# Identification and characterization of novel *SUMO* genes in bread wheat

Eid I. Ibrahim[1], Kotb A. Attia[2], Abdelhalim Ghazy[1], Itoh Kimiko[3] and Abdullah Al-Doss[1]

[1] Plant Production Department, College of Food and Agriculture Sciences, King Saud University, Riyadh, Saudi Arabia
[2] Center of Excellence in Biotechnology Research, King Saud University, Riyadh, Saudi Arabia
[3] Institute of Science and Technology, Niigata University, Niigata, Japan

## ABSTRACT

The small-ubiquitin-like-modifier (SUMO), composed of approximately 100 amino acid residues, regulates the cellular activity of several proteins by posttranslational modification. Almost all plant species express a family of SUMO isoforms. Nevertheless, the *SUMO* genes in wheat (*TaSUMOs*) have not undergone complete characterization, and the roles of *TaSUMOs* remain unidentified. The study identified four new *SUMO* genes in wheat, named *TaSUMO4-7*, in addition to the previously known *TaSUMO1-3*. These genes are part of the conserved SUMO family, as indicated by phylogenetic analysis. The genes contain the characteristic SUMO-acceptor site motif and the essential C-terminal diglycine motif for processing. Expression analysis showed that *TaSUMO4-7* genes are expressed in various wheat tissues. Bioinformatics analysis predicted the biochemical properties and structures of the proteins, which were found to localize in the cytoplasm and nucleus. The study confirms that the new *TaSUMO4-7* genes are functional members of the wheat SUMO family and lays the groundwork for further research into their specific roles.

**Subjects** Agricultural Science, Bioinformatics, Genetics, Genomics, Molecular Biology
**Keywords** SUMO genes, Wheat, Bioinformatics, Gene/Protein expression

Submitted 25 February 2025
Accepted 30 October 2025
Published 28 November 2025

Corresponding authors
Eid I. Ibrahim,
eidibrahim1986@gmail
Abdelhalim Ghazy,
aghazy@ksu.edu.sa

Academic editor
Diaa Abd El-Moneim

Additional Information and Declarations can be found on page 18

DOI **10.7717/peerj.20432**



## INTRODUCTION

Among the various posttranslational modifications (PTMs) crucial for cellular function in eukaryotes, SUMOylation stands out for its similarity and significance alongside the widely studied ubiquitination pathway. However, diverse PTMs, including phosphorylation, lipidation, glycosylation, acetylation, and ubiquitination, all contribute to regulating various cellular processes (*Kerscher, Felberbaum & Hochstrasser, 2006*; *Clague, Coulson & Urbé, 2012*; *Isono & Nagel, 2014*; *Wang, Peterson & Loring, 2014*). In contrast to ubiquitination, which mainly aims to degrade substrates, SUMOylation controls substrate functions primarily by changing protein protein interactions, intracellular localization or other forms of PTMs (*Zhao, 2007*). SUMOylation is a vital controlling tool in eukaryotic cells that influences diverse processes, such as enzyme stability and activity, subnuclear localization of proteins, cell cycle regulation, and DNA repair (*Miura, Jin & Hasegawa, 2007a*). SUMOylation is achieved through a multistep enzymatic pathway involving three key players: E1 (a small-ubiquitin-like-modifier (SUMO)-activating-enzyme), E2 (a

SUMO-conjugating-enzyme), and E3 (a SUMO-ligating-enzyme) (*Miura, Jin & Hasegawa, 2007a*).

SUMOylation refers to the modification of a substrate using a small-ubiquitin-like modifier. SUMO belongs to the category of low-molecular-weight proteins, which typically contain between 100 and 130 amino acids (*Kim, Baek & Chung, 2002*). In 1999, SUMO was first discovered in tomato plants as an interaction partner with the plant pathogenic fungus *Trichoderma viride*'s ethylene-inducing-xylanase (EIX) (*Hanania et al., 1999*). Ubiquitous across eukaryotic cells, the SUMO protein is known as a crucial controlling mechanism that regulates various cellular pathways through targeted protein modification (*Morrell & Sadanandom, 2019*; *Le Roux et al., 2019*). Unlike worms, fruit flies, and yeasts, plants express numerous SUMO isoforms that may indicate SUMO crosstalk and/or substrate specificity (*Flotho & Melchior, 2013*; *Van den Burg et al., 2010*). Seven *SUMO*-genes have been discovered in *Oryza sativa* (*Ikarashi et al., 2012*; *Rosa et al., 2018*; *Ibrahim et al., 2021*; *Teramura et al., 2021*), and three *SUMO*-genes have been recognized in wheat (*Wang et al., 2018*); while, in the *Arabidopsis thaliana* genome, eight *SUMO*-genes have been reported (*Kurepa et al., 2003*; *Novatchkova et al., 2004*; *Hay, 2005*).

The SUMO-precursor is shacked in the C terminal GG motif *via* a SUMO-specific-protease and a subsequent pathway involving E1 (SAEs), E2 (SCEs), and E3 (SLEs), which tie mature SUMO to the certain protein by covalently binding to lysine (K) residues (*Bernier-Villamor et al., 2002*; *Lois & Lima, 2005*; *Yunus & Lima, 2006*; *Miura et al., 2007b*; *Yaeno & Iba, 2008*; *Nayak & Müller, 2014*). The attachment of SUMO to a certain protein (substrate) is not permanent but rather a dynamic process controlled by SUMO proteases. These enzymes are believed to be crucial regulators of SUMOylation activity (*Le Roux, 2016*). Most SUMO-targeted proteins possess a specific recognition motif known as the consensus pattern, typically represented as ΨKXE/D. This pattern includes a hydrophobic residue (Ψ) followed by a lysine (K), then any amino acid (X), followed by either glutamic acid (E) or aspartic acid (D) (*Lois, Lima & Chua, 2003*; *Rosa et al., 2018*; *Morrell & Sadanandom, 2019*).

*SUMO* genes are believed to be vital for regulating a range of environmental interactions. It has been shown to control a variety of plant biological mechanisms, like abiotic (*Van den Burg et al., 2010*; *Augustine & Vierstra, 2018*; *Han et al., 2022*) and biotic (*Flick & Kaiser, 2009*) stresses tolerance, plant development (*Castro et al., 2018*), root growth (*Zhang, Qi & Yang, 2010*) and plant reproduction (*Augustine et al., 2016*). While the importance of SUMOylation and its related genes is well established, significant gaps remain in our understanding of *SUMO* genes in bread wheat. These gaps encompass their specific functions, evolutionary connections to other crops, and potential involvement in managing diverse environmental stresses. Given the crucial role of SUMOs in agriculturally relevant traits, elucidating the specific functions of SUMOs in wheat has immense potential for enhancing both grain quality and yield.

While the study by *Wang et al. (2018)* successfully identified three *TaSUMO* genes (*TaSUMO1–3*) in wheat through homology-based searches using Arabidopsis and rice sequences, their approach did not fully capture the diversity of the *SUMO* gene family in the wheat genome. Our study expands upon their findings by employing more stringent

bioinformatic criteria ($E$-value $<1.0^{e-5}$, % identity $\geq 50$) and additional validation steps, including Pfam and InterPro domain analysis, to ensure the accuracy of gene identification. Crucially, we combined computational predictions with experimental validation, using PCR amplification, sequencing, and tissue-specific expression profiling to confirm the existence of four new *TaSUMO* genes (*TaSUMO4–7*). Furthermore, we conducted functional characterization through subcellular localization assays (*via* fluorescent protein tagging and confocal microscopy) and interaction studies with the SUMO-conjugation enzyme *SCE1a*, providing mechanistic insights that were absent in the earlier work. This study lays the groundwork for future in-depth investigations of the *TaSUMO* gene family by utilizing powerful bioinformatic tools and insights.

# MATERIALS & METHODS

## Identification of *SUMO* genes in wheat

Three wheat *SUMO*-genes have been reported (*Wang et al., 2018*) specifically *TaSUMO1* (TraesCS3A02G424000.1, TraesCS3B02G460000.1, TraesCS3D02G419300.1), *TaSUMO2* (TraesCS3B02G459900.1, TraesCS3D02G419200.1), and *TaSUMO3* (TraesCS3A02G366700.1, TraesCS3D02G359600.1). We obtained their sequences from the Ensembl Plants (http://plants.ensembl.org/). To identify new *SUMO* genes in *Triticum aestivum* L. (*TaSUMOs*), *SUMO* sequences in *Arabidopsis thaliana* (*AtSUMOs*) were obtained from TAIR (http://www.arabidopsis.org, accessed on 16 January 2022) and rice *SUMO* sequences (*OsSUMOs*) derived from the Rice Annotation Project (https://rice.uga.edu/, accessed on 16 January 2022) were used. These sequences served as queries for exploring the wheat database within the Ensembl Plants datasets (http://plants.ensembl.org/Triticum_aestivum/Info/Index, accessed on 17 January 2022). To ensure the accuracy of subsequent analyses, stringent filtering criteria were applied (E-val $< 1.0^{e-5}$, %ID $\geq 50$). This involved removing redundant genes and selecting the longest transcript representative of each gene family member. Further validation of the identified TaSUMO family members was conducted using the InterPro website (https://www.ebi.ac.uk/interpro/, accessed on 19 January 2022) by submitting the predicted protein sequences.

## Plant material growth and DNA extraction

Seeds of *Triticum aestivum* L. cv. VEERY (Chilean variety) were germinated in a controlled environment chamber under the following conditions: 23 °C ± 2 °C throughout the day, 16 °C ± 2 °C at night, 70% relative humidity, 16 h of light and 8 h of darkness (long-day conditions), and a light intensity of approximately 60 μmol $m^{-2}$ $s^{-1}$ at the Plant Production Department, College of Food and Agriculture Sciences, King Saud University, Riyadh, Saudi Arabia. The cultivar grains were obtained from the International Maize and Wheat Improvement Center (CIMMYT). The molecular experiments were conducted at the Biotechnology Laboratory within the Plant Production Department, KSU. Total DNA was extracted from young wheat leaves *via* the Wizard® Genomic DNA Purification Kit manufactured by Promega (Madison, WI, USA).

## Candidate gene amplification and sequencing of putative *TaSUMOs*

The putative *TaSUMO4-7* sequences were amplified by the definite primers presented in Table S1. These primers were designed based on the genomic data of the *SUMO*-region in the *Triticum aestivum* genome database. The PCR amplification involved the addition of 10 μL of GoTaq® Green Master Mix (1×) (Promega Corporation, Madison, WI, USA), two μL of mixed primers, six μL of water, and two μL of DNA. The cycling procedure was 95 °C for 5 min (initial denaturation), followed by 35 cycles of 95 °C for 30 s (denaturation), 52−64 °C for 30 s (annealing), 72 °C for 1 min (extension), and a final extension at 72 °C for 7 min. Amplified DNA (10 μL) was visualized on a 3% agarose gel. The expected bands were purified with a QIAquick® PCR Purification Kit and sequenced by Macrogen® (Seoul, South Korea). The coding regions were analysed for SUMO family characteristics.

