Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Pangenome-Wide Identification, Evolutionary Characterisation and Stress-Responsive Dynamics of SIZ1-Type SUMO E3 Ligase Gene Family in Bread Wheat (Triticum aestivum L.)

Hemant Sharma
Indian Agricultural Research Institute
Mallana Gowdra Mallikarjuna
mgrpatil@iari.res.in

Indian Agricultural Research Institute
Garudapalya Muniswamy Keerthi
Indian Agricultural Research Institute
Niharika Mallick
Indian Agricultural Research Institute
Niranjana Murukan
Indian Agricultural Research Institute
Lekshmy Sathee
Indian Agricultural Research Institute
Shailendra Kumar Jha
Indian Agricultural Research Institute

Research Article

Keywords: Post-translational modification, SUMOylation, SUMO E3 ligase, Evolution, Expression dynamics, Stress

Posted Date: April 7th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9114016/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background:

SUMOylation, a key post-translational modification in plants, modulates diverse stress responses through SUMO E3 ligases (SIZ1). The SIZ1 gene family remains uncharacterized in bread wheat (*Triticum aestivum* L.), despite its proven functional relevance in model systems.

Results:

The pangenome survey of bread wheat identified 15 TaSIZ1 genes, unevenly distributed across the genomes. The reference genome Chinese Spring harboured 14 TaSIZ1 genes distributed across homoeologous groups 1, 3, 4, 5, and 7. These intron-rich TaSIZ1 genes with 14–17 exons encode nuclear-localised, hydrophilic SIZ proteins (pI 4.85–6.60) that carry a conserved zf-MIZ1 domain. The TaSIZ1 family expansion occurred predominantly through whole-genome and segmental duplications under purifying selection (Ka/Ks < 1). The synteny and orthology among wheat, barley, rice, maize, and sorghum SIZ1 genes further confirmed evolutionary conservation, with strong purifying selection (Ka/Ks: 0.21–0.46). Promoter analysis revealed abundant stress- and hormone-responsive cis-elements (ABRE, MBS, ARE, GT1), alongside major transcription factor binding sites for ERF, WRKY, MIKC_MADS, and MYB families. Expression profiling showed higher basal activity of TaSIZ1-1A, 1B, 1D, 3B, and 3D in vegetative tissues and induction by heat, drought, powdery mildew, and leaf rust.

Conclusions:

Our study showed that TaSIZ1 genes are structurally conserved but show functional divergence across homoeologs and genotypes. The strong, specific induction of selected copies underscores their utility in enhancing stress resilience. This study provides the first comprehensive framework for understanding TaSIZ1 regulation in wheat and nominates candidate genes for multi-stress engineering and wheat breeding.

Background

Post-translational modifications (PTMs) regulate the function, stability, and interaction networks of proteins, thereby influencing numerous biological processes. SUMOylation is one of the crucial PTMs, where a Small Ubiquitin-like Modifier (SUMO) protein is covalently attached to target proteins, modulating their activity, localisation, or stability. SUMOylation modifications are conserved among eukaryotes and are involved in critical cellular processes, including transcription, DNA repair, response to abiotic and biotic stresses, and cell cycle regulation [1, 2]. SUMOylation involves a cascade of enzymatic reactions, including the activation (E1), conjugation (E2), and ligation (E3) of SUMO to target proteins. The E3 SUMO ligases facilitate substrate specificity and enhance the efficiency of SUMO transfer. Among known E3 ligases, the SAP and Miz1 (SIZ1)-type SUMO E3 ligase is particularly well-studied. SIZ1 plays a pivotal role in plant stress and developmental processes [3].

In *Arabidopsis thaliana*, the SIZ1 mediate the SUMOylation of key proteins of pathways controlling photomorphogenesis, flowering, phosphate starvation, and freezing tolerance, showing responses to cold, drought, and heat stress conditions, highlighting its significance in stress adaptation [4–7]. Functional studies across species demonstrate that overexpression of SIZ1 homologs, such as *OsSIZ1* in rice and cotton, significantly enhances tolerance to abiotic stresses, including drought, heat, and salt, while also improving photosynthesis, biomass accumulation, and yield under adverse conditions [8, 9]. Recently, alternative splicing of SIZ1 has been shown to generate distinct isoforms with stress-specific localisation and activity, underscoring the dynamic regulation of SUMOylation in response to environmental fluctuations [10].

Furthermore, SIZ1 plays a key role in enhancing salicylic acid (SA) signalling, which is essential for systemic acquired resistance (SAR) against biotrophic pathogens. It facilitates the SUMOylation of key defence regulators, such as NPR1 (Non-Expressor of Pathogenesis-Related Genes 1), thereby stabilising NPR1 and promoting the expression of defence-related genes [2, 3, 11]. Additionally, SIZ1 helps balance growth and immune responses, preventing the detrimental effects of constitutive defence activation. In *Arabidopsis thaliana*, the *siz1* mutants exhibit enhanced resistance to pathogens, such as *Pseudomonas syringae*, but suffer from growth retardation due to excessive salicylic acid (SA) signalling [12]. SIZ1 also interacts with the jasmonic acid (JA) pathway, which fine-tunes the crosstalk between SA- and JA-mediated defence to optimise plant immunity [12–14]. Similarly, SIZ1 provides temperature-sensitive immunity, where the SIZ1 function is essential for maintaining disease resistance across different environments [14]. These evidences underline SIZ1 as a key regulator of plant immunity, making it a valuable target for improving crop disease resistance.

Bread wheat (*Triticum aestivum* L.) is a major staple crop of global importance for food and nutrition with a large and complex genome [15]. The identification and characterisation of SIZ1-like genes help in understanding the SUMOylation pathway, which is crucial for improving stress resilience and yield stability. Previous studies have identified components of the SUMOylation machinery in model plants, such as Arabidopsis and rice (*Oryza sativa*), but the comprehensive identification, evolutionary and functional characterization of SUMO E3 ligases in wheat remains limited [8, 9, 16–20].

Recent advancements in bioinformatics tools and the availability of high-quality wheat genome assemblies (IWGSC RefSeq v2.1) have facilitated the identification of gene families across the wheat genome [15]. Utilising wheat genomic resources, a genome-wide identification and analysis of SIZ1-type SUMO E3 ligases in wheat can provide insights into their gene structure, evolutionary and regulatory relationships and expression responses. This information can serve as a foundation for functional studies and the development of wheat varieties with enhanced stress tolerance through molecular breeding or genetic engineering. Thus, the present investigation aimed to identify and characterise SIZ1-like SUMO E3 ligase genes in the wheat genome, determine the evolutionary relationships between wheat SIZ1 homologs and those from other cultivars, and decipher the regulatory and functional response of the identified SIZ1 genes to various stresses.

Materials and Methods

Sequence Retrieval and Identification of SIZ1 Genes and Proteins

The whole-genome sequences and annotation files of common wheat (*Triticum aestivum*) and the sequences of SIZ1 from *Arabidopsis thaliana* (AtSIZ1, AT5G60410) and rice (OsSIZ1, Os05t0125000) were downloaded from the EnsemblPlants database (https://plants.ensembl.org/). The wheat pangenomes were retrieved from the wheat pangenomes database (https://wheatgenomes.com/). The protein sequences (CDSs) of *AtSIZ1* and *OsSIZ1* were used to perform BLASTp search (E-Value ≤ 1e-5) against the wheat proteome assembly (IWGSC RefSeq v2.1). HMMSearch (E-value ≤ 1e-5) was performed across the whole wheat proteome for pfam profile PF02891. The non-redundant hits from both the BLASTp and HMMSearch were again subjected to HMMScan to confirm the presence of the PF02891 domain and finalise the number of SIZ1 genes in wheat. Further, to mine SIZ1 in related species for evolutionary analysis, the identified wheat SIZ1 (TaSIZ1) proteins were used as query sequences to perform BLASTp searches (E-Value ≤ 1e-5) against the rice, sorghum, maize and barley proteomes downloaded from EnsemblPlants Database (https://plants.ensembl.org/), and the HMMSearch and HMMScan were performed as explained above.

Analysis of Chromosome Location, Gene Structures, and Motifs

The physical locations and gene structures of the identified SIZ1 genes were retrieved from the GFF file downloaded from the EnsemblPlants database and mapped using TBtools (https://github.com/CJ-Chen/TBtools) [21]. Conserved motifs within SIZ1 proteins were identified using the MEME Suite (http://meme-suite.org/tools/meme), and motif annotation was verified using the HMMER tool (https://www.ebi.ac.uk/Tools/hmmer/search/phmmer).

Physico-chemical properties and Gene Ontology (GO) Terms Analysis

Physicochemical properties, including molecular weight, isoelectric point (pI), and amino acid composition, were calculated using the ProtParam tool (http://web.expasy.org/protparam) [22]. Subcellular localization and Gene Ontology (GO) annotations of SIZ1 proteins were predicted using the WoLFPSORT (https://wolfpsort.hgc.jp) server and the ShinyGO tool (http://bioinformatics.sdstate.edu/go), respectively.

Evolutionary Analyses of SIZ1 Family in Wheat and Related Species

Multiple Sequence Alignment and Phylogenetic Analysis

The SIZ1 proteins from wheat, rice, maize, barley and sorghum were aligned with MUSCLE [23]. The multiple alignment was trimmed with the trimAL v1.2 tool (http://trimal.cgenomics.org) [24]. The phylogenetic tree was constructed using the maximum likelihood method. The best model identified using Bayesian information criterion and phylogenetic tree was constructed using 5000 ultrabootstraps using IQ-TREE version 2 [25] software. The tree was realised in iTOL [26].

Analysis of SIZ1 family expansion and selection pressure

Identification of TaSIZ1 paralogs was performed through self-BLASTp search (e-value < 1e-5) within the wheat proteome using BLAST + 2.10.1 (https://www.ncbi.nlm.nih.gov/). Subsequently, DupGen_finder software was used to harvest the various duplications in the wheat genome [27]. For identification of orthologs across the target species, viz., wheat, rice, maize, barley, and sorghum, the SIZ1 sequences were subjected to OrthoFinder software (https://github.com/davidemms/OrthoFinder) [28]. The nonsynonymous rate (Ka), synonymous rate (Ks) and evolutionary constraint (Ka/Ks) were calculated for each of the paralog and ortholog pairs of SIZ1 using KaKs calculator v.3.0 (https://ngdc.cncb.ac.cn/tools/kaks) [29]

Synteny and Collinearity Analysis

The whole proteomes of related cereals, *viz*., rice, maize, barley and sorghum, were aligned with the wheat proteome using the BLASTp with an e-value < 1e-10 and the best five hits per query. The MCScanX software was used to identify syntenic and collinear blocks [30] and visualised with TBtools v2.052 (https://bio.tools/tbtools) [21].

Regulatory analysis SIZ1 Genes

Prediction of Cis-acting Elements in SIZ1 Promoters

Sequences of 1.5 kb upstream regions of SIZ1 genes were retrieved and scanned for cis-acting elements using the PlantCARE server (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Only response elements with a matrix value > 5 on the sense strand were fetched for subsequent analysis and interpretation [31].

SIZ1-Regulatory Network Analysis

Putative microRNA (miRNA) targets in SIZ1 genes were predicted using the psRNATarget server (https://plantgrn.noble.org/psRNATarget/) with default parameters. The Transcription factor (TF) was predicted using PlantRegMap (https://plantregmap.gao-lab.org/binding_site_prediction.php). Further, the regulatory network of *TaSIZ1*s was visualised with Cytoscape [32].

Gene Expression Analysis

In-silico Expression Analysis

Transcriptome data from the Wheat Expression Database (http://www.wheat-expression.com/) were retrieved to analyse the expression of SIZ1 genes across different tissues and stress conditions. The heatmaps were generated using TBtools.

Quantitative real-time PCR analysis of TaSIZ1 genes

Plant material and treatments

The leaf rust–susceptible and drought-tolerant cultivar C306 and the leaf rust–resistant but comparatively drought-sensitive cultivar HD2888 were used to study the TaSIZ1s expression (Supplementary Table 1). Seeds were surface-sterilised with 0.1% $HgCl_2$ for 2 minutes, rinsed thoroughly with sterile distilled water, and treated with 0.5% Bavistin before sowing. Plants were grown in a controlled growth chamber under greenhouse conditions in soil-filled pots. For the leaf rust expression analysis, 14-day-old seedlings were inoculated with the 77 − 5 strain of *Puccinia triticina*. Leaf samples were collected at 0, 24, 48, and 96 hours post-inoculation. For drought stress treatment, 14-day-old seedlings of both genotypes were transferred to a nutrient solution supplemented with 10% polyethene glycol (PEG-6000) to simulate osmotic stress. Leaf tissues were harvested after 7 days of treatment. All collected samples were immediately flash-frozen in liquid nitrogen, stored at − 80°C, and later used for total RNA isolation.

RNA extraction and quantitative real-time PCR

Six highly expressed TaSIZ1 gene copies (based on in-silico expression analysis) were selected for qRT-PCR validation. The primers used for RT-qPCR of TaSIZ1 genes were synthesised using Primer Express Software v3.0.1 (Thermo Fisher Scientific; Supplementary Table 2). Total RNA was isolated using the TRIzol reagent (Ambion, Naugatuck) as per the manufacturer's specifications and treated with RNase-free DNase I (Invitrogen, California, United States) for 15 minutes to degrade any residual genomic DNA. First-strand cDNA was synthesised from DNase I-treated total RNA using Revert Aid First Strand cDNA Synthesis Kit (Thermo Scientific, Massachusetts, United States) according to the manufacturer's instructions. RT-qPCR was performed in optical 96-well plates using a CFX96 (Bio-Rad) and SYBR Green (Bio-Rad). The Wheat β-actin gene (accession number AB181991) was used as an endogenous control. The CT values were analysed through the $2^{-\Delta\Delta CT}$ method [33].

Results

Mining and Copy Number of SIZ1 genes in Wheat Pangenome Cultivars and Related Species

A pangenome-wide identification of the SIZ1 gene family, using the hexaploid wheat reference genome (cv. Chinese Spring) along with 12 high-quality cultivar assemblies from the pangenome project, revealed 15 *SIZ1* genes (Table 1; Supplementary Table 3; Fig. 1). Genes were systematically renamed based on their chromosomal positions in Chinese Spring as *TaSIZ1-1A* to *TaSIZ1-7D*, with corresponding copies in each cultivar assigned the same names. In Chinese Spring, 14 TaSIZ1 genes were identified with three homoeologs each on chromosomes 1, 4, 5, and 7 of the A, B, and D genomes, and two on chromosome 3 (*TaSIZ1-3B* and *TaSIZ1-3D*). Similarly, the genotypes Lancer, Landmark, and Mattis showed 14 TaSIZ1 genes on chromosomes 1, 4, 5, and 7 of the A, B, and D genomes, and two on chromosome 3 (*TaSIZ1-3A* and *TaSIZ1-3B*). The lowest number of TaSIZ1 genes (12) is reported in the Alchemy and Kariega genotypes, with non-significant hits on 1B, 4D, 7D, and 7A, 7B, 7D chromosomes, respectively (Fig. 1). Further, mining SIZ1s in related species, *viz*., barley, maize, rice, and sorghum, yielded 5, 4, 3, and 4 SIZ1s, respectively (Supplementary Table 4). All the predicted SIZ1 proteins had the characteristic zinc-finger MIZ (zf-MIZ; PF02891) domain, with Zn^{2+}-coordinating Cys/His residues located in the central region of the protein (Fig. 1A), which confirms that the identified sequences represent SIZ1-type SUMO E3 ligases. Additionally, the target species, including wheat, showed non-uniform distributions of SIZ1 genes across their genomes (Supplementary Tables 3–4; Fig. 1).

