

# Genome-Wide Identification and Expression Analysis of Sucrose Nonfermenting-1-Related Protein Kinase (*SnRK*) Genes in *Triticum aestivum* in Response to Abiotic Stress

Shefali Mishra, Pradeep Sharma\*, Rajender Singh, Ratan Tiwari, GP Singh

ICAR-Indian Institute of Wheat and Barley Research, Karnal

\*Email ID: [Pradeep.Sharma@icar.gov.in](mailto:Pradeep.Sharma@icar.gov.in)

## Abstract

The *SnRK* gene family is a key regulator that modulates signalling pathways by phosphorylating the target protein in response to stress in plant. Although, the function of the the *SnRK* genes has been described in a variety of species, heitherto not reported in *Triticum aetivum*. The *SnRK* gene family was discovered in the wheat genome and its structural features were reported in this work. Identified 174 *SnRK* genes dispersed across 21 chromosomes were detected in the genome of *Triticum aestivum* which were divided into three subgroups (*SnRK1/2/3*) based on phylogenetic analyses and domain types. The gene intron-exon structure and protein-motif composition of *SnRKs* were similar within each subgroup but different amongst the groups. Gene duplication between the wheat, Arabidopsis, rice and barley genomes was also investigated in order to get insight into the evolutionary aspects of the *TaSnRK* family genes. *SnRK* genes showed differential expression patterns in leaves, roots, spike, and grains. There were additional redundant stress-related cis-elements identified in the promoters of 129 *SnRK* genes, and their expression levels varied greatly in response to ABA, drought, and light regulated elements. In particular,

*TaSnRK2.11* had higher and increased expression under the abiotic stresses and can be a candidate gene for the abiotic stress tolerance. The findings will help to characterise the functional properties of TaSnRK genes for future research

**Keywords:** *T. aestivum*, Genome-wide, SnRK and abiotic stress.

## 1. Introduction

Bread wheat (*T. aestivum* L.) is a major cereal crop and an important source of carbohydrates and protein in the human diet, accounting for 20% of daily calorie consumption. The most significant economic feature is grain yield, which is influenced by a variety of biotic and abiotic stressors. By 2050, it is expected that the demand for wheat would have increased by 60% [1]. To cope with abiotic challenges such as heat, salt and drought, plants generate a number of molecular defence mechanisms. Plants respond with these stressors in two ways: gene expression regulation and protein modification [2]. Protein kinase-mediated phosphorylation and dephosphorylation are critical steps in protein modification [3]. Sucrose nonfermenting 1 (SNF1)-related protein kinases (SnRKs) are a family of protein kinase genes involved in a wide range of physiological activities [4]. Based on their sequence similarities and gene architectures, SnRKs may be split into three subfamilies in plants: SnRK1, SnRK2, and SnRK3 [4,5]. The SnRK1 subfamily has a conserved N-terminal protein kinase domain, which is similar to a AMPKs in mammals and SNF1 in yeast [6]. The SnRK2 and SnRK3 subfamilies are distinct in plants, with both showing more variability than the SnRK1 subfamily members in terms of plant diversity whereas SnRK2 family members contain a conserved P kinase domain and a C-terminal variable adjusting domain [7]. SnRK3, also known as CIPK (CBL-interacting protein kinases), includes N-terminal protein kinase domains and NAF domains, and PPI domains at the C-terminus [8,9].

Plant cells respond to hunger and metabolic stress via the SnRK1 family of genes. SnRK1 kinases are catalytic components of heterotrimeric complexes, that interact with two other subunits [10]. SnRK1 was found to be involved in activation of sucrose synthase expression and to play an important role in the regulation of carbohydrate metabolism control in *S. tuberosum* [11]. Low energy stress may also cause SnRK1 nuclear translocation and subsequent induction of SnRK1 target genes, allowing for the replenishment of cellular energy for plant development [12]. According to the study, SnRK1 genes were also shown to be hotspots in a network of signalling pathways that comprised included cell cycle regulation, pathogen responses, and meristem development [13].

SnRK2 genes, on the other hand, are important in plants' responses to abiotic stresses, notably osmotic and salt stress. To cope with osmotic stress, SnRK2.10 phosphorylates several target genes in *A. thaliana*, including the dehydrins ERD10 and ERD14 [14]. In *N. tabacum*, SnRK2.1 regulates salt tolerance positively [15,16].

There are three types of SnRK2 subfamily genes in *A. thaliana*: ABA-independent kinases, kinase-responsive genes, and drought stress genes that are significantly activated by ABA [7,17]. The most comprehensive study on ABA-dependent group 3 kinases is now underway. For example, one of the SnRK2 family genes, AtSnRK2.6 (OST1), is involved in the ABA signalling pathway in stomatal guard cells, and OST1 protein stability in Arabidopsis may be controlled by the E3-ubiquitin ligase HOS15 to reduce ABA signal sensitivity. [18,19].

CBL- interacting protein kinases (CIPKs), also known as SnRK3 kinases, play a crucial role in plant stress tolerance [4,20]. The SOS (salt excessively sensitive) system, for example, was the first CBL-CIPK route found in *A. thaliana*. SOS3 (AtCBL4) on the cell membrane recognised the calcium signal created by salt stress, and then SOS3 joined with SOS2 (AtCIPK24) to phosphorylate SOS1 (Na<sup>+</sup>/H<sup>+</sup> antiporter) to remove excess Na<sup>+</sup> from root cells

[21,22]. MdCIPK13 and MdCIPK22 also improved drought and salt tolerance in apple by phosphorylating the sucrose transporter MdsUT2.2 [23,24]. In *B. napus*, overexpression of BnCBL1-BnCIPK6 may improve high salinity tolerance and low potassium tolerance [25]. Finally, growing data has highlighted the relevance of SnRKs function in nutrient consumption and stress response, and researchers may be able to increase plant stress resistance by genetic modification of these genes.