## Bioinformatic characterization of *TaSUMOs*

Characterization of the *TaSUMO* gene family included features such as amino acid count, chromosomal location, gene ID, length of coding sequences (CDSs), and exon number extracted from the Ensembl Plants database. The biochemical characteristics of the SUMO proteins were analysed using the ExPASy ProtParam server tool (https://web.expasy.org/protparam/, accessed on 20 September 2025) (*Gasteiger et al., 2003*). This included calculating various parameters of the primary structure, like the amino acid composition, molecular weight (MW), and theoretical isoelectric point (T-pi). The grand average of hydropathicity (GRAVY) was calculated using the method described by *Kyte & Doolittle (1982)*, while the instability index (Ii) was calculated using the method described by *Guruprasad, Reddy & Pandit (1990)*. This latter analysis considers the frequency of specific dipeptides as an indicator of protein stability, with proteins having an Ii less than 40 considered stable and those above 40 considered unstable. Finally, the aliphatic index (Ai) was also calculated using the formula proposed by *Ikai (1980)*.

The Protein Homology/analogY Recognition Engine V 2.0, Phyre2 web portal (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index/, accessed on 21 January 2022) (*Kelley et al., 2015*) was employed to computationally predict the secondary and three-dimensional (3D) structures of the proteins. Computational prediction of SUMO-binding sites within the TaSUMO genes was performed using the SUMOplot™ Analysis Program (https://www.abcepta.com/sumoplot/, accessed on 17 January 2022). Additionally, the pairwise sequence identity among the TaSUMO protein sequences was detected *via* Vector NTI® Software (Invitrogen, Waltham, MA, USA).

## TaSUMO sequence comparison

Phylogenetic analysis incorporated diverse data sets, including the three identified TaSUMO protein sequences from *Wang et al. (2018)*, eight AtSUMO sequences reported by *Kurepa et al. (2003)*, *Novatchkova et al. (2004)*, and *Hay (2005)*, and seven OsSUMO sequences described by *Ikarashi et al. (2012)*, *Rosa et al. (2018)*, *Ibrahim et al. (2021)*, *Teramura et al. (2021)*, and *Ibrahim (2022)*. SUMO protein sequences from humans, yeast, and other species were got from the National Center for Biotechnology Information (NCBI) and the UniProt database. Multiple sequence alignments were performed through the UniProt

database website (https://www.uniprot.org/align/, accessed on 17 January 2022) using the ClustalO function. The SUMO protein sequences were queried against the InterPro website (*Finn et al., 2007*; *Larkin et al., 2007*) to identify conserved domains characteristic of the SUMO family. A phylogenetic tree was constructed using the neighbor−joining method (*Saitou & Nei, 1987*) with the Poisson correction model and 1,000 bootstrap replicates (*Felsenstein, 1985*) implemented in MEGA X software (*Kumar et al., 2018*). This method was subsequently applied to generate a specific phylogenetic tree of TaSUMO protein sequences.

## RNA extraction and RT–PCR

To determine the transcription levels of the *TaSUMO4-7* genes, various parts of the plants were collected, frozen in liquid nitrogen, and ground. RNA was subsequently extracted with a commercial kit (RNeasy® Plant Mini Kit, Qiagen, Hilden, Germany). The extracted RNA was used to create first-strand cDNA copies (SuperScript II, Invitrogen, Waltham, MA, USA). RT−PCR was subsequently performed by a specific enzyme (KOD Dash DNA polymerase, Toyobo, Osaka, Japan) and a thermal cycler (Veriti 96-well, Applied Biosystems, Waltham, MA, USA). Gene-specific primers for the putative *TaSUMO* genes (Table S1) were used to amplify their transcripts, and wheat actin served as an internal control for normalization.

## Construction of expression vectors harboring *TaSUMOs*

We constructed expression vectors containing *TaSUMO* components using the *pUC119* vector (*Vieira & Messing, 1989*) and protocols from *Ikarashi et al. (2012)*. These vectors expressed fluorescent protein tags, either red (DsRFP) or green (GFP), under the regulation of the cauliflower mosaic virus 35S (CaMV35S) promoter (Fig. S1). All primers used are presented in Table S1. The new genes were introduced into pDsRFP *via* an In-Fusion HD Cloning Kit (Takara Bio) to create vectors named *pDsRFP:TaSUMO4-7*, which were used to study the localization and expression of the new TaSUMO proteins. Additionally, *pDsRFP:TaSUMO4-7ΔGG* vectors with deleted GG motifs were constructed to understand their role in protein localization. To study the coexpression and localization with TaSUMO4-7, the *Arabidopsis thaliana* SUMO-conjugation-enzyme (SCE1a) was inserted into the *pGFP:SCE1a* vector. Furthermore, cell organelle markers (peroxisome, mitochondria, plastid, cis-Golgi) were fused with GFP in vectors (*pPTS2:GFP, pmt:GFP, pWxTP:GFP, pGFP:SYP31*) and cotransfected with *DsRFP:TaSUMO4-7* in onion cells to determine whether TaSUMO4-7 proteins localize to specific organelles (Fig. S1).

## TaSUMOs, vectors transformation and protein expression visualization

The vectors containing the *TaSUMO4-7* genes were introduced into onion epidermal cells with a particle bombardment technique (Biolistic® PDS-1000/He, BioRad, Hercules, CA, USA). The bombarded cells were incubated in the dark for 24 h at room temperature. The expression of the TaSUMO4-7 proteins was then visualized by confocal laser scanning microscopy (FV300-BX61; Olympus, Tokyo, Japan) following protocols explored by *Kitajima et al. (2009)* and *Ikarashi et al. (2012)*.

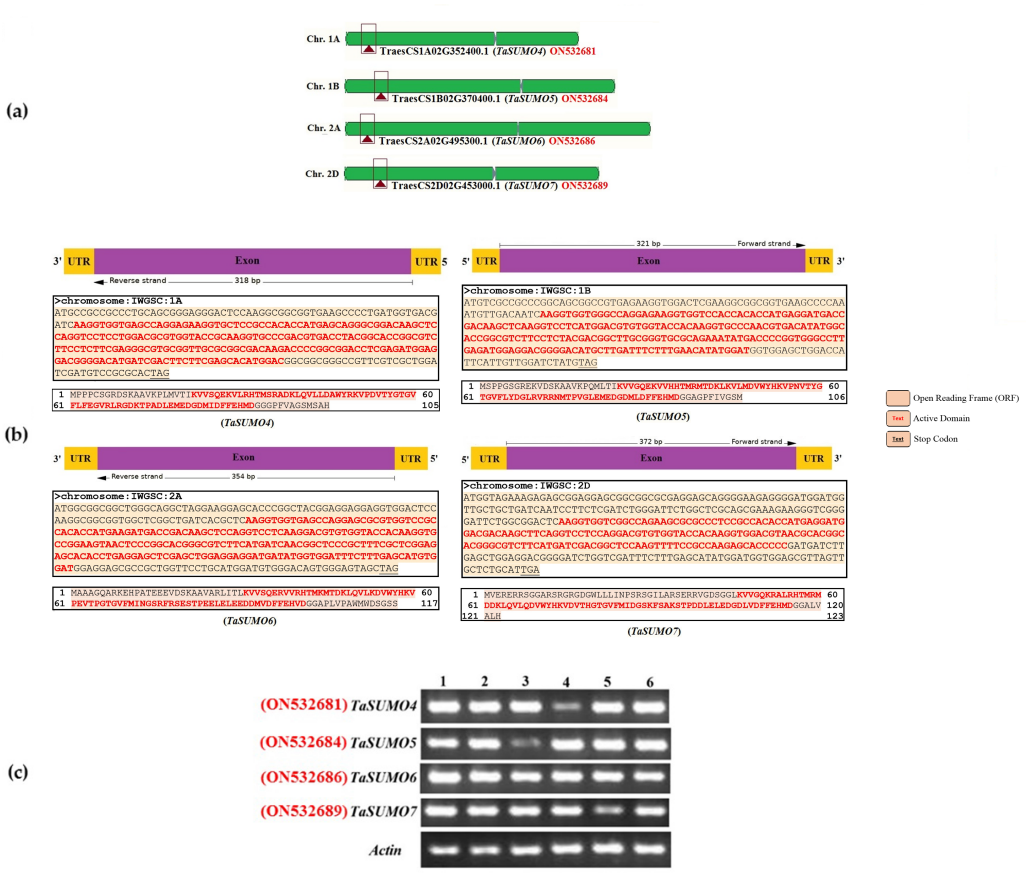



**Figure 1** **Characterization and expression analysis of *TaSUMO4-7* genes in wheat.** (A) Wheat chromosomes and locations of the *TaSUMO4, 5, 6, and 7* genes. (B) *TaSUMO4, 5, 6, and 7* gene structure, coding regions and proteins sequences. (C) Transcripts level of *TaSUMO4, 5, 6, and 7* genes. Each transcript was amplified by RT–PCR. 1: young leaves; 2: roots on young plants; 3: mature leaves; 4: flag leaves; 5: spikelets; 6: seeds. As a control, the transcript of the wheat Actin gene was analysed (bottom panel).
Full-size DOI: 10.7717/peerj.20432/fig-1

# RESULTS

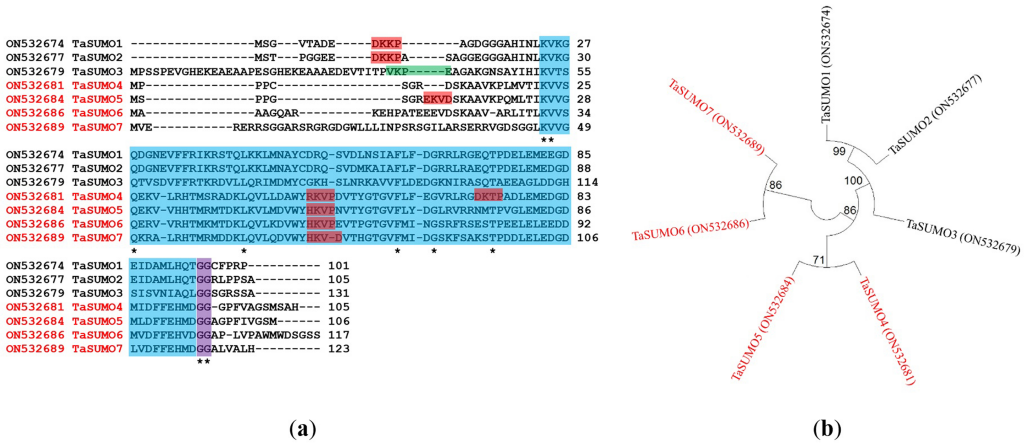
## New *TaSUMO* genes molecular characterization

The sequences of the novel *TaSUMO* genes were amplified using PCR with specifically designed primers (presented in Table S1). These primers were created based on the DNA sequences of potential *SUMO* regions identified within the wheat genome database (Fig. 1A). Sequencing results revealed that the 318, 321, 354, and 372 bp fragments of the novel *TaSUMO4-7* genes were located on the long arms of chromosomes 1A, 1B, 2A and 2D, respectively, encoding 105, 106, 117, and 123 amino acid SUMO proteins, respectively (Table 1, Fig. 1B).