Table 1
The distribution and range of various physicochemical properties of TaSIZ1 sequences mined from thirteen pangenome species.

Gene	No. of genomes*	Gene Length (bp)	cDNA Length (bp)	CDS Length (bp)	Gene % GC content	Protein Length (aa)	GRAVY	Aliphatic Index	Instability Index	Molecular Weight (kDa)	Aromaticity	Isoelectric Point (pI)
TaSIZ1-1A	*13/13*	14874–16412	3105–4826	2619–2625	41.12–42.15	872–874	-0.61–0.60	70.76–71.22	45.69–46.19	96.91–97.02	0.07	4.97
TaSIZ1-1B	*12/13*	14559–16368	3001–4042	2622–2625	40.57–41.29	873–874	-0.59–0.59	72.12–72.35	44.97–45.84	96.99–97.11	0.07	4.90
TaSIZ1-1D	*13/13*	15318–21517	3070–5037	2622–4512	40.82-47.00	873–930	-0.6–0.58	69.55–73.11	45.00-48.70	96.91-103.32	0.07	4.92–5.06
TaSIZ1-3A	*6/13*	7520–7878	3057–3572	2787–2805	42.48–43.06	928–934	-0.54–0.48	79.08–81.73	46.39–51.31	102.44-103.32	0.05	4.85–4.99
TaSIZ1-3B	*12/13*	6580–8015	2775–3705	2775–2793	42.13–42.98	924–930	-0.53–0.52	78.47–79.01	45.93–46.46	102.26-103.04	0.05	4.87–4.93
TaSIZ1-3D	*4/13*	6490–7634	2787–3363	2787–2901	42.06–42.57	928–966	-0.5–0.49	79.39–80.61	45.68–47.93	102.64-106.72	0.05	4.92–4.94
TaSIZ1-4A	*13/13*	8317–9057	2334–3074	2334–2334	40.54–41.55	777–777	-0.55–0.55	73.87-74.00	38.93–39.17	85.16–85.20	0.06	5.16
TaSIZ1-4B	*13/13*	8764–12143	2334–4667	2334–3972	40.67–41.52	777–837	-0.57–0.56	72.79–74.13	38.25–43.56	85.12–91.88	0.06	5.02–5.78
TaSIZ1-4D	*12/13*	8358–9442	2334–3021	2334–2517	40.63–41.50	777–838	-0.58–0.56	73.28–73.87	38.57–43.42	85.26–91.99	0.06	5.20–5.84
TaSIZ1-5A	*13/13*	8258–9238	2379–3359	2379–2583	40.99–41.98	792–860	-0.61–0.59	79.03–80.97	40.71–48.5	87.73–94.79	0.05	5.27–5.56
TaSIZ1-5B	*13/13*	8939–9504	2986–3154	2385–2661	40.59–41.96	794–886	-0.63–0.46	74.28–80.54	41.88–62.01	87.97–97.16	0.05	5.15–6.60
TaSIZ1-5D	*13/13*	8852–10028	2904–3613	2355–2622	41.50-42.26	784–873	-0.62–0.59	78.29–79.45	40.74–50.43	86.54–96.34	0.05	5.20–5.69
TaSIZ1-7A	*12/13*	8688–10327	2788–3319	2352–2739	38.65–40.71	783–912	-0.47–0.44	72.45–74.06	62.29–64.13	85.01–98.89	0.05	6.36–7.02
TaSIZ1-7B	*11/13*	9003–9723	2941–3373	2385–2661	40.18–41.86	794–886	-0.6–0.46	74.28–80.54	41.88–62.01	87.97–96.28	0.05	5.15–6.60
TaSIZ1-7D	*12/13*	8075–8782	2825–3222	2316-2652	38.57–40.61	771–883	-0.45–0.40	73.44–74.98	60.3-61.88	83.9-95.97	0.05	6.11–6.45

Physico-Chemical Properties of SIZ1 Genes and Protein Sequences

Genomic lengths range from about 6.6 kb (*TaSIZ1-3B* in Kariega) to over 21 kb (*TaSIZ1-1D* in Chinese Spring) owing to differences in intron number and size rather than changes in the core CDS. In contrast, cDNA lengths are generally 3.0-3.6 kb, with a few longer transcripts up to 5,037 bp (Alchemy: *TaSIZ1-1D*), suggesting alternative splicing or longer UTRs in certain genotypes. CDS lengths are mostly 2,316–2,811 bp, with some longer variants up to 4,512 bp, resulting in protein lengths of 771–966 amino acids. TaSIZ1s from chromosome 4 (TaSIZ1-4A/4B/4D) consistently showed shorter proteins of about 777 amino acids (aa), while TaSIZ1s from chromosomes 3 and 5 code longer proteins (930–966 aa), likely due to domain extensions or low-complexity region insertions. The GC content of TaSIZ1 genes is moderate and highly conserved, typically ranging from 38.5% to 47.0%. Chromosome-1D *TaSIZ1s* have the highest GC content (46–47%) compared to 7A and 7D *TaSIZ1s* (38–39% GC) (Supplementary Table 3).

Physicochemical analysis showed that TaSIZ1 proteins are highly similar across all cultivars, reinforcing the strong functional conservation within the SIZ1 SUMO E3 ligases across the target pangenome species. Predicted molecular weights range from approximately 83.90 kDa to 106.7 kDa, consistent with CDS length variation. Aromaticity values for TaSIZ1 were low (about 0.05–0.07), typical of regulatory proteins rich in charged and polar residues. All TaSIZ1 proteins have strongly negative GRAVY scores (− 0.63 to − 0.40), reflecting a hydrophilic nature, as observed in nuclear-localised proteins, and allowing extensive protein–protein interactions. Despite this, they display relatively high aliphatic indices (69–82), suggesting reasonable thermostability. Most proteins have instability indices above 40 and are predicted to be unstable, a common property of signalling and scaffold proteins. However, chromosome-4 homoeologs (TaSIZ1-4A, TaSIZ1-4B, and TaSIZ1-4D) have instability indices below 40 and are classified as stable, indicating this clade may form more rigid or long-lived structural components within the SIZ1 network. Predicted isoelectric points (pI) are mostly in the acidic to near-neutral range (pI 4.85–7.02), so nearly all TaSIZ1 proteins carry a negative net charge at pH 7, with only a few group-7A/7D members approaching neutrality or slight basicity (Supplementary Table 3–4).

Gene structure and motif analysis of TaSIZ1 Genes

Structural analysis of the TaSIZ1 genes showed that all members have multiple exons and introns, with a highly conserved exon-intron organisation across homoeologs. The genes are intron-rich, with exon counts ranging from 14 to 17 and intron counts from 13 to 16. *TaSIZ1-1A* and *TaSIZ1-1B* are the most complex, each with 17 exons and 16 introns, while *TaSIZ1-1D* has 16 exons and 15 introns. Most other genes (*TaSIZ1-3B*, *TaSIZ1-4A*, *TaSIZ1*-4B, *TaSIZ1*-4D,

TaSIZ1-5A, *TaSIZ1-5B*, *TaSIZ1-5D*, *TaSIZ1-7B*) share 16 exons and 15 introns. *TaSIZ1-3D* and *TaSIZ1-7A* have 15 exons and 14 introns, and *TaSIZ1-7D* has 14 exons and 13 introns. Phase 0 introns are most common, accounting for about 43.48% of the total, followed by phase 2 with 32.61% and phase 1 with 23.91% (Fig. 2A, B). Motif analysis showed that TaSIZ1 proteins contain multiple conserved motifs, with each gene having 4 to 10 motifs (Fig. 2). The majority of *TaSIZ1* members possess nearly all motifs (9–10), while TaSIZ1-3D, TaSIZ1-7A, and TaSIZ1-7D have fewer, indicating structurally simplified variants. Motif 1, along with Motifs 4, 6, and 9, was found across the TaSIZ1 genes and was the most conserved, whereas Motifs 8 and 10 were found less commonly in the TaSIZ1 family (Fig. 2A, C).

Gene Ontology (GO) Terms and Sub-Cellular Localisation Analysis

GO annotation indicates that all 14 TaSIZ1 proteins from cv. Chinese spring was significantly enriched for the molecular function GO term zinc ion binding (GO:0008270). Further, the 12 of 14 TaSIZ1s (excluding TaSIZ1-3D; TaSIZ1-7A) were significantly enriched for Sumo ligase activity (GO:0061665), Sumo transferase activity (GO:0019789), Aminoacyltransferase activity (GO:0016755), Ubiquitin-like protein transferase activity (GO:0019787), Transition metal ion binding (GO:0046914) and Acyltransferase activity (GO:0016746). Similarly, for biological function category the these 12 TaSIZ1s were showed significant enrichments for Protein sumoylation (GO:0016925), Peptidyl-lysine modification (GO:0018205), Peptidyl-amino acid modification (GO:0018193), Protein modification by small protein conjugation (GO:0032446), Post-translational protein modification (GO:0043687), Protein modification by small protein conjugation or removal (GO:0070647) under the biological function category (Supplementary Table 5). Further, the subcellular localisation predictions indicate that TaSIZ1 proteins are predominantly nuclear localised, with confidence scores of 0.89–0.96. (Supplementary Table 6).

Evolutionary Analyses of the SIZ1 Family in Wheat and Related Species

Multiple Sequence Alignment and Phylogenetic Analysis

The multiple sequence alignment of SIZ1 proteins using MUSCLE produced an intact alignment of the PFAM domain PF02891 from 345 to 395 amino acids, indicating sequence conservation and suitability for phylogenetic analysis (Fig. 3A). The IQ-Tree model find option identified JTT+G4 based on the Bayesian Information Criterion (BIC) as best model for phylogenetic tree construction. The phylogenetic analysis grouped SIZ1 proteins from wheat, barley, rice, maize, and sorghum into three major clades based on percent identity (> 50%) among the protein sequences (Supplementary Table 7), corresponding to wheat homoeologous chromosome groups 1/3, 4/5, and 7 (Fig. 3B).

Clade I contains wheat TaSIZ1 proteins from chromosomes 4A, 4B, 4D, 5A, 5B, and 5D, barley HvSIZ1-4H/5H, OsSIZ1b, ZmSIZ1b, SbSIZ1a, and SbSIZ1d from barley, rice, maize, and sorghum, respectively. Wheat chromosome 5 homoeologs show high per cent identity (95.42–95.71%), and HvSIZ1-5H shares 92.29–93.19% identity with them. Wheat chromosome-4 proteins are nearly identical (91.98–99.05% identity), with HvSIZ1-4H showing 89.12–93.55% identity. OsSIZ1b is moderately similar (67–72%) to Triticeae SIZ1 proteins in the clade. The Panicoid pair ZmSIZ1b and SbSIZ1d is closely related (85.54%) but only 62.61–63.38% identical to TaSIZ1-5A, TaSIZ1-5B and TaSIZ1-5D, indicating an earlier divergence between Triticeae and Panicoideae. SbSIZ1a has emerged as the most distant and outlier member of this phylogenetic tree. The clade II (yellow) includes wheat TaSIZ1 proteins from chromosomes 1A, 1B, 1D, 3B, and 3D, as well as HvSIZ1-1H/3H from barley and OsSIZ1a, ZmSIZ1a, ZmSIZ1c, and SbSIZ1c from rice, maize, and sorghum. Within this clade, *Triticeae* proteins are highly conserved: HvSIZ1-1H shares 97.03–98.05% identity with TaSIZ1-1A/1B/1D, and HvSIZ1-3H shares 91.41–91.53% identity with TaSIZ1-3B/3D. Maize and sorghum sequences shared 73–77% identity with wheat and barley but remained highly similar to each other (> 90.00%). Clade III (green) includes TaSIZ1-7A, 7B and 7D proteins from chromosome 7, HvSIZ1-7H from barley, OsSIZ1c and OsSIZ1d from rice, ZmSIZ1d from maize, and SbSIZ1b from sorghum. The three wheat homoeologs are again highly conserved (96.12–97.40%) and share high identity (90.41–92.45%) with HvSIZ1-7H (Supplementary Table 7, Fig. 3B).

Duplication and selection pressure analysis of TaSIZ1

Gene duplications drive the expansion of gene families by generating new members that contribute to functional diversity and enable the evolution of new functions. We analysed the types of duplications and the selection pressures that influenced the expansion of *TaSIZ1* genes in the wheat genome. The results indicate that single-gene duplications play a significant role in the expansion of SIZ1 (Supplementary Table 8 and Fig. 4a). In total, 24 paralog pairs were found among the 14 TaSIZ1s of the Chinese Spring cultivar, encompassing > 50% of whole-genome duplications (13), followed by segmental (9) and dispersed duplications (2) (Fig. 4b).

Selection pressure was assessed by analysing the ratio of non-synonymous (Ka) to synonymous (Ks) substitutions, which indicates the direction and strength of natural selection on protein-coding genes. Ka/Ks analysis of 24 wheat paralogous pairs showed Ka/Ks < 1, indicating strong purifying selection and functional conservation during the expansion of TaSIZ1 sequences in the wheat genome (Supplementary Table 8; Fig. 4c). Pairs with very low Ks (< 0.10) among the 14 duplication pairs observed within chromosomal group but between the homoeologous genomes A, B, D, except the duplication pair TaSIZ1-3D-TaSIZ1-7D (Supplementary Table 8), suggesting their recent origin. On the other hand, the paralogs between the chromosomal groups showed higher synonymous mutation rates (Ks > 0.45), suggesting that these pairs arose from ancient polyploidy events (Supplementary Table 8; Fig. 4c).