Finally, the role of SnRKs function in nutrient absorption and stress response is becoming apparent, and scientists may be able to increase plant stress resistance by genetically modifying these genes. However, there was no information available on *SnRk* genes in wheat, one of the most important cereals. In the current study, we identified 174 *SnRK* gene family members in the genome of *Triticum aestivum*. We also detected their evolutionary relationships, gene architecture, protein motifs, chromosomal position, and *cis*-elements in promoter regions in a systematic way. qRT-PCR was also used to study the differential expression of the *SnRK* gene under abiotic stress. The findings of this study provide light on the molecular mechanisms that support stress tolerance and molecular breeding.

## 2. Material and methods

### 2.1 *SnRKs* identification and characterization

TAIR (<https://www.Arabidopsis.org/>) was used to get the amino acid sequences of the SnRKs gene family in *A. thaliana* whereas, RGAP (<http://rice.plantbiology.msu.edu>) was used to obtain rice SnRK genes. The similar AtSnRK genes were blasted as query against the *T. aestivum* proteome databank (<http://plants.ensembl.org/index.html>) in local NCBI blast. Hidden Markov Model (HMM) was used for identification of TaSnRK proteins. To validate the above method, SnRK domains were assessed by the SMART (<http://smart.embl-heidelberg.de/>) [26], PFAM [27] and the NCBI Conserved Domain [28] databases. Following

these strict criteria, sequences containing the SnRK domain were chosen for further research. Additionally, tools from ExPASy tools ([http://www.expasy.ch/tools/pi\\_tool.html](http://www.expasy.ch/tools/pi_tool.html)) were also utilized to calculate the number of amino acids, isoelectric point (pI) and molecular weights (MW) of each SnRK protein.

## 2.2 *SnRK genes family phylogenetic analysis*

ClustalW was used to construct a multiple sequence alignment of 174 non-redundant TaSnRK amino acids using with default settings [29,30]. A phylogenetic tree was constructed using MEGA X and the Neighbor-Joining (NJ) method [30], with the poisson model, pairwise deletion, and 1000 replications of bootstrap. MEGA X was used to generate an unrooted NJ tree containing all *A.thaliana*, *O. sativa* and *H. vulgare* SnRKs protein sequences.

## 2.3 *The SnRK gene family's motif composition and gene structural analysis*

The online tool MEME (Multiple Expectation Maximization for Motif Elicitation) [31] (<http://meme.sdsc.edu/meme/itro.html>) was used to identify conserved motifs in TaSnRK proteins using the following parameters: The number of repeats is limitless, the maximum number of motifs is restricted to ten, and the ideal motif length is six to one hundred residues. The TaSnRK family gene's exon-intron structures were examined using the Gene Structure Display Server online tool (GSDS v.2.0: <http://gsds.cbi.pku.edu.cn>) and the gff3 data file [32].

## 2.4 *Analysis of Chromosomal location and Orthologous genes in T. aestivum*

Based on physical location information from the Plant ensembl database, the chromosomal coordinates of all *TaSnRK* genes were mapped to 21 chromosomes of the wheat genome using MapChart version 3.0 [33]. All the *T. aestivum* gene sequences were aligned to identify gene duplication using BLASTP with an e value of  $1e^{-10}$ . The pattern of duplicated

SnRK were classified as segmental, tandem duplications using MCScanX with default parameters [34]. A tandem duplication is defined as a chromosomal area of less than 200 kb that contains two or more genes [35]. Each duplicated TaSnRK gene's synonymous (Ks) and non-synonymous (Ka) substitution was determined using TBtools [36].

### *2.5 cis-elements in promoter regions of TaSnRKs*

The wheat genome database was used to extract upstream sequences (1500 bp) from the start codon of each *TaSnRK* gene, which were then subsequently utilised to calculate *cis*-element distributions in promoter regions with PlantCARE software (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) [37].

### *2.6 MIR genes that targeting TaSnRK*

The *TaSnRK* sequences were retrieved from the database of wheat genome. The psRNATarget service was used to examine the matured miRNA sequences [38,39] and the *TaSnRK* transcript sequences with default settings [40,41].

### *2.7 Gene expression profiling using RNA-seq data*

The expVIP database (<http://www.wheat-expression.com/>) was used to acquire the TPM (transcripts per million) value for each *TaSnRK*. Heatmaps were created using Clustvis (<https://biit.cs.ut.ee/clustvis/>) [42].

### *2.8 Plant material and growth conditions*

In this experiment, seeds of bread wheat genotypes (C-306, WL-711, RAJ3765, HS240, Kharchia 65, and HD2687) were obtained from the Germplasm Unit, ICAR-Indian Institute of Wheat and Barley Research, Karnal, India. Seeds were sterilised with 1% sodium hypochloride for 10 minutes before being washed three times with distilled water and germinating in petri plates at 22°C under controlled conditions.

After 5 days of germination, seedlings were transferred to a culture bottle filled with full-strength Hoagland's solution and left to develop for seven days. In two sets of three biological replications, each genotype was seeded. For drought stress, two contrasting wheat genotypes, C306-drought tolerant and WL711, drought susceptible were used.

After 14 days of development in Hoagland's solution, one group of seedlings was subjected to osmotic shock using 25% (v/v) polyethylene glycol (PEG) 6000 for 24 hours, while the other was left untreated [43]. Leaf samples from control and stressed seedlings were collected at the time periods specified above for expression analysis. For heat stress, two contrasting wheat genotypes were chosen: RAJ3765- heat resistant, and HS240- heat sensitive. These plants were maintained for 3h at 42°C (at Basal), followed by 3h at 37°C (at Normal), and finally 3h at 42°C (at Acquired). Leaf samples from the basal and acquired stress levels were harvested at the time interval stated above. For salt stress, Kharchia 65-salt resistant and HD2687- salt sensitive wheat genotypes were used. Two week old seedlings of both genotypes were exposed to 150 mM NaCl treatment. The leaf samples were taken 12h and 24 h the following treatment. For total RNA extraction, all obtained samples were wrapped in foil and promptly frozen in liquid nitrogen and kept at -80°C.