RT-PCR was employed to assess the expression levels of the *TaSUMO4-7* genes across various wheat tissues, encompassing young leaves, roots from young plants, mature leaves, flag leaves, spikelets, and seeds. The new *TaSUMO*-genes were found to be expressed in all

**Table 1** Characterization of the putative proteins of the identified wheat *SUMO* genes.

| TaSUMO | GenBank accession | No. of AA | Gene ID | Chr. No. | No. Exon | CDS length (bp) | SUMOplotTM prediction of SUMOylation sites | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Motifs with high probability | | | Motifs with low probability | | |
| | | | | | | | Motif | Pos. | Score | Motif | Pos. | Score |
| TaSUMO1 | ON532674 | 101 | TraesCS3A02G424000.1 | 3A | 3 | 743 | - | - | - | DKKP | K10 | 0.39 |
| | ON532675 | | TraesCS3B02G460000.1 | 3B | | 677 | | | | | | |
| | ON532676 | | TraesCS3D02G419300.1 | 3D | | 851 | | | | | | |
| TaSUMO2 | ON532677 | 105 | TraesCS3B02G459900.1 | 3B | 3 | 767 | - | - | - | DKKP | K10 | 0.39 |
| | ON532678 | 106 | TraesCS3D02G419200.1 | 3D | | 709 | | | | | | |
| TaSUMO3 | ON532679 | 131 | TraesCS3A02G366700.1 | 3A | 1 | 396 | VKPE | K37 | 0.93 | - | - | - |
| | ON532680 | 123 | TraesCS3D02G359600.1 | 3D | | 478 | | | | | | |
| TaSUMO4 | ON532681 | 105 | TraesCS1A02G352400.1 | 1A | 1 | 318 | - | - | - | DKTP<br>RKVP | K72<br>K50 | 0.39<br>0.34 |
| | ON532682 | 104 | TraesCS1B02G369700.1 | 1B | | 315 | | | | HKVP | K49 | 0.41 |
| | ON532683 | 105 | TraesCS1D02G357900.1 | 1D | | 318 | | | | HKVP<br>DKTP | K50<br>K72 | 0.41<br>0.39 |
| TaSUMO5 | ON532684 | 106 | TraesCS1B02G370400.1 | 1B | 1 | 321 | - | - | - | EKVD<br>HKVP | K10<br>K53 | 0.50<br>0.41 |
| | ON532685 | 106 | TraesCS1D02G358000.1 | 1D | | 321 | | | | | | |
| TaSUMO6 | ON532686 | 117 | TraesCS2A02G495300.1 | 2A | 1 | 354 | - | - | - | HKVP | K59 | 0.41 |
| | ON532687 | 113 | TraesCS2B02G523500.1 | 2B | | 342 | | | | | | |
| | ON532688 | 112 | TraesCS2D02G495600.1 | 2D | | 471 | | | | | | |
| TaSUMO7 | ON532689 | 123 | TraesCS2D02G453000.1 | 2D | 1 | 372 | – | – | – | HKVD | K74 | 0.52 |




(a) (b)

**Figure 2 Sequence alignment and phylogenetic analysis of *TaSUMO* genes: identification of conserved Motifs, SUMOylation sites, and evolutionary relationships.** (A) *TaSUMO* genes alignment. Di-glycine motifs at the ends of the mature TaSUMO proteins are indicated by the purple box. Potential SUMOylation sites (Ψ KXE/D) are boxed in green for a high probability and with red for a low probability. The cyan box indicates the conserved ubiquitin-like domain. Identical amino acids in TaSUMO proteins are marked with an asterisk. (B) Phylogenetic tree analysis of the TaSUMO family.

Full-size DOI: 10.7717/peerj.20432/fig-2

of the wheat tested tissues (Fig. 1C), demonstrating that they are expressed in a constitutive manner.

Proteins with 105, 106, 117, and 123 amino acid residues were predicted for the unique *TaSUMO4-7* genes according to a BLAST investigation performed built on protein sequences in the wheat genome database. All the identified proteins possessed a C terminal di.glycine (GG) motif and shared consensus motifs (ΨKXE/D), indicative of a high resemblance degree of sequence with other known TaSUMO members (Fig. 2A). These findings revealed that *TaSUMO4-7* contains SUMO protein family features.

A phylogenetic tree was established utilizing SUMO protein sequences from diverse species, encompassing wheat (*Triticum aestivum*, Ta), *Arabidopsis thaliana* (At), rice (*Oryza sativa*, Os), *Brachypodium distachyon* (Bd), human (*Homo sapiens*, Hs), sorghum (*Sorghum bicolor*, Sb), foxtail millet (*Setaria italica*, Si), and baker's yeast (*Saccharomyces cerevisiae*, Sc). The phylogenetic tree exposed a high similarity degree between the four novel *TaSUMO* genes since they were clustered in the same tree, and the degree of similarity was greater between each of the TaSUMO4 and TaSUMO5 proteins and between the TaSUMO6 and TaSUMO7 proteins because they were found in the same sub-tree (Figs. 2B, 3A).

These findings were supported by the results obtained from Vector NTI® Software (Invitrogen, Waltham, MA, USA) for the percentage of amino acid identity between TaSUMO proteins; however, the percentages were 71% and 72% between TaSUMO4 and TaSUMO5 and between TaSUMO6 and TaSUMO7, respectively (Table 2). Monocotyledons including rice (*Oryza sativa*, Os) and sorghum (*Sorghum bicolor*, Sb) had similarity with TaSUMO4-7 proteins (Fig. 3A).

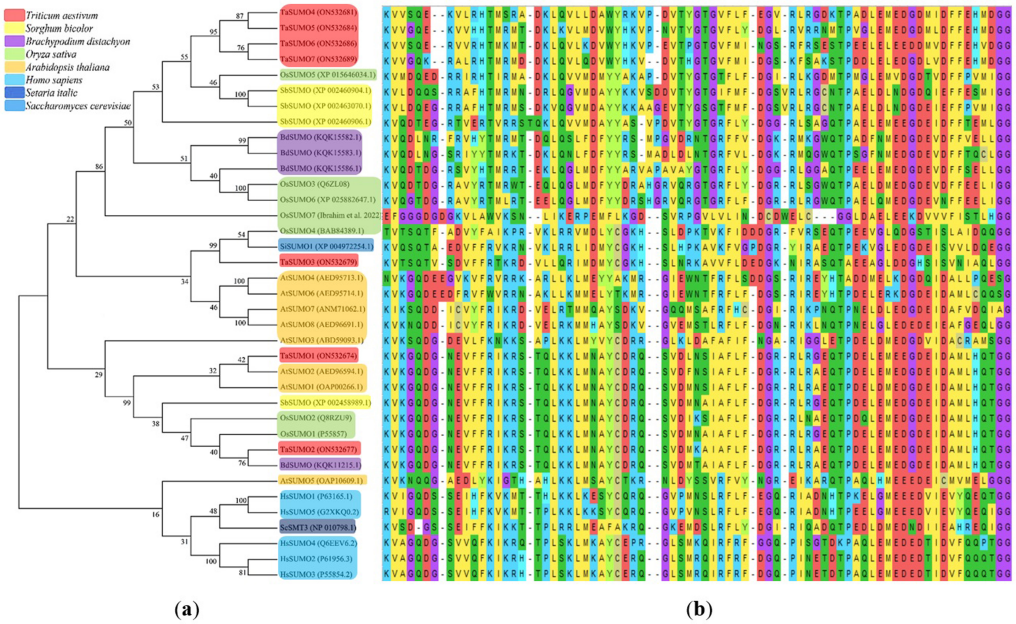



**Figure 3 Multiple sequence alignment results of SUMOs from ClustalW.** (A) Phylogenetic tree analysis of *SUMO* genes. (B) SUMOs alignment of the conserved ubiquitin-like domain.
Full-size DOI: 10.7717/peerj.20432/fig-3

**Table 2 Amino acid sequence identities (%) among wheat SUMO proteins were calculated by using the Vector NTI program.**

| TaSUMO | SUMO1 (*Wang et al., 2018*) | SUMO2 (*Wang et al., 2018*) | SUMO3 (*Wang et al., 2018*) | SUMO4 (*Ibrahim, 2022*) | SUMO5 (*Ibrahim, 2022*) | SUMO6 (*Ibrahim, 2022*) | SUMO7 (*Ibrahim, 2022*) |
|---|---|---|---|---|---|---|---|
| SUMO1 | 100 | | | | | | |
| SUMO2 | 84 | 100 | | | | | |
| SUMO3 | 36 | 42 | 100 | | | | |
| SUMO4 | 38 | 38 | 33 | 100 | | | |
| SUMO5 | 36 | 38 | 28 | 71 | 100 | | |
| SUMO6 | 35 | 34 | 24 | 63 | 67 | 100 | |
| SUMO7 | 31 | 34 | 30 | 64 | 65 | 72 | 100 |

TaSUMOs and SUMOs from diverse species have similar protein sequences (Fig. 2A) and conserved domains (Fig. 3B), indicating that *TaSUMO4-7* are evolutionarily conserved SUMO family members.

