Journal of Advanced Research 61 (2024) 19–33

Contents lists available at ScienceDirect

Journal of Advanced Research

journal homepage: www.elsevier.com/locate/jare

Integrative gene duplication and genome-wide analysis as an approach to facilitate wheat reverse genetics: An example in the *TaCIPK* family

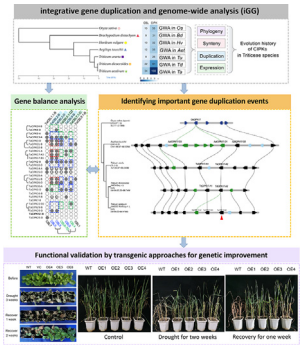
Ya'nan Wu [a,1], Jialu Feng [b,1], Qian Zhang [a,1], Yaqiong Wang [a], Yanbin Guan [a], Ruibin Wang [a], Fu Shi [a], Fang Zeng [a], Yuesheng Wang [a], Mingjie Chen [a], Junli Chang [a], Guangyuan He [a,*], Guangxiao Yang [a,*], Yin Li [a,*]

[a] *The Genetic Engineering International Cooperation Base of Chinese Ministry of Science and Technology, The Key Laboratory of Molecular Biophysics of Chinese Ministry of Education, College of Life Science and Technology, Huazhong University of Science & Technology, Wuhan 430074, China*
[b] *Hubei Provincial Key Laboratory of Occupational Hazard Identification and Control, School of Medicine, Wuhan University of Science and Technology, Wuhan 430065, China*

HIGHLIGHTS

- The iGG analysis has been proposed to utilize evolutionary information between the crop and several its relatives.
- The evolution of *CIPK* and *CBL* genes have been characterized in seven species to prove the iGG method.
- The divergent expression is important for keeping the stoichiometric balance of TaCBL-TaCIPK.
- Our iGG analysis reveals a Triticeae-specific *TaCIPK* duplicate responding to drought.
- Transgenic plants of *TaCIPK17*-OE in tobacco and wheat showed higher resistance to drought.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 13 April 2023
Revised 25 August 2023
Accepted 6 September 2023
Available online 9 September 2023

Keywords:
Integrative gene duplication and genome-wide analysis
Evolutionary expansion
Reverse genetics
Triticeae species
CBL
CIPK

ABSTRACT

Introduction: Reverse genetic studies conducted in the plant with a complex or polyploidy genome enriched with large gene families (like wheat) often meet challenges in identifying the key candidate genes related to important traits and prioritizing the genes for functional experiments.
Objective: To overcome the above-mentioned challenges of reverse genetics, this work aims to establish an efficient multi-species strategy for genome-wide gene identification and prioritization of the key candidate genes.
Methods: We established the integrative gene duplication and genome-wide analysis (iGG analysis) as a strategy for pinpointing key candidate genes deserving functional research. The iGG captures the evolution, and the expansion/contraction of large gene families across phylogeny-related species and integrates spatial–temporal expression information for gene function inference. Transgenic approaches were also employed to functional validation.
Results: As a proof-of-concept for the iGG analysis, we took the wheat calcineurin B-like protein-interacting protein kinases (CIPKs) family as an example. We identified *CIPK*s from seven monocot species, established the orthologous relationship of *CIPK*s between rice and wheat, and characterized Triticeae-specific *CIPK* duplicates (*e.g.*, *CIPK4* and *CIPK17*). Integrated with our analysis of CBLs and CBL-CIPK interaction, we revealed that divergent expressions of *TaCBL*s and *TaCIPK*s could play an important role in keeping the stoichiometric balance of CBL-CIPK. Furthermore, we validated the function of

* Corresponding authors.
E-mail addresses: hegy@hust.edu.cn (G. He), ygx@hust.edu.cn (G. Yang), yinli2021@hust.edu.cn (Y. Li).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jare.2023.09.005

TaCIPK17-A2 in the regulation of drought tolerance by using transgenic approaches. Overexpression of *TaCIPK17* enhanced antioxidant capacity and improved drought tolerance in wheat.
Conclusion: The iGG analysis leverages evolutionary and comparative genomics of crops with large genomes to rapidly highlight the duplicated genes potentially associated with speciation, domestication and/or particular traits that deserve reverse-genetic functional studies. Through the identification of Triticeae-specific *TaCIPK17* duplicates and functional validation, we demonstrated the effectiveness of the iGG analysis and provided a new target gene for improving drought tolerance in wheat.

Introduction

As a staple food, wheat (*Triticum aestivum* L.) is of great importance in facilitating the development of human civilization [1]. Nowadays, wheat is among the top-three most produced cereal crops, serving as the staple for about 40% of the population in the world. Wheat evolution features allopolyploidization, reticulate evolution and gene flow with the wheat relatives, and a recent burst of gene duplication in Triticeae species [2–4]. These features have shaped the hexaploid bread wheat genome as one of the most complexed in cereals, with a genome size of 14.5 Gb containing over 85% repetitive DNA [2].

Forward genetics in wheat has long been challenging owing to several features of wheat biology and genomics [5,6]: **(1)** Functional redundancy within homoeologous syntenic genes or between duplicated gene copies often hinders the screening of mutant phenotypes by using mutagenesis or gene editing approaches; **(2)** The large and complex wheat genome with long linkage disequilibrium distances poses challenges to fine map the causal genetic loci of a given trait and to determine causal mutations [7,8]; **(3)** Self pollination requires more tedious work to construct large mapping populations for forward and quantitative genetics purposes compared to other cereals (*e.g.*, rice and maize). Therefore, reverse genetics represents an essential approach to conduct gene functional studies in wheat and to facilitate genetic improvement.

Reverse genetic studies have become popular in wheat and relied on homologous cloning and genome-wide analysis (GWA) of a given gene family based on the known functional and bioinformatics knowledge mostly achieved in model species (*e.g.*, Arabidopsis and rice). Such studies usually adopt the following workflow: **(1)** Identification of gene family members from the wheat genome based on sequence similarity and/or the presence of certain protein domains; **(2)** Sequence analysis to indicate possible functional conservation of the members between the model species and wheat; **(3)** Phylogenetic analysis to assign family members to different groups, linking the members with certain homologs in the model species; **(4)** Expression analysis using RNA-seq data to predict the biological functions of given family members. These analyses greatly contribute to the translation of established gene functional knowledge from the model species to wheat. Follow-up experiments (such as validating the predicted gene functions using transgenic plants or mutants) are needed to confirm the gene functions in wheat and to apply to genetic improvement and breeding. However, it is apparent that gene families usually expand in the wheat genome, and members often get lost, duplicated or functionally divergent during wheat evolution [3,9]. Thus, to accurately predict the function of candidate gene involved in agronomically important traits by genome-wide analysis (GWA) is not seemingly straight-forward in wheat. Although wheat transformation techniques with particle bombardment and *Agrobacterium* have been improved [10–14], accurate identification of the candidate genes is still one of the bottlenecks in wheat reverse genetics.

To overcome such limitations and to efficiently identify and prioritize the candidate genes controlling important agronomic traits, we propose that the genome-wide analysis should be done in a series of Triticeae species, including the hexaploid wheat and their diploid relatives, so as: **(1)** to capture the evolutionary history of the gene family; **(2)** to identify those members duplicated or functionally divergent during wheat evolution with the aids from combining multiple data sources, such as expression data, protein-DNA or protein–protein interaction data and selection footprints. We think that such gene members evolve along with the evolution of Triticeae species (even the speciation of wheat), and thus may be involved in the developmental processes, environmental adaptation and the traits particularly important to wheat. This integrative gene duplication and genome-wide analysis (herein abbreviated as iGG analysis, with the strategy depicted in Fig. 1A) will likely identify the genes deserving further experimental efforts in functional and reverse genetic studies.

In the present study, as a proof-of-concept, we identified a large gene family encoding the calcineurin B-like protein-interacting protein kinases (CIPKs) across seven monocot species and identified the Triticeae- duplicated *CIPK* genes (*TaCIPK17* copies) specifically responding to drought stress. CIPKs are plant-specific serine-threonine kinase proteins that interacts with CBL to form the CBL-CIPK protein complex, which involve in Ca^{2+} sensing and signaling pathways [15,16]. The structure of CIPK proteins is conserved and the kinase domain in the N terminal consisting of ATP binding site and activation segment with three phosphorylation switches (Thr, Ser, Tyr), the regulatory domain comprising of auto-regulatory NAF/FISL motif and the phosphatase interaction motif (PPI), and the NAF motif is the docking for CBLs [15]. The CBL proteins contain four EF-hands for binding to Ca^{2+} and a conserved PFPF/FPSF motif [17]. When Ca^{2+} signal changed by stresses, CBL binds to Ca^{2+} and subsequently interacts with CIPK to relief the autoinhibition of the catalytic kinase domain of CIPKs, and CBL-CIPK complex could active or repress downstream proteins through complex regulatory networks [17–21]. The CBL-CIPK pathway play a critical role in regulating plant growth and responding to stresses (salinity, dehydration, cold, low K^+, etc.) and other signals, such as abscisic acid (ABA) and reactive oxygen species (ROS) [15,18–22].