### *2.9 Isolation of RNA and Real time PCR analysis*

RNA was extracted using TRIzol Reagent according to the manufacturer's instructions. The isolated RNA was processed with DNase I (NEB, USA) to get DNA-free RNA. Superscript-III reverse transcriptase (Invitrogen, USA) was used to isolate first-strand cDNA from 1g of total RNA. For qRT-PCR analysis, the cDNA was diluted to 1:2 and 1.5 $\mu$ L of the diluted cDNA was used as a template in a 20 $\mu$ L reaction volume according to the manufacturer's instructions. The BIO-RAD CFX96 was used for Real time quantitative RT-PCR analysis was carried out by using SYBR Green (Bio-Rad). The endogenous control  $\beta$ -actin was used

to standardise the expression data [44]. Using the  $2^{-\Delta\Delta Ct}$  method, the expression was measured as a relative fold change [45]. The error bars represents the standard deviation of the expression across three biological replicates.

### 3. Results

#### 3.1 Identification of SnRK genes of wheat

A total of 174 SnRK genes were discovered and classified into three groups: SnRK 1, SnRK 2 and SnRK3 with 14, 65 and 95 genes, respectively (Table S1). Wheat has more SnRK genes than other plant species, such as *Arabidopsis thaliana* and rice. It might be owing to the hexaploid structure of the wheat genome. The longest amino acid sequence of each protein was chosen for further study. Gene IDs, amino acid quantities, molecular weights (MW), isoelectric points (pI), GRAVY and sub-cellular localization were among the metrics examined (Table 1). The molecular weight of all these TaSnRK proteins spans from 26.8 to 146.1 kDa, while the amino acid length spans from 239 to 1335. The isoelectric point of SnRKs family ranged from 4.3 to 9.4 indicating basic nature of proteins. The hydrophobicity and hydrophilicity are revealed by the GRAVY scores. SnRK proteins from *T. aestivum*, on the other hand, show a negative GRAVY score, implying that they are hydrophilic. The degree of hydrophilicity, on the other hand, is proportionate to the increased variability. According to the expected subcellular localization data, most of the *TaSnRK* genes are expressed in the nucleus and cytoplasm, followed by chloroplast and mitochondria (Table 1).

#### 3.2 Phylogenetic analysis of SnRK genes family

To investigate the evolutionary relationships and categorization of SnRK proteins, we did an unrooted phylogenetic tree analysis utilising the full-length amino acid sequences of 174 SnRKs genes from *T. aestivum*, 38 from *A. thaliana*, 50 from *O. sativa* and 50 from *H. vulgare* (Fig. 1). The clustering of 38 AtSnRKs into three groups was reported earlier [3].

Based on the phylogenetic analysis and domains presence, 174 TaSnRKs were also divided into three groups in this study. Of these, 14 proteins in the TaSnRK1 subfamily have Pkinase (PF00069 of Pfam), KA1 (PF02149) , and UBA (PF00627) domains, whereas 65 proteins in the TaSnRK2 subfamily have pkinase domains with strong resemblance to AtSnRK2 subfamily, and 95 proteins in the SnRK3 subfamily have Pkinase and NAF (PF03822) domain (Fig. 1).

### 3.3 The SnRK gene family's motif composition and gene structural analysis

The exon-intron structure of *TaSnRK* genes showed that the TaSnRK1 subfamily genes have more than 10 exons, while the TaSnRK2 subfamily genes have 2 to 34 exons followed by TaSnRK3 subfamily which have 1 to 33 exons (Fig 2). Notably, 14 genes are intron-less. In addition, the TaSnRK3 family is divided into two subgroups. SnRK3 subgroup 1 genes had more than ten exons, but SnRK3 subgroup 2 genes had less than four exons. MEME analysis showed that 10 conserved motifs are identified in TaSnRK proteins (Fig 3A). The conserved motif's sequence and length details have been listed in Table S2. The conserved P kinase domain including the pattern 1, 2, 3, 5, 6 was discovered in all TaSnRK genes in this investigation (Fig. 3B). Furthermore, TaSnRK genes from the same subfamily have comparable pattern compositions, but TaSnRK genes from different subfamilies have varied motif compositions. TaSnRK1 subfamily genes have 9 motif (motifs 1, 2, 3, 4, 5, 6, 7, 8 and 10) while TaSnRK2 genes had either motifs 1, 2, 4, 6, 8 or motifs 1, 2, 3, 4, 5 and 10. TaSnRK3 genes had 10 motifs while a few of them do not have any motif (Fig. 3A). In summary, the phylogenetic analysis categorization is strongly supported by the similar structures and the conserved motif compositions in the same subspace of SnRK genes

### 3.4 Analysis of Chromosomal location and Orthologous genes in *T. aestivum*

The chromosomal distribution of all *TaSnRK* genes across the genome was investigated which provides useful information on the genomic regions (Fig.4A). The A sub-genome had 56 *TaSnRK* genes, comprising 2 genes from *TaSnRK1* subfamily, 21 genes from *TaSnRK2* subfamily, and 33 genes from *TaSnRK3* subfamily, whereas the B sub-genome had 61 genes, comprising 8 *TaSnRK1* genes, 21 of *TaSnRK2* genes, and 32 of *TaSnRK3* genes. However, 4 of *TaSnRK1* genes, 23 of *TaSnRK2* genes, and 30 of *TaSnRK3* genes were found on the D sub-genome (Fig. 4B). These findings indicated that *TaSnRK* genes has been distributed randomly dispersed throughout A, B, and D chromosomes (Fig. 4C).

In this study, 57 orthologous pairings were identified between *T. aestivum* and *H. vulgare*, while 6 between *T. aestivum* and *A. thaliana*. However, 53 orthologous genes were found within wheat species for instance between *T. aestivum* and *T. urartu*, while 102 orthologous genes between *T. aestivum* and *Ae. dicoccoides* and 63 between *T. aestivum* and *Ae. tauschii* (Table S3).

