv, M.-M., Li, X., Ye, T.-Z., He, X., Rong, S.-H., Dong, Y.-L., Guan, Y., Gao, X.-L., and Zhu, J.-Q. (2018). The rice OsDUF810 family: OsDUF810. 7 may be involved in the tolerance to salt and drought. *Molecular Biology* 52, 489–496.
24. Li, M., Guo, L., Guo, C., Wang, L., and Chen, L. (2016). Over-expression of a DUF1644 protein gene, SIDP361, enhances tolerance to salt stress in transgenic rice. *Journal of Plant Biology* 59, 62–73.
25. Li, S. (2015). The Arabidopsis thaliana TCP transcription factors: a broadening horizon beyond development. *Plant signaling & behavior* 10, e1044192.
26. Liu, H.-H., Tian, X., Li, Y.-J., Wu, C.-A., and Zheng, C.-C. (2008). Microarray-based analysis of stress-regulated microRNAs in Arabidopsis thaliana. *Rna* 14, 836–843.
27. Luo, C., Guo, C., Wang, W., Wang, L., and Chen, L. (2014). Overexpression of a new stress-repressive gene OsDSR2 encoding a protein with a DUF966 domain increases salt and simulated drought stress sensitivities and reduces ABA sensitivity in rice. *Plant cell reports* 33, 323–336.
28. Maruyama-Nakashita A, Nakamura Y, Saito K, Takahashi H (2007) Identification of a novel cis-acting element in SULTR1; 2 promoter conferring sulfur deficiency response in Arabidopsis roots. In "PLANT AND CELL PHYSIOLOGY", Vol. 48, pp. S34-S34. OXFORD UNIV PRESS GREAT CLARENDON ST, OXFORD OX2 6DP, ENGLAND
29. Mattick, J. S., and Gagen, M. J. (2001). The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. *Molecular biology and evolution* 18, 1611–1630.
30. Moore, R. C., and Purugganan, M. D. (2005). The evolutionary dynamics of plant duplicate genes. *Current opinion in plant biology* 8, 122–128.
31. Osakabe, Y., Yamaguchi-Shinozaki, K., Shinozaki, K., and Tran, L. S. P. (2014). ABA control of plant macroelement membrane transport systems in response to water deficit and high salinity. *New Phytologist* 202, 35–49.
32. Quandahor, P., Gou, Y., Lin, C., Coulter, J. A., and Liu, C. (2021). Comparison of root tolerance to drought and aphid (*Myzus persicae* Sulzer) resistance among different potato (*Solanum tuberosum* L.) cultivars. *Scientific reports* 11, 1–13.
33. Sattar, F. A., Hamooh, B. T., Wellman, G., Ali, M., Shah, S. H., Anwar, Y., and Mousa, M. A. A. (2021). Growth and Biochemical Responses of Potato Cultivars under In Vitro Lithium Chloride and Mannitol Simulated Salinity and Drought Stress. *Plants* 10, 924.
34. Villanueva, M. A., Islas-Flores, T., and Ullah, H. (2016). signaling through WD-repeat proteins in plants. *Frontiers in plant science* 7, 1157.
35. Wang, Y., Zhang, Z., Liu, H., An, Y., Han, B., Wu, Y., Chang, L., Hu, T., and Yang, P. (2018). Overexpression of an alfalfa (*Medicago sativa*) gene, MsDUF, negatively impacted seed germination and response to osmotic stress in transgenic tobacco. *Plant Cell, Tissue and Organ Culture (PCTOC)* 132, 525–534.