Here, we chose the *CIPK* family as an example to demonstrate the power of duplication-oriented genome-wide analysis for several reasons: **(1)** The family is large, usually containing ~30 genes in a diploid species [17,22]; **(2)** The functions of several CBLs and CIPKs have been well studied in Arabidopsis with extensive knowledge regarding its protein–protein interaction and regulatory mechanisms [15,18–21]; **(3)** The reference genomes of several Triticeae species have been available recently owing to the advances in genomics and bioinformatics, laying the foundation for comparative genome-wide analysis among these species [23–29]; **(4)** CBL and CIPK represent a pair of families following the gene balance hypothesis [30–32], and hence their identification may provide evolutionary insights into gene dosage balance and duplication-driven functional divergence; **(5)** The genome-wide *CIPK* identification studies have been mostly done in eudicot species [33–38],

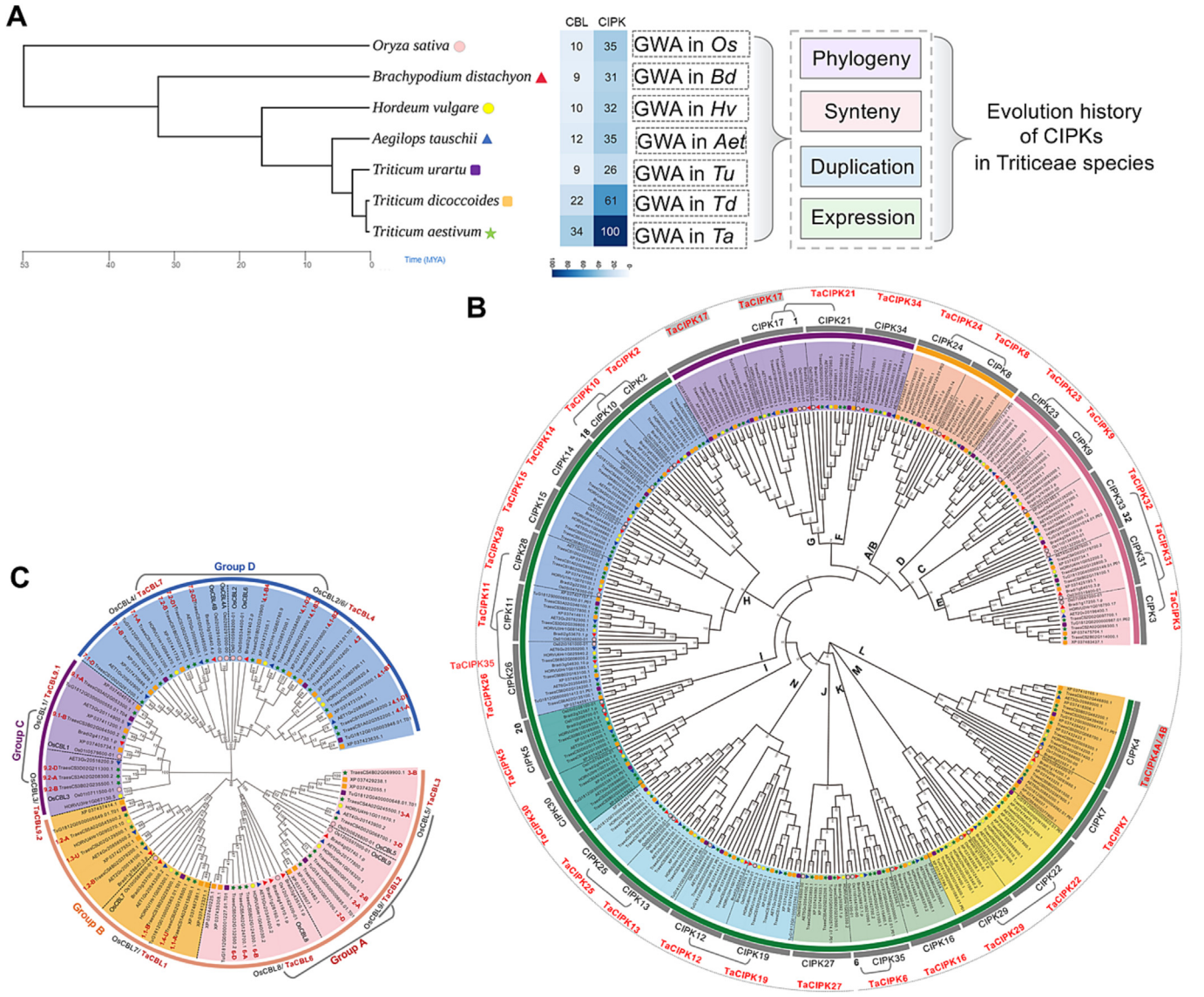


Fig. 1. IGG analysis portraits the phylogeny and evolutionary history of CBLs and CIPKs in the seven *Poaceae* species. **(A)** The workflow for the iGG analysis and the number of *CBL* and *CIPK* genes identified in each of the seven Poaceae species, indicating the expansion of *CBL* and *CIPK* families in bread wheat (Tables S2, S3). The phylogeny of the species was obtained from Timetree (http://timetree.org/). **(B)** Phylogenetic tree of the identified CIPK proteins. The 320 CIPK proteins are divided into 14 groups with the group number labeled in bold and colored in the background. The orthologous CIPKs from the seven species are grouped in each of the same branches with the rice and wheat CIPK numbers indicated in the outer circles in black and red, respectively. The *OsCIPKs* derived from WGD events are indicated with brackets. **(C)** Phylogenetic tree of the identified CBL proteins showing the four groups of CBLs (groups A, B, C and D, indicated with colored backgrounds). The names of OsCBLs and TaCBLs are labeled in black and red, respectively.

but not many reported in agronomically important grasses [39–41]. Further functional validation in both transgenic tobacco and wheat proved that *TaCIPK17* improves drought resistance. CIPKs, together with the interacting partner CBLs, belongs to one of the key protein machineries for Ca^{2+} sensing and signaling transduction [15,16].

Materials and methods

Identification of CIPK and CBL genes

To identify the *CIPK* and *CBL* genes, the genome assemblies and corresponding annotated protein-coding genes and proteins of *Aegilops tauschii*, *Oryza sativa*, *Triticum aestivum*, and *Triticum urartu*, respectively, were obtained from the EnsemblPlant database (https://plants.ensembl.org/index.html) (data accessed by 23 July 2021) [42]. The genomes and gene annotations of *Brachypodium distachyon* and *Hordeum vulgare*, respectively, were downloaded from the Phytozome v12 (https://phytozome-next.jgi.doe.gov/) (data accessed by 23 July 2021) [43], while the published genome of *Triticum dicoccoide* was obtained from NCBI [23]. The Hidden Markov Model (HMM) profiles of the CIPK protein kinase domain (PF00069) (PKC hereafter), NAF domain (PF03822), EF-hand (PF00036, PF13202, PF13833, PF13499) were downloaded from Pfam (https://pfam.xfam.org/). The reported CIPKs and CBLs from *Arabidopsis* and rice were used as inquiries of BLAST search to retrieve potential CIPK (identity >90%, E-value < $1E^{-5}$) [32] and CBL (E-value < $1E^{-10}$) -encoding genes in the genomes of *B. distachyon*, *H. vulgare*, *A. tauschii*, *T. urartu*, *T. dicoccoides*, and *T. aestivum* [17,22,41], followed by the search of protein sequences with particular domain and/or motif using HMMER v3.0 (Fig. S1).

CIPK proteins should have the PKC and NAF domains, while CBLs should have the EF hands confirmed with the NCBI Conserved Domain Search. The pI and molecular weight (MW) of CIPKs and CBLs were predicted by ExPASy (https://www.expasy.org).

The identification of homeologs between the subgenomes of *T. aestivum* was based on previous studies [3,44]. The rice *CIPK*s and *CBL*s that were evolved from whole genome duplication (WGD) events were identified according to rice comparative genomic analyses [45,46]. The genes recently duplicated in the Triticeae species have been identified and categorized into different types of small-scale duplication (SSD) events by Wang et al. [3]. Briefly, the paralogous gene pairs were firstly identified with the best-reciprocal blast approach (BLASTP, using the parameter "-outfmt 6 -evalue 1e-5") within several diploid monocot genomes (*i.e.*, maize, sorghum, rice, barley, *Ae. tauschii* and *T. urartu*). Self-genomic comparisons were performed to identify syntenic blocks (containing syntenic genes) with MCScanX (using the parameter "-e 1e-5 –m 25 –w 5") as having arisen from WGD. The remaining duplicated genes were classified into types of SSD (tandem duplicates (TD), proximal duplicates (PD) and dispersed duplicates (DD)), which were further supported by the synonymous substitution (K_s) analysis. This analysis has identified numerous SSD gene pairs arisen in the Triceae species, termed as a recent burst of gene duplications (RBGD) [3].

Sequence analysis and phylogenetic analysis

Multiple sequence alignment of full-length protein sequences of CIPK and CBL were performed with MUSCLE. Only the protein sequences deduced from primary transcripts of *CIPK*s or *CBL*s were used. Phylogenetic trees were constructed by the maximum-likelihood (ML) method with 1000 bootstrap replicates using MEGA-X software [47]. Phylogenetic trees were visualized and edited by EvolView (https://www.evolgenius.info/evolview-v2/). Conservative motifs of TaCIPKs were analyzed by the MEME tool (https://meme-suite.org/meme/tools/meme) [48]. The percentages of sequences identity between the CIPK17 proteins was calculated by using the Clustal-Omega tool on the EMDL-EBI database (https://www.ebi.ac.uk/Tools/msa/clustalo/), which does not only provide sequence alignment results, but also produce the 'Percent Identity Matrix' showing the percent of sequence identity between each pairwise comparison. The sequence identity matrix produced by the Clustal-Omega tool between the CIPK17 proteins from rice, barley, *Ae. tauschii*, *T. urartu*, *T. dicoccoides*, and *T. aestivum* was used to infer the ancient and newly duplicated copies at the *CIPK17* locus and shown in heat map.

Chromosome Location, collinearity analysis and gene structure analysis

The chromosomal location of *TaCIPKs* were visualized by TBtools [49]. Inter-species gene synteny analysis of *CIPK*s was calculated using MCScanX with the homoeologous gene information obtained elsewhere [44]. MicroCollinearity of the *CIPK17* loci among the monocot species was analyzed and visualized by TGT website (https://wheat.cau.edu.cn/TGT/) [50].

