Genome-Wide Mining of Wheat DUF966 Gene Family Provides New Insights Into Salt Stress Responses

Xiaoyi Zhou1,2†, Xiaoguo Zhu2†, Wenna Shao1,2, Jinghan Song1, Wenqiang Jiang1, Yiqin He1, Junliang Yin1, Dongfang Ma1* and Yongli Qiao2*

1 Hubei Collaborative Innovation Center for Grain Industry/Engineering Research Center of Ecology and Agricultural Use of Wetland, Ministry of Education/College of Agriculture, Yangtze University, Jingzhou, China, 2 Shanghai Key Laboratory of Plant Molecular Sciences, College of Life Sciences, Shanghai Normal University, Shanghai, China

Domain of unknown function (DUF) proteins constitute a great deal of families of functionally uncharacterized proteins in eukaryotes. The DUF966 gene family is found in monocotyledons, dicotyledons, mosses, and other species. However, little is known about the functions of DUF966 genes in wheat (Triticum aestivum L.). In this study, we identified and characterized the TaDUF966 gene family members in wheat by in silico analysis. A total of 28 TaDUF966 proteins were identified in wheat. Phylogenetic analysis divided these proteins into two groups (Groups I and II). Proteins in each group showed a highly conserved DUF966 domain and conserved motif distribution, implying their functional conservation. Analysis of gene expression profiling data showed that some TaDUF966 genes were induced by salt stress. We further confirmed the role of TaDUF966-9B in salt stress using virus induced gene silencing (VIGS) assay. Compared with the empty vector control, the TaDUF966-9B knockdown plants exhibited severe leaf curling at 10 days post-inoculation with BSMV under salt stress, suggesting that TaDUF966 genes play a vital role in salt stress tolerance in wheat. Taken together, these results expand our knowledge of the evolution of the DUF966 gene family in wheat and promote the potential application of these genes in wheat genetic improvement.

Keywords: TaDUF966, phylogenetic analysis, Triticum aestivum L., transcriptome analysis, abiotic stresses, virus induced gene silencing

INTRODUCTION

Wheat (Triticum aestivum L.) is one of the most important grain crops in the world and the main source of food for the majority of the human population (Yin J. et al., 2018). Abiotic stresses, such as drought, salinity, and extreme temperatures, negatively impact the growth, development, quality, and yield of wheat and other crops (Zhu et al., 2019a). Domain of unknown function (DUF) is a general term used for many domains in the Pfam database that have not been confirmed. These domains have two distinct characteristics: a relatively conserved amino acid sequence and the function of this domain is unknown (Bateman et al., 2010). DUF proteins are reportedly involved in abiotic stress tolerance in plants. For example, the wheat TaWTF1 gene,
which encodes the Domain of Unknown Function 860 (DUF860) protein, contains stress- and hormone-responsive elements in its promoter region and is induced by heat stress at the seedling and flowering stages (Qin et al., 2013). In rice (Oryza sativa L.), the expression of OsDUF936.3, OsDUF936.5, and OsDUF936.6 is significantly increased under salt stress conditions, and overexpression of OsDUF936.6 in Escherichia coli significantly improves tolerance to salt stress (Li et al., 2017). Additionally, OsSIDP366, a DUF1644 gene, is a positive regulator of drought and salt stress tolerance in rice (Guo et al., 2016). In Ammopiptanthus mongolicus, the AmDUF1517 gene is potentially involved in the regulation of cold tolerance (Gu and Cheng, 2014).

Increasing studies show that genes containing DUF domains are involved in plant growth and developmental processes. Rice REL2 (ROLLED and ERECT LEAF 2 (OsREL2)) is the first DUF domain-containing protein shown to be involved in the regulation of leaf morphology (Yang et al., 2016). Additionally, rice BEAK-SHAPED GRAIN 1 (BSG1), a DUF640 gene, determines grain shape and size, probably by controlling cell division and expansion in the grain hull (Yan et al., 2013). In Arabidopsis thaliana, silencing of the At2g23470 gene, which encodes a member of the DUF647 protein family, results in severe infertility (Li and Zhao, 2012). Li et al. (2016) determined the role of DUF1313 gene family members in photoperiod sensitivity in maize (Zea mays L.). In the woody perennial plant Populus deltoids, PdDUF231A, a DUF231 family protein, is involved in xylan acetylation and cellulose biosynthesis (Yang et al., 2017). In addition to their role in plant growth, development, and abiotic stress response, DUF proteins are also involved in the biotic stress response. For example, silencing of the OsDUF500 gene in rice increases resistance to bacterial blight strain PXO99 (Li et al., 2012). The plant DUF538 protein is predicted as a partial structural homolog of the bactericidal/permeability-increasing (BPI) protein of the mammalian innate immune system, which provides the first line of defense against different pathogens including bacteria, fungi, viruses, and parasites. Experimental evidence shows that exogenous application of the purified fused product of Celosia DUF538 affects bacterial growth, possibly by binding to the bacterial membrane, similar to the BPI protein (Gholizadeh and Kohnehrouz, 2013).