36. Waqas, M., Azhar, M. T., Rana, I. A., Azeem, F., Ali, M. A., Nawaz, M. A., Chung, G., and Atif, R. M. (2019). Genome-wide identification and expression analyses of WRKY transcription factor family members from chickpea (*Cicer arietinum* L.) reveal their role in abiotic stress-responses. *Genes & genomics* 41, 467–481.
37. Wu, J., Wang, J., Pan, C., Guan, X., Wang, Y., Liu, S., He, Y., Chen, J., Chen, L., and Lu, G. (2014). Genome-wide identification of MAPKK and MAPKKK gene families in tomato and transcriptional profiling analysis during development and stress response. *PLoS one* 9, e103032.
38. Xin, Z., Mandaokar, A., Chen, J., Last, R. L., and Browse, J. (2007). Arabidopsis ESK1 encodes a novel regulator of freezing tolerance. *The Plant Journal* 49, 786–799.
39. Yang, Q., Niu, X., Tian, X., Zhang, X., Cong, J., Wang, R., Zhang, G., and Li, G. (2020). Comprehensive genomic analysis of the DUF4228 gene family in land plants and expression profiling of ATDUF4228 under abiotic stresses. *BMC genomics* 21, 1–16.
40. Yang, Z., and Bielawski, J. P. (2000). Statistical methods for detecting molecular adaptation. *Trends in ecology & evolution* 15, 496–503.
41. Zaynab, M., Fatima, M., Abbas, S., Sharif, Y., Umair, M., Zafar, M. H., and Bahadar, K. (2018). Role of secondary metabolites in plant defense against pathogens. *Microbial pathogenesis* 124, 198–202.
42. Zaynab M, Hussain A, Sharif Y, Fatima M, Sajid M, Rehman N, Yang X, Khan KA, Ghranh HA, Li S (2021a) Mitogen-Activated Protein Kinase Expression Profiling Revealed Its Role in Regulating Stress Responses in Potato (*Solanum tuberosum*). *Plants* 10, 1371
43. Zaynab, M., Kanwal, S., Furqan, M., Islam, W., Noman, A., Ali, G. M., Rehman, N., Zafar, S., Sughra, K., and Jahanzab, M. (2017). Proteomic approach to address low seed germination in *Cyclobalopsis gilva*. *Biotechnology letters* 39, 1441–1451.
44. Zaynab M, Sharif Y, Abbas S, Afzal MZ, Qasim M, Khalofah A, Ansari MJ, Khan KA, Tao L, Li S (2021b) Saponin toxicity as key player in plant defence against pathogens. *Toxicon*
45. Zaynab, M., Sharif, Y., Fatima, M., Afzal, M. Z., Aslam, M. M., Raza, M. F., Anwar, M., Raza, M. A., Sajjad, N., and Yang, X. (2020). CRISPR/Cas9 to generate plant immunity against pathogen. *Microbial pathogenesis* 141, 103996.
46. Zaynab M, Wang Z, Hussain A, Bahadar K, Sajid M, Sharif Y, Azam M, Sughra K, Raza MA, Khan KA (2021c) ATP-binding cassette transporters expression profiling revealed its role in the development and regulating stress response in *Solanum tuberosum*. *Molecular Biology Reports*, 1–14