We identified 11 segmental events across different chromosomes and 2 tandem duplication occurrences in the same chromosome using the BLAST and MCScanX techniques. The findings revealed that gene duplication may have produced some *TaSnRK* genes, and that segmental duplication events were important in the growth of *TaSnRK* genes in the wheat genome. We also looked into the frequency of tandem duplication occurrences. There were 45 *TaSnRK* gene pairs found in this area all of which were strongly related. (Table S4). However, the identities of of these were >80%, indicating that they were included into tandem duplication occurrences.

The evaluation of the evolutionary limitations of Ks values, Ka/Ks ratios and the divergence time for the paralogous and orthology genes in the SnRK family has been computed (Table. S4). The  $K_a/K_s$  ratios of the majority of segmentally duplicated *TaSnRK* gene pairs were <1,

the mean values of TaSnRK3 gene pairs ( $K_a/K_s = 0.30$ ) and TaSnRK2 ( $K_a/K_s = 0.35$ ) were lower than TaSnRK1 ( $K_a/K_s = 0.41$ ). Furthermore, segmental gene divergence time spans from 18.76 Mya to 34.97 Mya. These findings showed that the *TaSnRK* gene family may have been subjected to significant purifying selection during evolution.

### 3.5 Promoter analysis

PlantCARE was employed to look at *cis*-regions (1.5 kb upstream from ATG to better understand the function and regulatory mechanisms of *TaSnRK* genes. Of 174 TaSnRKs, we discovered 129 genes with *cis*- elements (Fig. 5, Table S5). MyB, ABRE, and LTRE *cis*-elements were found to be involved in drought, ABA, and low-temperature responses. Auxin (9.77 %), MeJA (51.72 %), and Gibberallin (11.49 %) *cis*-elements were found in phytohormones. It was also shown that most genes have many *cis*-element types. In addition, the TaSnRK1 (30.50) family had more *cis*-elements than the TaSnRK2 (19.29) or TaSnRK3 (23.81) families (Table S5). Finally, the *cis*-elements study revealed that most *TaSnRK* genes can respond to a variety of environmental challenges, and that distinct subfamily genes can be regulated in various ways.

### 3.6 MicroRNAs targeting *TaSnRK*

The psRNATarget server was used to examine the role of miRNAs in regulating the expression of *TaSnRK* genes. We estimated 16 *TaSnRK* genes as potential targets for eight distinct miRNAs (Table 2). Tae-miR319 implicated in the regulation of seven *TaSnRK* genes (TaSnRK3.33, TaSnRK3.36, TaSnRK2.37, TaSnRK2.35, TaSnRK2.36, TaSnRK3.38, TaSnRK2.39). Tae-miR408 accounted for regulating expression of three *TaSnRK* genes (TaSnRK3.63, TaSnRK3.64, TaSnRK3.62). The expression of *TaSnRK* genes (TaSnRK1.3, TaSnRK1.2, and TaSnRK2.39, TaSnRK2.38) may be influenced by the Tae-miR164 and

Tae-miR167b (Table 2). Tae-miR1119 and Tae-miR160 were predicted to influence the expression of TaSnRK3.38 and TaSnRK3.74, respectively.

### 3.7 Expression profile of *SnRK* genes in different tissues under abiotic stress

To understand more about how TaSnRKs play a role in development and stress responses, we used the expVIP database to calculate TPM values for all TaSnRKs from experiments that included abiotic stress and different developmental stages. Heatmaps were created using the TPM data. The expression patterns of 174 *TaSnRK* genes were studied in diverse wheat tissues. Based on their expression patterns, TaSnRKs were divided into three categories. Group 1 had 16 genes with high expression levels ( $\log_2$ -based values  $> 5$ ) across all tissues investigated. For example, SnRK3.40 was found to be highly expressed in all vegetative organs, with  $\log_2$ -based average values. In Group 2, the expression levels of 55 *TaSnRK* genes were significantly lower across all detected tissues ( $\log_2$ -based values  $> 2$ ). In group 3, 103 *TaSnRK* genes were involved with the lowest expression levels in diverse tissues at various stages. Meanwhile, Group 1 had one TaSnRK1, 6 TaSnRK2 genes and 9 TaSnRK3 genes; Group 2 had 8 TaSnRK1, 21 TaSnRK2, and 26 TaSnRK3 genes and Group 3 had 5 TaSnRK1, 38 TaSnRK2, and 69 TaSnRK3 genes (Fig. 6A). The expression pattern of *TaSnRK* genes was nonetheless studied using drought, salinity and heat stresses [46]. The expression levels of *TaSnRK* genes have been significantly changed under various abiotic stresses (Fig. 6). TaSnRK3.17 and TaSnRK3.18, for example, were highly stimulated by all treatments, whereas TaSnRK3.15 and TaSnRK3.16 responded to dehydration by increasing their levels of expression. But, TaSnRK 3.37, TaSnRK3.45, and TaSnRK3.46 have demonstrated high expression levels in response to drought and heat stress (Fig. 6B). These results demonstrated that TaSnRKs have a large variety of patterns of expression and even genes within the same subfamily had different patterns of expression.

### 3.8 Validation of *TaSnRK* for abiotic stresses using qRT-PCR

SnRK family members play critical roles in plant abiotic stress response. However, the precise mechanism underpinning SnRK function in plants is not been well understood. The expression analysis of SnRK2s and SnRK3s in leaf tissue was identified by qRT-PCR following different stress treatments at different time intervals. This was to study the function of *TaSnRK2* and *TaSnRK3* genes in responding to salt, heat and drought stress (Fig. 7).

After 3 hours of heat stress treatment, the expression level of four genes namely *TaSnRK2.16*, *TaSnRK2.11*, *TaSnRK3.23*, and *TaSnRK2.8* was up-regulated at basal thermotolerance as compared to acquired thermotolerance stress treatment in wheat genotype Raj3765, while expression of *TaSnRK2.16* was increased 5 fold in wheat genotype HS240 (Fig. 7). Furthermore, to investigate the effect of SnRK during salt stress treatment, expressions of SnRK genes were carried out at different time intervals. At 24 hours following salt stress, the expression levels of two genes, *TaSnRK2.11* and *TaSnRK2.12*, were found to be raised 32-fold and 20-fold, respectively, in wheat genotype Kharchia 65 (Fig 7). Whereas *TaSnRK3.87* and *TaSnRK2.16* expression level was increased 4- and 8- fold at 12 h, respectively in wheat genotype HD2687. Under drought stress, we selected two contrasting genotypes i.e. C306 and WL711. *TaSnRK 2.11* expression value were 10-fold up-regulated whereas *TaSnRK3.87* were 4-fold in wheat genotype C306 while, in WL711 the expression of *TaSnRK2.3*, *TaSnRK2.11*, *TaSnRK3.18* and *TaSnRK2.8* was more than 4-folds. In conclusion, expression level of *TaSnRK2.11* was higher under salt, drought and heat stresses and it could be further utilized to enhancing abiotic tolerance in wheat.