## TaSUMO4-7 *in silico* characterization

TaSUMO4-7 proteins were characterized using bioinformatic analysis and compared with other TaSUMO, AtSUMO and OsSUMO proteins. The *TaSUMO4* and *TaSUMO5* genes were found on the long arm of chromosome one; the *TaSUMO6* and *TaSUMO7* genes on the long arm of chromosome two; and the *TaSUMO1*, *TaSUMO2*, and *TaSUMO3* genes on chromosome three, according to the findings (Table 1). The results showed that the *TaSUMO4-7* genes have one exon, similar to the *TaSUMO3* gene, whereas the *TaSUMO1* and *TaSUMO2* genes have three exons (Table 1). Furthermore, the TaSUMO4 protein has

**Table 3** TaSUMOs biochemical properties determined using ProtParam.

| TaSUMO | GenBank accession | Formula | MW (kDa) | T. Pi | TNNCR | TNPCR | Ii | Stability | Ai | GRAVY |
|---|---|---|---|---|---|---|---|---|---|---|
| TaSUMO1 | ON532674 | $C_{479}H_{764}N_{142}O_{156}S_6$ | 11.20 | 5.11 | 18 | 14 | 45.37 | Unstable | 63.76 | −0.75 |
| TaSUMO2 | ON532677 | $C_{489}H_{788}N_{146}O_{160}S_6$ | 11.46 | 5.36 | 18 | 15 | 52.46 | Unstable | 61.43 | −0.77 |
| TaSUMO3 | ON532679 | $C_{599}H_{964}N_{176}O_{203}S_4$ | 14.01 | 5.33 | 21 | 15 | 52.61 | Unstable | 72.29 | −0.53 |
| TaSUMO4 | ON532681 | $C_{512}H_{809}N_{139}O_{151}S_8$ | 11.58 | 5.85 | 15 | 13 | 31.91 | Stable | 74.19 | −0.21 |
| TaSUMO5 | ON532684 | $C_{524}H_{832}N_{142}O_{150}S_{10}$ | 11.84 | 6.90 | 13 | 13 | 27.71 | Stable | 74.25 | −0.21 |
| TaSUMO6 | ON532686 | $C_{576}H_{903}N_{157}O_{180}S_6$ | 13.10 | 4.91 | 21 | 13 | 50.80 | Unstable | 73.25 | −0.40 |
| TaSUMO7 | ON532689 | $C_{591}H_{952}N_{184}O_{180}S_5$ | 13.68 | 6.98 | 19 | 19 | 47.52 | Unstable | 81.54 | −0.49 |

**Notes.**

MW, Molecular weight; T. Pi, Theoretical pI; TNNCR, Total number of negatively charged residues (Asp + Glu); TNPCR, Total number of positively charged residues (Arg + Lys); Ii, Instability index; Ai, Aliphatic index; GRAVY, Grand average of hydropathicity.

two motifs for SUMOylation sites at the K72 and K50 positions, while TaSUMO5 has two motifs at K10, similar to TaSUMO1, 2 and K53 positions; however, TaSUMO6, 7 have only one motif at the K59 and K74 positions, respectively (Table 1).

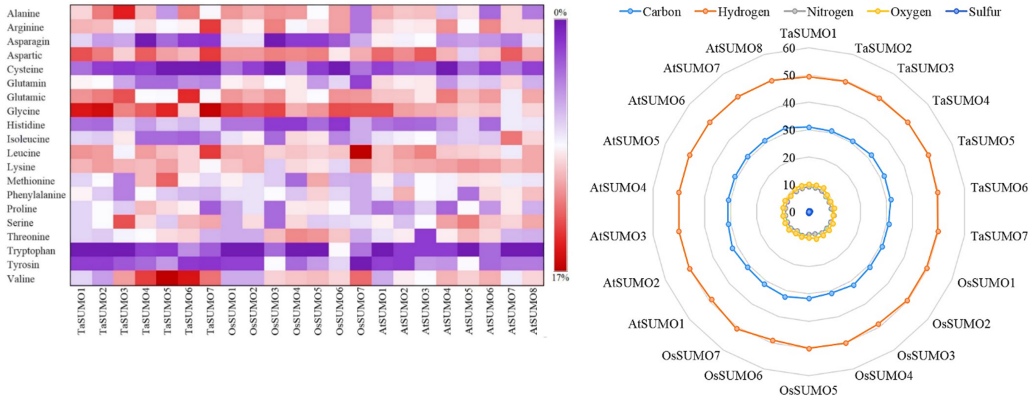
## Investigations of biochemical properties

Table 3 and Table S2 and Figs. 4 and 5 show the results of the biochemical features analysis of the SUMO proteins performed by ProtParam. The MWs of the TaSUMO4-7 proteins were determined to be 11.58, 11.84, 13.10, and 13.68 kDa. The MW of the TaSUMO6 protein was the same as the MW of the AtSUMO2 protein. The T-Pi values for the TaSUMO4-7 proteins were 5.85, 6.90, 4.91, and 6.98. The TaSUMO5 protein has a T-Pi value that is similar to that of TaSUMO7 and AtSUMO4.

Furthermore, the TaSUMO4-6 proteins had the same TNPCR (Arg + Lys) values (13) as OsSUMO1, OsSUMO2, AtSUMO1, and AtSUMO1. For TaSUMO5, seven proteins, TNNCR (Asp + Glu) and TNPCR (Arg + Lys), had equal values of 13 and 19, respectively. Moreover, the TaSUMO4 protein, like AtSUMO7, had a TNNCR (Asp + Glu) value (15), and the TaSUMO6 protein had a value (21) similar to that of TaSUMO3, OsSUMO3 and AtSUMO4 for the same parameter. Additionally, the TaSUMO7 protein had a value (19) identical to that of the OsSUMO6, AtSUMO2 and AtSUMO3 proteins (Table 3, Table S2).

The instability index (Ii) values for TaSUMO4 and five proteins were less than 40 (31.91, 27.71), suggesting that the proteins were stable; however, for TaSUMO6, seven proteins were more than 40 (50.80, 47.52), showing that the proteins were unstable, similar to the TaSUMO1-3 proteins. Moreover, the Ai values for the TaSUMO4-6 proteins were fairly comparable (74.19, 74.25, 73.25), whereas the TaSUMO7 protein was slightly different (81.54).

In terms of GRAVY, TaSUMO4 and TaSUMO5 proteins had equal values (−0.21), while TaSUMO6 and TaSUMO7 proteins had extremely close values (−0.40 and −0.49, respectively), as shown in Table 3 and Table S2. The amino acid composition analysis revealed that the TaSUMO5-7 proteins sequences had all 20 amino acids except cysteine (C), whereas the TaSUMO4 protein sequences contained all 20 amino acids except asparagin (N) (Fig. 4A, Table S3). Among the analysed SUMO proteins, TaSUMO4 had the highest percentage of the amino acid Pro (P, 6.7%), while the TaSUMO5 protein had the highest percentage of the amino acids Met (M, 9.4%) and Val (V, 13.2%). Additionally, TaSUMO6




(a) (b)

**Figure 4 Compositional profiling of TaSUMOs, OsSUMOs, and AtSUMOs: amino acid and atomic distribution patterns.** (A) Heatmap demonstrating the amino acid compositions of TaSUMOs, OsSUMOs and AtSUMOs according to percentage. A common amino acid found in all SUMOs is depicted. (B) Radar graph showing the atomic composition by percentage (number of each atom divided by the number of total atoms) of TaSUMOs, OsSUMOs and AtSUMOs.

Full-size DOI: 10.7717/peerj.20432/fig-4

had the largest amount of the amino acid Glu (E, 11.1%) and the lowest percentage of the amino acid Ile (I, 1.7%).

Additionally, the TaSUMO7 protein possessed high proportions of Arg (R, 10.6%) and Gly (G, 13%) amino acids and the lowest ratio of Pro (P, 1.6%) amino acid. Among the basic 20 amino acids, valine was the component with the largest percentage (10.2%, 13.2%, 12%) of the TaSUMO4-6 proteins, while it was a glycine (13%) in the TaSUMO7 protein (Fig. 4A, Table S4).

The amino acid's atomic composition (AC) includes (carbon, C), (hydrogen, H), (nitrogen, N), (oxygen, O), and (sulfur, S). OsSUMO6 had the highest TNA value (2,034) among the analysed SUMO proteins, while TaSUMO3 had the highest TNA value (1,946) among the TaSUMO proteins, and TaSUMO7 had the highest TNA value among the four novel wheat SUMO proteins (Table S3). Analysis of the atomic composition using a radar graph (Fig. 4B) revealed a highly conserved elemental profile across all tested SUMO proteins from wheat, rice, and Arabidopsis, with only minor visual variations.

## Secondary and tertiary structure prediction of TaSUMO4-7 proteins

The secondary structure was obtained using the PSIPRED method (*Jones, 1999*), which is based on sequence-based prediction. The predicted secondary structure of the protein contained three main states: α-helices, β-strands, and coils. These are visually represented by green helices, blue arrows, and faint lines, respectively. The confidence level in the prediction is also displayed, with red representing high confidence and blue demonstrating low confidence. The predicted secondary structure of all the novel TaSUMO proteins included six β-strands intervened by nine coils and three helixes, with the exception of TaSUMO7, which has four helixes (Fig. 6A). Using the Phyre2 engine, the three-dimensional model of TaSUMO5,6 were obtained based on the crystal structure of the

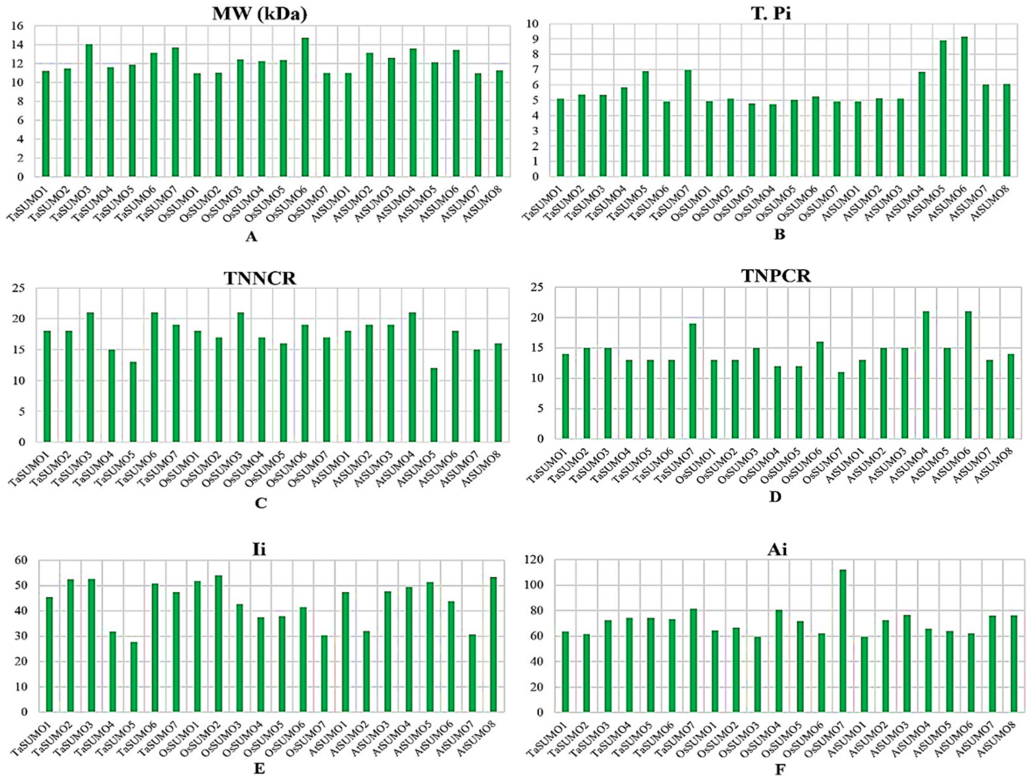



**Figure 5 Chart represents the biochemical properties of TaSUMOs, OsSUMOs, and AtSUMOs determined by ProtParam.** (A) Molecular weight; (B) Theoretical pI; (C) Total number of negatively charged residues (Asp + Glu); (D) Total number of positively charged residues (Arg + Lys); (E) Ii Instability index; (F) Aliphatic index.