Ortholog analysis of SIZ1 in wheat and related species

Ortholog analysis of TaSIZ1 genes and their counterparts in barley, rice, maize, and sorghum identified many-to-one relationships except two pairs of one-to-one orthologs between wheat and barley (TaSIZ1-3D–HvSIZ1-3H; TaSIZ1-4A–HvSIZ1-4H) and many-to-many ortholog clade group between wheat and maize (TaSIZ1-1A/1B/1D/3D–ZmSIZ1a/c). Among the orthologs between wheat and barley, three many-to-one (TaSIZ1-1A/1B/1D–HvSIZ1-1H; TaSIZ1-7A/7B/7D–HvSIZ1-7H; TaSIZ1-5A/5B/5D–HvSIZ1-5H) orthologs with mean Ks values of 0.11, suggesting their recent divergence as compared to the rest of the species (Supplementary Table 9; Fig. 5a). The many-to-one orthologs observed between wheat-rice orthologs (TaSIZ1-1A/1B/1D/3D–OsSIZ1a; TaSIZ1-7A/7B/7D–OsSIZ1c; TaSIZ1-4A/5A/5B/5D–OsSIZ1b) and the mean Ks value of 0.61 was found among these orthologs (Supplementary Table 9; Fig. 5b). Interestingly,

wheat-maize orthologs showed a mean Ks value of 0.74 for many-to-many and many-to-one (TaSIZ1-4A/5A/5B/5D–ZmSIZ1b; TaSIZ1-7A/7B/7D–ZmSIZ1d) orthologs (Supplementary Table 9; Fig. 5c). The highest range (0.60–4.30) and mean (2.56) of Ks values among all the many-to-one ortholog pairs (TaSIZ1-1A/1B/1D/3D–SbSIZ1c; TaSIZ1-7A/7B/7D–SbSIZ1b; TaSIZ1-4A/5A/5B/5D–SbSIZ1d) of wheat-sorghum suggesting varying divergence times owing to various evolutionary events like episodic bursts of gene duplication or biased gene retention etc (Supplementary Table 9; Fig. 5d). All these wheat TaSIZ1s diverged under strong purifying selection (Ka/Ks < 1) for all the orthologs from rice (0.31), barley (0.27), maize (0.29), and sorghum (0.30).

Synteny and Collinearity Analysis

To further understand deep evolutionary insights, the syntenic and collinearity was worked out between the wheat SIZ1 genes with barley, rice, maize and sorghum (Supplementary Table 10; Fig. 6). Synteny analysis of the SIZ1 gene families in wheat (Ta), barley (Hv), rice (Os), and maize (Zm) showed a highly conserved collinear relationship (Supplementary Table 10; Fig. 6). Nineteen SIZ1-containing syntenic blocks harboured 19 pairs of collinear SIZ1s between wheat and barley, indicating the strongest syntenic signals owing to their evolutionary lineage. Following barley, maize, sorghum, and rice, wheat also showed synteny with SIZ1 genes, with wheat spanning 13, 7, and 6 synteny blocks, each containing a respective number of SIZ1 gene pairs (Supplementary Table 10; Fig. 6).

Regulatory analysis SIZ1 Genes

Cis-regulatory elements analysis

Scanning of 1.5 kb upstream of the TaSIZ1 promoter identified numerous cis-acting regulatory elements, suggesting diverse regulatory potential (Fig. 7; Supplementary Tables S11–S13). A total of 45 distinct cis-element types were detected across all TaSIZ1 promoters. Each promoter harboured multiple elements, including abundant core promoter motifs such as the *TATA and CAAT* boxes. Light-responsive elements such as *G-box, Box 4*, and *I-box* occurred in many promoters. Among the hormone-responsive elements, the ABA-responsive *ABRE motif* was observed most frequently (71) across the sub-genomes. The MeJA-responsive *CGTCA-motif* (37) and *TGACG-motif* (13) were also found in the promoter sequences of TaSIZ1s across the sub-genomes. Gibberellin-responsive elements (*P-box, GARE-motif*) and auxin-responsive elements (*AuxRR-core, TGA-element*) were detected at lower frequencies. Multiple stress-related cis-elements were present as well, including *ARE* (*anaerobic response element*), *LTR* (*low-temperature response element*), *MBS* (*MYB-binding site*), *WUN-motif* (wound-response element), and *STRE* (stress response element) (Fig. 7; Supplementary Tables S11–S13). Interestingly, some *cis* elements are confined to specific homoeologs or sub-genomes. The auxin-responsive *TGA-element* was found only in B-genome promoters, and a *circadian* rhythm-related element appeared in just two promoters (*TaSIZ1-1B; TaSIZ1-7D*). Genome-specific differences in cis-element abundance were also observed: D-genome promoters harboured roughly twice as many cis-elements (392 total) as A-genome (176) or B-genome (177) promoters (Fig. 7; Supplementary Tables S11–S13). TaSIZ1-7D exhibited the highest motif density, containing 23 *ABREs* and 21 *G-box* elements, 9 *AREs* and 7 *LTRs* in addition to *TATA* (25) and *CAAT* (28) boxes, indicating complex regulation in response to stresses. Genes such as *TaSIZ1-1B* and *TaSIZ1-1D* also showed rich promoter compositions. *TaSIZ1-1B* had 12 *ABREs* and 12 G-boxes, while *TaSIZ1-1D* harboured nine *CGTCA-motifs* (MeJA-responsive) and six *LTRs* (low-temperature responsive). Notably, TaSIZ1-5A featured a rare *3-AF3 binding site* and a *GCN4 Motif, suggesting a specialised regulatory* pattern. Rare or unique cis-elements were gene-specific. Auxin-responsive motifs (*TGA-element, AuxRR-core*) appeared exclusively in *TaSIZ1-4B*, while salicylic acid-related *TCA-elements* and the *CAG-motif* were confined to *TaSIZ1-4D*. The *AE-box* was observed only in *TaSIZ1-5D*. Moreover, motifs related to *circadian* regulation (circadian), the *LAMP-element*, and *Gap-boxes* were limited to *TaSIZ1-1B* and *TaSIZ1-7D*. The *WUN-motif* (wound response) appeared only in *TaSIZ1-3B* and *TaSIZ1-3D*. This cis-element landscape revealed that, while specific motifs, such as *ABRE*, were universally distributed, many genes exhibited distinct combinations of hormone-, stress-, or development-related elements. These patterns suggest functional divergence in transcriptional regulation among TaSIZ1 genes (Fig. 7; Supplementary Tables S11–S13).

Regulation of TaSIZ1s through MicroRNAs and Transcription Factors

A total of 63 unique miRNAs were predicted to regulate 14 TaSIZ1 genes in the Chinese Spring cultivar. After removing duplicates, the number of unique miRNA targets per TaSIZ1 gene ranges from 7 to 24. *TaSIZ1-1B* and *TaSIZ1-1D* have the highest number of unique miRNAs (24), while *TaSIZ1-3D* has the fewest (7). The most common miRNAs across the family are tae-miR1120, tae-miR1122, tae-miR1127, tae-miR1130, tae-miR1137, tae-miR1139, tae-miR9655, tae-miR9773, tae-miR9780, and tae-miR6197, each targeting multiple TaSIZ1 copies in groups 1, 3, 4, 5, and 7. In contrast, low-frequency miRNAs include rare or single targets such as tae-miR9664 (TaSIZ1-1D), tae-miR9667 (TaSIZ1-1A), tae-miR1136 and tae-miR5175 (TaSIZ1-7A), and tae-miR1119 (TaSIZ1-1D). Other sparsely distributed miRNAs, such as tae-miR6201 and tae-miR7757, are limited to group 1 copies (Supplementary Table 14; Fig. 8a,b,f).

The prediction of transcription factor binding sites for TaSIZ1 genes revealed broad transcriptional regulation by 37 transcription factors at 3684 binding sites (Supplementary Table 15). Among the predicted TFs, ERF (1327), BBR-BPC (388), MIKC_MADS (222), C2H2 (203), MYB (187), and bZIP (119) were found to regulate all 14 TaSIZ1 genes in the Chinese spring cultivar (Supplementary Table 15; Fig. 8c,f). Interestingly, the TF E2F/DP were predicted only in the homoeologs of chromosome 5 (*TaSIZ1-5A; TaSIZ1-5B; TaSIZ1-5D*), where the TF AAR-B was found in the homoeologs of chromosomes 3 (*TaSIZ1-53; TaSIZ1-3D*) and five only (Supplementary Table 15; Fig. 8c). Further, among the genes, *TaSIZ1-1A* and *TaSIZ1-1D* showed the highest number of transcription factors (491), followed by TaSIZ1-3B (465) and *TaSIZ1-1B* (330). In contrast, *TaSIZ1-3D* (102) showed a lower number of TFs (Fig. 8d). The *TaSIZ1* genes of all the sub-genomes showed a majority of the TFs (34) identified in one or other gene (Fig. 8e). Interestingly, TFs SRS and RAV were limited to A and B sub-genomes (*TaSIZ1-4A,4B*), and TALE to sub-genome B (*TaSIZ1-4B*) (Fig. 8c,e).

The regulatory network was constructed using unique TFSs and miRNAs to capture a regulatory snapshot of TaSIZ1 genes in the Chinese Spring cultivar (Fig. 8f). The TaSIZ1 network comprised 114 nodes, 596 unique edges, and an average of 10.46 neighbours, with a density of 0.046. The network confirmed

that TFs play a major regulatory role in the expression and regulation of TaSIZ1, compared with miRNAs, by targeting multiple TaSIZ1 genes, although there are more unique regulatory miRNAs than unique TFs (Supplementary Table 15; Fig. 8f).

Gene Expression Analysis

In-silico Expression Analysis

Tissue-specific expression of TaSIZ1 genes

Tissue-specific expression analysis of TaSIZ1 genes revealed apparent variation among homoeologs. The root (The *TaSIZ1-1A, TaSIZ1-1B*, and *TaSIZ1-1D* copies displayed the highest expression levels, especially in roots (2.84–4.05 FPKM), stems (1.91–3.24 FPKM), grains (2.46–3.66), followed by spike (2.15–2.96). Similarly, *TaSIZ1-3B* (root: 3.64; stem: 2.90; grain: 2.92; spike, 2.18)and *TaSIZ1-3D* (root: 3.35; stem: 2.57; grain: 2.92; spike, 2.07) showed a similar trend of expression across all target tissues, except leaves. Furthermore, the remaining *TaSIZ1s* also showed moderate expression. However, leaf tissue showed lower *TaSIZ1* expression levels (0.34–1.35) (Fig. 9a).

The expression analysis of TaSIZ1 genes under Powdery Mildew and stripe rust infection

Expression analysis of TaSIZ1 genes based on FPKM values revealed apparent differences among homoeologs in response to pathogen infection. *TaSIZ1-3B* showed high expression among all genes under non-inoculated conditions (3.12); higher expression was observed in response to powdery mildew infection at 48 h (4.13) and 72 h (4.17). However, under leaf rust inoculation, *TaSIZ1-3B* expression did not increase with post-inoculation hours (Fig. 9b,c). A similar expression pattern was observed for TaSIZ1-3D in response to powdery mildew and leaf rust infections (Fig. 9b,c). Similarly, the homoeologs of chromosome 1 showed enhanced expression in response to powdery mildew infection over post-inoculation hours (Fig. 9b), whereas no elevated or differential expression patterns were found for leaf rust infection (Fig. 9c). The expression patterns suggest that TaSIZ1 genes are found to be more responsive to powdery mildew infection rather than leaf rust.

The expression analysis of TaSIZ1 genes under Heat and Drought

Expression profiling of TaSIZ1 genes under drought, heat, and combined stress conditions revealed strong, differential regulation among homoeologs. Under control conditions, *TaSIZ1-1A, TaSIZ1-1B, TaSIZ1-1D*, and *TaSIZ1-3B* exhibited moderate expression levels (2.0–2.5 FPKM), whereas most other members showed low expression (< 1.0 FPKM). During drought stress, expression of these chromosome 1 and 3 copies increased slightly at one hour (2.5–3.0 FPKM) and remained steady at six hours. Under heat stress, a sharp induction was observed, particularly in *TaSIZ1-3B* reaching approximately 4.5 FPKM at six hours, while the chromosome 1 copies also showed moderate increases (2.5–3.0 FPKM). Under combined drought and heat stress, expression patterns resembled those seen under heat alone, with *TaSIZ1-3B* being the most highly expressed gene (4.0 FPKM), followed by moderate expression of chromosome 1 homoeologs (2.5 FPKM) (Fig. 9d).

Under control conditions, *TaSIZ1-1B* (2.77), *TaSIZ1-1D* (1.94), *TaSIZ1-3B* (2.90), and *TaSIZ1-3D* (2.30) showed relatively high expression (2.0–2.9 FPKM) as compared to the rest of the *TaSIZ1* genes (Fig. 9d). However, these genes showed elevated expression in response to prolonged heat (2.63–4.30), and combined heat and drought (2.22–3.68) stresses of 6h (Fig. 9d), suggesting that these TaSIZ1 genes are responsive to prolonged heat stress conditions in a given set of genotypes in the target experiment. Interestingly, *TaSIZ1-1A* showed the least expression across the stress conditions and control. The remaining TaSIZ1 expression showed only minor changes, largely independent of stress treatments (Fig. 9d).

qRT-PCR Expression Analysis of TaSIZ1 genes in wheat

The genotypes C-306 and HD2888 showed contrasting responses to leaf rust resistance and drought tolerance (Fig. 10a-c), indicating the underlying changes in the expression patterns at the genome level. Expression profiling of six TaSIZ1 genes across three leaf rust post-inoculation time points (24 h, 48 h, and 96 h) revealed transcriptional differences between the resistant genotype HD2888 and the susceptible genotype C306. For *TaSIZ1-1A*, C306 showed slight upregulation at 24 h (1.1-fold) and 96 h (1.35-fold), but downregulation at 48 h (–1.95-fold). In contrast, HD2888 showed strong, consistent upregulation, reaching a 4.1-fold increase at 96 hours. For *TaSIZ1-1B*, C306 maintained low expression, with a downregulation at 24 h (–1.11-fold) and a marginal increase at later stages, whereas HD2888 showed pronounced upregulation at 48 h (3.6-fold). Similarly, *TaSIZ1-1D* showed low expression in C306 (1.05–1.2-fold), whereas HD2888 showed marked upregulation at 24 h (5.2-fold) and 96 h (6.0-fold). For *TaSIZ1-3B*, C306 showed downregulation at both 24 h (–0.25-fold) and 96 h (–1.85-fold), while HD2888 consistently exhibited upregulation (2.0–3.1-fold). Finally, *TaSIZ1-3D* was weakly expressed in C306 (1.0–1.45-fold), whereas HD2888 was strongly upregulated at all stages, with the highest fold change at 24 h (4.6-fold) (Fig. 10d).

Expression analysis of TaSIZ1 genes in the drought-stressed plants for seven days revealed distinct differences between the drought-tolerant genotype C306 and the comparatively drought-sensitive genotype HD2888. Drought stress resulted in upregulation of all target TASIZ1 genes in both tolerant and sensitive genotypes (Fig. 10e). However, the tolerant genotype C306 showed significantly higher fold changes for *TaSIZ1-1A* (C306: 3.24; HD2888: 2.12), *TaSIZ1-1B* (C306: 2.50; HD2888: 1.80), *TaSIZ1-3B* (C306: 4.21; HD2888: 3.09), and *TaSIZ1-3D* (C306: 4.60; HD2888: 2.45); whereas *TaSIZ1-1D* showed the fold changes in C306 as compared to HD2888 (C306: 2.91; HD2888: 3.70) expression (Fig. 10e).