## 4. Discussion

The *TaSnRK* gene family had 174 members of the *T. aestivum* genome, which were designated as *TaSnRK1*, *TaSnRK2*, and *TaSnRK 3* based on their subfamily classification.

The *TaSnRK* gene family was carefully searched, including evolutionary linkages, protein patterns, gene architectures, gene duplication, distributions of chromosomal, and the promoter *cis*-elements. This research progresses towards the future functionality of *SnRK* genes in order to improve the abiotic stress adaptive of plant. *SnRK* genes were found in *A. thaliana* [4], *O. sativa* [47], *B. distachyon* [48], and *E. grandis* [49]] having 39, 48, 44, and 34, respectively. The number of *TaSnRK* genes in the genome of *T. aestivum* is substantially higher than in diploid plants. There were 14 *TaSnRK1* genes identified, 65 *TaSnRK2* genes, and 95 *TaSnRK3* genes were discovered and divided into three subfamilies. According to more detailed description, *T. aestivum* and other species have similar member proportions for each subfamily. *T. aestivum* is a naturally occurring amphidiploid evolved from *T. urartu* (AA), *A. speltoides* (BB), and *A. tauschii* (DD). There were 56, 61, and 57 *TaSnRK* genes discovered in the A, B, and D sub-genomes demonstrating that *SnRK* genes had a similar functional role in progenitor species.

The *SnRK* subfamily genes' conserved domains vary, but the N-terminal protein kinase domain is maintained. The NAF domain of *SnRK3* subfamily genes has been found to interact with CBLs in a calcium-dependent manner. Furthermore, the NAF domain identifies a group of heterologous kinases that CBL calcium sensor protein targets and engage in a variety of signalling cascades [8]. This study discovered that various *TaSnRKs* subfamily genes shared different types of conserved domains. This might imply that the *TaSnRK* genes family is functionally diverse based on the domains contained.

In the *AtSnRK* and *TaSnRK* genes, certain subfamily genes showed substantial structural exon-intron divergences and gene length differences. In genes with fewer introns, increased expression in plants has been previously reported [50,51]. A compact gene structure with minimal introns has enabled fast activation and response to changing environmental

conditions [51]. However, we found no evidence that *TaSnRK* gene with fewer introns had greater expression levels than the other TaSnRKs genes.

According to accumulating data, gene activity was often linked to discrepancies in the promoter region [52]. *cis*-elements played an important function in regulating gene expression throughout development and environmental changes [53,54]. TaSnRKs had several *cis*-elements, including growth hormones, MyB, ABRE, and LTR according to promoter analysis in this investigation. The presence of at least one of these components in the majority of gene promoters demonstrated that many TaSnRKs were capable of reacting to a range of abiotic stressors while also promoting growth. When gene expression profiles from TaSnRKs with MyB and ABRE *cis*-elements were combined under drought stress, TaSnRKs with MyB and ABRE *cis*-elements grew by an average of 6.3-fold, but TaSnRKs without *cis*-elements only grew by 3.5-fold. As a result, *cis*-elements analysis can aid in gene function, particularly gene expression patterns under various stress conditions.

*TaSnRK* gene expression levels in *T. aestivum* tissues and organs were investigated using transcriptome data [46]. The study's findings revealed that expression patterns of these genes are classified into three categories (Fig. 7). In turn, we found that Group 2 TaSnRKs contain less *cis*-elements than TaSnRKs in Groups 1 and 3 in their promoters. Every gene in Group 1 has an average of 5.08 Auxin, 13.55 MeJA, and 8.47 ABRE, 1.69 MYB, 3.38 LTR compared with each gene in Group 3 and averages of 2.05 in Auxin, 11.08 of MeJA, 1.43 of Gibberellins and 6.67 ABRE, 1.43 MYB, 0.61 LTR, each gene in Group 2 with a total of 0.35 Auxin, 9.12 MeJA, 4.56 Gibberellin, 3.50 ABRE, 1.40 MYB, 0.35 LTRE. These data show that TaSnRK activity is linked to promoter region differences.

In addition, the roles and activities of numerous TaSnRKs in response to various abiotic stressors were determined. Drought stress findings show that ABA production and *A. thaliana*

signals in response to drought were orthologous to AtSnRK2.3 [39] and imply that TaSnRK2.11 and AtSnRK2.3 play the same similar function in drought stress. The extreme expression changes in *T. aestivum* against to effects of drought, salt, thermal stress and ABA induction were reported in TaSnRK2.8, TaSnRK2.12, TaSnRK3.11, TaSnRK2.16, TaSnRK3.23 and TaSnRK3.83 whereas its orthologs, AtSnRK2.2, could also respond to ABA induction and osmotic stress in *A. thaliana*. This shows that under different conditions the TaSnRK2.11 gene may be activated substantially.

Findings of this study provide a thorough description of the *SnRK* gene family in *T. aestivum*. It helps us to better understand the biological role of specific *TaSnRK* specific genes in *T. aestivum*. The study presented just a fundamental characterisation of the *TaSnRK* genes and a comprehensive functional validation would be required to hold the importance of the SnRK family.