Analysis of gene expression patterns

Five publicly available RNA-seq data sets were used to capture the expression profiles of TaCBLs and TaCIPKs across developmental tissues and stages as well as the responses to biotic and abiotic stresses, including drought, heat, cold, stripe rust (Str) and powdery mildew (Pwd) (in Supplementary Method) [51–55]. The RNA-seq reads were mapped to IWGSC RefSeq v1.0 using STAR v2.7.3, and the uniquely mapped reads were used for identifying differentially expressed genes within each dataset by DEseq (q values <0.05, $|\log_2(\text{FoldChange})| > 1$). The gene expression levels were quantified in transcripts per million (TPM) at the gene level with the Salmon package [55]. The expression profiles were visualized in $\log_2(\text{TPM} + 0.1)$ in heat maps. The homoeologous syntenic genes from the A, B, and D subgenomes, respectively, (also known as triads) of *TaCBL*s and *TaCIPK*s were used for the analysis of homoeolog expression bias (HEB) based on the previous method [44].

Protein-protein interaction assays

Y2H-based interactions between multiple TaCIPKs and TaCBLs were performed with the Matchmaker Gold Yeast Two-Hybrid System (Clontech). Multiple *TaCIPK* genes (*e.g.*, *TaCIPK2*, *9*, *12*, *14*, *16*, *17*, *23*, *29* and *30*) were cloned to examine their interactions with several CBLs representing each of the triads (*TaCBL1*, *2*, *3*, *4*, *6*, *7*, and *9*). Recombinant plasmids of pGADT7-CIPK and pGBKT7-CBL were co-transformed into yeast strain AH109. Positive transformants were spotted on medium DDO (SD/-Trp/-Leu), TDO (SD/-Trp/-Leu/-His), or QDO (SD/-Trp/-Leu/-His/-Ade). Together with the previous study from our group [56], TaCBL-TaCIPK interactions were comprehensively analyzed, containing 27 TaCIPKs and 7 TaCBLs with the interaction relationships simplified to a binary plot drawn by ChiPlot. The full-length of TaCIPK17-A2 and three truncated forms (*i.e.*, M1, M2, and M3) were cloned to examine the protein regions affecting CBL-CIPK interaction. In addition, the TaCBL1.1, 2, 3, 4.1, 6, 7.1, and 9.1, as well as TaCIPK17 and its truncated forms (M1, M2 and M3) were tested for the auto- activity. For bimolecular fluorescence complementation (BiFC), *TaCIPK17* was cloned into the SpYNE vector, while *TaCBL1, 2, 3*, and *6* were cloned into the SpYCE vector. Onion epidermal cells were co-infiltrated with mixtures of an equal amount of SpYNE/SpYCE culture. YFP signal were checked in the epidermal cells from the infiltrated onion tissue after 48 h with fluorescence microscopy (OLYMPUS LX71, Japan).

RNA isolation and quantitative real-time PCR (qRT-PCR)

Total RNA was extracted from the leaf tissues of wheat or tobacco plants with the plant total RNA kit (Zomanbio, China) and the cDNA was reverse transcribed with All-in-One RT SuperMix (Vazyme, China). The qRT-PCR was performed with qPCR SYBR Green Master Mix (Vazyme, China) and *NtGAPDH* and *TaActin1* as the internal reference gene in tobacco and wheat, respectively. Quantitative-PCR data were analyzed by using the $2^{-\Delta\Delta CT}$ method and the statistical differences were determined with the Student's *t* test using results from three biological replicates with technical duplicates. All primers are provided in Table S1.

Plant transformations of wheat and tobacco

For the particle bombardment-mediated wheat transformation, *TaCIPK17-A2* was cloned into the pAHC25 plasmid, and the recombinant plasmid was transformed into immature embryo-derived calli of wheat cultivar L88-31 as described elsewhere [14]. The positive transgenic events were identified from the transgenic plants survived from selection followed by PCR examination of the presence and expression of the transgene for the consecutive generations. PCR-positive transgenic events were propagated and selected to obtain non-segregant lines [57]. For PCR examination of the foreign gene, DNA was extracted from leaf tissues with the CTAB method and the PCR were performed with the primers provided in Table S1.

The gene sequence of *TaCIPK17-A2* was cloned into plasmid of *pBI121*. The construct was transformed into tobacco using *Agrobacterium*-mediated leaf discs protocol [58]. The transgenic tobacco

seeds were selected with 50 mg/L kanamycin on Murashige-Skoog (MS) medium and DNA extracted from tobacco leaves was used to identify transgenic positive tobaccos. The detection primers were listed in Table S1.

Drought phenotype analysis of transgenic plants

The wheat seeds of *TaCIPK17-A2* over-expression transgenic lines and WT were germinated in the dark under 23 ℃, then moved to light for a week after the buds grew, and vernalized for 10 days under 4 ℃, then the seedings were planted in soil mix (nutrient soil: vermiculite = 3:1). The wheat plants of transgenic and WT grew for a month in the greenhouse, and subsequently subjected to drought for two weeks. Drought-treated plants were re-watered for one week.

The tobacco plants of three independent transgenic lines and control (WT and VC) were germinated on MS medium for five days and then the tobacco plants were planted in vermiculite under well-watered condition. The four-week-old tobacco plants were treated under drought for three weeks and re-watered for two weeks. The survival rates of tobacco plants were calculated after recovery.

Measurement of physiological indices

To measure the physiological indicators of transgenic wheat plants, the four-week-old plants were treated under drought for two weeks. The leaves under normal growth conditions of WT and three *TaCIPK17-A2* overexpression lines were sampled to measure physiological indices, including the contents of H_2O_2, proline (PRO), malondialdehyde (MDA) and soluble sugar, the activities of catalase (CAT), total superoxide dismutase (T-SOD), peroxidase (POD) and restraining ability to hydroxyl free radicals (RAHFR) [59].

To measure the physiological indicators of transgenic tobacco plants, the plants were grown in vermiculite for four weeks and treated under drought condition for three weeks, and then re-watered for seven days. The tobacco leaves of WT, VC and transgenic lines at four stages, including before drought, seven days after drought, 14 days after drought and re-watered after seven days, were sampled to measure physiological indices, including the contents of H_2O_2, MDA and soluble sugar, and the activities of CAT, T-SOD, POD, and RAHFR. The physiological indices were measured by using the corresponding detection kits (NJJCBIO (Nanjing Jiancheng Bioengineering Institute) Ltd., Nanjing, China) with the MultiSkan GO 1510 microplate reader (Thermo Fisher Scientific Ltd., U.S.A.) [24,60].

Oxidative stress treatment of tobacco leaves

The four-week-old tobacco leaf discs of transgenic and control lines were treated by 20 μM and 50 μM MV. Chlorophyll content of leaf discs under control and MV treatment conditions were measured by colorimetric method. The 0.2-g leaf discs were soaked in buffer (Acetone: ethyl alcohol: ddH_2O = 4.5: 4.5:1) in the dark until the leaf discs turned white. The absorbance of each sample at 645 nm and 663 nm were measured using a spectrophotometer and calculated the chlorophyll content.

Statistical analysis

All experiments were performed in three biological replicates in the present study and the data were analyzed with the SPSS software. The statistical differences of gene expression or physiological parameters between wildtype and transgenic lines were calculated by Student's *t* test ($^*P < 0.05$, $^{**}P < 0.01$).

Results

Genome-wide analysis of CBLs and CIPKs

The *CIPK* gene family was chosen to exemplify the integrative gene duplication and genome-wide analysis (iGG analysis) based on its family size, biological importance (in both signaling transduction and as a classical example for evaluating the gene balance hypothesis) and the lack of comprehensive analysis in the post-genomic era of Triticeae species. To capture the evolution of *CIPK*s during wheat speciation, *CIPK* genes have been genome-wide identified in seven species, including five in the Triticeae Tribe, yielding 320 *CIPK*s and 100 *TaCIPK*s (Fig. 1A-B, S2). Phylogenetic analysis separated CIPKs into 13 groups (*i.e.*, A/B, C, D, E, F, G, H, I, J, K, L, M, and N), and OsCIPKs and Triticeae CIPKs present in each of the 13 phylogenetic groups. The ancestor of major cereals crops has experienced multiple whole-genome duplication (WGD) events [45,46]. Among the 35 identified *OsCIPK*s, we found ten pairs of *OsCIPK*s derived from the WGD events, including *OsCIPK1*/*17*/*21*, *OsCIPK2*/*10*/*18*, *OsCIPK3*/*31*/*33*, *OsCIPK6*/*35*, *OsCIPK8*/*24*, *OsCIPK9*/*23*, *OsCIPK11*/*26*/*28*, *OsCIPK12*/*19*, *OsCIPK13*/*25*, *OsCIPK22*/*29* [18,33]. Among these WGD-derived *CIPK*s, a few may have been lost during Triticeae evolution, such as *OsCIPK1*, *OsCIPK6* and *OsCIPK18*. For the remaining *CIPK* genes, phylogenetic clustering and micro-syntenic results matched well and therefore the orthologous relationship of *CIPK*s between rice and wheat was established here. Most *TaCIPK*s contain three homoeologous triads except for *TaCIPK10* and *TaCIPK35* (Table S2). Particularly, *CIPK17* and *CIPK22* have been expanded by tandem duplication, while *CIPK4* experienced segmental duplication (Fig. 1B; Table S2).