## Figures

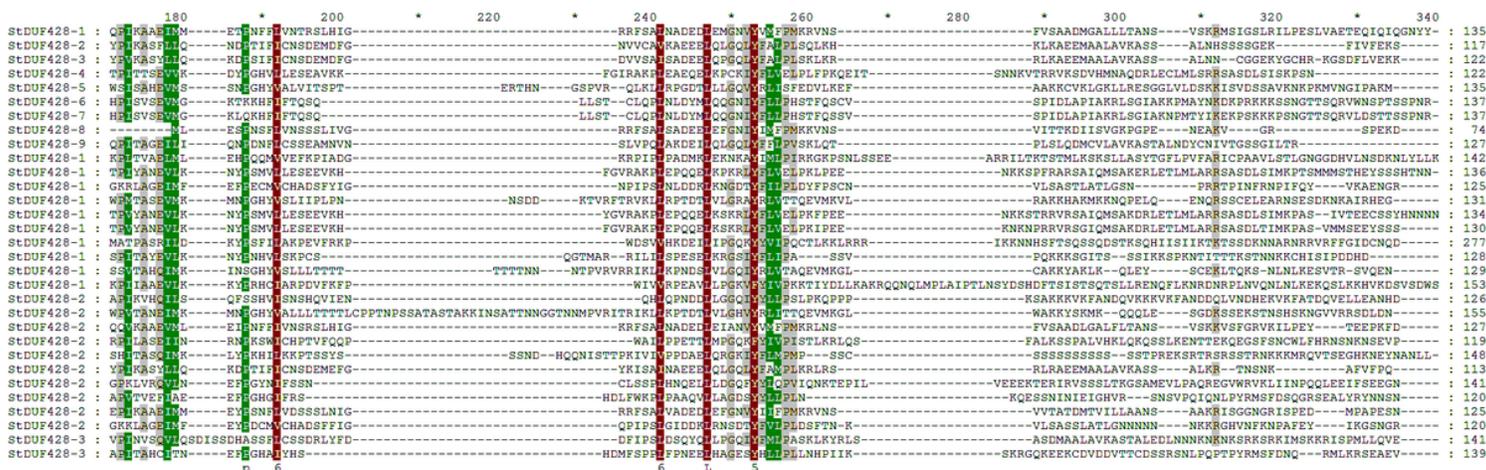


Figure 1

Multiple sequence alignment of StDUF4228

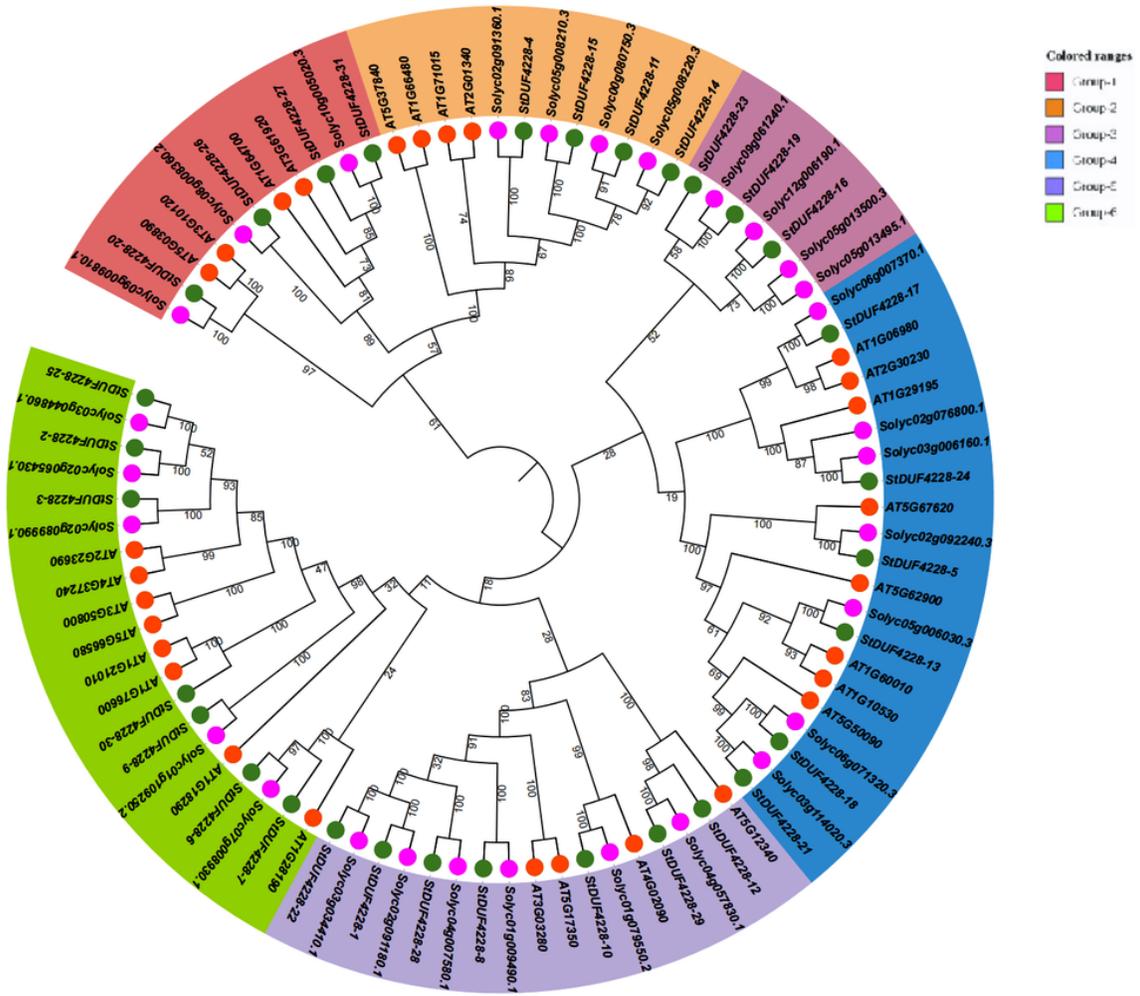


Figure 2

An accumulative phylogenetic tree containing AtDUF4228 (*Arabidopsis thaliana*), SIDUF4228 (*Solanum lycopersicum*), and StDUF4228 (*Solanum tuberosum*). Numbers on the nodes represent bootstrap values.