Discussion

Expansion and Evolution of TaSIZ1 Gene Family

The TaSIZ1 gene family in wheat has expanded significantly due to ancient genome duplications and polyploidy. We identified multiple TaSIZ1 loci distributed across five homoeologous chromosome groups (1, 3, 4, 5, and 7) in bread wheat. This distribution mirrors the presence of SIZ1 genes on five chromosomes in diploid barley (1H, 3H, 4H, 5H, 7H), suggesting that the ancestral grass genome already harboured several SIZ1 paralogs. Certainly, cereals have experienced at least two rounds of whole-genome duplication [34, 35], which likely generated multiple SIZ1 loci that were retained over evolution. Subsequent allopolyploidisation in bread wheat (AABBDD) tripled these loci, resulting in a complement of homoeologous TaSIZ1 copies from each progenitor genome. Most of these homoeologs have been maintained, indicating purifying selection and an essential conserved function. Supporting this, the duplicated TaSIZ1 pairs showed Ka/Ks < 1.0, suggesting strong selective constraint on TaSIZ1 coding sequences; none of the TaSIZ1 duplicates showed evidence of positive selection. Hence, the retention of TaSIZ1 copies is supported by the theory of polyploid redundancy, which buffers the loss of critical stress-responsive genes in wheat [36, 37].

The presence of TaSIZ1 genes across the pangenome of species further indicates that they are an essential part of the wheat core gene set. However, we did observe subtle presence/absence and copy number variations. Notably, a SIZ1 homoeolog on chromosome 3A is missing or pseudogenized in the Chinese Spring reference, whereas it is present in other cultivars, reflecting intraspecific variation in gene content. Such presence/absence variation (PAV) is not uncommon in wheat and can arise from structural rearrangements during breeding programmes [35]. For example, we detected a cultivar-specific duplication on 4A in one genotype, suggesting rare tandem duplication or gene copy fragmentation events. Nevertheless, the fact that all three sub-genomes contribute at least four TaSIZ1 genes emphasises the expansion of this family relative to diploid models. In a dicot model, *Arabidopsis thaliana*, only a single SIZ1 gene exists (*AtSIZ1*) [9]. Diploid cereals also generally carry more than one SIZ1-type gene. The rice genome encodes two to three SIZ1 homologs [9] (Supplementary Table S4) and sorghum and maize (paleotetraploids) possess ~ 4 SIZ1 each (Supplementary Table S4). These findings indicate that SIZ1-type SUMO E3 ligases form a small multigene family in plants, with lineage-specific expansion. Phylogenetic analysis clustered wheat SIZ1 proteins with their cereal orthologs, which is consistent with speciation. The wheat SIZ1 proteins share ~ 59–81% identity with rice (Supplementary Table S7) and > 30% identity with Arabidopsis SIZ1, demonstrating that SIZ1 is highly conserved yet has diversified within the grass lineage. Interestingly, *OsSIZ1* and *OsSIZ2* can functionally complement an Arabidopsis *siz1* mutant, indicating that the core SUMO ligase activity of SIZ1 has been conserved from 140–150 million years of monocot-dicot divergence [19, 38]. The expansion of TaSIZ1 copies in wheat, therefore, likely provides additional regulatory scope or specialisation rather than entirely novel functions. The presence of intact coding sequences across the TaSIZ1 paralogs, with no premature stop codons or frame-shifts, and their promoter regions showing no evidence of transposon-driven degeneration, supports the notion of conserved constraint with potential sub-functionalization. Further, maintenance of multiple TaSIZ1 genes despite potential redundancy implies that each copy may contribute in distinct contexts, *viz*., developmental stage, tissue, or stress type (Fig. 9a), buffering the plant against losing this critical SUMO E3 ligase function and offering an additional level of resilience in wheat [35].

Structural Features and Regulatory Divergence of SIZ1 Sequences

Despite their expansion, the TaSIZ1 proteins are structurally conserved, characteristic of the Siz/PIAS family of SUMO E3 ligases. All identified TaSIZ1 proteins contain the canonical domains known from Arabidopsis and yeast SIZ1, an N-terminal SAP domain (a DNA-binding module), the PINIT domain, a centrally located Siz/PIAS-specific RING (SP-RING) domain, and a C-terminal region harbouring SUMO-interaction motifs (SIMs) and nuclear localisation signals (NLS) [39, 40]. We confirmed the presence of the Cys_3His-Cys_4 SP-RING motif in every TaSIZ1 through assuring the whole domain of SIZ1, which is essential for recruiting the SUMO-conjugating enzyme and catalysing SUMO transfer [40]. Thus, each TaSIZ1 protein is biochemically capable of catalysing SUMOylation, much like *OsSIZ1* or *AtSIZ1* genes.

Most TaSIZ1 genes span large genomic regions, ranging from 7.62 to 21.51 kb, and contain 14–17 exons per gene. The exon/intron structure is conserved mainly among homoeologs. The conserved gene structure further supports a common origin. Nonetheless, some divergence exists in non-coding regions and untranslated regions, which could influence gene regulation. We identified numerous *cis*-regulatory elements within the 1.5 kb promoters of TaSIZ1 genes, pointing to divergent regulatory controls. All TaSIZ1 promoters are enriched in stress-and hormone-responsive motifs, but the composition and abundance of these elements vary by gene (Supplementary Table S11). For example, many TaSIZ1 promoters contain several ABRE motifs, MBS elements, AREs, and low-temperature-responsive elements. Elements associated with defence and pathogen response, such as the WUN motif, TGACG motifs (elicitor-responsive), and TC-rich repeats, are also present, indicating potential regulation by biotic stress signals. Such differential cis-element profiles imply that although TaSIZ1 genes may be functionally redundant at the protein level, they could be differentially regulated at the transcriptional level in response to specific cues. This kind of sub-functionalization is common in duplicated stress-responsive genes, allowing finer control over when and where each copy is expressed [35, 41].

Beyond promoter motifs, we investigated post-transcriptional and transcriptional regulators of TaSIZ1 by predicting several microRNA (miRNA) binding sites in TaSIZ1 mRNAs and Transcription factors (Supplementary Table S14-15). Notably, some homoeologs have unique miRNA target sites that others lack, suggesting homoeolog-specific miRNA-mediated regulation. For example, a conserved miR9664 target site is present in TaSIZ1-D copy but not in its A or B counterparts, hinting that miR9664 might selectively attenuate that homoeolog's expression under stress [41, 42]. Similarly, 37 distinct TF families were predicted for 14 TaSIZ1 promoters (Supplementary Table S15). Among these, ERF/AP2, BBR-BPC, MIKC-type MADS, C2H2 zinc-finger, MYB, and bZIP transcription factors appear to bind all TaSIZ1 promoters, as anticipated, since these TF families include many master regulators of stress and development. For example, bZIP and MYB factors mediate ABA and drought responses, and AP2/ERF factors mediate ethylene/jasmonate and abiotic stress signals [6, 7]. Differences in TF binding motifs also mirror the subgenome bias. E2F/DP binding sites were found only in the promoters of the chromosome 5 genes (TaSIZ1-5A/5B/5D), hinting that these copies might be cell-cycle regulated [43]. Similarly, a plant-specific TF family, SRS (STY-related SHI gene regulators), is found only for TaSIZ1-4A/4B, and RAV, an AP2/B3 hybrid TF, is found only for those same two genes, suggesting a unique developmental or stress regulation of the chromosome-4 SIZ1 [44]. Further, the variety of cis-elements indicates that each TaSIZ1 gene integrates a distinct combination of upstream signals, including hormones such as ABA, gibberellin, and salicylic acid, as well as environmental cues such as heat shock or drought-induced dehydration. This

regulatory divergence reconfirms that the expansion of SIZ1 copies has been accompanied by neo- or sub-functionalization of regulatory regions, enabling wheat to deploy SUMOylating capacity in a modular fashion under diverse stress conditions.

Functional Implications of TaSIZ1 genes in Stress Adaptation

In-silico and quantitative real-time PCR (qRT-PCR) expression analyses under various stress conditions demonstrate that the TaSIZ1 family is mostly stress-inducible, indicating a significant role in wheat's adaptation to abiotic and biotic challenges. SUMOylating activity in plants is known to surge during stress as a protective response [45, 46]and, correspondingly, we observed upregulation of multiple TaSIZ1 genes in response to heat, drought, and pathogen attack. The *in-silico* expression of *TaSIZ1-1B, TaSIZ1-1D, TaSIZ1-3B* and *TaSIZ1-3D* was markedly induced by heat stress as compared to drought, aligning with the need for enhanced SUMOylation to counteract protein destabilisation (Fig. 9d) [9]. Additionally, qRT-PCR analysis of six genes, viz., TaSIZ1-1A, TaSIZ1-1B, TaSIZ1-1D, TaSIZ1-3B, and TaSIZ1-3D, showed significant increases in expression in drought-sensitive genotypes. However, the tolerant genotype C-306 showed higher expression levels than the drought-sensitive genotype HD2888 (Fig. 10e). This suggests that efficient SIZ1 induction may correlate with better drought resilience. The Arabidopsis *siz1* mutants show reduced SUMO conjugates and heat sensitivity [47, 48], highlighting SIZ1's role in thermotolerance. SUMOylation is dynamically regulated by heat stress and water status [49, 50]. One homoeolog, TaSIZ1-1A, showed low *in-silico* expression, implying sub-functionalization with constitutive or alternative roles in the target genotype [12].

The role of SIZ1 in drought tolerance is well documented in Arabidopsis, where siz1 mutants are hypersensitive to drought and fail to induce ~ 300 drought-responsive genes [4]. In Arabidopsis, SIZ1 SUMOylates ABA INSENSITIVE 5 (ABI5), a bZIP transcription factor, attenuating ABA signalling during seed germination. Under drought, however, SIZ1 appears to promote stress-responsive gene expression, for instance, SIZ1 facilitates the expression of many ABA-responsive genes and proline biosynthesis genes (*P5CS*) that help in Osmo-protection [4, 51]. Our finding that TaSIZ1 induction accompanies drought stress (and is stronger in a drought-hardy line) aligns with this protective function. It is likely that in wheat, SIZ1-mediated SUMOylation stabilises or activates positive regulators of drought tolerance, such as DREB transcription factors or enzymes of osmolyte pathways. SUMOylation might also suppress negative regulators of stress responses, striking a balance that favours survival under water limitation [9, 52]. Therefore, the TaSIZ1 genes contribute to drought adaptation in wheat by boosting SUMOylation capacity, which in turn stabilises proteins and stress-response pathways that mitigate drought damage. Variation in TaSIZ1 response across genotypes may partly explain differences in drought resilience. This hypothesis could be tested by comparing SIZ1 allele function in tolerant vs. sensitive wheat varieties or differential SUMOylation patterns.

Public transcriptome data and our infection assays indicate that one or more TaSIZ1 genes are induced during early infection by these pathogens. For instance, 48 hours after powdery mildew inoculation, transcripts of a TaSIZ1 on chromosome 5 were upregulated compared to an uninfected control (Fig. 9b). Likewise, upon leaf rust infection, the rust-resistant cultivar (HD2888) showed a transient spike in TaSIZ1 expression at 24, 48 and 96 h post-inoculation, whereas a susceptible cultivar (C-306) had a low or non-significant expression folds (Fig. 10d). In stark contrast, the stripe rust infection where there was no-significant induction over the infection time course (Fig. 9b-c). This differential reaction suggests that wheat plants discriminate the pathogens at the signalling level, and that SIZ1-mediated SUMOylation is more engaged in the powdery mildew and leaf rust response than in the stripe rust response in a given genetic background. This pattern suggests TaSIZ1 may be involved in orchestrating or fine-tuning the defence response. The loss of SIZ1 leads to constitutive activation of salicylic acid (SA)-dependent defences and enhanced resistance to biotrophic pathogens [12, 53]. This implies that *AtSIZ1* normally restrains SA accumulation under non-stress conditions to prevent autoimmunity. However, SIZ1 also provides temperature-dependent disease resistance. At the elevated temperatures, *siz1* mutants cannot effectively resist pathogens because SIZ1 is needed to stabilise specific NB-LRR immune receptors or other defence machinery [12, 54]. In wheat, our observation of TaSIZ1 induction during pathogen challenge, especially in resistant backgrounds, suggests a positive contribution to defence under actual attack. One possible explanation is that upon pathogen detection, wheat elevates SIZ1 levels to enhance SUMOylation of defence regulators, thereby potentiating immune signalling or protecting key defence proteins from degradation [55]. SUMOylation in plant immunity can both suppress negative regulators and support positive regulators of defence, depending on context [11, 56]. Further experiments, i.e. silencing or overexpressing TaSIZ1 genes in wheat will be required to clarify their exact role in disease resistance. Nonetheless, the induction patterns and the known functions of SIZ1 in model plants strongly suggest that TaSIZ1 proteins contribute to immune responses of wheat, likely by maintaining a balance between activating defences and preventing hyperactivation that could be detrimental to growth.

Our current results highlight the biological significance of the *TaSIZ1* family for stress adaptation. By having multiple copies of the SIZ1 genes, wheat can deploy the SUMOylation machinery across various stress conditions, with each *TaSIZ1* optimised for specific signals. These multi-copies may provide combinatorial control and robustness, *i.e.*, if one pathway is compromised, alternative SIZ1 regulation can uphold the SUMOylation capacity. The fact that SIZ1 influences diverse processes in Arabidopsis [40, 52, 57] suggests that TaSIZ1 genes could have pleiotropic effects on wheat development and stress physiology. We observed basal expression of TaSIZ1 under non-stress conditions, suggesting roles in growth or reproduction warranting investigation. For instance, OsSIZ*1* in rice is involved in nitrogen and phosphate signalling and anther dehiscence [20, 58, 59]. Importantly, our results, together with prior studies, highlight TaSIZ1 genes as promising targets for crop improvement. SIZ1 is already considered a candidate gene for engineering stress-tolerant crops, as enhancing its expression can improve tolerance to drought, heat, and even multi-stress conditions without obvious downsides. In wheat, the presence of multiple homoeologs raises the possibility of modifying one copy using tools such as genome editing to boost stress responsiveness while retaining others for essential functions. One could envision editing TaSIZ1 promoters to create hyper-responsive TaSIZ1 alleles that turn on more strongly under stress, thereby ramping up protective SUMOylation when needed. Alternatively, breeding programs might exploit natural variation in *TaSIZ1* regulatory regions that confer higher drought- or pathogen-induced SIZ1 expression. Any such strategies must consider the delicate balance SIZ1 strikes between growth and defence; however, the reports suggest that increasing SIZ1 activity tends to enhance stress resilience.

Conclusion

Our pangenome-wide analysis identified 15 TaSIZ1 genes in bread wheat, grouped into three major evolutionary clades, whose expansion was primarily driven by whole-genome and segmental duplications and maintained under strong purifying selection (Ka/Ks < 1). Despite high structural conservation, TaSIZ1 genes exhibit pronounced regulatory diversification, with promoter regions enriched in stress- and hormone-responsive cis-elements (ABRE, MBS, ARE, GT1) and binding sites for key transcription factor families, including ERF, WRKY, MIKC_MADS, and MYB, enabling fine-tuned transcriptional control. Expression profiling across tissues and stress conditions, together with validation in genotypes with contrasting phenotypic responses, revealed strong, genotype-dependent induction of TaSIZ1-1A, TaSIZ1-1D, TaSIZ1-3B, and TaSIZ1-3D under leaf rust infection in the resistant genotype HD2888 and under drought stress in the tolerant genotype C306. Thus, the results demonstrate that SIZ1-type SUMO E3 ligases have evolved into a small yet functionally powerful multigene family in wheat, acting as master regulators of stress-responsive SUMOylation to amplify stress-signalling pathways and protect the proteome. This study provides the first comprehensive genomic and regulatory framework for TaSIZ1 genes in wheat and establishes a foundation for dissecting TaSIZ1-mediated molecular mechanisms, prioritizing *TaSIZ1* genes for multi-stress engineering, and accelerating the breeding of climate-resilient wheat cultivars.