In addition, 106 CBLs, which interact with CIPK proteins as one of the key modules of calcium signaling transduction in plants, have been identified in the seven species (Fig. 1C; Table S3). The identified CBLs fall into four phylogenetic groups (*i.e.*, group A, B, C and D) with three sets of CBLs raised from ancient WGD events, *OsCBL1* (*TaCBL9.1*)-*OsCBL3* (*TaCBL9.2*), *OsCBL2* (*TaCBL4*) -*OsCBL4* (*TaCBL7*), and *OsCBL5* (*TaCBL3*) -*OsCBL9* (*TaCBL2*) - *OsCBL8* (*TaCBL6*). In wheat, *TaCBL1*, *TaCBL4* and *TaCBL7* have evolved into multiple copies by tandem duplication.

During genome evolution, duplicated genes may become divergent and finally gain different functions, which is known as sub- and neo-functionalization. Thus, we analyzed the motif composition of TaCIPK proteins to address whether the polyploidization and tandem duplication are associated with different motif composition. Indeed, we observed the differences in motif composition in a number of TaCIPK triads (2, 4B, 19, 25, 27, 34) (Fig. S3, S4). Also, tandem duplicated *TaCIPK* genes encodes the proteins with motif-composition differences, as demonstrated in proteins encoded by the copies of *TaCIPK17* and *TaCIPK22*, respectively. Notably, many TaCIPK proteins differ in their C-terminal including the NAF motif (corresponding to the motifs 9 and 10 in Fig. S4), which has been known to affect the specificity of CBL-CIPK interactions [56], implying potential functional relevance of these differences in motif composition. Taken together, wheat allopolyploidization and the recent gene duplication events among the Triticeae Tribe account for the expansion of TaCIPK and TaCBL families.

Expression patterns of TaCBLs and TaCIPKs

We profiled the expression patterns of *TaCBL*s and *TaCIPK*s by using publicly available RNA-seq datasets in wheat to obtain insights into the expression and potential functions of *TaCIPK*s (Supplementary Method). Firstly, several *TaCIPK*s (*i.e.*, *TaCIPK2*, *9*, *23*, *31*, *32*) exhibited widely expressed patterns, whereas many

other *TaCIPKs* (*e.g.*, *TaCIPK8*, *25*, *29*) are preferentially expressed in certain tissues or developmental stages, or specifically expressed at a few stages or under certain conditions (*e.g.*, *TaCIPK4*, *5*, *14*, *19*, *30*) (Fig. 2A). These expression patterns suggest that *TaCIPK2*, *9*, *23*, *31*, *32* may have pivotal roles across plant developmental processes and abiotic stress response, while other *TaCIPKs* likely are involved in particular functions. Second, most WGD-derived *TaCIPK* pairs show distinct expression patterns across developmental stages and/or in response to biotic and abiotic stresses, indicating that these *TaCIPK* pairs could be functionally divergent due to distinct expression profiles. For example, *TaCIPK31* and *TaCIPK32* were highly expressed in up-ground tissues of wheat, while *TaCIPK3* triads were expressed in leaves and spikes. *TaCIPK2* triads were ubiquitously expressed and has the highest expression level among all *TaCIPKs*, whereas *TaCIPK10*'s expression was primarily induced by stresses (Fig. 2A). Third, duplicated *TaCIPK* genes tend to show lower expression levels and specific patterns compared to the non-duplicated *TaCIPK* triads, and the expression profiles among the gene duplicates of a given *TaCIPK* are quite different. For instance, *TaCIPK4* was expressed in stem tissue and up-regulated by cold stress, while the expression of *TaCIPK4b* was related to biotic stresses.

The trends of expression divergence between *TaCBLs* are similar to those seen in *TaCIPKs*: **(1)** some *TaCBLs* (*TaCBL2* and *TaCBL6*) were highly and widely expressed, while the remaining *TaCBLs* exhibited the expression patterns specific to certain tissues, stages

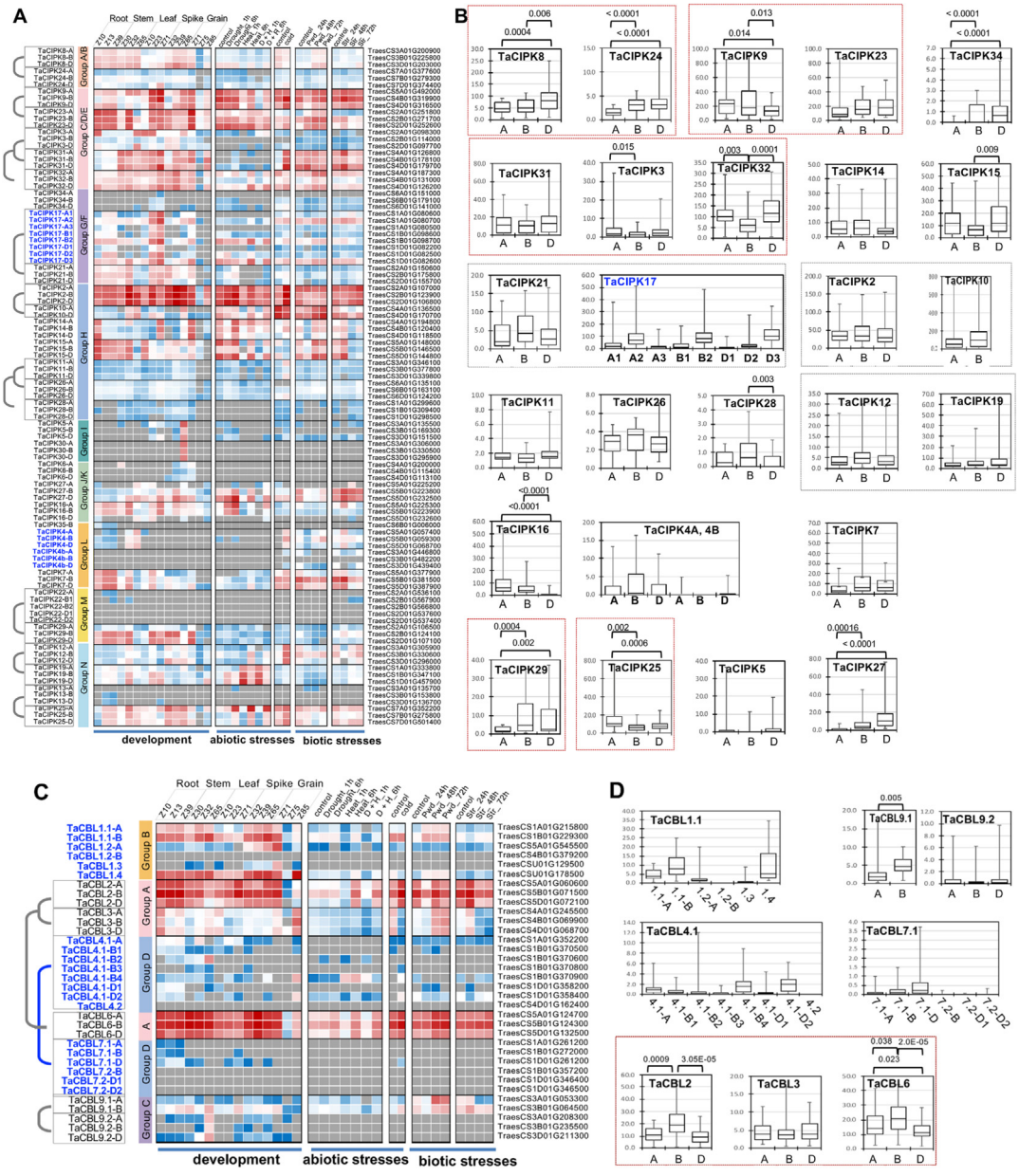


Fig. 2. Expression profiles of TaCIPKs and TaCBLs and the analysis of homoeolog expression bias (HEB). (A, C) The expression profiles of TaCIPKs (A) and TaCBLs (C) across developmental stages in five tissues (root, stem, leaf, spike and grain) and under biotic and abiotic stresses shown in the heatmap [2,49,52,53]. For Fig. 2A and 2C, colors indicate gene expression levels (log_2(TPM + 0.1) with grey meaning not expressed; phylogenetic groups are indicated. TaCIPKs, TaCBLs and their corresponding geneIDs are labeled with the tandem duplicated gene names shown in blue; the WGD-derived pairs of TaCIPKs or TaCBLs are indicated in brackets. (B, D) Boxplots showing HEB for some TaCIPKs (B) and TaCBLs (D), with the A, B, and D on the x-axis representing the A, B and D sub-genomes of T. aestivum, respectively. Significant differences in the expression between homoeologous triads (Student's *t*-test) are indicated with P-values. The WGD-derived TaCIPK or TaCBL gene pairs are indicated in boxes with the gene pair having HEB showing in red boxes.

and stress treatments (*e.g.*, *TaCBL1*, *3*, *4*, *9*) (Fig. 2C); **(2)** WGD-derived *TaCBL* pairs exhibited apparently different expression profiles, such as *TaCBL2* and *TaCBL3*; **(3)** duplicated copies of *TaCBL* have generally lowered expression levels, with specific expression profiles seen between the duplicates (for example, the copies of *TaCBL4.1*).

Homoeolog expression bias (HEB) is a feature known for the gene expression in allopolyploid species and has been found to affect the expression of numerous genes in bread wheat [44]. Many *TaCIPK* triads derived from ancestral WGD events have HEB phenomenon (Fig. 2B), such as *TaCIPK3*, *TaCIPK8*, *TaCIPK9*, *TaCIPK24*, *TaCIPK25* and *TaCIPK29*. Within a pair of WGD-derived *TaCIPKs*, they have distinct HEB categories: for instance, *TaCIPK9* belongs to the D-suppressed HEB, while *TaCIPK23* has a balanced HEB pattern. We extended our HEB analysis to *TaCBL* genes (Fig. 2D). The WGD-derived *TaCBL* pairs also differed in the HEB categories: *TaCBL2* has a B-dominant HEB pattern, while *TaCBL3* expression is balanced. Overall, these results support that HEB contributes to differentiated expression of *TaCBLs* and *TaCIPKs*.