The DUF966 protein superfamily is found in monocotyledons, dicotyledons, mosses, and other species (Luo and Tian, 2017). Proteins in this superfamily contain one or two highly conserved DUF966 domains. In tomato (Solanum lycopersicum), IWS19 (a DUF966 gene) was transiently and rapidly induced by salt stress and was isolated from salt-treated roots (Tirajo, 2005). Although the molecular function of DUF966 genes is not fully understood, the available evidence strongly suggests that DUF966 family genes are involved in abiotic stress tolerance. For example, in rice, overexpression of the new stress suppressor gene, DUF966-stress repressive gene 2 (OsDSR2), dramatically increased the sensitivity to salt and drought stresses and reduced sensitivity to abscisic acid (ABA) (Luo et al., 2014). In Arabidopsis, the AtST39 gene (At3g46110) plays a positive regulatory role in salt and drought stress resistance pathways (Wang, 2014).

Although the DUF966 genes play vital roles in plants, to the best of our knowledge, their functions remain unexplored in wheat. In this study, we conducted a comprehensive analysis of TaDUF966 family genes in abiotic stress tolerance in wheat, based on their genomic sequence and transcriptome data. A total of 28 TaDUF966 family genes were identified, and their conserved motifs, genomic structure, evolutionary relationship, and functional classification were investigated. Moreover, expression analysis of TaDUF966 genes was conducted, and cis-acting elements in their promoter regions were examined. Our results provide valuable information for understanding the classification and putative functions of DUF966 genes.

MATERIALS AND METHODS

Plant Material and Stress Treatment

Common wheat (Triticum aestivum L.) cultivar Zhengmai was used in this study. Plants were grown at 26°C under a 16 h day/8 h dark photoperiod. To conduct salt stress treatments, plants were treated with 150 mM sodium chloride (NaCl) at the stages of seedling with one and half leaves. Leaf and root samples were collected at 2, 4, 8, 12, 24, and 96 h after the salt treatment (Jiang et al., 2019). The experiment was performed with three independent biological replicates.

Identification of TaDUF966 Genes and Sequence Analysis

The TaDUF966 gene family members were identified by in silico analysis of the wheat reference genome assembly (RefSeq v1.1; International Wheat Genome Sequencing Consortium [IWGSC]) (https://wheat-urgi.versailles.inra.fr/Seq-Repository/Assemblies) (He et al., 2020). To obtain the amino acid sequences of the putative TaDUF966 family proteins, BLASTp (Basic Local Alignment Search Tool protein) analysis was conducted using seven rice DUF966s (OsDSRs) (http://rice.plantbiology.msu.edu/index.shtml) and five Arabidopsis (https://www.arabidopsis.org/index.jsp) DUF966s (AtUOFs) as query sequences against the wheat reference IWGSC v1.1 (Jiang et al., 2020a). To ensure the reliability of TaDUF966 protein sequences, a cut-off value (E < 1 × 10^-21) was used. The TaDUF966 family members were further screened and validated using the SMART (Letunic et al., 2004) and Pfam databases (version 32.0, PF06136) (http://pfam.xfam.org/).

Phylogenetic Analysis, Chromosomal Locations, and Duplication Patterns

Phylogenetic analysis of TaDUF966 proteins was conducted using the LG model (Le and Gascuel, 2008) based maximum likelihood (ML) method in MEGA 7.0 (Kumar et al., 2016). The test guidance was repeated 1,000 times (Felsenstein, 1985) to calculate the support for each node. A midpoint-rooted base tree was drawn using the Interactive Tree of Life (IToL, version 3.2.317, http://itol.embl.de) (Zhu et al., 2019b). The GFF3 gene annotation file was obtained from the wheat database IWGSC v1.1,
and TaDUF966 gene annotations were extracted. Chromosomal start and stop sites and the chromosomal location of each TaDUF966 gene were used to draw the physical map with the MapInspect software. All possible homolog of the DUF966 start and stop sites and the chromosomal location of each in each subgenome of common wheat were determined using “all-against-all” BLAST, with an E-value cut-off of $1 \times 10^{-10}$ and identity $> 75\%$ (Letunic et al., 2004). The R package “circize” was used to draw a graph showing its position and homology relationship (Fang et al., 2020). The non-synonymous (K$_a$) and synonymous (K$_s$) substitution rates were calculated using DnaSP 6.0 (Rozas et al., 2017) to diagnose the form of sequence evolution.

**Gene Structure and Predicted Protein Feature Analyses**

The predicted TaDUF966 protein sequences were analyzed using protein identification and analysis tools available online. The protein length, isoelectric point (pI), molecular weight (MW), instability, and index were predicted using the ProtParam (Gasteiger et al., 2005) tool (https://web.expasy.org/protparam/). The subcellular localization of these proteins was predicted using the Plant-mPlLoc online tool (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/) (Chou and Shen, 2010).

The MEME (v5.0.3) (Bailey and Elkan, 1994) suite analysis tool and MAST (http://meme-suite.org/tools) motif search tool were used to identify conserved motifs in TaDUF966 proteins. The known DUF966 protein sequences, including AtDUF966 and OsDUSR, were used for training the parameter settings. Then, the trained parameters were used to identify conserved TaDUF966 motifs as follows: each sequence may contain any number of non-overlapping repeats of each motif; number of different motifs (1–10); motif width = 6–50 amino acids (aa). The functions of predicted motifs were analyzed using InterPro (http://www.ebi.ac.uk/interpro) and SMART motif search programs (http://smart.embl-heidelberg.de/SMART). In addition, multiple sequence alignment analysis of TaDUF966 proteins was conducted using ClustalX2 (Kim and Joo, 2010).