Abbreviations

ABA
Abscisic acid
ABRE
ABA-responsive element
BLAST
Basic Local Alignment Search Tool
CDS
Coding sequence
GO
Gene Ontology
HMM
Hidden Markov Model
IWGSC
International Wheat Genome Sequencing Consortium
Ka/Ks
Ratio of non-synonymous to synonymous substitution rates

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing Interest:

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this manuscript.

Funding

The work is supported by the Department of Biotechnology, Government of India, through a project entitled “Mapping and transfer of novel resistance genes for multiple biotic stresses in wheat (*Triticum aestivum* L.) (BT/PR33041/AGIII/103/1166/2019)”. The funding agencies had no role in designing the study, data collection and analyses, the decision to publish, or the preparation of the manuscript.

Author Contribution

Hemant Sharma: Data curation, Formal Analysis, Writing- Original draft preparation. **Mallana Gowdra Mallikarjuna:** Conceptualization, Methodology, Formal Analysis, Supervision, Writing- Original draft preparation, Writing – review & editing. **Garudapalya Muniswamy Keerthi:** Methodology, Data curation. **Niharika Malik:** Data curation, Writing – review & editing. **Niranjana Murukan:** Data curation, Writing – review & editing. **Lekshmi Sathee:** Data curation, Writing – review & editing. **Shailendra Kumar Jha:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Acknowledgement:

We are thankful to “The National Phytotron Facility, Indian Agricultural Research Institute, New Delhi” for extending the controlled environment glasshouse facility.

Data Availability

All raw datasets were downloaded from publicly available databases. The remaining supporting data sets are included in the supplementary files.

References

1. Gill G. Something about SUMO inhibits transcription. Curr Opin Genet Dev. 2005;15:536–41. https://doi.org/10.1016/j.gde.2005.07.004.
2. Miura K, Hasegawa PM. Sumoylation and other ubiquitin-like post-translational modifications in plants. Trends Cell Biol. 2010;20:223–32. https://doi.org/10.1016/j.tcb.2010.01.007.
3. Miura K, Rus A, Sharkhuu A, Yokoi S, Karthikeyan AS, Raghothama KG et al. The *Arabidopsis* SUMO E3 ligase SIZ1 controls phosphate deficiency responses. Proceedings of the National Academy of Sciences. 2005;102:7760–5. https://doi.org/10.1073/pnas.0500778102
4. Catala R, Ouyang J, Abreu IA, Hu Y, Seo H, Zhang X, et al. The *Arabidopsis* E3 SUMO Ligase SIZ1 Regulates Plant Growth and Drought Responses. Plant Cell. 2007;19:2952–66. https://doi.org/10.1105/tpc.106.049981.
5. Seo PJ, Park JM, Kang SK, Kim SG, Park CM. An Arabidopsis senescence-associated protein SAG29 regulates cell viability under high salinity. Planta. 2011;233:189–200. https://doi.org/10.1007/S00425-010-1293-8.
6. Park HJ, Kim W-Y, Park HC, Lee SY, Bohnert HJ, Yun D-J. SUMO and SUMOylation in Plants. Mol Cells. 2011;32:305–16. https://doi.org/10.1007/s10059-011-0122-7.
7. Zhang X, Huai J, Liu S, Jin JB, Lin R. SIZ1-Mediated SUMO Modification of SEUSS Regulates Photomorphogenesis in Arabidopsis. Plant Commun. 2020;1:100080. https://doi.org/10.1016/j.xplc.2020.100080.
8. Balasubramaniam T, Mathangadeera R, Sugumar T, Wijewardene I, Sun L, Smith J, et al. Co-overexpression of OsSIZ1 and LtRCA in Arabidopsis thaliana improves heat, drought, and salt tolerance. Plant Sci. 2025;357:112536. https://doi.org/10.1016/j.plantsci.2025.112536.
9. Mishra N, Srivastava AP, Esmaeili N, Hu W, Shen G. Overexpression of the rice gene OsSIZ1 in Arabidopsis improves drought-, heat-, and salt-tolerance simultaneously. PLoS ONE. 2018;13:e0201716. https://doi.org/10.1371/journal.pone.0201716.
10. Kwak JS, Song JT, Seo HS. E3 SUMO ligase SIZ1 splicing variants localize and function according to external conditions. Plant Physiol. 2024;195:1601–23. https://doi.org/10.1093/plphys/kiae108.
11. Saleh A, Withers J, Mohan R, Marqués J, Gu Y, Yan S, et al. Posttranslational Modifications of the Master Transcriptional Regulator NPR1 Enable Dynamic but Tight Control of Plant Immune Responses. Cell Host Microbe. 2015;18:169–82. https://doi.org/10.1016/j.chom.2015.07.005.
12. Lee J, Nam J, Park HC, Na G, Miura K, Jin JB, et al. Salicylic acid-mediated innate immunity in Arabidopsis is regulated by SIZ1 SUMO E3 ligase. Plant J. 2007;49:79–90. https://doi.org/10.1111/j.1365-313X.2006.02947.x.
13. Li C, Yang C, Lai J, SUMOylation. Interplaying with the plant disease triangle. PLoS Pathog. 2025;21:e1013327. https://doi.org/10.1371/journal.ppat.1013327.
14. Jin JB, Jin YH, Lee J, Miura K, Yoo CY, Kim W, et al. The SUMO E3 ligase, *AtSIZ1*, regulates flowering by controlling a salicylic acid-mediated floral promotion pathway and through affects on *FLC* chromatin structure. Plant J. 2008;53:530–40. https://doi.org/10.1111/j.1365-313X.2007.03359.x.
15. Appels R, Eversole K, Stein N, Feuillet C, Keller B, Rogers J et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science (1979). 2018;361. https://doi.org/10.1126/science.aar7191
16. Augustine RC, Vierstra RD. SUMOylation: re-wiring the plant nucleus during stress and development. Curr Opin Plant Biol. 2018;45:143–54. https://doi.org/10.1016/j.pbi.2018.06.006.
17. Zhang C, Wu Y, Liu J, Song B, Yu Z, Li J-F, et al. SUMOylation controls peptide processing to generate damage-associated molecular patterns in Arabidopsis. Dev Cell. 2025;60:696–e7054. https://doi.org/10.1016/j.devcel.2024.11.010.
18. Liu S, Lenoir CJG, Amaro TMMM, Rodriguez PA, Huitema E, Bos JIB. Virulence strategies of an insect herbivore and oomycete plant pathogen converge on host E3 SUMO ligase SIZ1. New Phytol. 2022;235:1599–614. https://doi.org/10.1111/nph.18184.
19. Park HC, Kim H, Koo SC, Park HJ, Cheong MS, Hong H, et al. Functional characterization of the SIZ/PIAS-type SUMO E3 ligases, OsSIZ1 and OsSIZ2 in rice. Plant Cell Environ. 2010;33:1923–34. https://doi.org/10.1111/j.1365-3040.2010.02195.x.
20. Pei W, Jain A, Ai H, Liu X, Feng B, Wang X, et al. OsSIZ2 regulates nitrogen homeostasis and some of the reproductive traits in rice. J Plant Physiol. 2019;232:51–60. https://doi.org/10.1016/j.jplph.2018.11.020.
21. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant. 2020;13:1194–202. https://doi.org/10.1016/J.MOLP.2020.06.009.
22. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. The Proteomics Protocols Handbook. Totowa, NJ: Humana; 2005. pp. 571–607. https://doi.org/10.1385/1-59259-890-0:571.
23. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7. https://doi.org/10.1093/NAR/GKH340.

24. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3. https://doi.org/10.1093/bioinformatics/btp348.
25. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020;37:1530–4. https://doi.org/10.1093/molbev/msaa015.
26. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res. 2024;52:W78–82. https://doi.org/10.1093/nar/gkae268.
27. Qiao X, Li Q, Yin H, Qi K, Li L, Wang R, et al. Gene duplication and evolution in recurring polyploidization–diploidization cycles in plants. Genome Biol. 2019;20:38. https://doi.org/10.1186/s13059-019-1650-2.
28. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20:238. https://doi.org/10.1186/s13059-019-1832-y.
29. Zhang Z. KaKs_calculator 3.0: Calculating selective pressure on coding and non-coding sequences. Genomics Proteom Bioinf. 2022. https://doi.org/10.1016/J.GPB.2021.12.002.
30. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40. https://doi.org/10.1093/NAR/GKR1293.
31. Sharma H, Batra R, Kumar S, Kumar M, Kumar S, Balyan HS, et al. Identification and characterization of 20S proteasome genes and their relevance to heat/drought tolerance in bread wheat. Gene Rep. 2022;27:101552. https://doi.org/10.1016/j.genrep.2022.101552.
32. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504. https://doi.org/10.1101/gr.1239303.
33. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2 – ΔΔCT method. Methods. 2001;25:402–8. https://doi.org/10.1006/meth.2001.1262.
34. Bellec A, Sow MD, Pont C, Civan P, Mardoc E, Duchemin W, et al. Tracing 100 million years of grass genome evolutionary plasticity. Plant J. 2023;114:1243–66. https://doi.org/10.1111/tpj.16185.
35. Walkowiak S, Gao L, Monat C, Haberer G, Kassa MT, Brinton J, et al. Multiple wheat genomes reveal global variation in modern breeding. Nature. 2020;588:277–83. https://doi.org/10.1038/s41586-020-2961-x.
36. Burrows S, Dorussen D, Crudgington J, Di Santolo G, Simmonds J, Catoni M et al. Partial redundancy buffers deleterious effects of mutating *DNA methyltransferase 1–1* (*MET1-1*) in polyploid wheat. J Exp Bot. 2025;76:2500–16. https://doi.org/10.1093/jxb/eraf016
37. Wang T, van Dijk ADJ, Cai X, Wu J, Bonnema G, Wang X. Brassica diversity through the lens of polyploidy: genomic evolution, introgression, and homoeologous exchange. Hortic Plant J. 2025;11:1777–90. https://doi.org/10.1016/j.hpj.2025.08.002.
38. Chang C-C, Chen H-L, Li W-H, Chaw S-M. Dating the Monocot?Dicot Divergence and the Origin of Core Eudicots Using Whole Chloroplast Genomes. J Mol Evol. 2004;58:424–41. https://doi.org/10.1007/s00239-003-2564-9.
39. Miura K, Jin JB, Hasegawa PM. Sumoylation, a post-translational regulatory process in plants. Curr Opin Plant Biol. 2007;10:495–502. https://doi.org/10.1016/j.pbi.2007.07.002.
40. Jmii S, Cappadocia L, Plant, SUMO E3 Ligases. Function, Structural Organization, and Connection With DNA. Front Plant Sci. 2021;12. https://doi.org/10.3389/fpls.2021.652170.
41. Vinodh Kumar PN, Mallikarjuna MG, Jha SK, Mahato A, Lal SK, Y KR, et al. Unravelling structural, functional, evolutionary and genetic basis of SWEET transporters regulating abiotic stress tolerance in maize. Int J Biol Macromol. 2023;229:539–60. https://doi.org/10.1016/j.ijbiomac.2022.12.326.
42. Ramachandran SR, Mueth NA, Zheng P, Hulbert SH. Analysis of miRNAs in Two Wheat Cultivars Infected With Puccinia striiformis f. sp. tritici. Front Plant Sci. 2020;10. https://doi.org/10.3389/fpls.2019.01574.
43. Berckmans B, De Veylder L. Transcriptional control of the cell cycle. Curr Opin Plant Biol. 2009;12:599–605. https://doi.org/10.1016/j.pbi.2009.07.005.
44. Karami M, Fatahi N, Lohrasebi T, Razavi K. RAV transcription factor regulatory function in response to salt stress in two Iranian wheat landraces. J Plant Res. 2022;135:121–36. https://doi.org/10.1007/s10265-021-01356-7.
45. Coleman D, Kawamura A, Ikeuchi M, Favero DS, Lambolez A, Rymen B, et al. The SUMO E3 Ligase SIZ1 Negatively Regulates Shoot Regeneration. Plant Physiol. 2020;184:330–44. https://doi.org/10.1104/pp.20.00626.
46. Zhang S, Wang S, Lv J, Liu Z, Wang Y, Ma N, et al. SUMO E3 Ligase SlSIZ1 Facilitates Heat Tolerance in Tomato. Plant Cell Physiol. 2018;59:58–71. https://doi.org/10.1093/pcp/pcx160.
47. Saracco SA, Miller MJ, Kurepa J, Vierstra RD. Genetic Analysis of SUMOylation in Arabidopsis: Conjugation of SUMO1 and SUMO2 to Nuclear Proteins Is Essential. Plant Physiol. 2007;145:119–34. https://doi.org/10.1104/pp.107.102285.
48. Miura K, Jin JB, Lee J, Yoo CY, Stirm V, Miura T, et al. SIZ1-Mediated Sumoylation of ICE1 Controls *CBF3/DREB1A* Expression and Freezing Tolerance in *Arabidopsis*. Plant Cell. 2007;19:1403–14. https://doi.org/10.1105/tpc.106.048397.
49. Rytz TC, Miller MJ, McLoughlin F, Augustine RC, Marshall RS, Juan Y, et al. SUMOylome Profiling Reveals a Diverse Array of Nuclear Targets Modified by the SUMO Ligase SIZ1 during Heat Stress. Plant Cell. 2018;30:1077–99. https://doi.org/10.1105/tpc.17.00993.
50. Kurepa J, Walker JM, Smalle J, Gosink MM, Davis SJ, Durham TL, et al. The Small Ubiquitin-like Modifier (SUMO) Protein Modification System in Arabidopsis. J Biol Chem. 2003;278:6862–72. https://doi.org/10.1074/jbc.M209694200.
51. Karan R, Subudhi PK. A stress inducible SUMO conjugating enzyme gene (SaSce9) from a grass halophyte Spartina alterniflora enhances salinity and drought stress tolerance in Arabidopsis. BMC Plant Biol. 2012;12:187. https://doi.org/10.1186/1471-2229-12-187.