Protein-protein interactions between TaCBLs and TaCIPKs

Since the CBL-CIPK protein complex functions in sensing Ca^{2+} and kinase-mediated signaling transduction responding to various biotic and abiotic stresses [15], information regarding TaCBL-TaCIPK interaction is key to understand their biological functions. Our group previously reported a number of TaCBL-TaCIPK interactions determined by the yeast-two-hybrid (Y2H) method based on the limited *TaCIPKs* identified from the early version of wheat genome draft sequence [51,56]. The present study allows a more com-

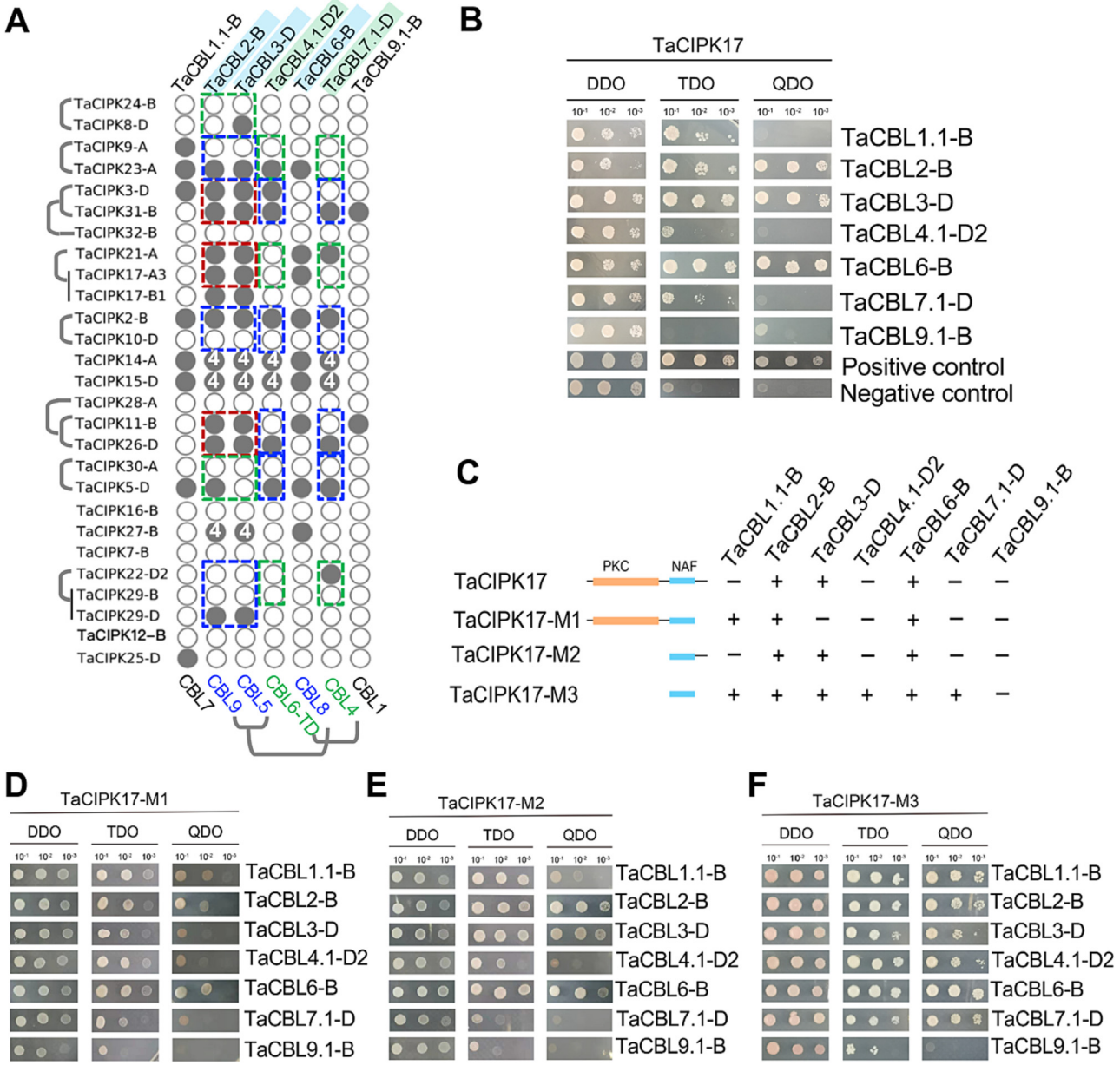
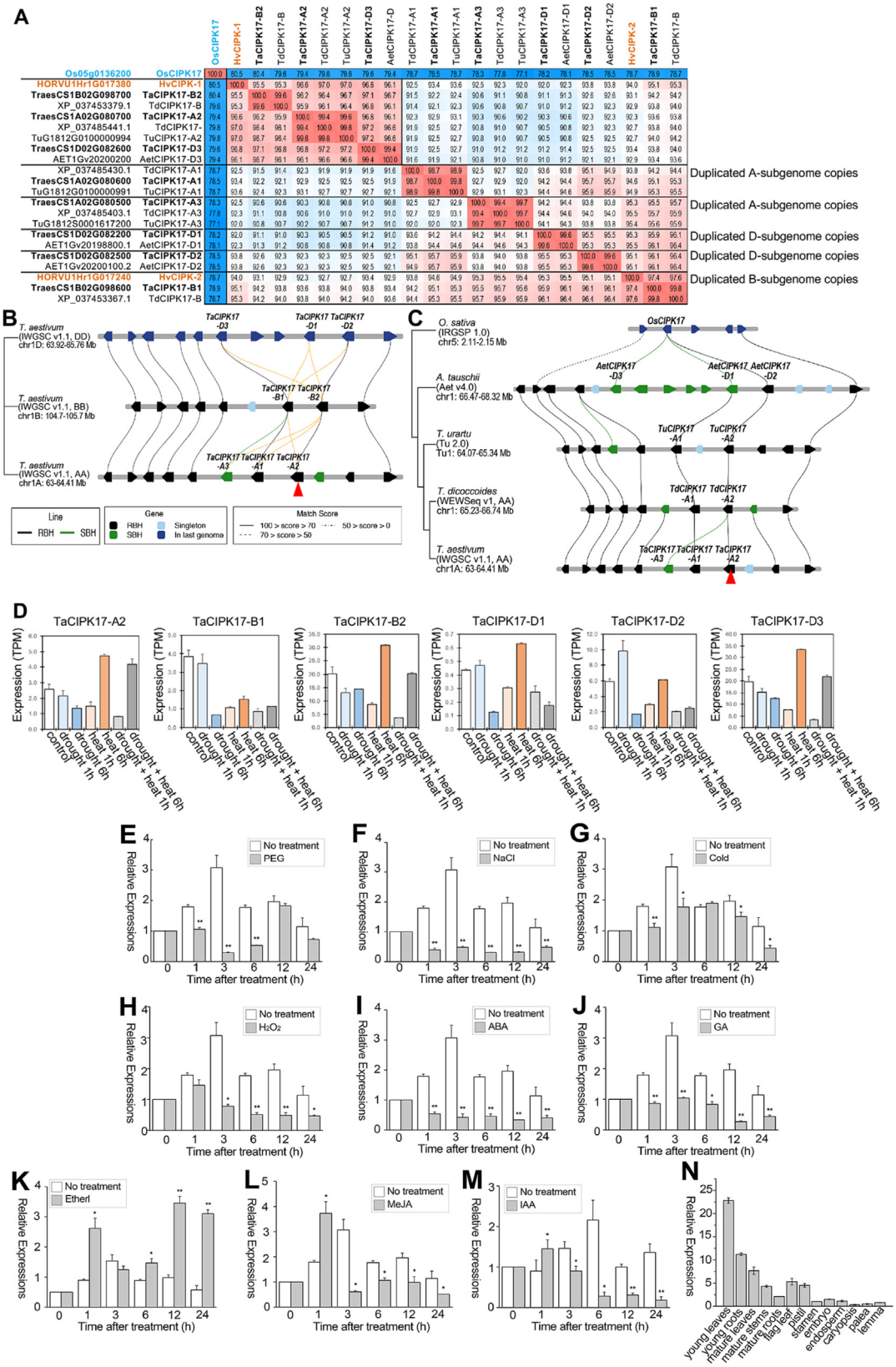


Fig. 3. Evolutionary history of the TaCIPKs and TaCBLs and their interaction patterns. **(A)** The binary plot indicates the interaction patterns between TaCIPKs and TaCBLs (grey circles meaning interactions). *TaCIPKs* derived from the WGD events during the ancestors of major cereal crops are labeled with brackets, while the WGD-derived *TaCBLs* are labeled with blue or green backgrounds. OsCBL orthologous to the corresponding TaCBLs are indicated at the bottom of the binary plot. **(B)** Y2H interactions between the full-length TaCIPK17 and TaCBLs. **(C)** Summary of the interaction patterns between TaCBLs, full-length TaCIPK17 and several truncated TaCIPK17s (namely, TaCIPK17-M1, TaCIPK17-M2, and TaCIPK17-M3). The Y2H-based interaction results between TaCBLs and TaCIPK17-M1 **(D)**, TaCIPK17-M2 **(E)**, and TaCIPK17-M3 **(F)**, respectively.