**Prediction of Cis-Acting Elements**

To analyze putative cis-elements in TaDUF966 gene promoters, the 1,500 bp sequence upstream of the start codon of each gene was searched in the Plant-CARE database (Bailey and Elkan, 1994). The identified motifs were analyzed using the Fisher test, and those with an adjusted $p$-value $< 0.05$ were considered significantly enriched.

**Analysis of Wheat Transcriptome Under Salt Stress**

Transcriptome responses of two wheat cultivars to salt stress RNA-seq datasets generated previously under an abiotic stress treatments were downloaded from NCBI and mapped to the wheat reference genome using HISAT2. Then, genes were assembled using cufflinks to inspect the expression levels of TaDUF966 genes (normalized by TPM, transcripts per kilobase of exon model per million mapped reads). The R package “heatmap” was used to draw the heatmap of TaDUF966 genes (Zhu et al., 2019c).

**RESULTS**

**Identification and Classification of TaDUF966 Genes**

A total of 28 TaDUF966 proteins were identified from wheat reference genome using seven rice and five Arabidopsis protein sequences as query (Table S4). These wheat proteins were named according to the following rules: (1) “TaDUF966” stands for the hexaploid wheat DUF966 gene family; (2) Arabic numerals represent the gene number; (3) “A/B/D” represents the wheat subgenome to which the gene belongs [Table 1; (Song et al., 2019)].
### Table 1 | Information of the DUF966 gene family in common wheat as well as its wild relatives.

| Designation | Protein ID | Chr | Location |
|-------------|------------|-----|----------|
| TaDUF966-1A | TraesCS1A02G3590300LC | 1A | 396288719-396292574 |
| TaDUF966-2B | TraesCS2A02G0335500 | 2B | 17040006-17041737 |
| TaDUF966-3A | TraesCSA02G3318800 | 3A | 561034605-561037042 |
| TaDUF966-3B | TraesCSB02G3472100 | 3B | 556645507-556648123 |
| TaDUF966-4B | TraesCSD02G3212500 | 4B | 426822130-426834648 |
| TaDUF966-4C | TraesCSA02G3567000 | 4C | 604342231-604346064 |
| TaDUF966-4D | TraesCSD02G3389500 | 4D | 612826563-612890662 |
| TaDUF966-5B | TraesCSB02G3612100 | 5B | 829508771-829510876 |
| TaDUF966-5D | TraesCSD02G3540500 | 5D | 611889001-611891152 |
| TaDUF966-6A | TraesCSA02G3908000 | 6A | 513678404-513683855 |
| TaDUF966-6B | TraesCSB02G4090800LC | 6B | 153226727-153227568 |
| TaDUF966-9B | TraesCSA02G3290800 | 9B | 30147071-301500991 |
| TaDUF966-9B | TraesCSB02G3267700 | 9B | 541363253-541385628 |
| TaDUF966-9D | TraesCSB02G2673900 | 9D | 437906505-437913999 |
| TaDUF966-9D | TraesCSB02G2908000 | 9D | 589102284-589116112 |
| TaDUF966-9D | TraesCSB02G3092000 | 9D | 23818716-23829288 |
| TaDUF966-9D | TraesCSB02G3297000 | 9D | 13196531-13201007 |
| TaDUF966-10A | TraesCSA02G311800 | 10A | 603496846-603499370 |
| TaDUF966-10B | TraesCSB02G0006000 | 10B | 556850-559568 |
| TaDUF966-10D | TraesCSB02G0011000LC | 10D | 1257755-1260502 |
| TaDUF966-11B | TraesCSB02G3196600LC | 11B | 34797620-347978519 |
| TaDUF966-12B | TraesCSB02G3681100LC | 12B | 64827617-648274459 |
| TaDUF966-13A | TraesCSA02G1170000LC | 13A | 49204982-492056642 |
| TaDUF966-13A | TraesCSA02G2818000 | 13A | 306345454-306348514 |
| TaDUF966-14A | TraesCSB02G1796000 | 14A | 265430055-265433243 |
| TaDUF966-14A | TraesCSB02G2803000 | 14A | 272597397-272620133 |
| TuDUF966-1A | TuG1812U00002574.01.T01 | 1A | 305051841-305065387 |
| TuDUF966-2A | TuG1812U000104100.01.T01 | 2A | 305051841-305065412 |
| TuDUF966-3A | TuG1812U00003957.01.T01 | 3A | 596116058-596119041 |
| TuDUF966-4A | TuG1812U00000777.01.T01 | 4A | 4169023-4170232 |
| TuDUF966-4B | TuG1812U00000309.01.T01 | 4B | 15533663-15538158 |
| TuDUF966-5A | TuG1812U00003116.01.T01 | 5A | 557235547-557238342 |
| TuDUF966-5B | TuG1812U000024800.01.T01 | 5B | 4271740-42729247 |
| TuDUF966-6A | TuG1812U00003981.01.T01 | 6A | 51135861-511369171 |
| TuDUF966-6A | TuG1812U00003981.01.T01 | 6A | 51135861-511369171 |
| TuDUF966-7A | TuG1812U0000575500.01.T01 | 7A | 40-4352 |
| TuDUF966-7A | TuG1812U0000575500.01.T01 | 7A | 40-4352 |
| TuDUF966-10A | TRIDC4A043530.1 | 10A | 595442094-595443469 |
| TuDUF966-11B | TRIDC4B0001500 | 11B | 4480304-4482866 |
| TuDUF966-12B | TRIDC4B0005350 | 12B | 22750429-22754174 |
| TuDUF966-13B | TRIDC4B0046350 | 13B | 540704740-540706829 |
| TuDUF966-14A | TRIDC7A0373900.1 | 14A | 303463007-303467206 |
| TuDUF966-15B | TRIDC7B0028580.1 | 15B | 275715194-275718556 |
| AeDUF966-1D | AET2G020621000.2 | 1D | 304624693-304647592 |
| AeDUF966-1D | AET3G020385200.1 | 1D | 156184974-156185692 |
| AeDUF966-3D | AET3G020731600.1 | 3D | 434177880-434179497 |
| AeDUF966-4D | AET3G02082400.1 | 4D | 49672272-496724099 |
| AeDUF966-5D | AET3G121480000.1 | 5D | 628342697-628345442 |
| AeDUF966-6D | AET4G020040000.1 | 6D | 894539-897603 |
| AeDUF966-7D | AET4G020659700.1 | 7D | 14460999-14465228 |
| AeDUF966-8D | AET4G020567800.1 | 8D | 44956872-449589873 |
| AeDUF966-9D | AET7G020681900.2 | 9D | 274570324-274576316 |
We identified 7 triads TaDUFs (Table S5). To better analyze the phylogenetic relationships among wheat and other plant DUF966 proteins, we constructed a phylogenetic tree using DUF966 sequences from wheat, rice, and Arabidopsis (Figure 1A). All DUF966 proteins were divided into two major groups (Groups I and II) (Figure 2A). Each group contained 14 members, and TaDUF966 sequences were present in both groups. Group II contained DUF966 proteins from only wheat and rice, indicating that DUF966 genes in wheat were more closely related to those in rice (Figure 1A).