Full-size DOI: 10.7717/peerj.20432/fig-5

human (*H. sapiens*) SUMO-1 domain (PDB code 1A5R) with a confidence of 99.9%, whereas the 3D structure of the TaSUMO4 protein was determined based on the crystal structure of the *Trypanosoma brucei* SUMO chain (PDB code 2K8H) with a confidence of 99.8%. However, for the TaSUOM7 protein, the 3D structural model was established by relying on the crystal structure of the *Mus musculus* SUMO chain (PDB code 3A4R) with a confidence of 99.7%. JSmol (JavaScript-Based Molecular Viewer From Jmol) (*Hanson et al., 2013*) was used to view 3D structural models of proteins (Fig. 6B).

## TaSUMO4-7 proteins cellular localization and expression

*DsRFP:TaSUMO4*, *DsRFP:TaSUMO5* and *DsRFP:TaSUMO6* are found mainly in the cytoplasm and nucleus of onion cells, while *DsRFP:TaSUMO7* is mostly localized in the nucleus (Fig. 7A). Additionally, the fusion proteins *DsRFP:TaSUMO4ΔGG, DsRFP:TaSUMO5ΔGG, DsRFP:TaSUMO6ΔGG,* and *DsRFP:TaSUMO7ΔGG* were observed in both the nucleus and cytoplasm, with significantly weaker signals in the nucleus (Fig. 7B). To further investigate the localization of these proteins, we used vectors expressing fluorescent markers (*pPTS2:GFP, pmt:GFP, pWxTP:GFP, pGFP:SYP31*) for different organelles (peroxisomes, mitochondria, plastids, and Golgi apparatuses). This allowed

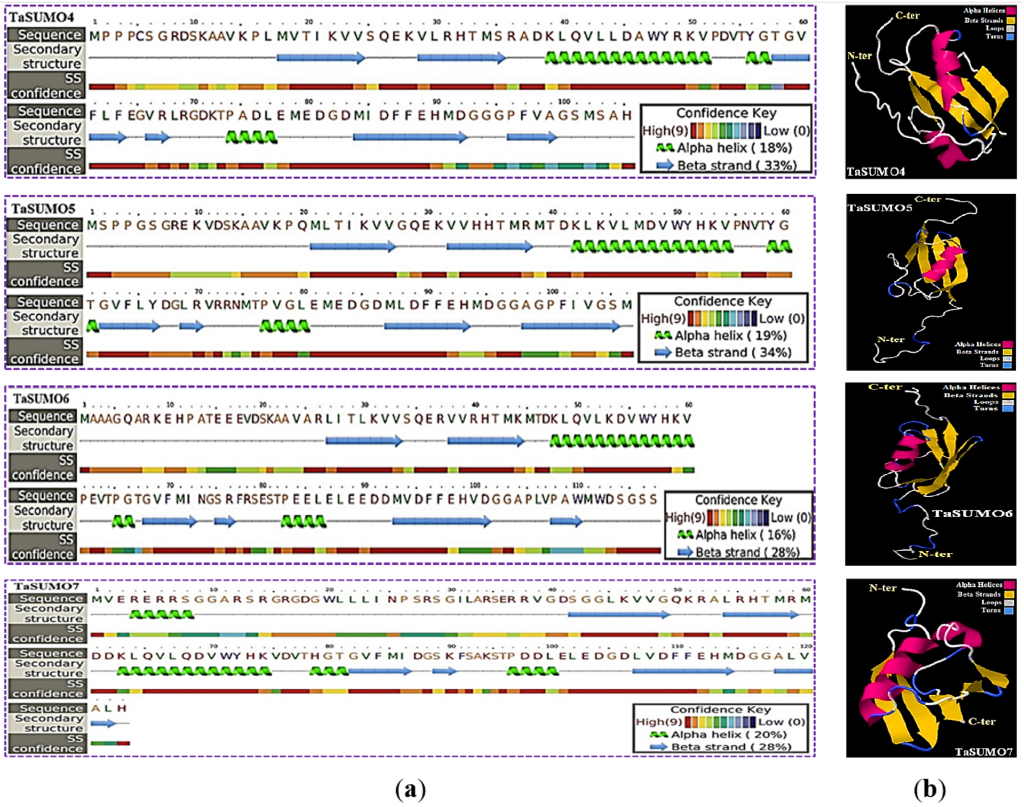



**Figure 6 Structural prediction and 3D modeling of TaSUMO4, TaSUMO5, TaSUMO6, and TaSUMO7 proteins using Phyre2.** (A) Prediction of the secondary structure of the putative TaSUMO4, 5, 6, and 7 proteins by the Phyre2 server. (B) The predicted 3D model of the TaSUMO4, 5, 6, and 7 proteins, designed by the Phyre2 server.

Full-size DOI: 10.7717/peerj.20432/fig-6

us to determine whether the SUMO proteins localize to any other cellular compartments. In onion epidermal cells, *pDsRFP:TaSUMO4, pDsRFP:TaSUMO5, pDsRFP:TaSUMO6* and *pDsRFPd:TaSUMO7* were transfected into cells. According to our findings, TaSUMO4-7 proteins were not found among any of the previously mentioned organelles under investigation (Fig. 8).

## *GFP:SCE1a* and TaSUMO4-7 co localization and co expression

The *GFP:SCE1a* (SUMO-conjugation-enzyme) and *DsRFP:TaSUMO4-7* fusion proteins colocalized and coexpressed in the nucleus, where they formed nuclear sub-domains. Unlike signals observed when expressed alone, signals were collected as dot-like structures inside the nucleus (Fig. 9A). In addition, high-magnification observation was performed to analyse the localization of the proteins in detail. High-magnification imaging (Fig. 9B) revealed colocalization of *DsRFP*-tagged TaSUMO4-7 proteins with the SUMO-conjugating enzyme *GFP:SCE1a* within the nucleus, suggesting that conjugation of SUMO proteins to their substrates likely occurs at these sites. However, it is important to note that the processing of the C-terminal di-glycine motif, essential for conjugation, may have already taken place in other cellular compartments prior to nuclear localization.

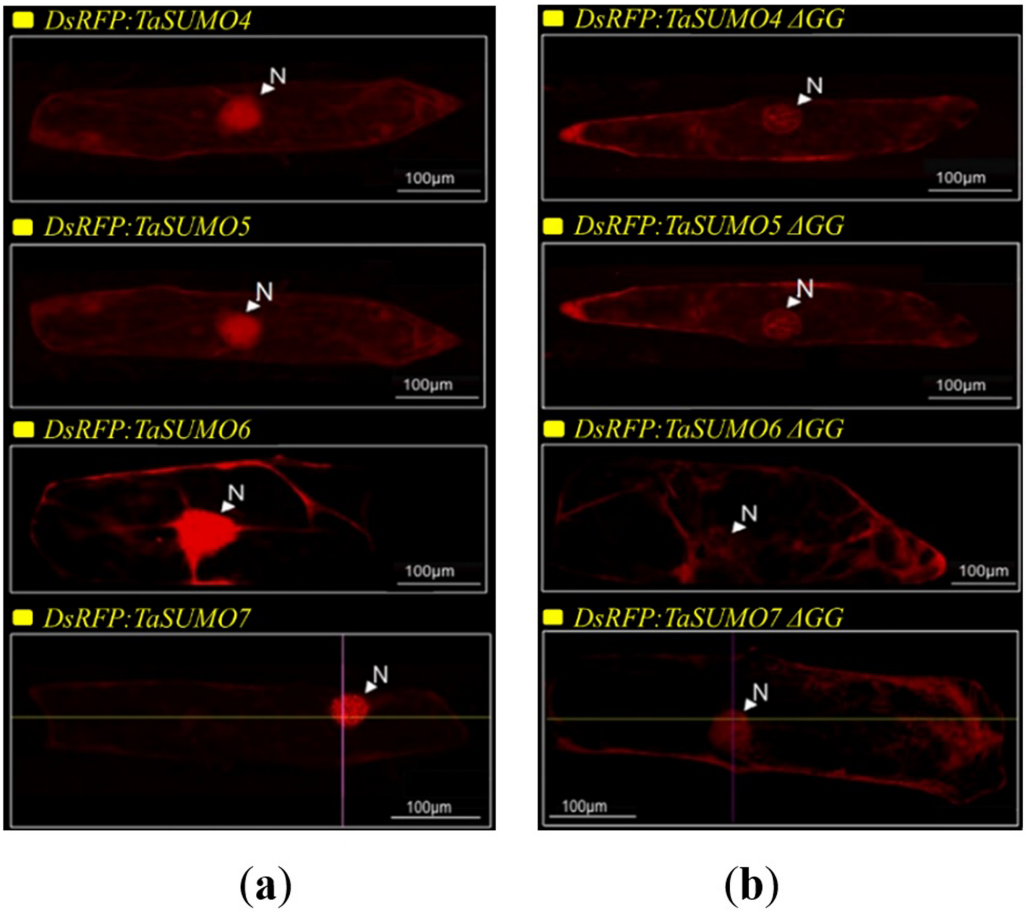



**Figure 7 Cellular localization of TaSUMO4-7 proteins in onion cells: role of the GG motif in subcellular targeting.** (A) Localization of *DsRFP:TaSUMO4,5,6,7* in onion cells. (B) Localization of *DsRFP:TaSUMO4,5,6,7* with a GG deletion in onion cells.

Full-size DOI: 10.7717/peerj.20432/fig-7

## DISCUSSION

This study identified and characterized four novel *SUMO* genes within the wheat genome, named *TaSUMO4*, *TaSUMO5*, *TaSUMO6*, and *TaSUMO7*, representing the first discovery of these genes in wheat. The sequencing analysis of the putative proteins of these genes demonstrated to have the SUMOylation features, including a C terminal diglycine (GG) motif and consensus motifs (ΨKXE/D), indicating a high degree of homology to other *TaSUMO* members (Fig. 2A). Many *SUMO*-genes have been described in several plant species, including rice (*Ikarashi et al., 2012*; *Rosa et al., 2018*; *Ibrahim et al., 2021*; *Teramura et al., 2021*; *Ibrahim, 2022*) and Arabidopsis (*Kurepa et al., 2003*). However, only one study of *SUMO*-genes in wheat, in which three genes were identified and characterized *in-silico*, has been found in the database (*Mayer et al., 2014*; *Wang et al., 2018*). Our findings defined four genes and were built on a previous report in the nomenclature and comparison.