## 5. Conclusion

*SnRK* genes involving in a number of signalling pathways, include biotic and abiotic stress responses. The SnRK gene family in *T. aestivum* has been extensively studied. A total of 174 *TaSnRK* genes were discovered and categorized into three subgroups based on motif composition and gene structural similarity within each subfamily. The evolutionary characteristics of the TaSnRK genes may be additionally derived from phylogenetic analyses of the SnRK genes in *A. thaliana* and *O. sativa* . In order to better understanding the biological rsignificance of *TaSnRK* genes in *T. aestivum*, the TaSnRK family's microRNA targeting *cis*-acting elements and gene expressions were also examined.

## Acknowledgments

This work was supported by Indian Council of Agricultural Research, New Delhi, India- core grant. We are thankful to ICAR-Indian Institute of Wheat and Barley Research for providing the facility and Dr SK Muthusamy for helpful discussion. This Research paper is IIWBR contribution No. 329.

## References

1. FAO *Crop Prospects and Food Situation - Quaterly Global Report; 2020*; ISBN 9789251322628.
2. Bohnert, H.J.; Gong, Q.; Li, P.; Ma, S. Unraveling abiotic stress tolerance mechanisms - Getting genomics going. *Curr. Opin. Plant Biol.* **2006**, *9*, 180–188, doi:10.1016/j.pbi.2006.01.003.
3. Hunter, T. Protein kinases and phosphatases: The Yin and Yang of protein phosphorylation and signaling. *Cell* **1995**, *80*, 225–236, doi:10.1016/0092-8674(95)90405-0.
4. Hrabak, E.M.; Chan, C.W.M.; Gribskov, M.; Harper, J.F.; Choi, J.H.; Halford, N.; Kudla, J.; Luan, S.; Nimmo, H.G.; Sussman, M.R.; et al. The Arabidopsis CDPK-SnRK superfamily of protein kinases. *Plant Physiol.* **2003**, *132*, 666–680, doi:10.1104/pp.102.011999.
5. Halford, N.G.; Hardie, D.G. SNF1-related protein kinases: Global regulators of carbon metabolism in plants? *Plant Mol. Biol.* 1998.
6. Celenza, J.L.; Carlson, M. A yeast gene that is essential for release from glucose

- repression encodes a protein kinase. *Science* (80-. ). **1986**, 233, 1175–1180, doi:10.1126/science.3526554.
7. Kulik, A.; Wawer, I.; Krzywińska, E.; Bucholc, M.; Dobrowolska, G. SnRK2 protein Kinases - Key regulators of plant response to abiotic stresses. *Omi. A J. Integr. Biol.* **2011**, 15, 859–872, doi:10.1089/omi.2011.0091.
  8. Albrecht, Â.; Ritz, O.; Linder, S.; Harter, K. Albrecht-2001-The NAF domain defin. **2001**, 20.
  9. Ohta, M.; Guo, Y.; Halfter, U.; Zhu, J.K. A novel domain in the protein kinase SOS2 mediates interaction with the protein phosphatase 2C ABI2. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, 100, 11771–11776, doi:10.1073/pnas.2034853100.
  10. Hardie, D.G.; Carling, D.; Carlson, M. The AMP-activated/SNF1 protein kinase subfamily: Metabolic sensors of the eukaryotic cell? *Annu. Rev. Biochem.* **1998**, 67, 821–855, doi:10.1146/annurev.biochem.67.1.821.
  11. Purcell, P.C.; Smith, A.M.; Halford, N.G. Antisense expression of a sucrose non-fermenting-1-related protein kinase sequence in potato results in decreased expression of sucrose synthase in tubers and loss of sucrose-inducibility of sucrose synthase transcripts in leaves. *Plant J.* **1998**, 14, 195–202, doi:10.1046/j.1365-313X.1998.00108.x.
  12. Ramon, M.; Dang, T.V.T.; Broeckx, T.; Hulsmans, S.; Crepin, N.; Sheen, J.; Rolland, F. Default activation and nuclear translocation of the plant cellular energy sensor SnRK1 regulate metabolic stress responses and development. *Plant Cell* **2019**, 31, 1614–1632, doi:10.1105/tpc.18.00500.
  13. Halford, N.G.; Hey, S.J. Snf1-related protein kinases (SnRKs) act within an intricate

- network that links metabolic and stress signalling in plants. *Biochem. J.* **2009**, *419*, 247–259, doi:10.1042/BJ20082408.
14. Maszkowska, J.; Dębski, J.; Kulik, A.; Kistowski, M.; Bucholc, M.; Lichočka, M.; Klimecka, M.; Sztatelman, O.; Szymańska, K.P.; Dadlez, M.; et al. Phosphoproteomic analysis reveals that dehydrins ERD10 and ERD14 are phosphorylated by SNF1-related protein kinase 2.10 in response to osmotic stress. *Plant Cell Environ.* **2019**, *42*, 931–946, doi:10.1111/pce.13465.
  15. Diédhiou, C.J.; Popova, O. V.; Dietz, K.J.; Golldack, D. The SNF1-type serine-threonine protein kinase SAPK4 regulates stress-responsive gene expression in rice. *BMC Plant Biol.* **2008**, *8*, 1–13, doi:10.1186/1471-2229-8-49.
  16. Zhang, H.; Jia, H.; Liu, G.; Yang, S.; Zhang, S.; Yang, Y.; Yang, P.; Cui, H. Cloning and characterization of SnRK2 subfamily II genes from *Nicotiana tabacum*. *Mol. Biol. Rep.* **2014**, doi:10.1007/s11033-014-3440-y.
  17. Kawa, D.; Meyer, A.J.; Dekker, H.L.; Abd-El-Haliem, A.M.; Gevaert, K.; Van De Slijke, E.; Maszkowska, J.; Bucholc, M.; Dobrowolska, G.; De Jaeger, G.; et al. SnRK2 protein kinases and mRNA decapping machinery control root development and response to salt. *Plant Physiol.* **2020**, *182*, 361–371, doi:10.1104/pp.19.00818.
  18. The, S.; Cell, P.; Dec, N. Arabidopsis OST1 Protein Kinase Mediates the Regulation of Stomatal Aperture by Abscisic Acid and Acts Upstream of Reactive Oxygen Species Production Author ( s ): Anna-Chiara Mustilli , Sylvain Merlot , Alain Vavasseur , Francesca Fenzi and Jérôme Publis. **2016**, *14*, 3089–3099, doi:10.1105/tpc.007906.This.
  19. Ali, A.; Kim, J.K.; Jan, M.; Khan, H.A.; Khan, I.U.; Shen, M.; Park, J.; Lim, C.J.; Hussain, S.; Baek, D.; et al. Rheostatic Control of ABA Signaling through HOS15-