Chromosomal Location of TaDUF966

The identified 28 TaDUF966 genes were distributed on 13 chromosomes, and each chromosome contained 1–4 TaDUF966 genes (Figure 1B and Table 1). The 24 genes of the TaDUF966 gene were located on chromosome 3 (3A, 3B, and 3D, 11 genes), chromosome 4 (4A, 4B, and 4D, 9 genes) and chromosome 7 (7A, 7B, and 7D, 4 genes). The other 4 chromosomes (Chr1A, Chr2B, Chr5B, and Chr6B) contained only one TaDUF966 gene for each chromosome. Most members of Group I were distributed on 3 chromosomes (3A, 3B, and 3D), while members of Group II were evenly distributed on 13 chromosomes.

Homologous Gene Pairs and Synteny Analysis

The homologous pairs of DUF966 were identified in the common wheat ancestral species, Triticum urartu (AA), Triticum dicocoides (AABB), and Aegilops tauschii (DD); 10, 15, and 9 homologous pairs of DUF966 were identified, respectively, named as TuDUF966, TdDUF966, and AeDUF966 based on the naming rules. The location and other details of these genes are summarized in Table 1. Comparisons between TuDUF966 vs. TdDUF966, TaDUF966 vs. TdDUF966, TaDUF966 vs. TaDUF966, TaDUF966 vs. AeDUF966, and TaDUF966 vs. TaDUF966 revealed 2, 21, 12, 6, and 21 pairs of putative paralogs, respectively (Figure 1C). To better understand the evolutionary factors affecting the DUF966 gene family, we calculated the Ka/Ks ratios among homologous pairs of TaDUF966, TuDUF966, TdDUF966, and AeDUF966 genes (Table S6). Most of the DUF966 gene pairs showed Ka/Ks ratios < 1, indicating that these genes contained more synonymous than non-synonymous changes. Moreover, these genes were under negative selection pressure to protect the state of the ancestor. Three homologous pairs, including TaDUF966-11B/TuDUF966-9A, TaDUF966-11B/TuDUF966-8A, and TaDUF966-11B/TaDUF966-13B, showed Ka/Ks ratios > 1, indicating positive selection, accelerated evolution, and neofunctionalization. Analysis of duplication events of TaDUF966 genes revealed that gene duplication occurred mainly among Group II genes (Figure 1B, Table S6). Of the 28 TaDUF966 genes, 17 (60.7%) genes showed segmental duplication events. Furthermore, TaDUF966 genes showed segmental and tandem duplications, while TdDUF966 genes showed segmental duplications (Table S6). These data suggest that segmental duplication events played a key role in the expansion of the DUF966 gene family in wheat.