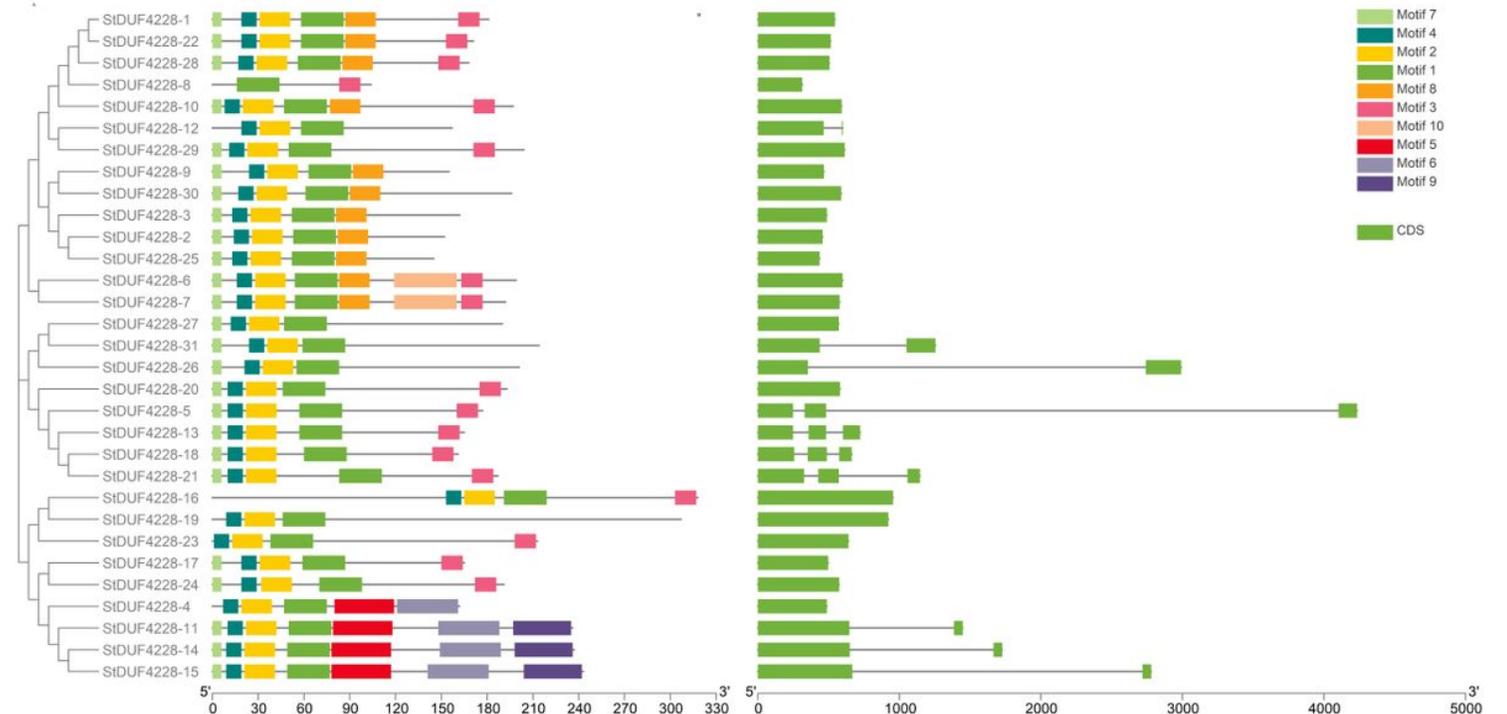


Figure 3

(A) De novo MEME motifs' distributions in the StDUF4228 proteins. B. Distribution of exons, introns, and UTR (untranslated regions) in StDUF4228 gene sequences.