A
Duplicated A-subgenome copies
Duplicated A-subgenome copies
Duplicated D-subgenome copies
Duplicated D-subgenome copies
Duplicated B-subgenome copies
B
C
D
TaCIPK17-A2
TaCIPK17-B1
TaCIPK17-B2
TaCIPK17-D1
TaCIPK17-D2
TaCIPK17-D3
Expression (TPM)
control
drought 1h
drought 6h
heat 1h
heat 6h
drought + heat 1h
drought + heat 6h
E
No treatment
PEG
F
NaCl
G
Cold
H
H2O2
I
ABA
J
GA
K
Ethrel
L
MeJA
M
IAA
N
Relative Expressions
Time after treatment (h)
young leaves
young roots
mature leaves
mature stems
mature roots
flag leaf
pistil
stamen
embryo
endosperm
caryopsis
palea
lemma

prehensive *TaCIPK* identification and TaCBL-TaCIPK interaction analysis. Here, we cloned seven *TaCBLs*, representing each of the *TaCBL* groups, for the Y2H experiments to study protein–protein interactions (PPIs) between TaCBLs and TaCIPKs (Fig. 3A, S5A). In most cases, one TaCIPK interact with multiple TaCBLs. Importantly, our cloned genes include many WGD-derived pairs of TaCBLs and TaCIPKs, allowing us to interrogate evolutionary insights into the dosage balance between CBLs and CIPKs. Four scenarios of CBL-CIPK stoichiometric relationship were supported by our results of PPI and expression data. If the duplicated copies of both CIPK and CBL have been retained, three different CBL-CIPK interaction patterns could be possible: **(1)** scenario 1, both duplicated CBLs interact with both duplicated CIPKs (2 CBLs: 2 CIPKs, indicated in red boxes, Fig. 3A); **(2)** scenario 2, only one of the duplicated CBLs interacts with both duplicated CIPKs (1 CBLs: 2 CIPKs, indicated in blue boxes, Fig. 3A); **(3)** scenario 3, only one of the duplicated CBLs interacts with one of the CIPK duplicates (1 CBLs: 1 CIPKs, indicated in green boxes, Fig. 3A). Besides, in the 4th scenario, both duplicated CBLs interact with a non-duplicated CIPK (indicated in white letter "4" in Fig. 3A), resulting in 2 CBLs: 1 CIPKs.

We particularly investigated the possible principles contributing to the gene dosage balance in scenarios 2 and 4, since the gene dosage may be imbalanced in these two scenarios. For example, TaCIPK23 interacted with both TaCBL2 and TaCBL3, but TaCBL3 showed a much lower and specific expression patterns when compared to TaCBL2. Similarly, either TaCIPK2 or TaCIPK29 interacted with both TaCBL2 and TaCBL3. Interestingly, *TaCIPK2*, *TaCIPK23* and *TaCIPK29* were generally highly expressed in the examined tissues (including the root, stem, leaf, spike and grain) and stages, but were expressed at lower levels at the beginning of first leaf growth (Z10), the stem tissue during anthesis (Z65), and grains during milk and dough stages (Z65, Z75; Fig. 2A). By contrast, the interacting CBLs (TaCBL2 and TaCBL3) have obtained contrasting expression profiles to potentially balance the gene dosage. For the other scenario-2 CBL-CIPK interactions (*e.g.*, TaCIPK2–TaCBL4/7, TaCIPK26–TaCBL4/7, TaCIPK5–TaCBL4/7), the duplicated CBLs (TaCBL4 and TaCBL7) have very specific expression patterns (Fig. 2C), helped to maintain 1 CBL: 1 CIPK balance. The scenario-4 CBL-CIPK interactions include TaCIPK14–TaCBL2/3, TaCIPK14–TaCBL4/7, TaCIPK15–TaCBL2/3, TaCIPK15–TaCBL4/7, and TaCIPK27–TaCBL2/3. For all of these theoretically unbalance CBL-CIPK scenarios, the divergently expression between duplicated copies of TaCBLs appears to help maintain the dosage balance of CBL and CIPK.

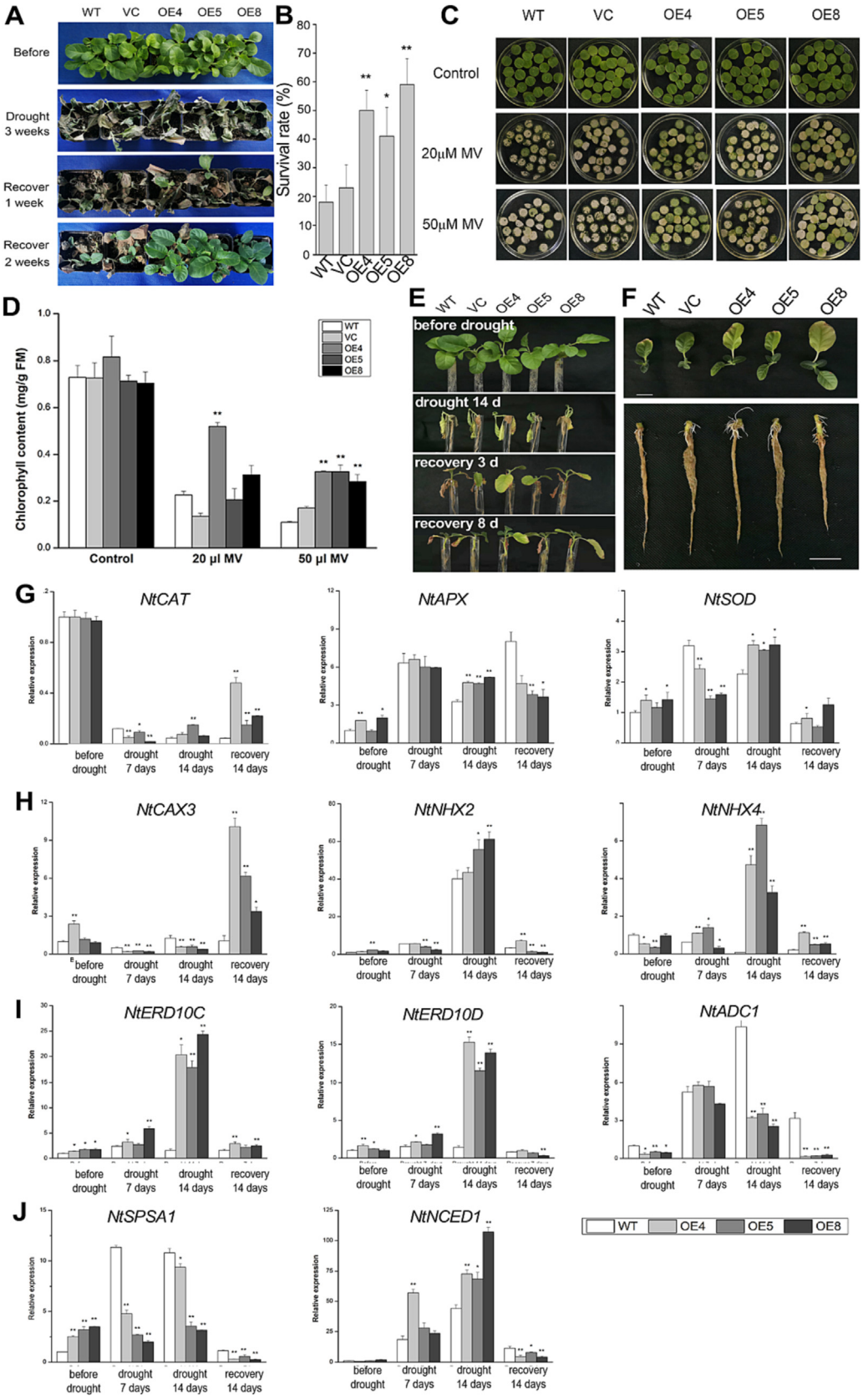
Tandemly duplicated TaCIPK17s potentially are involved in response to and regulation of drought stress

The genomes of the Triticeae species harbors many gene raised from a recent burst of gene duplication event (RBGD), mostly caused by tandem and dispersed duplications. These RBGD genes in wheat are enriched in certain biological functions, suggesting that RBGD serve as a potential approach to provide evolutionary innovations and is likely associated with important agronomic traits in wheat [3]. Leveraging the previously reported RBGD information, we found that *TaCIPK17-A2*, *TaCIPK17-B2* and *TaCIPK17-D3* are probably the ancient copies among the *TaCIPK17* duplicates, which is consistent with our sequence analysis and synteny results (Fig. 4A-C). Interestingly, the ancient copies of *TaCIPK17* have higher expression levels during development (Fig. 1A) and abiotic stress treatments (Fig. 4D) when compared to the newly duplicated copies (*e.g.*, *TaCIPK17-A1*, *-B1*, *-D1*, and *-D2*). The ancient *TaCIPK17* copies (*i.e.*, A2, B2, D3) were down-regulated by the drought treatment but up-regulated by heat stress, while *TaCIPK17-D2* was induced by drought stress (Fig. 4D). Besides, *TaCIPK17-B1* was repressed by both drought and heat treatments (Fig. 4D). This result demonstrated that the newly duplicated *TaCIPK* genes have adopted distinct expression patterns, suggesting these *TaCIPK*s may have new functions in wheat, possibly in response to the abiotic stresses. Furthermore, we searched published expression data under drought treatments and found that OsCIPK17 does not respond to drought stress [61–64], suggesting that *TaCIPK17* might be evolved to be drought-responsive after the divergence between rice and wheat ancestors. Thus, our duplication-oriented analysis of *CIPK* genes highlights *TaCIPK17* copies as the candidates for drought-stress regulation in wheat.