Analysis of Gene Structure and Motif Composition in TaDUF966 Genes

To investigate the structure of TaDUF966 genes, we extracted their detailed information from the GFF3 file of the reference genome (Jiang et al., 2020b). The results showed that all TaDUF966 genes contain introns. Comparative exon–intron structure analysis revealed the presence of 2–6 exons in Group I members and 4–8 exons in Group II members (Figure 2B). Motif usually refers to the basic structure that constitutes any kind of characteristic sequence. It is a subset of the structural domain, and its function is to reflect the various biological functions of the structural domain. The prediction of protein motifs is a useful protein analysis method. Analysis of 28 TaDUF966 proteins using MEME revealed 10 motifs (Figures 2C, D). Details of these 10 motifs are shown in Table S7. Each TaDUF966 protein contains 3–10 motifs, and most TaDUF966 proteins contain motifs 1, 2, and 3. In the phylogenetic tree, we found that members with relatively similar genetic relationships have more similar motifs, which indicates that DUF966 members gathered in the same subgroup may have more similar biological effects. In addition, all TaDUF966 proteins contain the DUF966 conserved Motif 1. Motif analysis using NCBI BLAST showed that seven of these 10 motifs (motifs 1–5, 9, and 10) belong to the DUF966 conserved domain, while the remaining three motifs were uncharacterized. Multiple sequence alignment analysis of TaDUF966 proteins revealed that the amino acid sequence of the DUF966 domain is highly conserved (Figure S1), suggesting that this domain plays a key role in protein function.

Analysis of TaDUF966 Protein Features

Protein feature analysis showed that TaDUF966 proteins contain an average of 422 aa (191 aa–644 aa) (Table 2). Other features of TaDUF966 proteins were as follows: average pI = 8.89 (range = 6.46–9.88); average MW = 46.1 kDa (range = 21.3–69.84 kDa); average instability index = 63.48 (range = 36.76–79.99) (Table S8). The instability index values indicate that TaDUF966 proteins are relatively unstable. The results of the subcellular localization prediction of all 28 TaDUF966 proteins showed that 27 proteins localized to the nucleus or chloroplast, whereas one protein localized to the cytoplasm.

Prediction of Cis-Acting Regulatory Elements in TaDUF966 Gene Promoters

Cis-acting elements play important roles in the initiation of gene transcription. The TaDUF966 gene promoters were predicted to possess cis-acting regulatory elements, which are involved in three different physiological and biochemical processes. (Figure 3, Tables S9, S10). Among the biotic/abiotic stress-responsive cis-elements, ARE (antioxidant response element) and G-box were highly enriched in TaDUF966 gene promoters. Secondly, the CAAT-box and TATA-box (cis-elements related to growth and development) were also enriched in TaDUF966 gene promoters. Among the phytohormone-related cis-acting elements, the ABA-responsive element (ABRE) was the most frequently identified in TaDUF966 gene promoters. In addition, Cis-acting regulatory elements involved in the response to methyl jasmonate (MeJA),
FIGURE 1 | Phylogenetic analysis of DUF966 gene family. (A) Phylogenetic tree of DUF966 proteins from wheat, Arabidopsis, and rice. Predicted amino acid sequences were aligned using ClustalW2 sequence alignment program. Different groups were marked by different colors. Blue, Group I; green, Group II. Green round, 28 DUF966 proteins from wheat; blue round, 7 DUF966 proteins from rice; red round, 5 DUF966 proteins from Arabidopsis. (B) Chromosomal localization of the 28 DUF966 genes in the wheat genome. Positions were determined in the map. Different groups of TaDUF966 are represented in different colors. Blue, Group I; green, Group II. (C) Genome-wide syntenic analysis among common wheat (Ta, AABBDD, red box) and progenitor species Aegilops tauschii (Ae, DD, blue box), T. urartu (Tu, AA, orange box), and T. dicoccoides (Td, AABB, green box). The lines between the two chromosomes link the pair of orthologous DUF966 genes.
CGTCA-box and TGACG-box, were also identified in most of the TaDUF966 gene promoters. In general, TaDUF966 gene promoter may respond to various endogenous and exogenous stimuli.

**Transcriptome Expression Profiling of TaDUF966 Genes Under Salt Stress**

Gene expression acts as a link between the transmission and realization of genetic information in organisms, and genome-wide analysis of gene expression patterns helps to identify the gene expression pathway and its regulation. In this study, RNA-seq data downloaded from NCBI were mapped to the wheat reference genome using HISAT2, and expression patterns of TaDUF966 genes under salt stress were analyzed. The pheatmap, based on standardized TPM values, showed that salt stress strongly induced TaDUF966-5, -9, and -14 (A, B, and D) but did not affect the expression of other genes (Figure 4A, Table S11) (Appels et al., 2018). The majority of TaDUF966 genes showed low expression, whereas 9 genes showed high expression in two different wheat varieties. The expression of TaDUF966 genes increased with the duration of salt treatment, reaching a peak at 6 h, followed by a gradual decline.