- Mediated OST1 Degradation. *Mol. Plant* **2019**, *12*, 1447–1462, doi:10.1016/j.molp.2019.08.005.
20. Kim, K.N.; Lee, J.S.; Han, H.; Choi, S.A.; Go, S.J.; Yoon, I.S. Isolation and characterization of a novel rice Ca<sup>2+</sup>-regulated protein kinase gene involved in responses to diverse signals including cold, light, cytokinins, sugars and salts. *Plant Mol. Biol.* **2003**, *52*, 1191–1202, doi:10.1023/B:PLAN.0000004330.62660.a2.
  21. Guo, Y.; Xiong, L.; Song, C.P.; Gong, D.; Halfter, U.; Zhu, J.K. A calcium sensor and its interacting protein kinase are global regulators of abscisic acid signaling in Arabidopsis. *Dev. Cell* **2002**, doi:10.1016/S1534-5807(02)00229-0.
  22. Liu, J.; Ishitani, M.; Halfter, U.; Kim, C.S.; Zhu, J.K. The Arabidopsis thaliana SOS2 gene encodes a protein kinase that is required for salt tolerance. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, doi:10.1073/pnas.97.7.3730.
  23. Ma, Q.J.; Sun, M.H.; Kang, H.; Lu, J.; You, C.X.; Hao, Y.J. A CIPK protein kinase targets sucrose transporter MdsUT2.2 at Ser 254 for phosphorylation to enhance salt tolerance. *Plant Cell Environ.* **2019**, *42*, 918–930, doi:10.1111/pce.13349.
  24. Ma, Q.J.; Sun, M.H.; Lu, J.; Kang, H.; You, C.X.; Hao, Y.J. An apple sucrose transporter MdsUT2.2 is a phosphorylation target for protein kinase MdCIPK22 in response to drought. *Plant Biotechnol. J.* **2019**, *17*, 625–637, doi:10.1111/pbi.13003.
  25. Greco, M.; Chiappetta, A.; Bruno, L.; Bitonti, M.B. In *Posidonia oceanica* cadmium induces changes in DNA methylation and chromatin patterning. *J. Exp. Bot.* **2012**, *63*, 695–709, doi:10.1093/jxb/err313.
  26. Letunic, I.; Doerks, T.; Bork, P. SMART 7: Recent updates to the protein domain annotation resource. *Nucleic Acids Res.* **2012**, *40*, 302–305, doi:10.1093/nar/gkr931.

27. Finn, R.D.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **2016**, *44*, D279–D285, doi:10.1093/nar/gkv1344.
28. Marchler-Bauer, A.; Lu, S.; Anderson, J.B.; Chitsaz, F.; Derbyshire, M.K.; DeWeese-Scott, C.; Fong, J.H.; Geer, L.Y.; Geer, R.C.; Gonzales, N.R.; et al. CDD: A Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **2011**, *39*, 225–229, doi:10.1093/nar/gkq1189.
29. Saitou, N.; Nei, M. ESCALA CIWA-AR Escala CIWA-Ar(Clinical Institute Withdrawal Assesment for Alcohol) Evaluación del Síndrome de Abstinencia Alcohólica. *Mol. Biol. Evol* **1987**, *4*, 406–425.
30. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549, doi:10.1093/molbev/msy096.
31. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, 1–7, doi:10.1093/nar/gkp335.
32. Hu, B.; Jin, J.; Guo, A.Y.; Zhang, H.; Luo, J.; Gao, G. GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* **2015**, *31*, 1296–1297, doi:10.1093/bioinformatics/btu817.
33. Voorrips, R.E. Mapchart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.* **2002**, *93*, 77–78, doi:10.1093/jhered/93.1.77.
34. Wang, Y.; Tang, H.; Debarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.H.; Jin, H.;

- Marler, B.; Guo, H.; et al. MCSScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **2012**, *40*, 1–14, doi:10.1093/nar/gkr1293.
35. Houb, E.B. The arms race is ancient history in Arabidopsis, the wildflower. *Nat. Rev. Genet.* **2001**, *2*, 516–527, doi:10.1038/35080508.
36. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202, doi:10.1016/j.molp.2020.06.009.
37. Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van De Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **2002**, *30*, 325–327, doi:10.1093/nar/30.1.325.
38. Kozomara, A.; Griffiths-Jones, S. MiRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **2014**, doi:10.1093/nar/gkt1181.
39. Sun, F.; Guo, G.; Du, J.; Guo, W.; Peng, H.; Ni, Z.; Sun, Q.; Yao, Y. Whole-genome discovery of miRNAs and their targets in wheat (*Triticum aestivum* L.). *BMC Plant Biol.* **2014**, doi:10.1186/1471-2229-14-142.
40. Dai, X.; Zhao, P.X. PsRNATarget: A plant small RNA target analysis server. *Nucleic Acids Res.* **2011**, doi:10.1093/nar/gkr319.
41. Katiyar, A.; Smita, S.; Muthusamy, S.K.; Chinnusamy, V.; Pandey, D.M.; Bansal, K.C. Identification of novel drought-responsive microRNAs and trans-acting siRNAs from *Sorghum bicolor* (L.) Moench by high-throughput sequencing analysis. *Front. Plant Sci.* **2015**, doi:10.3389/fpls.2015.00506.