The phylogenetic tree showed a high homology degree among the four novel *TaSUMO* genes, and there was a close similarity between *TaSUMO4* and *TaSUMO5*, and between

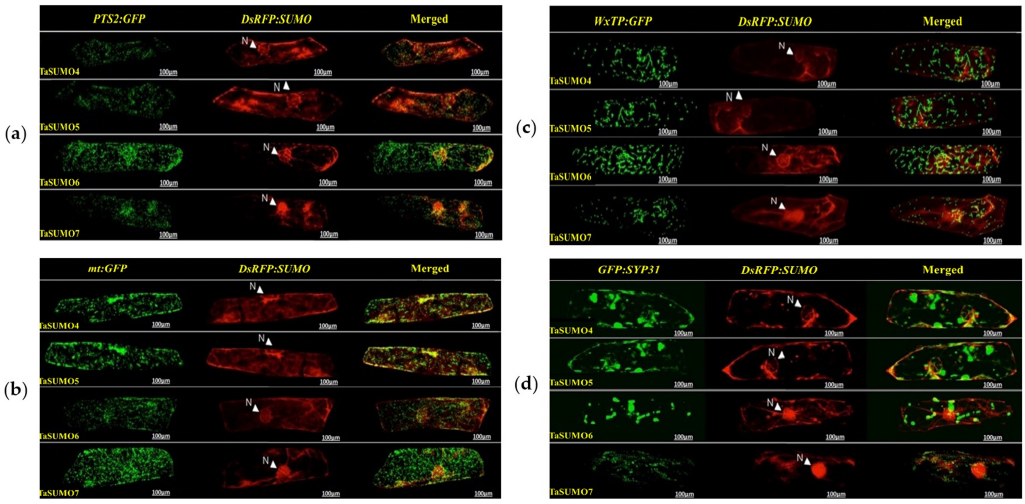



**Figure 8 Organelle-specific localization of TaSUMO4-7 proteins in onion cells: insights into peroxisomal, mitochondrial, plastidial, and Golgi targeting.** (A) Coexpression of *DsRFP:TaSUMO4,5,6,7* and *PTS2:GFP* in onion cells. PTS2 peroxisomal marker. (B) Coexpression of *DsRFP:TaSUMO4,5,6,7* and *mt:GFP* in onion cells. mt; mitochondrial marker. (C) Coexpression of *DsRFP:TaSUMO4,5,6,7* and *WxTP:GFP* in onion cells. WxTP, plastidial marker. (D) Coexpression of *DsRFP:TaSUMO4,5,6,7* and *SYP31:GFP* in onion cells. SYP31, cis-Golgi marker.

Full-size DOI: 10.7717/peerj.20432/fig-8

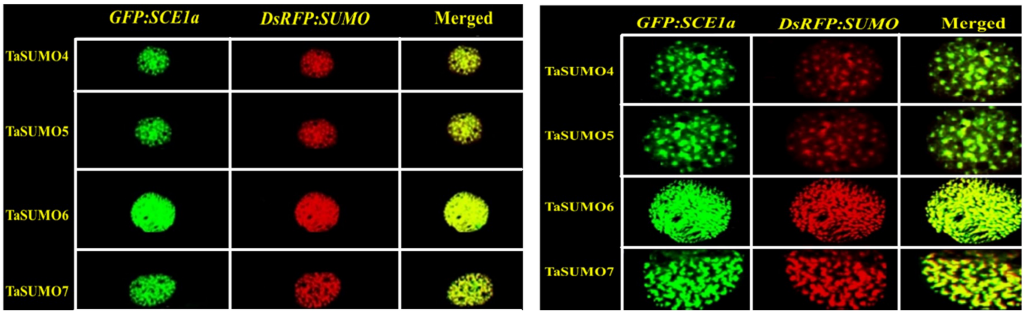



**Figure 9 Nuclear co-localization of DsRFP-Tagged TaSUMO4, TaSUMO5, TaSUMO6, and TaSUMO7 with GFP:SCE1a in onion epidermal cells.** (A) Co-expression and co-localization of *DsRFP:TaSUMO4,5,6,7* with *GFP:SCE1a* in the nucleus of onion cells. (B) Co-expression and co-localization of *DsRFP:TaSUMO4,5,6,7* with *GFP:SCE1a* in the nucleus of onion cells at high magnification.

Full-size DOI: 10.7717/peerj.20432/fig-9

*TaSUMO6* and *TaSUMO7* (Figs. 2B, 3A). Coincidentally, monocotyledons including (*Oryza sativa*, Os) and (*Sorghum bicolor*, Sc) showed high resemblance with TaSUMO4-7 proteins (Fig. 3A), which might share several important functions. The study revealed that the newly discovered TaSUMO4-7 proteins are evolutionarily well-preserved members of the SUMO family and exhibit close similarity to their counterparts in other species. Evolutionary tree analysis shown that the homologous genes in wheat, Arabidopsis, and rice are more closely related at the proximal level (*Wang et al., 2018*; *Ibrahim et al., 2021*). The current study revealed that the *TaSUMO4* and *TaSUMO5* genes were found on the long

arm of chromosome one, while the *TaSUMO6* and *TaSUMO7* genes were mapped on the long arm of chromosome two. However, the *TaSUMO1*, *TaSUMO2*, and *TaSUMO3* genes are located on chromosome three (Table 1) (*Wang et al., 2018*). These results suggested that the identified *TaSUMO*-genes might be distributed and clustered as tandem sets in the wheat genome.

Like findings in Arabidopsis (*Hammoudi et al., 2016*; *Joo et al., 2020*) and rice (*Wang et al., 2018*), this study identified tandem gene pairs among the newly discovered wheat SUMO genes (*TaSUMO4-7*). This finding suggested that tandem duplication, a well-established mechanism for generating genetic diversity and novel gene functions (*Lallemand et al., 2020*; *Ohno, 2013*), plays a similar role in the diversification of *SUMO* genes across diverse plant species. The study revealed that the new *TaSUMO4-7* genes have one exon, similar to the *TaSUMO3* gene, whereas the *TaSUMO1* and *TaSUMO2* genes have three exons (Table 1). Furthermore, the *TaSUMO4-7* proteins have different SUMOylation site motifs at different positions, which indicates that these genes are similar and have several functions. These results are consistent with the observations documented in previous studies by *Wang et al. (2018)*, *Miura & Hasegawa (2010)*, and *Ibrahim et al. (2021)*. The transcriptional levels of *TaSUMO4-7* were detected in diverse wheat tissues, including roots of young plants, young leaves, mature leaves, seeds, spikelets, and flag leaves (Fig. 1C), indicating that these genes are constitutively expressed. While our study suggested the ubiquitous expression of *TaSUMO4-7*, previous research by *Choulet et al. (2014)* and *Wang et al. (2018)* reported tissue- and time-specific expression patterns for *TaSUMO1-3* genes, highlighting potential functional diversity within the wheat SUMO family. In contrast to our findings, *Teramura et al. (2021)* reported tissue-specific expression of *OsSUMO4/6* genes in rice, suggesting functional diversification among individual SUMO family members.

This finding aligns with the essential roles of SUMOylation in various cellular processes like hormone signalling, stress response, and development (*Rodriguez, Dargemont & Hay, 2001*; *Srivastava, Zhang & Sadanandom, 2016*; *Augustine & Vierstra, 2018*). Therefore, understanding the expression patterns and cellular localization of *SUMO* transcripts is crucial for deciphering the functional mechanisms of plant SUMOylation. This study explored the subcellular localization of newly discovered TaSUMO proteins in onion cells. *DsRFP-tagged TaSUMO4, 5,* and *6* proteins were found in both the nucleus and cytoplasm, while *TaSUMO7* localized primarily to the nucleus in subdomain structures (Fig. 7A), consistent with previous reports (*Lois, Lima & Chua, 2003*). These outcomes suggest that the SUMO-activating system processes, activates, and conjugates these proteins within the cell. Removal of the GG motif reduced the nuclear and cytoplasmic accretion of *DsRFP-tagged TaSUMO4-7 ΔGG* proteins and resulted in weakened nuclear signals (Fig. 7B), highlighting the crucial role of the GG motif in SUMO protein processing, activation, and accumulation. Similar findings have been stated elsewhere (*Murtas et al., 2003*; *Ikarashi et al., 2012*; *Ibrahim et al., 2021*). Additionally, our study revealed no localization of the TaSUMO4-7 proteins to other cell organelles (Fig. 8). Previous research has shown that rice OsSUMO1/2/3 proteins are solely nuclear in rice root cells, while OsSUMO4 is localized to specific nuclear compartments compared to GFP alone (*Joo et al., 2020*). These

results underscore the significance of SUMO proteins in mediating protein interactions, localization, and function through the SUMOylation process.

The SUMO modification system includes SUMO modified substrates, activation enzymes (E1), binding enzymes (E2), ligases (E3), and proteases of SUMO modifiers (*Miura et al., 2007b*; *Park et al., 2011*). The SUMO-conjugation enzyme (*GFP:SCE1a*) and *DsRFP:TaSUMO4-7* fusion proteins were colocalized and coexpressed in the nucleus, where they formed nuclear sub-domains (Fig. 9A). These results suggest that *DsRFP:TaSUMO4-7* are processed, activated, conjugated with *GFP:SCE1a*, and transported to localize in nuclear foci *via* the SUMOylation process in onion cells, where it may bind to protein substrates. The wheat genome, like the rice and Arabidopsis genomes, encodes the activation enzymes SAE1a, SAE2, and SAE1B (*Wang et al., 2018*). Wheat and rice harbor three gene conjugation enzymes, namely, SCE1, SCE2 and SCE3, in addition to the four gene ligation enzymes SIZ1, SIZ2, SIZ3, and SIZ4 (*Kurepa et al., 2003*; *Jin et al., 2008*; *Castaño-Miquel, Seguí & Lois, 2011*; *Cai et al., 2017*; *Wang et al., 2018*). Although the conjugation enzyme-encoding genes SCE1, SCE2, and SCE3 have been extensively studied (*Chaikam & Karlson, 2010*; *Joo et al., 2019*), they have not yet been well characterized in wheat.

This study pioneered the investigation of the functional localization of TaSUMO4-7 using the *SCE1:GFP* fusion proteins, revealing its presence in distinct nuclear compartments. The SCE1 gene has been stated to show a powerful role in promoting SUMO conjugation to target proteins through the SUMOylation system (*Nurdiani, Widyajayantie & Nugroho, 2018*; *Ibrahim et al., 2021*). Increased levels of SUMO-conjugated-proteins (SUMO-conjugates) are hypothesized to play a role in how plants respond to challenges from their environment (abiotic stress) and other organisms (biotic stress), as evidenced by previous research (*Castro et al., 2016*; *Joo et al., 2019*). Maturation, activation, binding, connection, and function are all part of the multi-gene biological process produced by SUMO-isoform modification (*Park et al., 2011*; *Wang et al., 2018*). Plants and humans contain many SUMO isoforms, unlike yeast (*Gareau & Lima, 2010*; *Novatchkova et al., 2012*). In humans, variations in SUMO-isoform expression levels are thought to contribute to their functional diversity (*Wang, Peterson & Loring, 2014*). While many plant genomes, including wheat, exhibit SUMO gene diversification, the specific functions of these genes in wheat have largely not been explored.