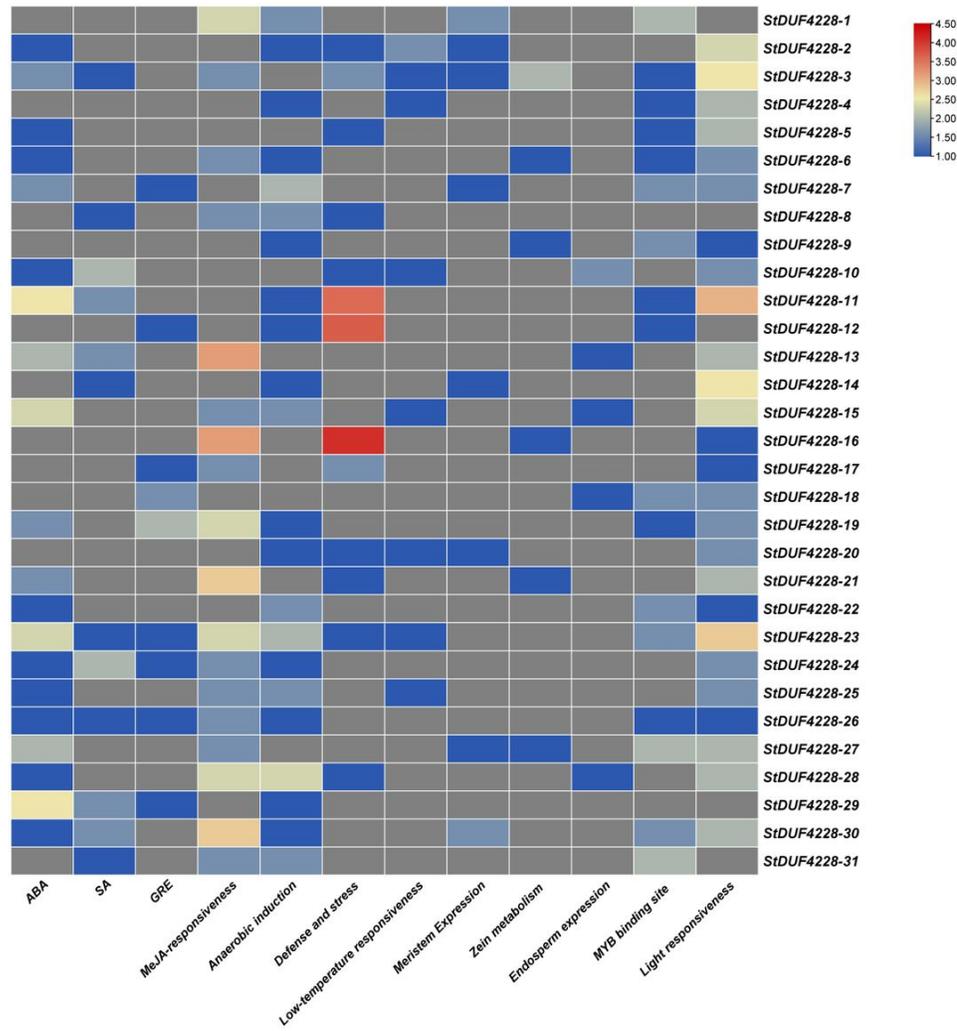


Figure 4

Cis-acting elements' distribution in the regulatory regions of StDUF4228.



Figure 5

Synten analysis of StDUF4228 among Sl, St, and At.