**Expression Analysis of TaDUF966 Genes**

To further reveal the potential functions of TaDUF966 genes in association with abiotic stress (salt stress), we analyzed the expression patterns of several TaDUF966 genes by qRT-PCR. According to transcriptome analysis, expression of TaDUF966-5, -9, and -14 (A, B, and D) was relatively higher in salt-treated

---

**TABLE 2 | Predicted amino acid sequence features of TaDUF966 proteins in wheat.**

| Designation | Domain Number | Alternative splices | Length (aa) | Cds (bp) | Exon Number |
|-------------|---------------|---------------------|-------------|----------|-------------|
| TaDUF966-1A | 2             | 1                   | 401         | 1,206    | 6           |
| TaDUF966-2B | 1             | 2                   | 292         | 879      | 5           |
| TaDUF966-3A | 1             | 2                   | 379         | 1,140    | 5           |
| TaDUF966-3B | 1             | 3                   | 387         | 1,164    | 5           |
| TaDUF966-3D | 2             | 1                   | 384         | 1,155    | 5           |
| TaDUF966-4A | 1             | 1                   | 406         | 1,211    | 4           |
| TaDUF966-4B | 1             | 1                   | 408         | 1,227    | 4           |
| TaDUF966-4D | 1             | 1                   | 405         | 1,218    | 4           |
| TaDUF966-5A | 1             | 1                   | 456         | 1,371    | 3           |
| TaDUF966-5B | 1             | 1                   | 454         | 1,365    | 3           |
| TaDUF966-5D | 1             | 1                   | 460         | 1,383    | 3           |
| TaDUF966-6B | 1             | 1                   | 420         | 1,283    | 8           |
| TaDUF966-7D | 1             | 1                   | 257         | 774      | 2           |
| TaDUF966-8A | 1             | 1                   | 396         | 1,191    | 5           |
| TaDUF966-8B | 1             | 1                   | 367         | 1,104    | 4           |
| TaDUF966-8D | 1             | 1                   | 395         | 1,188    | 5           |
| TaDUF966-9A | 1             | 2                   | 513         | 1,542    | 5           |
| TaDUF966-9B | 1             | 1                   | 515         | 1,548    | 5           |
| TaDUF966-9D | 1             | 1                   | 522         | 1,569    | 6           |
| TaDUF966-10A| 1             | 1                   | 508         | 1,527    | 6           |
| TaDUF966-10B| 1             | 2                   | 494         | 1,485    | 8           |
| TaDUF966-10D| 1             | 1                   | 565         | 1,698    | 8           |
| TaDUF966-11B| 1             | 1                   | 218         | 657      | 4           |
| TaDUF966-12B| 1             | 1                   | 191         | 576      | 4           |
| TaDUF966-13A| 1             | 1                   | 235         | 708      | 4           |
| TaDUF966-14A| 1             | 4                   | 644         | 1,935    | 7           |
| TaDUF966-14B| 1             | 3                   | 595         | 1,768    | 6           |
| TaDUF966-14D| 1             | 3                   | 568         | 1,707    | 7           |
roots compared with control roots. The results of qRT-PCR analysis of TaDUF966-5D, TaDUF966-9B, and TaDUF966-14D were consistent with RNA-seq data. These three genes showed higher expression levels in the root and leaf tissues of wheat plants treated with NaCl compared with the control. Expression levels of these three genes increased in leaf tissues at 4 or 8 h after NaCl treatment but decreased in root tissues at 8 h (Figure 4B). In particular, the expression level of TaDUF966-9B increased in root tissues at 2 h after NaCl treatment, decreased in root tissues at 8 h, and was induced by salt stress 24 h later. Therefore, TaDUF966-9B may be better induced by salt stress. In subsequent trials we focused on this gene.

Silencing of TaDUF966-9B Reduces Salt Tolerance in Wheat

To further confirm the function of TaDUF966-9B, we performed a virus-induced gene silencing (VIGS) assay. Silencing of the TaDUF966-9B gene was achieved using two constructs (BSMV: TaDUF966-9B-1 and BSMV: TaDUF966-9B-2), each carrying an approximately 250 bp fragment (Figure S2) of the corresponding TaDUF966-9B gene, respectively, applied as independent treatments. All plants inoculated with BSMV showed mild chlorotic mosaic symptoms at 10 days post-inoculation (dpi), while leaves inoculated with BSMV: TaPDS showed photobleaching (Figure 5A), indicating that the BSMV-induced gene silencing system was functional. The expression of each homolog was also analyzed by qRT-PCR. At 10 days after friction-implanted virus, the leaf curling phenotype was significantly stronger in plants expressing either construct than in plants carrying the empty vector control (Figures 5B, C), and the level of BSMV: TaDUF966-9B-1/2 mRNA in NaCl-treated plants was significantly lower than that in the empty vector control (Figure 5D). These results indicate that the TaDUF966-9B gene is involved in salt stress tolerance in wheat.
FIGURE 4 | Expression analysis of wheat TaDUF966 genes. (A) Expression profiles of TaDUF966 genes in root tissues, development stages, and NaCl treatment of wheat. The heat-map was established based on the R package “Pheatmap” and R function “boxplot”. The scale represents signal intensity of TPM values. (B) The qRT-PCR results. 10-day-old seedling leaves and roots were sampled after 0, 3, 6, 9, 12, and 24 h under stress conditions comprising 150 mM NaCl. t-tests were used to determine significant differences in expression patterns for CK and NaCl treatments (***p < 0.001, *p < 0.05).
DISCUSSION

Advances in high-throughput sequencing over the past two decades have increased the gap between genome sequence data and gene function information (Yin et al., 2020). Additionally, the number of DUF gene families has increased in the Pfam database (Bateman et al., 2010). Although the origin, diversity, and preliminary biological functions of these DUF gene families have been investigated in many studies, elucidating the biological functions of these genes, particularly in wheat, remains challenging.