While this study successfully identified four novel *TaSUMO* genes, some limitations should be noted. First, the functional characterization was limited to *in vitro* localization studies, and future work should include in planta validation through transgenic approaches. Second, the expression profiles were examined under normal growth conditions; abiotic/biotic stress treatments could reveal additional regulatory roles. Finally, the interaction partners of these novel TaSUMOs remain to be identified through proteomic approaches. Future studies should focus on (1) generating knockout mutants to determine phenotypic consequences, (2) examining SUMOylation dynamics under stress conditions, and (3) identifying substrate proteins through mass spectrometry-based approaches.

## CONCLUSIONS

This study identified and characterized four novel *SUMO* genes (*TaSUMO4-7*) in the wheat genome, confirming their SUMOylation features. All four genes exhibited constitutive expression across wheat tissues. DsRFP-tagged TaSUMO4, 5, and 6 proteins were detected in both the nucleus and cytoplasm, while TaSUMO7 was exclusively expressed in the nucleus. The presence of essential GG motifs in these genes suggested their involvement in SUMO protein processing, activation, and accumulation. Furthermore, the study demonstrated that TaSUMO4-7 proteins undergo processing, activation, conjugation with the *SCE1a* gene, and localization in nuclear foci through the plant cell's SUMOylation system, indicating their successful processing, activation, and nuclear transport within the plant cell. Although limited to *in vitro* characterization, this study provides the foundation for future investigations into the stress-responsive roles and molecular mechanisms of these newly identified *TaSUMO* genes in wheat.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This work was supported by the Ongoing Research Funding Program, (ORF-2025-954), King Saud University, Riyadh, Saudi Arabia. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:
The Ongoing Research Funding Program, King Saud University, Riyadh, Saudi Arabia: ORF-2025-954.

### Competing Interests

The authors declare there are no competing interests.

### Author Contributions

- Eid I. Ibrahim conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Kotb A. Attia conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Abdelhalim Ghazy conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Itoh Kimiko performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Abdullah Al-Doss conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.

## Data Availability

The following information was supplied regarding data availability:

The SUMO genomic/cDNA/protein sequences are available at NCBI: ON640912, ON640913, ON640914, ON640915, ON640916, ON640917, ON640918, ON640919, ON640920, ON640921, ON640922, ON640923, ON640924, ON640925, ON640926, ON640927, ON532674, ON532675, ON532676, ON532677, ON532678, ON532679, ON532680, ON532681, ON532682, ON532683, ON532684, ON532685, ON532686, ON532687, ON532688, ON532689, WAK97207, WAK97208, WAK97209, WAK97210, WAK97211, WAK97212, WAK97213, WAK97214, WAK97215, WAK97216, WAK97217, WAK97218, WAK97219, WAK97220, WAK97221, WAK97222.

## Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.20432#supplemental-information.

# REFERENCES

**Augustine RC, Vierstra RD. 2018.** SUMOylation: re-wiring the plant nucleus during stress and development. *Current Opinion in Plant Biology* **45**:143–154 DOI 10.1016/j.pbi.2018.06.006.

**Augustine RC, York SL, Rytz TC, Vierstra RD. 2016.** Defining the SUMO system in maize: SUMOylation is up-regulated during endosperm development and rapidly induced by stress. *Plant Physiology* **171**:2191–2210 DOI 10.1104/pp.16.00353.

**Bernier-Villamor V, Sampson DA, Matunis MJ, Lima CD. 2002.** Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* **108**:345–356 DOI 10.1016/S0092-8674(02)00630-X.

**Cai B, Kong X, Zhong C, Sun S, Zhou XF, Jin YH, Wang Y, Li X, Zhu Z, Jin JB. 2017.** SUMO E3 Ligases GmSIZ1a and GmSIZ1b regulate vegetative growth in soybean. *Journal of Integrative Plant Biology* **59**:2–14 DOI 10.1111/jipb.12504.

**Castaño-Miquel L, Seguí J, Lois LM. 2011.** Distinctive properties of Arabidopsis SUMO paralogues support the *in vivo* predominant role of AtSUMO1/2 isoforms. *Biochemical Journal* **436**:581–590 DOI 10.1042/BJ20101446.

**Castro PH, Couto D, Freitas S, Verde N, Macho AP, Huguet S, Botella MA, Ruiz-Albert J, Tavares RM, Bejarano ER. 2016.** SUMO proteases ULP1c and ULP1d are required for development and osmotic stress responses in *Arabidopsis thaliana*. *Plant Molecular Biology* **92**:143–159 DOI 10.1007/s11103-016-0500-9.

**Castro PH, Santos MÂ, Freitas S, Cana-Quijada P, Lourenço T, Rodrigues MA, Fonseca F, Ruiz-Albert J, Azevedo JE, Tavares RM. 2018.** *Arabidopsis thaliana* SPF1 and SPF2 are nuclear-located ULP2-like SUMO proteases that act downstream of SIZ1 in plant development. *Journal of Experimental Botany* **69**:4633–4649 DOI 10.1093/jxb/ery265.

**Chaikam V, Karlson DT. 2010.** Response and transcriptional regulation of rice SUMOylation system during development and stress conditions. *BMB Reports* **43**:103–109 DOI 10.5483/BMBRep.2010.43.2.103.

**Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J, Pingault L, Sourdille P, Couloux A, Paux E, Leroy P, Mangenot S, Guilhot N, Gouis JL, Balfourier F, Alaux M, Jamilloux V, Poulain J, Durand C, Bellec A, Gaspin C, Safar J, Dolezel J, Rogers J, Vandepoele K, Aury J-M, Mayer K, Berges H, Quesneville H, Wincker P, Feuillet C. 2014.** Structural and functional partitioning of bread wheat chromosome 3B. *Science* **345**:1249721 DOI 10.1126/science.1249721.

**Clague MJ, Coulson JM, Urbé S. 2012.** Cellular functions of the DUBs. *Journal of Cell Science* **125**:277–286 DOI 10.1242/jcs.090985.

**Felsenstein J. 1985.** Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**:783–791 DOI 10.1111/j.1558-5646.1985.tb00420.x.

**Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz H-R, Ceric G, Forslund K, Eddy SR, Sonnhammer EL. 2007.** The Pfam protein families database. *Nucleic Acids Research* **36**:D281–D288 DOI 10.1093/nar/gkm960.

**Flick K, Kaiser P. 2009.** Proteomic revelation: SUMO changes partners when the heat is on. *Science Signaling* **2**:pe45-pe45 DOI 10.1126/scisignal.281pe45.

**Flotho A, Melchior F. 2013.** Sumoylation: a regulatory protein modification in health and disease. *Annual Review of Biochemistry* **82**:357–385 DOI 10.1146/annurev-biochem-061909-093311.

**Gareau JR, Lima CD. 2010.** The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nature Reviews Molecular Cell Biology* **11**:861–871 DOI 10.1038/nrm3011.

**Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. 2003.** ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* **31**:3784–3788 DOI 10.1093/nar/gkg563.

**Guruprasad K, Reddy BB, Pandit MW. 1990.** Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. *PEDS* **4**:155–161 DOI 10.1093/protein/4.2.155.

**Hammoudi V, Vlachakis G, Schranz ME, Van den Burg HA. 2016.** Whole-genome duplications followed by tandem duplications drive diversification of the protein modifier SUMO in angiosperms. *New Phytologist* **211**:172–185 DOI 10.1111/nph.13911.

**Han D, Yu Z, Lai J, Yang C. 2022.** Post-translational modification: a strategic response to high temperature in plants. *ABiotech* **3**:49–64 DOI 10.1007/s42994-021-00067-w.

**Hanania U, Furman-Matarasso N, Ron M, Avni A. 1999.** Isolation of a novel SUMO protein from tomato that suppresses EIX-induced cell death. *TPJ* **19**:533–541 DOI 10.1046/j.1365-313X.1999.00547.x.

**Hanson RM, Prilusky J, Renjian Z, Nakane T, Sussman JL. 2013.** JSmol and the next-generation web-based representation of 3D molecular structure as applied to proteopedia. *Israel Journal of Chemistry* **53**:207–216 DOI 10.1002/ijch.201300024.

**Hay RT. 2005.** SUMO: a history of modification. *Molecular Cell* **18**:1–12 DOI 10.1016/j.molcel.2005.03.012.

**Ibrahim EI. 2022.** *Structural and functional characterization of SUMO genes in main food crops*. Riyadh: King Saud University, 138.

**Ibrahim EI, Attia KA, Ghazy AI, Itoh K, Almajhdi FN, Al-Doss AA. 2021.** Molecular characterization and functional localization of a novel SUMOylation gene in *Oryza sativa*. *Biology* **11**:53 DOI 10.3390/biology11010053.

**Ikai A. 1980.** Thermostability and aliphatic index of globular proteins. *JB* **88**:1895–1898 DOI 10.1093/oxfordjournals.jbchem.a133168.

**Ikarashi Y, Noguchi N, Attia K, Kitajima-Koga A, Mitsui T, Itoh K. 2012.** Cloning, expression, and intracellular localization of rice SUMO genes. *Bulletin of the Faculty of Agriculture, Niigata University* **65**:77–83.

**Isono E, Nagel M-K. 2014.** Deubiquitylating enzymes and their emerging role in plant biology. *Frontiers in Plant Science* **5**:56 DOI 10.3389/fpls.2014.00056.

**Jin JB, Jin YH, Lee J, Miura K, Yoo CY, Kim WY, Van Oosten M, Hyun Y, Somers DE, Lee I. 2008.** The SUMO E3 ligase, AtSIZ1, regulates flowering by controlling a salicylic acid-mediated floral promotion pathway and through affects on FLC chromatin structure. *TPJ* **53**:530–540 DOI 10.1111/j.1365-313X.2007.03359.x.

**Jones DT. 1999.** Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology* **292**:195–202 DOI 10.1006/jmbi.1999.3091.

**Joo J, Choi DH, Kim SH, Song SI. 2020.** Cellular localization of rice SUMO/SUMO conjugates and *in vitro* sumoylation using rice components. *Rice Science* **27** DOI 10.1016/j.rsci.2019.02.003.