Then, we selected *TaCIPK17-A2* as an example of the ancient copies and confirmed the expression changes under abiotic-stress treatments. Indeed, PEG treatment triggered the down-regulation of *TaCIPK17-A2* (Fig. 4E). Expression analysis further showed that *TaCIPK17-A2* was up-regulated by cold treatment but down-regulated by salt (NaCl) or H_2O_2 treatment (Fig. 4F-G). In addition, *TaCIPK17-A2* expression was regulated by several phytohormones: ethylene (Eth) and MeJA induced the expression of *TaCIPK17-A2*, while ABA and GA suppressed its expression (Fig. 4). Notably, time-series expression analysis demonstrated that Eth and MeJA could rapidly induce *TaCIPK17-A2* expression in one hour, whereas its most dramatic down-regulation was detected after six hours of the PEG or ABA treatment, implying that the PEG- or drought-mediated expression changes of *TaCIPK17-A2* might be ABA dependent. Besides, our qRT-PCR result validated that *TaCIPK17-A2* was highly expressed in leaves and roots but was expressed at a much lower level in some flower organs and seed tissues, such as the stamen, palea, lemma, embryo and endosperm (Fig. 4N).

In addition, we examined the TaCBLs that could interact with TaCIPK17-A2. Y2H results showed that full-length of TaCIPK17-A2 interact with TaCBL2, TaCBL3, and TaCBL6 (Fig. 3B). TaCIPK17-A2 and its three truncated forms (M1, M2, and M3) and several TaCBLs (*i.e.*, TaCBL1.1, 2, 3, 4.1, 6, 7.1, and 9.1) have been checked for their auto-activation ability, and our results showed that they did not exhibit any auto-activities (Fig. S5B). Detailed Y2H analysis demonstrated that the C-terminal regulatory region of TaCIPK17-A2, but not the N-terminal PKC domain, is key to the PPI specificity with TaCBLs (Fig. 3C-F).The interactions between TaCIPK17 and

Fig. 4. Tandemly duplicated copies of *TaCIPK17* differ in the expression levels and responses to abiotic-stress treatments. **(A)** Analysis of CIPK17 protein identity suggests that *TaCIPK17-A2*, *TaCIPK17-B2* and *TaCIPK17-D3* are probably the ancient copies, from which tandem duplication events occurs in the diploid ancestors. OsCIPK17 and HvCIPK17 are highlighted in blue and orange fonts, respectively. The analysis was calculated by using the Clustal-Omega tool on the EMDL-EBI database (https://www.ebi.ac.uk/Tools/msa/clustalo/) and shown in heat map (blue indicates low sequence identity and red indicates high sequence identity for each pair of CIPK proteins). **(B)** Syntenic analysis of the *TaCIPK17* locus between the wheat A, B, and D subgenomes. **(C)** Syntenic analysis of the *TaCIPK17-A* locus between rice, and other diploid and tetraploid wheat relatives. The figure legends are shared between Fig. B and C with *TaCIPK17-A2* indicated by red arrowheads. **(D)** Bar plots show the distinct expression changes in response to drought and/or heat treatments between *TaCIPK17-A2*/*TaCIPK17-B2*/*TaCIPK17-D3*, *TaCIPK17-B1*, and *TaCIPK17-D1*. **(E-M)** The expression changes of *TaCIPK17-A2* in response to various abiotic-stress and phytohormone treatments. The treatments include four abiotic stresses (20% PEG6000 **(E)**, 200 mM NaCl **(F)**, cold **(G)**, 10% H_2O_2 **(H)**) and five phytohormones (100 μM ABA **(I)**, 100 μM GA **(J)**, 100 μM etherl **(K)**, 100 μM MeJA **(L)**, 100 μM IAA **(M)**). For each treatment, 10-day-old wheat seedlings were treated and collected at 0, 1, 3, 6, 12, and 24 h after the treatment. Expression levels were determined by qRT-PCR with three biological replicates and are shown as means ± SEM. Within each treatment, significant differences in the expression level were calculated with Student's *t*-test by comparing the treated and corresponding untreated (control) samples ($^{*}P < 0.05$, $^{**}P < 0.01$).


A
WT VC OE4 OE5 OE8
Before
Drought 3 weeks
Recover 1 week
Recover 2 weeks
B
Survival rate (%)
WT VC OE4 OE5 OE8
C
WT VC OE4 OE5 OE8
Control
20μM MV
50μM MV
D
Chlorophyll content (mg/g FM)
WT VC OE4 OE5 OE8
Control
20 μl MV
50 μl MV
E
WT VC OE4 OE5 OE8
before drought
drought 14 d
recovery 3 d
recovery 8 d
F
WT VC OE4 OE5 OE8
G
NtCAT
NtAPX
NtSOD
Relative expression
before drought
drought 7 days
drought 14 days
recovery 14 days
H
NtCAX3
NtNHX2
NtNHX4
I
NtERD10C
NtERD10D
NtADC1
J
NtSPSA1
NtNCED1
WT OE4 OE5 OE8

TaCBL2/3/6 were also confirmed by the bimolecular fluorescence complementation (BiFC) analysis (Fig. S5C). Interestingly, TaCBL1 is highly expressed in multiple tissues and stages with apparent up-regulation upon the powdery mildew infection, while another TaCIPK17-interacting partner TaCBL3 shows relatively specific expression in roots and grains with distinct responses to biotic stresses (Fig. 2C), implying that interacting partner of TaCIPK17 might be of biological relevance.

Functional validation of TaCIPK17 in improving drought resistance

With the guidance of iGG analysis of *TaCIPK*s, we sought to validate the function of *TaCIPK17* in drought resistance *in planta*. We produced transgenic tobacco plants overexpressing *TaCIPK17-A2* (*TaCIPK17*-OE) and identified three independent transgenic lines (*i.e.*, OE4, OE5, OE8) with different expression levels of exogenous *TaCIPK17-A2*. OE4 had the highest expression level, followed by OE8 and OE5 (Fig. S6). Phenotypic analysis revealed that *TaCIPK17*-OE lines had significantly higher survival rates than those of the control lines (Fig. 5A, 5B). OE5 exhibited the lowest survival rates among the transgenic lines, suggesting that a certain amount of *CIPK* expression may be needed for drought resistance. Analyses of several physiological parameters showed that the ROS scavenging capacity has been enhanced in the *TaCIPK17*-OE lines: **(1)** relative hydroxyl radical scavenging ability (RAHFR) was slightly higher in the transgenic lines before drought treatment when compared to the controls; **(2)** POD activity was higher in the *TaCIPK17*-OE lines during drought and recovery stages; **(3)** CAT activity was significantly increased after seven days of recovery (Fig. S7). To further confirm the enhanced ROS scavenging capacity in the *TaCIPK17*-OE lines, methyl viologen (MV) was used to induce ROS accumulation in tobacco leaf discs (Fig. 5C). The leaf discs from *TaCIPK17*-OE lines appeared greener under MV treatment with significantly higher chlorophyll contents (Fig. 5D). Moreover, qRT-PCR analysis of several stress-related genes helped gain molecular insights into the improvement in ROS scavenging and drought recovery in tobacco: **(1)** *NtSPSA1*, responsible for sucrose metabolism, was increased in the *TaCIPK17*-OE lines but decreased during drought treatment, indicating its role in maintaining sugar homeostasis; **(2)** *NtNCED1*, encoding the key protein in ABA biosynthesis, dramatically up-regulated during drought stress in the *TaCIPK17*-OE lines, correlated with the induced expression of stress-related genes, *NtERD10C* and *NtERD10D*; **(3)** stress responsive genes encoding ion-channels were up-regulated during drought stress (*NtNHX2*/*4*) or the recovery stage (*NtCAX3*); **(4)** the genes encoding antioxidant enzymes (*NtAPX* and *NtSOD*) were also up-regulated after 14 days of drought treatment, consistent with the enhanced ROS scavenging ability. Additionally, a hydroponic culture system was employed to gain more understanding in the physiological differences in drought recovery between the *TaCIPK17*-OE and control lines. The hydroponically cultured seedlings of *TaCIPK17*-OE recovered better than the control plants, reflected in significant re-growth of leaves and lateral roots (Fig. 5E, 5F).

To test if *TaCIPK17* has the potential to genetically improve drought resistance in wheat, we generated transgenic lines of wheat overexpressing *TaCIPK17-A2* by using particle bombardment transformation (Fig. S8A). Several non-segregant lines of *TaCIPK17*-OE have been identified and qPCR analysis confirmed that the transgenic lines had significant higher expression levels of TaCIPK17 (Fig.S8B). Phenotypic analysis of the drought resistance demonstrated that four *TaCIPK17*-OE lines of wheat had better resistance to the drought treatment and recovered well. By contrast, the non-transgenic plants could not survive for two-week drought (Fig. 6A, B). Physiological indices were measured at the non-drought condition (labeled as control in Fig. 6C-J) and after two-week drought treatment to understand the physiological changes in the *TaCIPK17*-OE lines. The antioxidant capacity of non-transgenic plants was significantly decreased, including CAT, SOD, POD and RAHFR activities (Fig. 6C-G), while the proline content was drastically increased (Fig. 6H). In contrast, the SOD, POD and RAHFR activities were higher in the *TaCIPK17*-OE lines than in the non-transgenic plants, and, moreover, POD and CAT activities were clearly higher in the transgenic lines. Consistently, proline content was lower in the transgenic lines than in the control. Particularly, OE2 exhibited significantly increased antioxidant indicators and the lowest proline content. Similar, MDA content, usually reflecting membrane damages during stress conditions, was apparently lower in the transgenic lines than in the control plants (Fig. 6K). Overall, *TaCIPK17-A2* overexpression is a viable approach to increase antioxidant capacity and thus enhance drought tolerance in both wheat and tobacco plants.