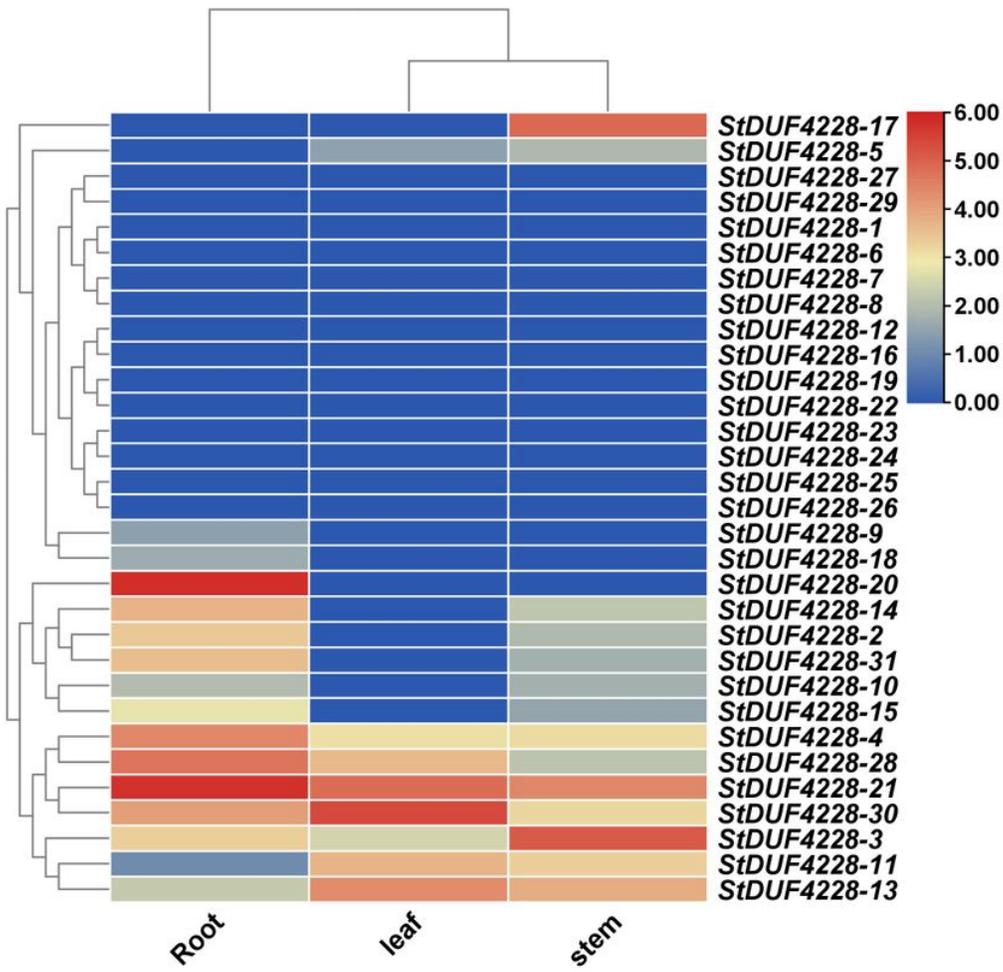


Figure 6

Expression profiling of StDUF4228 genes in leaves, roots, and stems

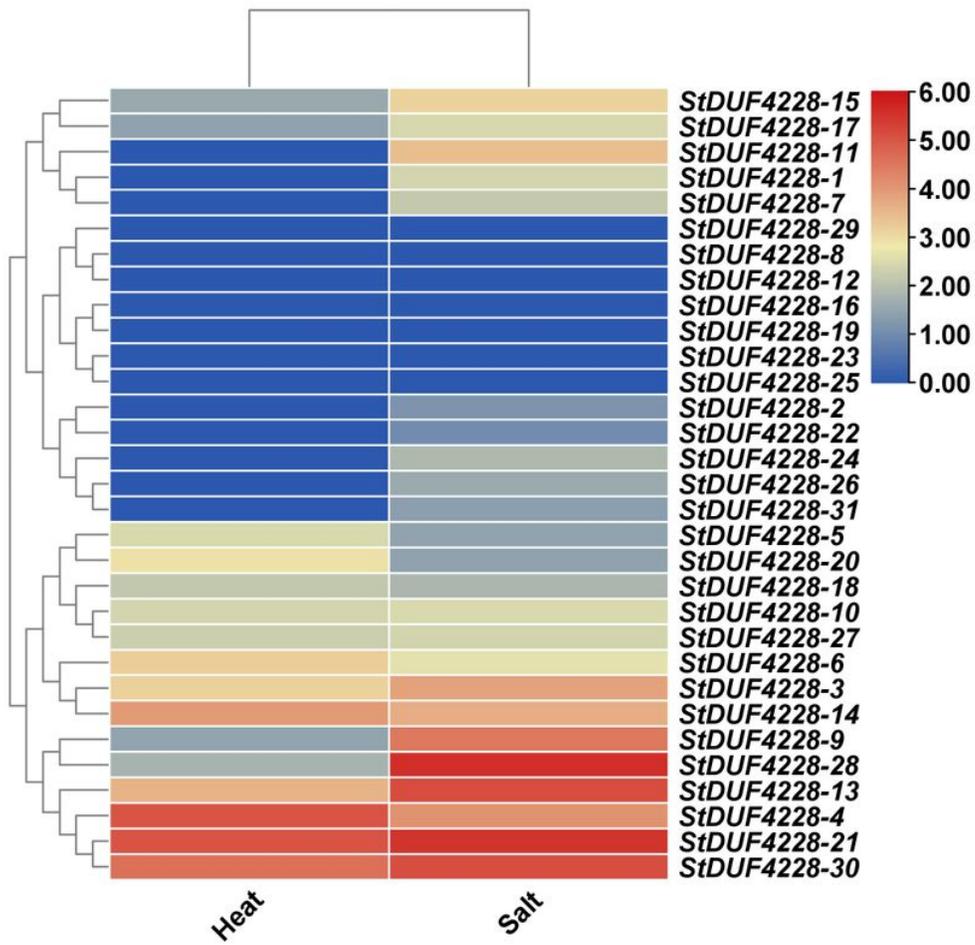


Figure 7

Expression profiling of StDUF4228 genes in response to heat and salt.