52. Miura K, Lee J, Jin JB, Yoo CY, Miura T, Hasegawa PM. Sumoylation of ABI5 by the *Arabidopsis* SUMO E3 ligase SIZ1 negatively regulates abscisic acid signaling. Proceedings of the National Academy of Sciences. 2009;106:5418–23. https://doi.org/10.1073/pnas.0811088106
53. Hammoudi V, Fokkens L, Beerens B, Vlachakis G, Chatterjee S, Arroyo-Mateos M, et al. The Arabidopsis SUMO E3 ligase SIZ1 mediates the temperature dependent trade-off between plant immunity and growth. PLoS Genet. 2018;14:e1007157. https://doi.org/10.1371/journal.pgen.1007157.
54. Hammoudi V, Beerens B, Jonker MJ, Helderman TA, Vlachakis G, Giesbers M, et al. The protein modifier SUMO is critical for integrity of the Arabidopsis shoot apex at warm ambient temperatures. J Exp Bot. 2021. https://doi.org/10.1093/jxb/erab262.
55. Gou M, Huang Q, Qian W, Zhang Z, Jia Z, Hua J. Sumoylation E3 Ligase SIZ1 Modulates Plant Immunity Partly through the Immune Receptor Gene *SNC1* in *Arabidopsis*. Mol Plant-Microbe Interactions®. 2017;30:334–42. https://doi.org/10.1094/MPMI-02-17-0041-R.
56. Saleh A, Withers J, Mohan R, Marqués J, Gu Y, Yan S, et al. Posttranslational Modifications of the Master Transcriptional Regulator NPR1 Enable Dynamic but Tight Control of Plant Immune Responses. Cell Host Microbe. 2016;19:127–30. https://doi.org/10.1016/j.chom.2015.12.008.
57. Lim S, Park J, Lee N, Jeong J, Toh S, Watanabe A, et al. ABA-INSENSITIVE3, ABA-INSENSITIVE5, and DELLAs Interact to Activate the Expression of *SOMNUS* and Other High-Temperature-Inducible Genes in Imbibed Seeds in *Arabidopsis*. Plant Cell. 2014;25:4863–78. https://doi.org/10.1105/tpc.113.118604.
58. Wang H, Sun R, Cao Y, Pei W, Sun Y, Zhou H, et al. *OsSIZ1*, a SUMO E3 Ligase Gene, is Involved in the Regulation of the Responses to Phosphate and Nitrogen in Rice. Plant Cell Physiol. 2015;56:2381–95. https://doi.org/10.1093/pcp/pcv162.
59. Thangasamy S, Guo C, Chuang M, Lai M, Chen J, Jauh G. Rice *SIZ1*, a SUMO E3 ligase, controls spikelet fertility through regulation of anther dehiscence. New Phytol. 2011;189:869–82. https://doi.org/10.1111/j.1469-8137.2010.03538.x.

Figures

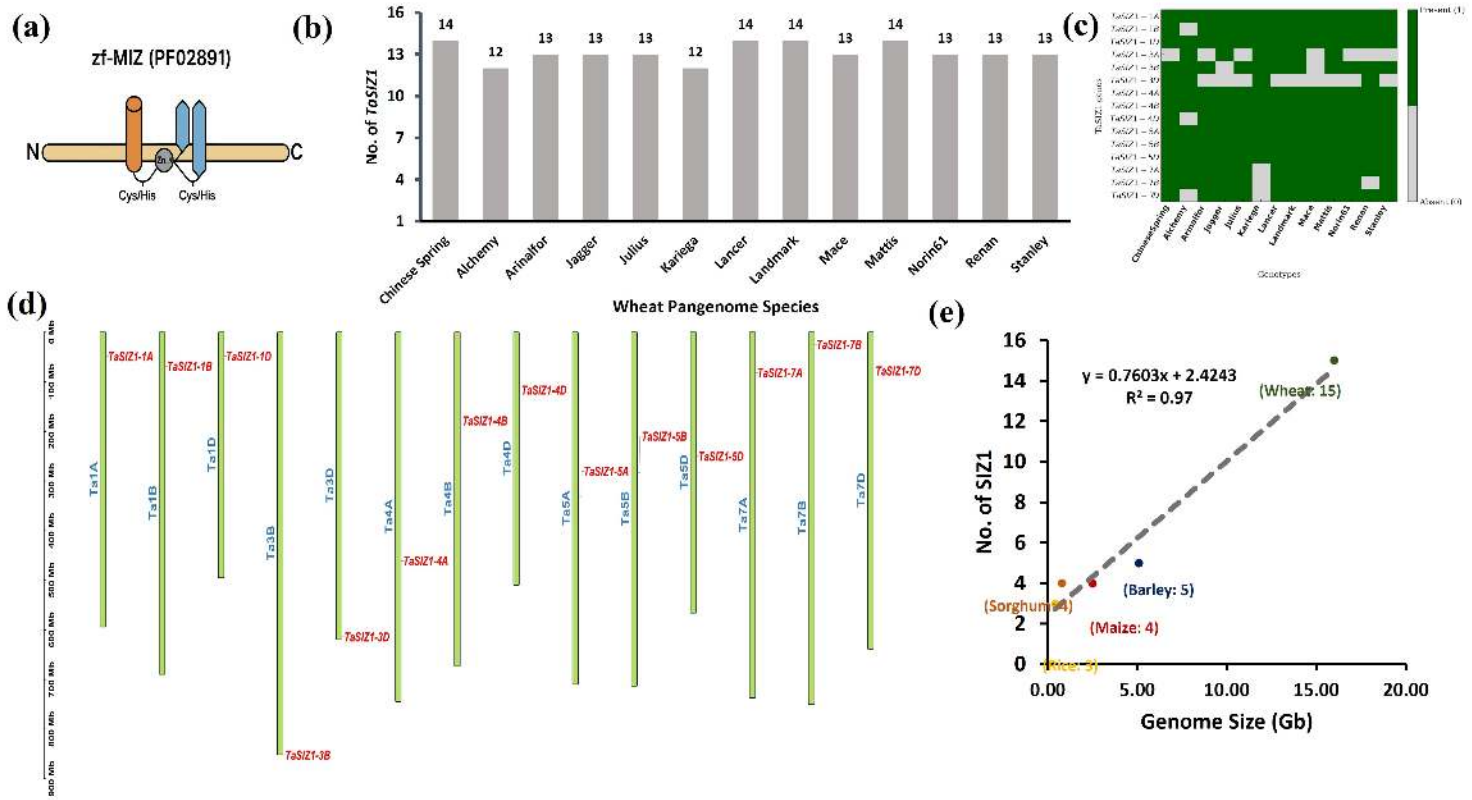

Figure 1. The feature and distribution of the *SIZ1* genes. **(a)** The zf-MIZ1 PFAM domain is considered for mining SIZ1 proteins in wheat and related species. **(b)** The number of TaSIZ1 genes identified through genome-wide search in the wheat pan-genome species. **(c)** The presence (green) – absence (grey) matrix of TaSIZ1 distribution among the thirteen wheat pangenome species. Each row represents one of the 15 TaSIZ1 genes, and each column represents a genotype; green cells indicate presence, while grey cells indicate absence. **(d)** The physical map of TaSIZ1 genes across the 21 wheat chromosomes of the reference genome species Chinese Spring. Vertical green bars represent chromosomes, the scale on the left shows chromosomal positions in Mb, and red labels indicate the positions of each TaSIZ1 gene. **(e)** The scatter plot shows the association between the number of SIZ1 genes and genome size in the target species.

Figure 1

See image above for figure legend

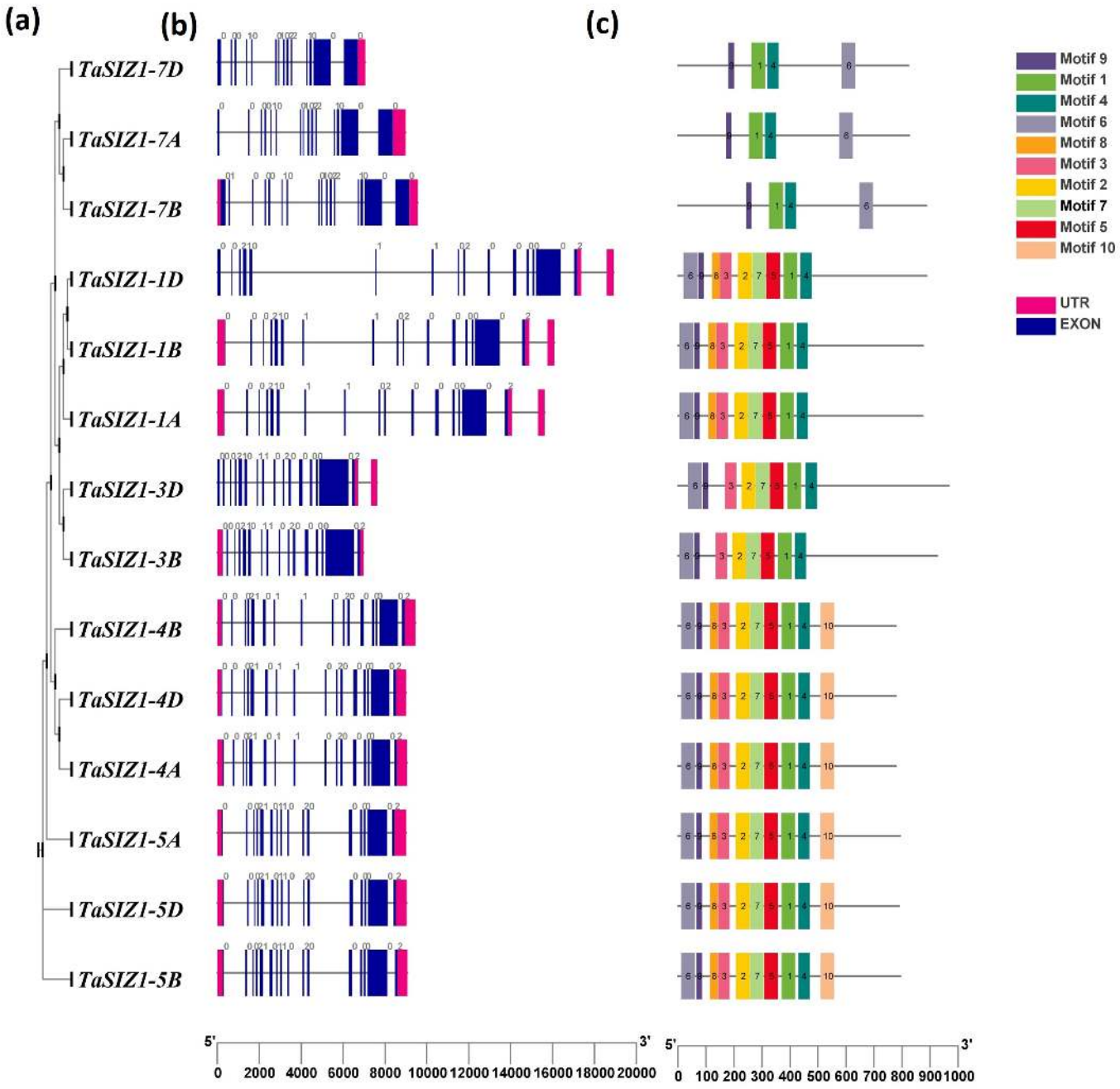

Figure 2. The phylogenetic association, gene structure and distribution of conserved motifs in *TaSIZ1* genes and protein sequences. **(a)** The gene architectures of *TaSIZ1* genes depict the distribution of exons, introns and intron phases. The black line, blue and pink boxes indicate introns, exons, and untranslated regions (UTRs), respectively. **(b)** Distribution of the best ten conserved motifs in the *TaSIZ1* proteins. Each motif is depicted with a specific colour.

Figure 2

See image above for figure legend

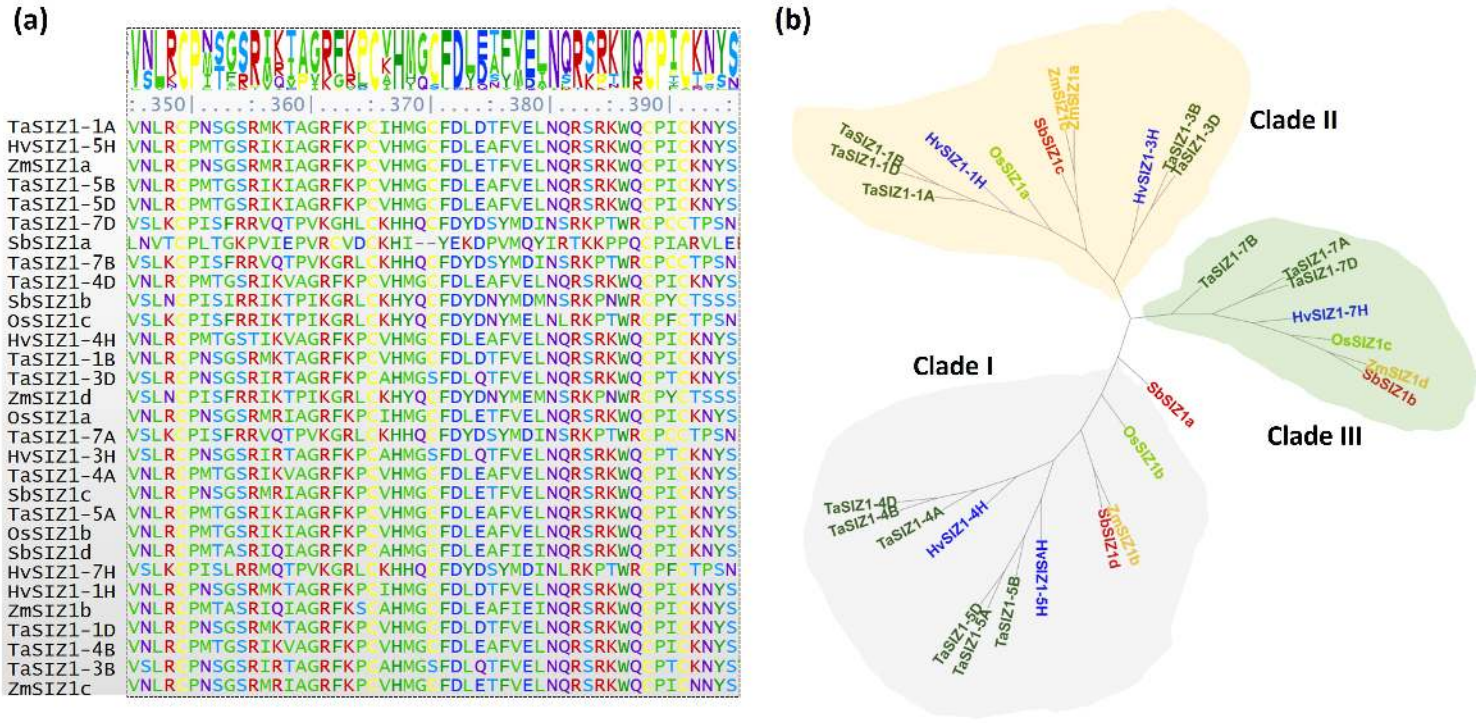

Figure 3. The multiple sequence alignment and phylogenetic tree analysis of SIZ1 proteins. **(a)** The multiple sequence alignment 30 SIZ1 sequences from wheat, barley, rice, maize, and sorghum showed alignment of the conserved zf-MIZ1 PFAM domain. **(b)** The phylogenetic association of SIZ1 proteins among the target species. The SIZ1 proteins from wheat, barley, rice, maize, and sorghum are represented in dark green, blue, light green, yellow, and red, respectively. The phylogenetic analysis was performed using the maximum likelihood method and the best-fit model JTT+G4, with 5000 ultra-bootstraps.