42. Metsalu, T.; Vilo, J. ClustVis: A web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* **2015**, *43*, W566–W570, doi:10.1093/nar/gkv468.
43. Kaur, A.; Gupta, O.P.; Meena, N.L.; Grewal, A.; Sharma, P. Comparative Temporal Expression Analysis of MicroRNAs and Their Target Genes in Contrasting Wheat Genotypes During Osmotic Stress. *Appl. Biochem. Biotechnol.* **2017**, *181*, 613–626, doi:10.1007/s12010-016-2236-z.
44. Muthusamy, S.K.; Dalal, M.; Chinnusamy, V.; Bansal, K.C. Differential regulation of genes coding for organelle and cytosolic ClpATPases under biotic and abiotic stresses in Wheat. *Front. Plant Sci.* **2016**, doi:10.3389/fpls.2016.00929.
45. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* **2001**, doi:10.1006/meth.2001.1262.
46. Coello, P.; Hirano, E.; Hey, S.J.; Muttucumar, N.; Martinez-Barajas, E.; Parry, M.A.J.; Halford, N.G. Evidence that abscisic acid promotes degradation of SNF1-related protein kinase (SnRK) 1 in wheat and activation of a putative calcium-dependent SnRK2. *J. Exp. Bot.* **2012**, *63*, 913–924, doi:10.1093/jxb/err320.
47. Kobayashi, Y.; Yamamoto, S.; Minami, H.; Kagaya, Y.; Hattori, T. Differential activation of the rice sucrose nonfermenting1-related protein kinase2 family by hyperosmotic stress and abscisic acid. *Plant Cell* **2004**, *16*, 1163–1177, doi:10.1105/tpc.019943.
48. Wang, L.; Hu, W.; Sun, J.; Liang, X.; Yang, X.; Wei, S.; Wang, X.; Zhou, Y.; Xiao, Q.; Yang, G.; et al. Genome-wide analysis of SnRK gene family in *Brachypodium distachyon* and functional characterization of BdSnRK2.9. *Plant Sci.* **2015**,

- doi:10.1016/j.plantsci.2015.05.008.
49. Wang, Y.; Yan, H.; Qiu, Z.; Hu, B.; Zeng, B.; Zhong, C.; Fan, C. Comprehensive analysis of SNRK gene family and their responses to salt stress in eucalyptus grandis. *Int. J. Mol. Sci.* **2019**, *20*, 1–21, doi:10.3390/ijms20112786.
  50. Chung, B.Y.W.; Simons, C.; Firth, A.E.; Brown, C.M.; Hellens, R.P. Effect of 5'UTR introns on gene expression in *Arabidopsis thaliana*. *BMC Genomics* **2006**, doi:10.1186/1471-2164-7-120.
  51. Jeffares, D.C.; Penkett, C.J.; Bähler, J. Rapidly regulated genes are intron poor. *Trends Genet.* **2008**, *24*, 375–378, doi:10.1016/j.tig.2008.05.006.
  52. Xue, T.; Wang, D.; Zhang, S.; Ehling, J.; Ni, F.; Jakab, S.; Zheng, C.; Zhong, Y. Genome-wide and expression analysis of protein phosphatase 2C in rice and *Arabidopsis*. *BMC Genomics* **2008**, *9*, 1–21, doi:10.1186/1471-2164-9-550.
  53. Ali, G.M.; Komatsu, S. Proteomic analysis of rice leaf sheath during drought stress. *J. Proteome Res.* **2006**, *5*, 396–403, doi:10.1021/pr050291g.
  54. Sarda, X.; Tusch, D.; Ferrare, K.; Legrand, E.; Dupuis, J.M.; Casse-Delbart, F.; Lamaze, T. Two TIP-like genes encoding aquaporins are expressed in sunflower guard cells. *Plant J.* **1997**, *12*, 1103–1111, doi:10.1046/j.1365-313X.1997.12051103.x.

## Legends of figure

**Fig. 1.** Phylogenetic tree of SnRK genes from *T. aestivum*, *A. thaliana*, *O. sativa* and *H. vulgare*. 174 SnRK genes from wheat, 38 from *Arabidopsis*, 50 from rice and 50 from barley were clustered into three subsets (SnRK1, SnRK2 and SnRK3). TaSnRK1, TaSnRK 2 and

TaSnRK 3 subgroups denoted blue, green and red respectively. The tree was made using the neighbor-joining method.

**Fig. 2.** Gene structure of TaSnRK. In yellow rectangles the exons and introns are shown in grey lines connected by two exons.

**Fig. 3.** Conserved motif of TaSnRK genes in *T.aestivum*. (A) Arrangement of ten conserved motifs in the *TaSnRK* genes. In each of the TaSnRKs sequences the varied coloured boxes are distinct motifs and positions; (B) the conserved residues are indicated by the logos..

**Fig.4 .** Distribution of TaSnRK genes in the 21 wheat chromosomes and in the three sub-genomes. Genes represent tandem duplications in highlighted boxes. (A) Physical chromosomal distribution map with left hand location and right hand gene name. (B) Distribution to 21 chromosomes of TaSnRK genes. (C) TaSnRK gene distribution in the three sub-genomes.

**Fig. 5.** *cis*-regulatory elements in *TaSnRK* genes . The promoter region of *TaSnRK* gene has been examined using PlantCare data to detect the presence of conserved *cis*-regulatory elements using.

**Fig. 6.** *TaSnRK* genes expression (A) tissue specific differential expression (B) abiotic stress specific tissue. The colour scale shows the values of the  $\log_2$

**Fig. 7.** Expression analysis of 10 *TaSnRK* genes at different time periods under abiotic stress (Heat, Salt and Drought) as determined assessed by qRT-PCR. The error bars indicate the standard deviation.

**Legends of table:**

**Table 1:** list of the newly known TaSnRK family members in *T. aestivum*

**Table 2.** Prediction TamiR genes and their targets

**Supplementary table:**

**Table S1.** Gene ID's of SnRK genes of Wheat, Barley, Rice and Arabidopsis and chromosomal location of genes in wheat genome.

**Table S2.** Details of the 10 motifs in the TaSnRK proteins of *T. aestivum*.

**Table S3.** Orthologous and paralogous genes of wheat with inter and intra-species.

**Table S4.** Ka/Ks ratio of paralogous genes

**Table S5.** Description of *cis*-regulatory elements on the basis of upstream region of the genes.