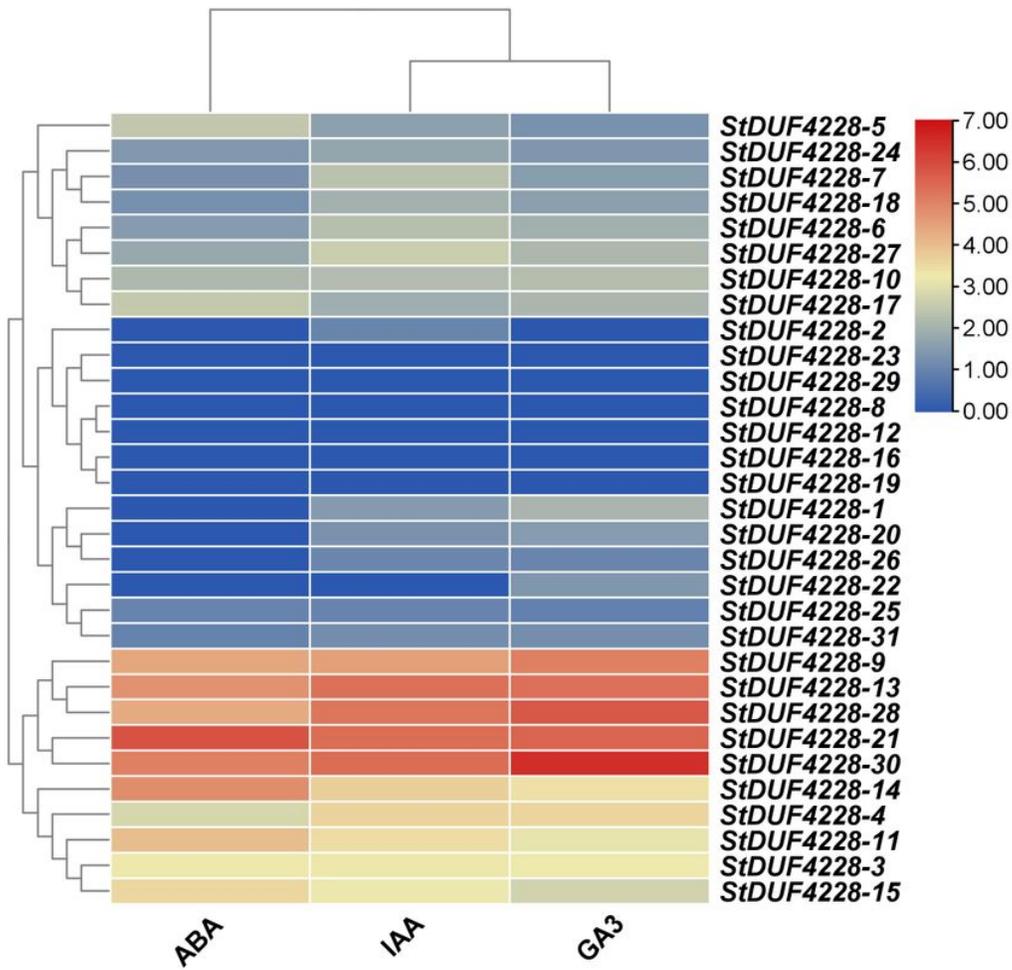


Figure 8

Expression profiling of StDUF4228 genes in response to IAA,GA3 and ABA.



Figure 9

Real-time relative expressions of StDUF4228 in leaves, roots, and stems.