**Joo J, Choi DH, Lee YH, Seo HS, Song SI. 2019.** The rice SUMO conjugating enzymes OsSCE1 and OsSCE3 have opposing effects on drought stress. *Journal of Plant Physiology/TD* **240**:152993 DOI 10.1016/j.jplph.2019.152993.

**Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. 2015.** The Phyre2 web portal for protein modelling, prediction and analysis. *Nature Protocols* **10**:845–858 DOI 10.1038/nprot.2015.053.

**Kerscher O, Felberbaum R, Hochstrasser M. 2006.** Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annual Review of Cell and Developmental Biology* **22**:159–180 DOI 10.1146/annurev.cellbio.22.010605.093503.

**Kim KI, Baek SH, Chung CH. 2002.** Versatile protein tag, SUMO: its enzymology and biological function. *Journal of Cellular Physiology* **191**:257–268 DOI 10.1002/jcp.10100.

**Kitajima A, Asatsuma S, Okada H, Hamada Y, Kaneko K, Nanjo Y, Kawagoe Y, Toyooka K, Matsuoka K, Takeuchi M. 2009.** The rice α-amylase glycoprotein is targeted from the Golgi apparatus through the secretory pathway to the plastids. *The Plant Cell* **21**:2844–2858 DOI 10.1105/tpc.109.068288.

**Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018.** MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**:1547 DOI 10.1093/molbev/msy096.

**Kurepa J, Walker JM, Smalle J, Gosink MM, Davis SJ, Durham TL, Sung D-Y, Vierstra RD. 2003.** The small ubiquitin-like modifier (SUMO) protein modification system in Arabidopsis: accumulation of SUMO1 and-2 conjugates is increased by stress. *Journal of Biological Chemistry* **278**:6862–6872 DOI 10.1074/jbc.M209694200.

**Kyte J, Doolittle RF. 1982.** A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157**:105–132 DOI 10.1016/0022-2836(82)90515-0.

**Lallemand T, Leduc M, Landès C, Rizzon C, Lerat E. 2020.** An overview of duplicated gene detection methods: why the duplication mechanism has to be accounted for in their choice. *Gene* **11**:1046 DOI 10.3390/genes11091046.

**Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R. 2007.** Clustal W and Clustal X version 2.0. *Bioinformatics* **23**:2947–2948 DOI 10.1093/bioinformatics/btm404.

**Le Roux ML. 2016.** *Engineering wheat (Triticum aestivum L.) for abiotic resilience by manipulating small ubiquitin-like modifiers.* Stellenbosch: Stellenbosch University.

**Le Roux ML, Kunert KJ, Van der Vyver C, Cullis CA, Botha A-M. 2019.** Expression of a small ubiquitin-like modifier protease increases drought tolerance in wheat (*Triticum aestivum* L.). *Frontiers in Plant Science* **10**:266 DOI 10.3389/fpls.2019.00266.

**Lois LM, Lima CD. 2005.** Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. *The EMBO Journal* **24**:439–451 DOI 10.1038/sj.emboj.7600552.

**Lois LM, Lima CD, Chua N-H. 2003.** Small ubiquitin-like modifier modulates abscisic acid signalling in Arabidopsis. *The Plant Cell* **15**:1347–1359 DOI 10.1105/tpc.009902.

**Mayer KFX, Rogers J, Doležel J, Pozniak C, Eversole K, Feuillet C, Gill B, Friebe B, Lukaszewski AJ, Sourdille P, Endo TR, Kubaláková M, Číhalíková J, Dubská Z, Vrána J, Šperková R, Šimková H, Febrer M, Clissold L, McLay K, Singh K, Chhuneja P, Singh NK, Khurana J, Akhunov E, Choulet F, Alberti A, Barbe V, Wincker P, Kanamori H, Kobayashi F, Itoh T, Matsumoto T, Sakai H, Tanaka T, Wu J, Ogihara Y, Handa H, Maclachlan PR, Sharpe A, Klassen D, Edwards D, Batley J, Olsen O-A, Sandve SR, Lien S, Steuernagel B, Wulff B, Caccamo M, Ayling S, Ramirez-Gonzalez RH, Clavijo BJ, Wright J, Pfeifer M, Spannagl M, Martis MM, Mascher M, Chapman J, Poland JA, Scholz U, Barry K, Waugh R, Rokhsar DS, Muehlbauer GJ, Stein N, Gundlach H, Zytnicki M, Jamilloux V, Quesneville H, Wicker T, Faccioli P, Colaiacovo M, Stanca AM, Budak H, Cattivelli L, Glover N, Pingault L, Paux E, Sharma S, Appels R, Bellgard M, Chapman B, Nussbaumer T, Bader KC, Rimbert H, Wang S, Knox R, Kilian A, Alaux M, Alfama F, Couderc L, Guilhot N, Viseux C, Loaec M, Keller B, Praud S. 2014.** A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* **345**:1251788 DOI 10.1126/science.1251788.

**Miura K, Hasegawa PM. 2010.** Sumoylation and other ubiquitin-like posttranslational modifications in plants. *Trends in Cell Biology* **20**:223–232 DOI 10.1016/j.tcb.2010.01.007.

**Miura K, Jin JB, Hasegawa PM. 2007a.** Sumoylation, a posttranslational regulatory process in plants. *Current Opinion in Plant Biology* **10**:495–502 DOI 10.1016/j.pbi.2007.07.002.

**Miura K, Jin JB, Lee J, Yoo CY, Stirm V, Miura T, Ashworth EN, Bressan RA, Yun D-J, Hasegawa PM. 2007b.** SIZ1-mediated sumoylation of ICE1 controls CBF3/DREB1A

expression and freezing tolerance in Arabidopsis. *The Plant Cell* **19**:1403–1414 DOI 10.1105/tpc.106.048397.

**Morrell R, Sadanandom A. 2019.** Dealing with stress: a review of plant SUMO proteases. *Frontiers in Plant Science* **10**:1122 DOI 10.3389/fpls.2019.01122.

**Murtas G, Reeves PH, Fu Y-F, Bancroft I, Dean C, Coupland G. 2003.** A nuclear protease required for flowering-time regulation in Arabidopsis reduces the abundance of SMALL UBIQUITIN-RELATED MODIFIER conjugates. *The Plant Cell* **15**:2308–2319 DOI 10.1105/tpc.015487.

**Nayak A, Müller S. 2014.** SUMO-specific proteases/isopeptidases: SENPs and beyond. *Genome Biology* **15**:1–7 DOI 10.1186/gb-2014-15-1-r1.

**Novatchkova M, Budhiraja R, Coupland G, Eisenhaber F, Bachmair A. 2004.** SUMO conjugation in plants. *Planta* **220**:1–8 DOI 10.1007/s00425-004-1370-y.

**Novatchkova M, Tomanov K, Hofmann K, Stuible HP, Bachmair A. 2012.** Update on sumoylation: defining core components of the plant SUMO conjugation system by phylogenetic comparison. *New Phytologist* **195**:23–31 DOI 10.1111/j.1469-8137.2012.04135.x.

**Nurdiani D, Widyajayantie D, Nugroho S. 2018.** OsSCE1 encoding SUMO E2-conjugating enzyme involves in drought stress response of *Oryza sativa*. *Rice Science* **25**:73–81 DOI 10.1016/j.rsci.2017.11.002.

**Ohno S. 2013.** *Evolution by gene duplication*. Berlin, Heidelberg: Springer Science & Business Media.

**Park HJ, Kim W-Y, Park HC, Lee SY, Bohnert HJ, Yun D-J. 2011.** SUMO and SUMOylation in plants. *Molecular Cell* **32**:305–316 DOI 10.1007/s10059-011-0122-7.

**Rodriguez MS, Dargemont C, Hay RT. 2001.** SUMO-1 conjugation *in vivo* requires both a consensus modification motif and nuclear targeting. *Journal of Biological Chemistry* **276**:12654–12659 DOI 10.1074/jbc.M009476200.

**Rosa MT, Almeida DM, Pires IS, Da Rosa Farias D, Martins AG, Da Maia LC, De Oliveira AC, Saibo NJ, Oliveira MM, Abreu IA. 2018.** Insights into the transcriptional and posttranscriptional regulation of the rice SUMOylation machinery and into the role of two rice SUMO proteases. *BMC Plant Biology* **18**:1–18 DOI 10.1186/s12870-017-1213-1.

**Saitou N, Nei M. 1987.** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**:406–425 DOI 10.1093/oxfordjournals.molbev.a040454.

**Srivastava AK, Zhang C, Sadanandom A. 2016.** Rice OVERLY TOLERANT TO SALT 1 (OTS1) SUMO protease is a positive regulator of seed germination and root development. *Plant Signalling & Behavior* **11**:e1173301 DOI 10.1080/15592324.2016.1173301.

**Teramura H, Yamada K, Ito K, Kasahara K, Kikuchi T, Kioka N, Fukuda M, Kusano H, Tanaka K, Shimada H. 2021.** Characterization of novel SUMO family genes in the rice genome. *Genes & Genetic Systems* **96(1)**:25–32 DOI 10.1266/ggs.20-00034.

**Van den Burg HA, Kini RK, Schuurink RC, Takken FL. 2010.** Arabidopsis small ubiquitin-like modifier paralogs have distinct functions in development and defense. *The Plant Cell* **22**:1998–2016 DOI 10.1105/tpc.109.070961.

**Vieira J, Messing J. 1989.** Production of single-stranded plasmid DNA. In: *Recombinant DNA methodology*. Netherlands: Elsevier, 225–233.

**Wang Y-C, Peterson SE, Loring JF. 2014.** Protein posttranslational modifications and regulation of pluripotency in human stem cells. *Cell Research* **24**:143–160 DOI 10.1038/cr.2013.151.

**Wang H, Tan J, Yao J, Ma H. 2018.** Characterization of the gene families of SUMOylation in wheat. *Jiangsu Agricultural Journal* **1**:1–10 DOI 10.3969/j.issn.1000-4440.2018.01.001.

**Yaeno T, Iba K. 2008.** BAH1/NLA, a RING-type ubiquitin E3 ligase, regulates the accumulation of salicylic acid and immune responses to *Pseudomonas Syringae* DC3000. *Plant Physiology* **148**:1032–1041 DOI 10.1104/pp.108.124529.

**Yunus AA, Lima CD. 2006.** Lysine activation and functional analysis of E2-mediated conjugation in the SUMO pathway. *Nature Structural & Molecular Biology* **13**:491–499 DOI 10.1038/nsmb1104.

**Zhang S, Qi Y, Yang C. 2010.** Arabidopsis SUMO E3 ligase AtMMS21 regulates root meristem development. *Plant Signalling & Behavior* **5**:53–55 DOI 10.4161/psb.5.1.10158.

**Zhao J. 2007.** Sumoylation regulates diverse biological processes. *Cellular and Molecular Life Science* **64**:3017–3033 DOI 10.1007/s00018-007-7137-4.