Figure 3

See image above for figure legend

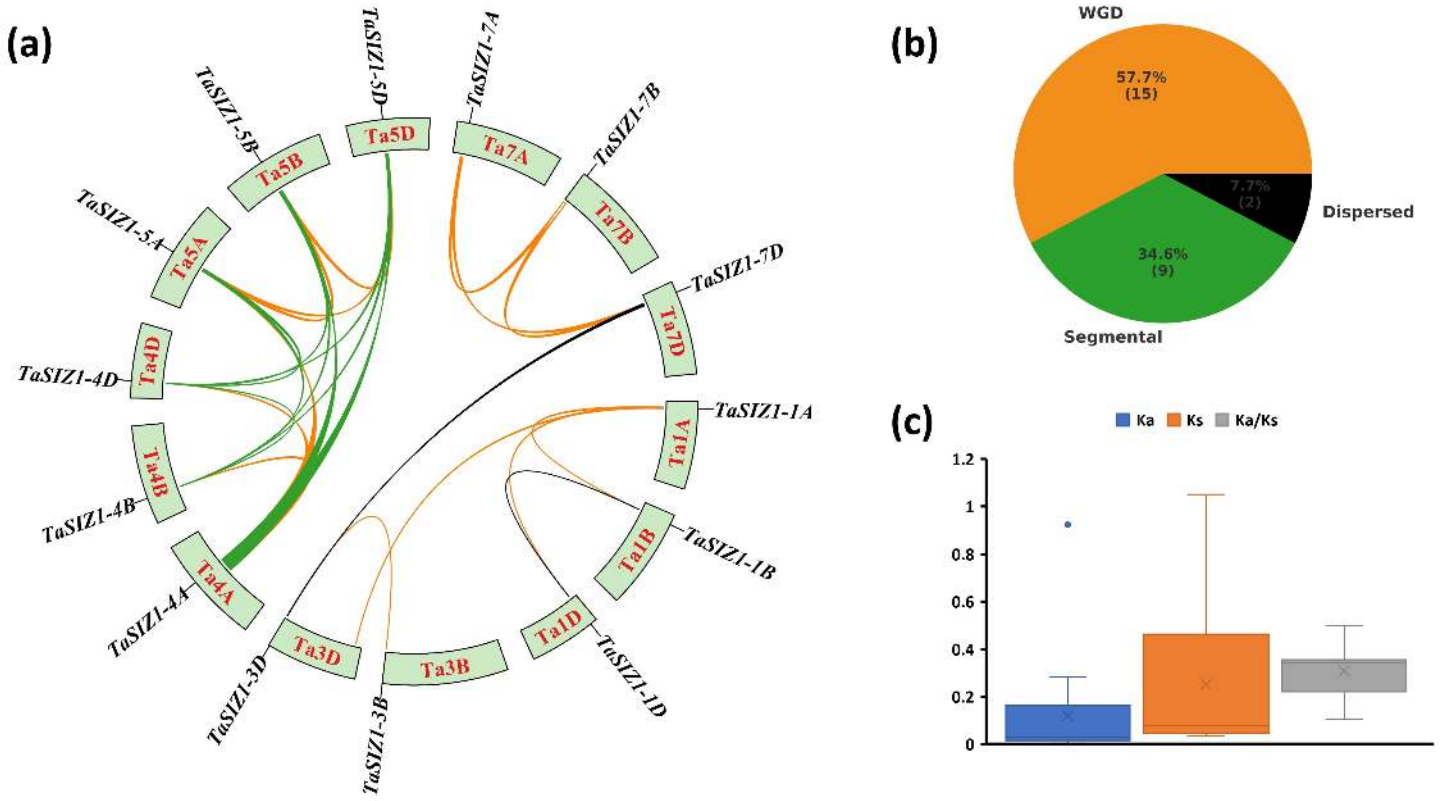

Figure 4. Duplication landscape and selection pressure on TaSIZ1 genes. **(a)** The chromosomal localisation and paralog association of TaSIZ1 genes. The orange, green, and black lines linking the paralogs represent whole-genome (WGD), segmental, and dispersed duplications, respectively. **(b)** Pie chart summarising the duplication types among 25 TaSIZ1 paralog pairs. **(c)** The extent of synonymous (Ks), non-synonymous (Ka) and non-synonymous by synonymous (Ka/Ks) mutations shaping the expansion of TaSIZ1s genes in the wheat genome.

Figure 4

See image above for figure legend

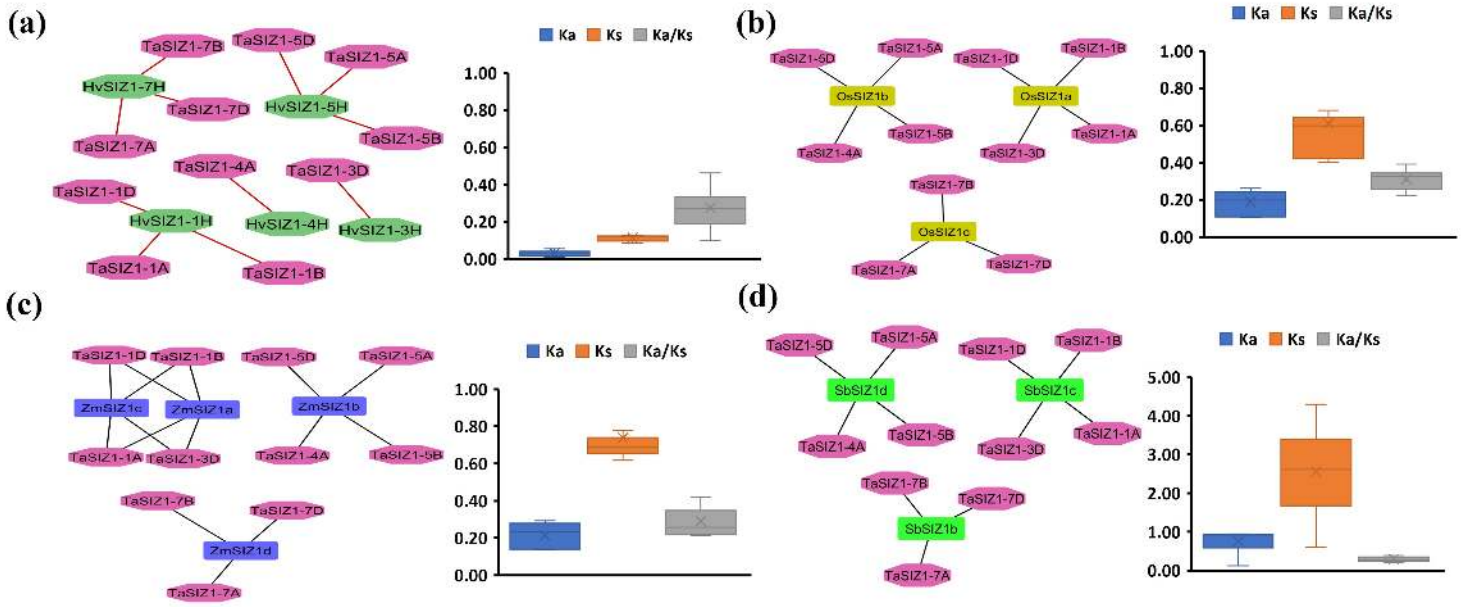

Figure 5. Orthologs and selection pressure analyses of SIZ1 genes. Ortholog associations and the extent of synonymous (Ks), non-synonymous (Ka), and non-synonymous-to-synonymous (Ka/Ks) mutations shaping the divergence of wheat SIZ1 from **(a)** barley (Hv), **(b)** rice (Os), **(c)** maize (Zm), and **(d)** sorghum (Sb). The boxplots display the Ka, Ks, and Ka/Ks distributions for each species ortholog pairs.

Figure 5

See image above for figure legend

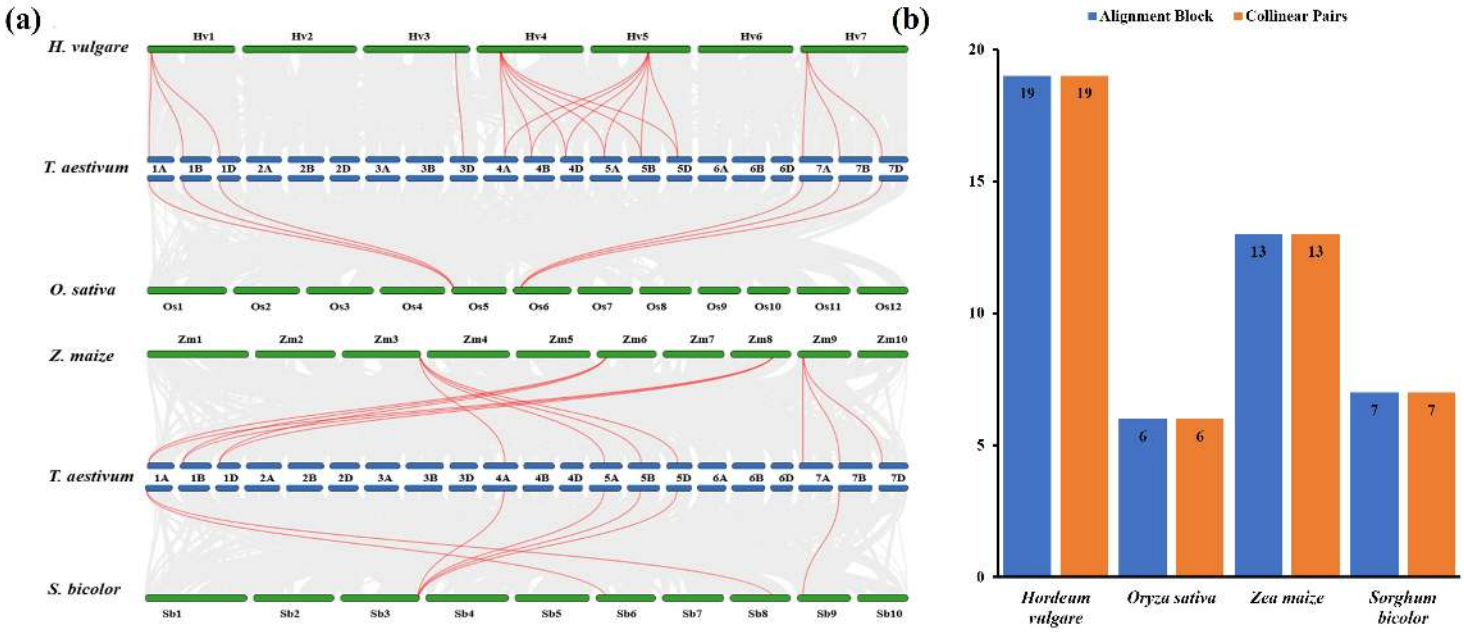

Figure 6. Syntenic association between wheat and related species for TaSIZ1 genes. **(a)** Syntenic relationships among the SIZ1 genes in wheat, barley, rice, and maize. Grey lines in the background represent colinear and syntenic blocks within the genomes of wheat, barley, maize, and rice, while red lines indicate syntenic TaSIZ1 gene pairs between two genomes. **(b)** The bar diagrams show the number of collinear SIZ1 genes and syntenic blocks between wheat and barley, maize, sorghum, and rice.

Figure 6

See image above for figure legend

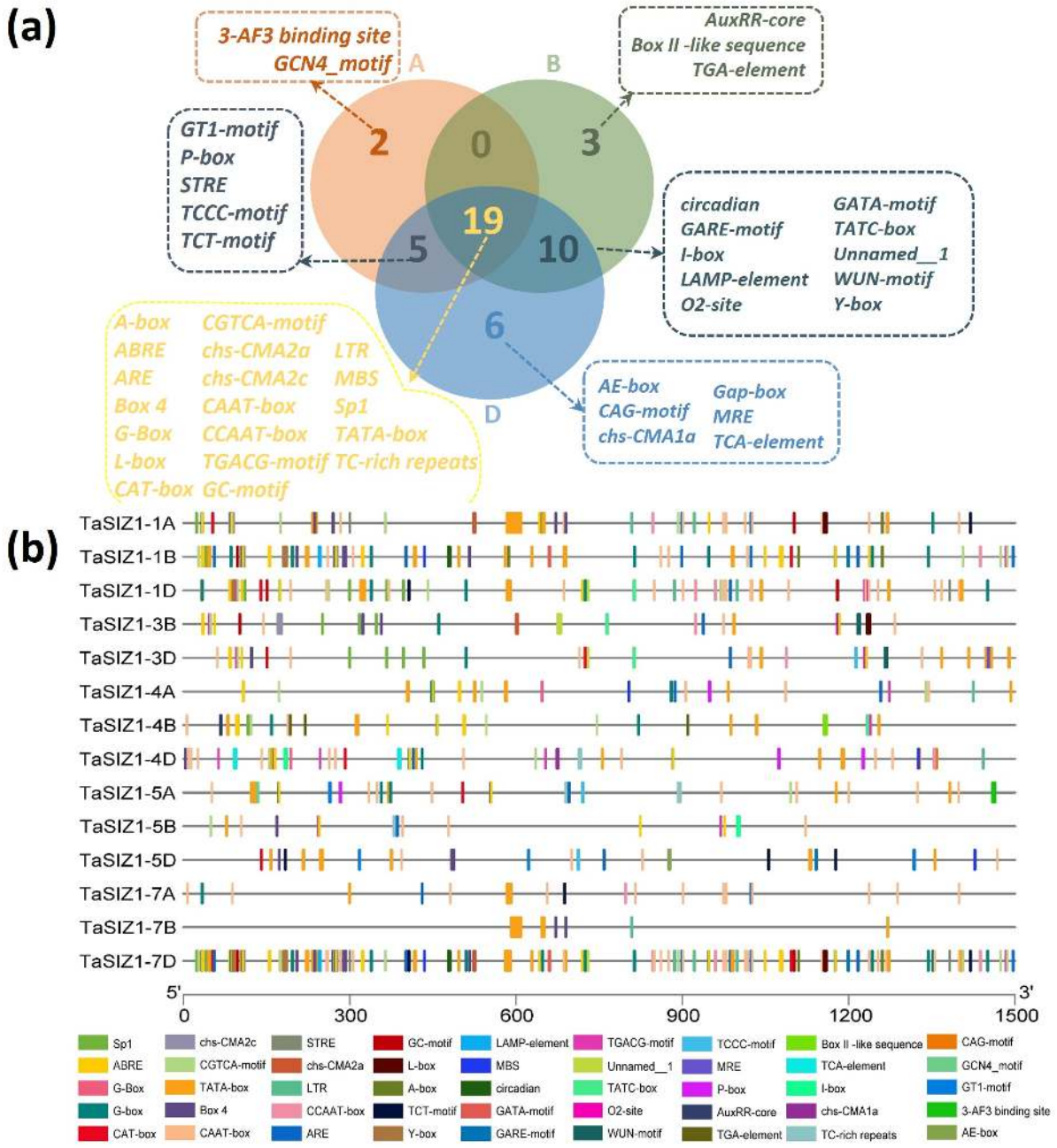

Figure 7. Distribution of *cis*-acting regulatory elements in *TaSIZ1* promoters. **(a)** Venn diagram summarising the distribution and overlap of TaSIZ1 *cis* elements among the wheat sub-genomes, and **(b)** Schematic representation of *cis*-element positions across the 1.5 kb promoters of all TaSIZ1 genes. Each horizontal bar represents one TaSIZ1 promoter and each coloured symbol denotes a specific *cis* element.