The current Pfam database 32.0 contains 17,939 gene families, of which 3,961 gene families (22%) represent DUFs (Chalupska et al., 2008). Bioinformatics analysis shows that the DUF966 gene family is widely distributed in rice, tomato, and Arabidopsis. In this study, a total of 28 TaDUF966 genes were identified in common wheat. These genes were distributed on 13 chromosomes, particularly on chromosomes 3, 4, and 7; there are six TaDUF966 genes, which are identified on A, B, and D subgenomes, respectively (Figure 1B). The other 10 TaDUF966 gene family members did not appear simultaneously on three homeologous chromosomes, indicating that TaDUF966 genes might have been amplified by replication during evolution, although a small number of genes could have been selectively deleted during evolution. This is consistent with previous research (Luo and Tian, 2017). Moreover, TaDUF966 proteins contained different numbers of motifs (3 to 10), but only one or two highly conserved DUF966 domains (Table 2), but functions of each domain remain to be investigated.

In addition to wheat, DUF966 genes have also been identified in other plant species such as Arabidopsis, rice, maize, soybean, sorghum, tomato, grape, pawpaw, and alfalfa (Luo et al., 2014). A previous study showed that the rice DUF966 gene, OsDSR2, and tomato DUF966 gene, JWS19, play a vital role in response to salt stress and simulated drought stress (Tirajo, 2005; Luo et al., 2014). Nevertheless, the function of these DUF966 proteins has not been characterized further. In this study, our qRT-PCR data showed that three TaDUF966 genes (TaDUF966-5D, TaDUF966-9B, and TaDUF966-14D) were up-regulated in the root and leaf tissues of wheat plants treated with NaCl, which is consistent with previous RNA-seq data (Figure 5). These results, together with phylogenetic analysis, strongly suggest that the DUF966 family members in wheat are also involved in salt stress tolerance.

To further verify the role of TaDUF966 family genes in salt stress tolerance, we silenced one wheat gene (TaDUF966-9B) using VIGS and then treated the plants with NaCl. The expression of TaDUF966-9B was not induced upon NaCl treatment, indicating that TaDUF966-9B responds to salt stress in wheat. Although gene-specific primers were designed for TaDUF966-9B, and sequences of cDNA fragments cloned into VIGS vectors were confirmed by Sanger sequencing, the possibility of knocking down the expression of TaDUF966-9B homologs could not be completely ruled out. Therefore, further studies are needed to
generate stable transgenic lines using the CRISPR/Cas technology to better understand the function of TaDUF966 genes. In conclusion, our results provide a comprehensive analysis of the biological function of DUF966 family genes in wheat.

DATA AVAILABILITY STATEMENT
All datasets presented in this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS
XGZ, DM, and YQ guided the design of the experiment. XYZ, WS, YH, WJ, JS, and JY directed the data analysis. XYZ conducted data analysis and manuscript writing. XGZ, DM and YQ supervised the experiment and confirmed the manuscript. XGZ is the guarantor of this work, so she can have full access to all the data in the research and is responsible for the integrity of the data and the accuracy of the data analysis. All authors contributed to the article and approved the submitted version.

REFERENCES
Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 361, eaar7191. doi: 10.1126/science.aar7191
Bailey, T. L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. 2, 28–36.
Bateman, A., Coggill, P., and Finn, R. D. (2010). DUFs: families in search of function. Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun. 66, 1148–1152. doi: 10.1107/S1744349110001685
Chou, K. C., and Shen, H. B. (2010). Plant-mPLoc: a top-down strategy to predicted the potential homolog of bactericidal/permeability-increasing protein in plant system. Protein J. 32, 163–171. doi: 10.1007/s10930-013-9473-6
Gu, L. J., and Cheng, H. M. (2014). Isolation, molecular cloning and characterization of a cold-responsive gene, AmDUF1517, from Ammopiptanthus mongolicus. Plant Cell Tissue Organ Culture (PCTOC) 117, 201–211. doi: 10.1007/s11240-014-0433-4

Thank all the above staff for the help in this study. The authors thank the reviewers for their valuable suggestions during the revision of the early manuscripts.

FUNDING
This research was funded by the National Key R&D Program of China, grant number 2018YFD0200506.

ACKNOWLEDGMENTS
The authors thank Professor Zhensheng Kang (Northwest A & F University, Yangling, China) for providing vector of VIGS.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.569838/full#supplementary-material
Li, L. H., Ye, T. Z., Guan, Y., Lv, M. M., Xie, C., Xu, J. H., et al. (2017). Genome-wide identification and analyses of the rice OsDUF936 family. Biotechnol. Biotechnol. Equip. 32, 309–315. doi: 10.1080/13102818.2017.1342411

Luo, C. K., and Tian, L. (2017). Bioinformatic analysis of the DUF966 gene family in rice. Mol. Plant Breed. 15, 4791–4796. doi: 10.13271/j.mpob.015.004791

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 Rep. 33, 323–336. doi: 10.1007/s00299-013-1532-0

Petty, I. T. D., French, R., Jones, R. W., and Jackson, A. O. (1990). Identification of barley stripe mosaic virus genes involved in viral RNA replication and systemic movement. EMBO J. 9, 3453–3457. doi: 10.1002/j.1460-2075.1990.tb07553.x

Qin, D. D., Xie, S. C., Liu, G., Ni, Z. F., Yao, Y. Y., Sun, Q. X., et al. (2013). Isolation and functional characterization of heat-stress-responsive gene TaWTF1 from wheat. Chin. Bull. Bot. 48, 34–41. doi: 10.3724/sp.j.1259.2013.00034

Rozas, J., Ferrer-Mata, A., Sanchez-Delbarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Mol. Biol. Evol. 34, 3299–3302. doi: 10.1093/molbev/msx2248

Song, J. H., Ma, D. F., Yin, J. L., Yang, L., He, Y. Q., Zhu, Z. W., et al. (2019). Genome-wide characterization and expression profiling of squamosa promoter binding protein-like (SBP) transcription factors in wheat (Triticum aestivum L.). Agronomy 9, 527. doi: 10.3390/agronomy9050527

Tirajo, A. (2005). Isolation and characterization of novel salt-responsive genes in tomato (Lycopersicon esculentum Mill.) roots. Biol. Sci. Dept.-Simon Fraser Univ. 1–304.

Wang, X. J. (2014). Functional analysis of AtST39 in salt and drought stress. Xiamen Univ. 1–100.

Yan, D. W., Zhou, Y., Ye, S. H., Zeng, L. J., Zhang, X. M., and He, Z. H. (2013). Beak-shaped grain 1/Triangular Hull 1, a DUF640 gene, is associated with grain shape, size and weight in rice. Sci. China Life Sci. 56, 275–283. doi: 10.1007/s11427-013-4449-5

Yang, S. Q., Li, W. Q., Miao, H., Gan, P. F., Qiao, L., Chang, Y. L., et al. (2016). REL2, A Gene Encoding An Unknown Function Protein which Contains DUF630 and DUF632 Domains Controls Leaf Rolling in Rice. Rice (N Y) 9, 37. doi: 10.1186/s12284-016-0105-6

Yang, Y., Yoo, C. G., Winkeler, K. A., Collins, C. M., Hinchee, M. A. W., Jawdy, S. S., et al. (2017). Overexpression of a Domain of Unknown Function 231-containing protein increases O-xylan acetylation and cellulose biosynthesis in Populus. Biotechnol. Biofuels 10, 311. doi: 10.1186/s13068-017-0998-3

Yin, J., Fang, Z. W., Sun, C., Zhang, P., Zhang, X., Lu, C., et al. (2018). Rapid identification of a stripe rust resistant gene in a space-induced wheat mutant using specific locus amplified fragment (SLAF) sequencing. Sci. Rep. 8, 3086. doi: 10.1038/s41598-018-21489-5

Yin, J. L., Liu, M. Y., Ma, D. F., Wu, J., Li, S. L., Zhu, Y. X., et al. (2018). Identification of circular RNAs and their targets during tomato fruit ripening. Postharvest Biol. Technol. 136, 90–98. doi: 10.1016/j.postharvbiot.2017.01.013

Yin, J. L., Tian, J. L., Li, G., Zhu, Y. X., Zhou, X. K., He, Y., et al. (2020). Carbohydrate, phytohormone, and associated transcriptome changes during storage root formation in Alternanthera philoxeroides. Weed Sci. 68, 1–39. doi: 10.1017/wsc.2020.37

Zhang, P., Jia, Y. J., Shi, J. X., Chen, C., Ye, W. W., Wang, Y. C., et al. (2019). The WY domain in the Phytophthora effector PSR1 is required for infection and RNA silencing suppression activity. NewPhytol. 223, 839–852. doi: 10.1111/nph.15836

Zhu, X. G., Jiao, M., Guo, J., Liu, P., Tan, C. L., Yang, Q., et al. (2018a). A novel MADS-box transcription factor PstMCM1-1 is responsible for full virulence of Puccinia striiformis f. sp. tritici. Environ. Microbiol. 20, 1452–1463. doi: 10.1111/1462-2920.14054

Zhu, X. G., Liu, W., Chu, X. L., Sun, Q. X., Tan, C. L., Yang, Q., et al. (2018b). The transcription factor PstST12 is required for virulence of Puccinia striiformis f. sp. tritici. Mol. Plant Pathol. 19, 961–974. doi: 10.1111/mpp.12582

Zhu, Y. X., Gong, H. J., and Yin, J. L. (2019a). Role of Silicon in Mediating Salt Tolerance in Plants: A Review. Plants (Basel) 8, 1–22. doi: 10.3390/plants8060147

Zhu, Y. X., Yang, L., Liu, N., Yang, J., Zhou, X. K., Xia, Y. C., et al. (2019b). Genome-wide identification, structure characterization, and expression pattern profiling of aquaporin gene family in cucumber. BMC Plant Biol. 19, 1–18. doi: 10.1186/s12870-019-1953-1

Zhu, Y. X., Yin, J. L., Liang, Y. F., Liu, J. Q., Jia, J. H., Huo, H. Q., et al. (2019c). Transcriptomic dynamics provide an insight into the mechanism for silicon-mediated alleviation of salt stress in cucumber plants. Ecotoxicol. Environ. Saf. 174, 245–254. doi: 10.1016/j.ecoenv.2019.02.075

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Zhou, Zhu, Shao, Song, Jiang, He, Yin, Ma and Qiao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.