Figure 7

See image above for figure legend

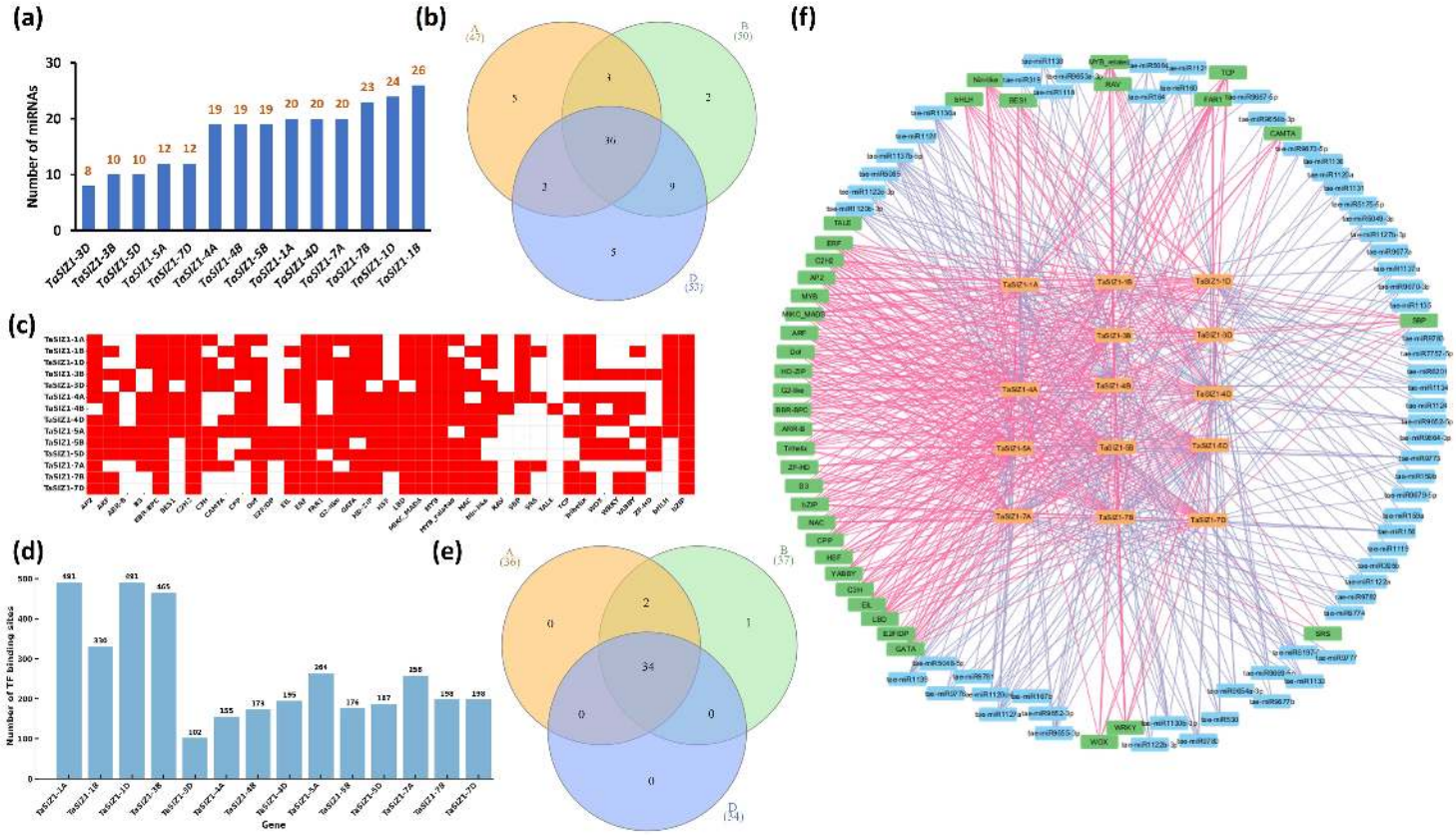

Figure 8. The regulatory snapshot of TaSIZ1 genes in wheat. **(a)** The bar graph showing the number of regulatory miRNAs predicted for each of the TaSIZ1 genes. **(b)** The Venn diagram showing sub-genome-wise distribution of predicted miRNAs for TaSIZ1 genes. **(c)** The heatmap showing the transcription factors predicted for the TaSIZ1 genes. The red and white colour squares represent the presence and absence of transcription factors with respective TaSIZ1 genes. **(d)** The bar diagram represents the total number of TF binding sites predicted in each of the TaSIZ1 genes. **(e)** The Venn diagram showing sub-genome-wise distribution of predicted TFs for TaSIZ1 genes. **(f)** The regulatory network of TaSIZ1 genes by miRNAs and TFs. The nodes coloured orange, green, and blue represent TaSIZ1 genes, TFs, and miRNAs. The pink and purple edges represent the interaction between TaSIZ1-TF and TF-miRNAs.

Figure 8

See image above for figure legend

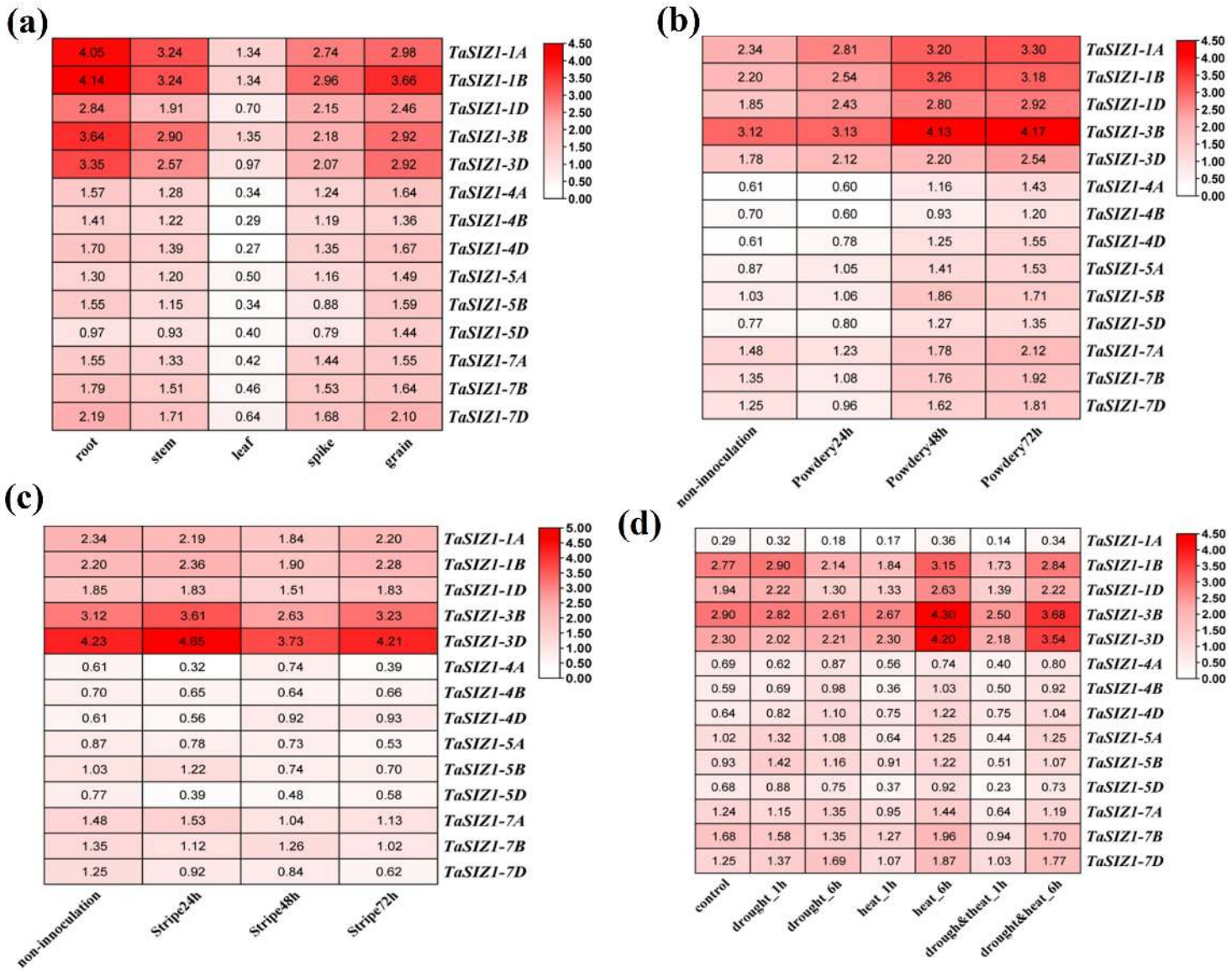

Figure 9. *In-silico* expression profiles of *TaSIZ1* genes in wheat across tissues and stress conditions. **(a)** Expression of *TaSIZ1* genes in root, stem, leaf, spike and grain tissues. **(b)** Expression of *TaSIZ1* genes in response to powdery mildew infection at 0, 24, 48, 72 h post-inoculation stages. **(c)** Expression of *TaSIZ1* genes in response to stripe rust infection at 0, 24, 48, 72 h post-inoculation stages. (d) under drought, heat, and combined drought–heat stress at multiple time points. Red indicates higher expression levels, while blue indicates lower expression.

Figure 9

See image above for figure legend

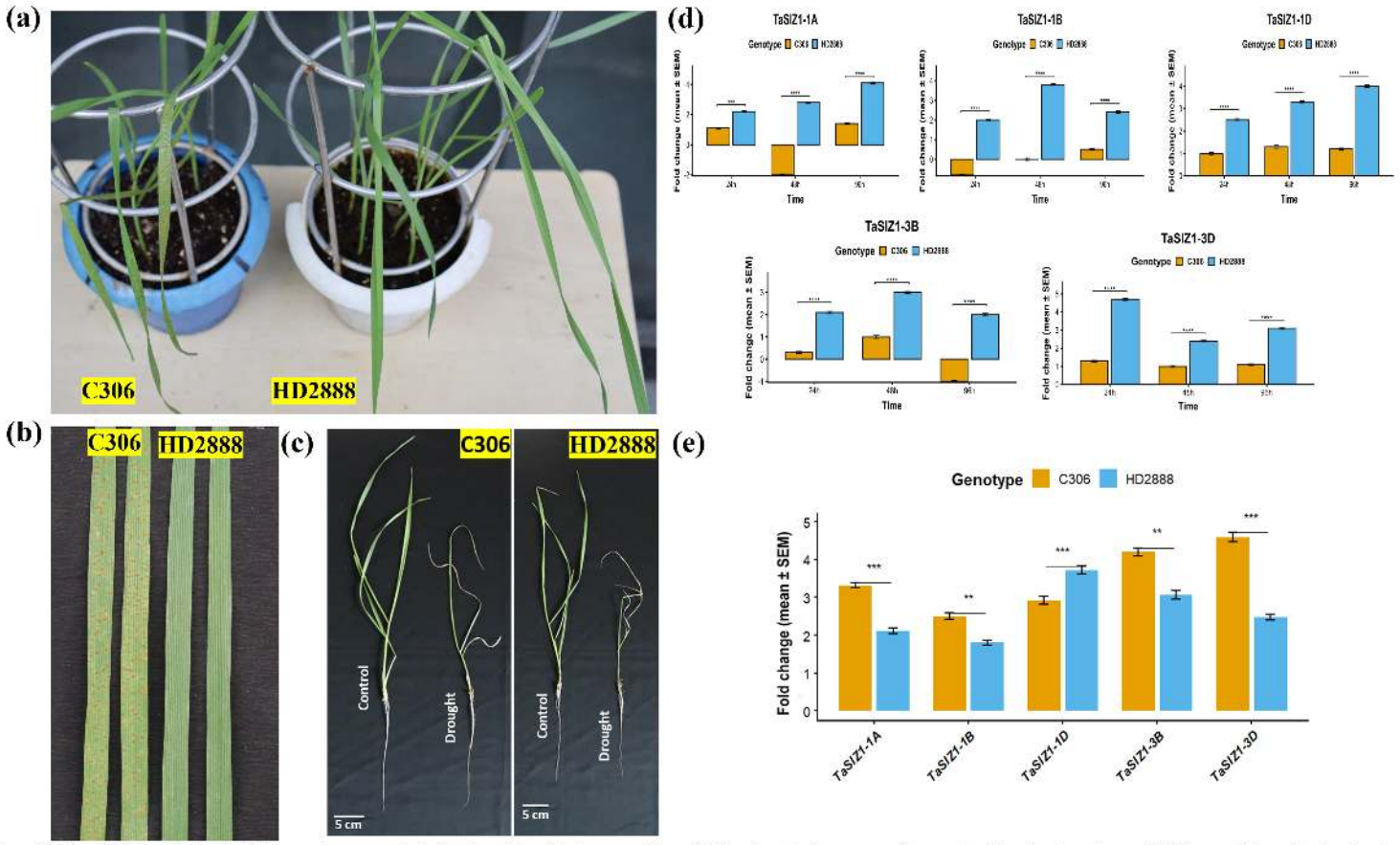

Figure 10. The qRT-PCR analysis of *TaSIZ1* genes in response to leaf rust and drought stress conditions. **(a)** The phenotypic response of two contrasting wheat genotypes, C306 (susceptible; yellow bars) and HD288 (resistant; blue bars), at three time points (24 h, 48 h, and 96 h post-inoculation) for leaf rust infection. **(b)** The leaf rust disease reaction on contrasting genotypes' leaves. **(c)** Expression profiles of *TaSIZ1* homoeologs in two wheat genotypes, C306 (susceptible; yellow bars) and HD288 (resistant; blue bars), at three time points (24 h, 48 h, and 96 h post inoculation) under leaf rust inoculation. **(d)** The phenotypic response of C306 (drought-tolerant; yellow bars) and HD288 (comparatively drought-sensitive; blue bars) lines in response to PEG-induced drought stress under hydroponic conditions for seven days. **(e)** Expression analysis of TaSIZ1 genes under 7-day drought stress conditions in wheat genotypes C306 and HD2888. *Note:* Expression levels are shown as fold change relative to control (0 h), with error bars indicating standard error of the mean (SEM). The **, *** denote statistically significant differences between genotypes at $p < 0.01$ and 0.001, respectively.

Figure 10

See image above for figure legend

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1SIZ1.docx
- STablesR.xlsx