Discussion

Expansion of TaCIPKs and TaCBLs and the factors contributing to the maintenance of CBL-CIPK gene dosage balance

In the present study, we demonstrate the integrative gene duplication and genome-wide analysis with a focus on the Triticeae Tribe as a proof-of-concept to facilitate rapid identification of important functional genes for wheat reverse genetics. With prior knowledge in the evolutionary history of speciation and divergence of a group of closely related species [3,45,61,65,66], we believe such duplication-oriented iGG analysis could easily identify the duplicated genes that may be associated with divergent expression and/or novel molecular networks and even traits important to a given species. In the present study, we identified 100 *TaCIPK*s and 34 *TaCBL*s (Fig. 1, S1, S2; Tables S2, S3). Owing to the high-quality wheat reference genome [2], the number of *TaCIPK*s and *TaCBL*s identified here apparently outnumbered those previously identified (79 *TaCIPK*s and 24 *TaCBL*s) by our group with the early version of wheat draft sequence [51,56]. More recently, an updated genome-wide analysis of the *TaCIPK* family identifies 123 predicted proteins encoded by 96 annotated wheat geneIDs [67], of which all geneIDs have been identified to encode TaCIPKs in our work. The study was focused on the *TaCIPK* gene expression in response to biotic and abiotic stresses and demonstrated

Fig. 5. *TaCIPK17*-overexpression improves drought recovery and ROS scavenging ability in transgenic tobacco plants. **(A)** The phenotypes of WT, VC and the TaCIPK-OE transgenic lines grown in pots after the drought treatment followed by a two-week re-watering. **(B)** The survival rates of WT, VC and the transgenic lines after re-watering. **(C)** Comparison of the phenotype of 4-week-old tobacco leaf discs between WT, VC and the transgenic lines (OE4, OE5, OE8) treated by 20 μM or 50 μM MV. **(D)** Chlorophyll contents of leaves after treatment with or without MV. **(E)** The phenotypes of WT, VC and the transgenic lines grown hydroponically after the drought treatment followed by an eight-day recovery. **(F)** *TaCIPK*-OE transgenic lines showed increased lateral root growth during the recovery of drought treatment (scale bar = 2 cm). A number of stress-related genes were quantified with qRT-PCR and compared between the transgenic and control lines before and after the drought treatment, including those encoding antioxidant enzymes (**G**, *NtCAT*, *NtAPX*, *NtSOD*), stress-related ion channels (**H**, *NtCAX3*, *NtNHX2*, *NtNHX4*), other stress-responsive genes (**I,** *NtERD10C*, *NtERD10D*, *NtADC1*) and key genes of sugar and ABA metabolism (**J**, *NtSPSA1*, *NtNCED1*). Data are means ± SE of three biological replicates ($^{*}P < 0.05$, $^{**}P < 0.01$).

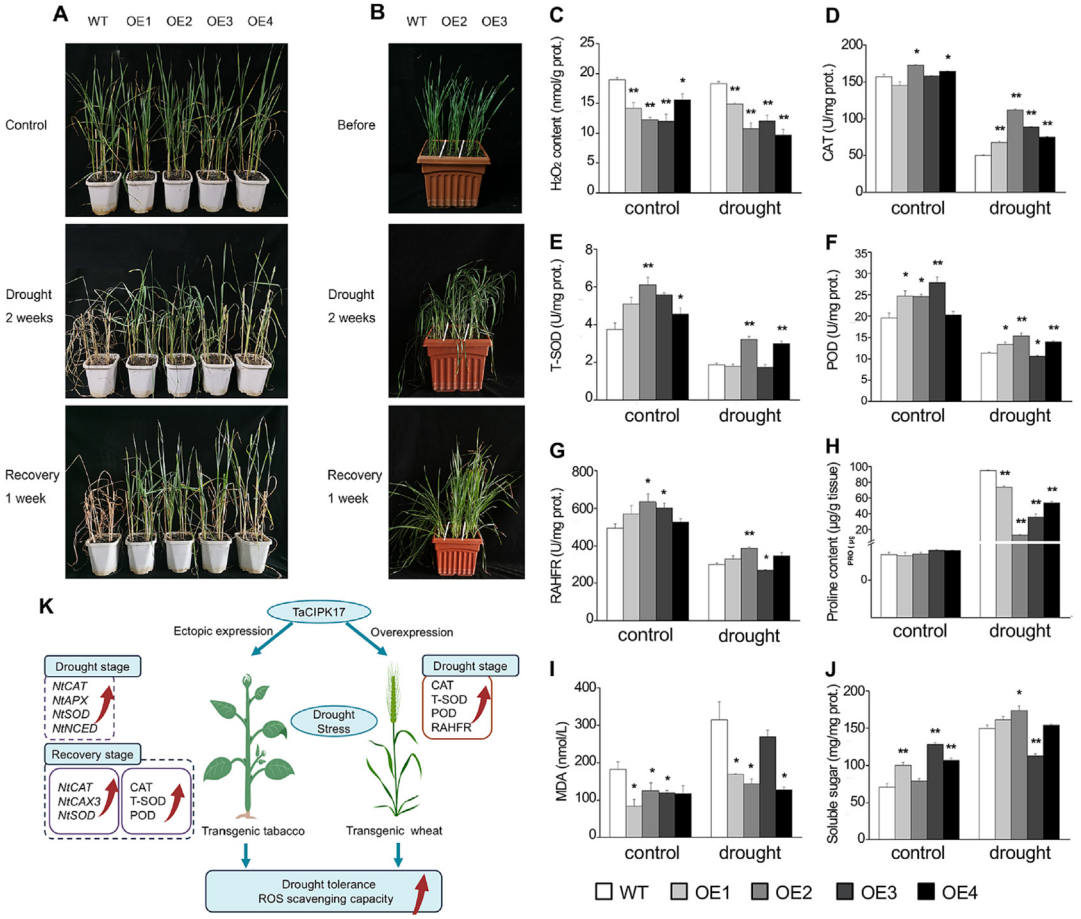


Fig. 6. *TaCIPK17*-overexpression improves drought resistance and ROS scavenging ability in transgenic wheat plants. **(A)** Comparison of the phenotypes of 4-week-old WT, VC and the transgenic lines of wheat (OE1, OE2, OE3, OE4) subject to drought and re-watering. **(B)** Comparison of the phenotypes of ten-day-old WT, VC and the transgenic lines of wheat subject to drought and re-watering. A number of physiological parameters, including H_2O_2 content **(C)**, activities of CAT **(D)**, T-SOD **(E)**, POD **(F)**, restraining ability to hydroxyl free radicals (RAHFR) **(G)**, proline content **(H)**, MDA content **(I)**, soluble sugar content (**J**), were compared between the *TaCIPK*-OE transgenic lines and the controls (WT and VC). Data are means ± S.E. of three biological replicates ($^{*}P < 0.05$, $^{**}P < 0.01$). **(K)** The proposed model of TaCIPK17-mediated drought stress tolerance in tobacco and wheat plants.

TaCIPK15-4A (*TaCIPK14-D* in our work, TraesCS4D02G118500) as a positive regulator for the resistance to *Blumeria graminis* f. sp. *tritici*. Notably, we identified several duplicated genes and loci with tandemly duplicated copies, allowing us to link *TaCIPKs*/*TaCBLs* evolution with functional relevance.

Both of the families have been expanded by two mechanisms, polyploidization and tandem duplication, with many gene duplicates derived from the latter mechanism are common for the Triticeae species. Thus, many of the tandemly duplicated *TaCIPKs* and *TaCBLs* were derived from the recent burst of small-scale duplications [3]. Another evolutionary feature of *TaCIPKs* and *TaCBLs* is that the gene pairs derived from ancient WGD events have been mostly retained in the Triticeae species (Fig. 1). Retention, but not gene loss, of *TaCIPKs* and *TaCBLs* after WGD and polyploidization suggests potential co-retention of the two families and is consistent with the biased functional categories of WGD-derived gene duplication, such as kinases, transporters, transcriptional regulators, and transcription factors [68,69]. Besides, the tandemly duplicates of *TaCIPKs* and *TaCBLs* were either preferentially expressed at certain tissues and/or stages or had highly specific expression patterns with obviously different expression contribution between the duplicates. This suggests that the specific expression patterns and biased expression contribution may further link the duplicates to different regulatory networks and ultimately distinct functions. Still, there are gene families that the spatial–temporal expression patterns do not correlate well with their functions, emphasizing the importance to integrate the information of expression, protein interaction and functional characterization. One such case is the rice *pyruvate kinase* (*OsPK*) family, of which many members were expressed at low level during seed development but their single mutants (produced through gene editing or RNAi) showed phenotypes of increased seed chalkiness and/or increased starch contents in a non-redundant manner [70].

CBL and *CIPK* are large gene families expanded during the evolution of land plants [40]. CBL binds to Ca^{2+} through its EF hands and further interact with the NAF domain of CIPK, with other parts of the CIPK C-terminal (*e.g.*, the autoinhibition domain and the phosphatase interaction motif) involved in the specificity of CBL-CIPK interaction [15]. As one of the core modules of Ca^{2+} signaling, CBL-CIPK has been shown to follow the gene balance hypothesis in some plants [32], in which genes encoding the members of protein complexes should maintain balanced after duplications to sustain an ideal stoichiometric range, whereas imbalanced concentrations of the components could have harmful effects [30,31]. It will be interesting to interrogate how *TaCBLs* and *TaCIPKs* maintain the dosage balance during the wheat polyploidization and the Triticeae RBGD. While we detected several CBL-CIPK interactions failed to maintain the ideal 1:1 ratio, we showed that the WGD-derived CBL or CIPK duplicates tend to have divergent spatial and temporal expression patterns, which could be an approach to maintain the

balanced dosage in a given tissue. Other factors may contribute to keep the balanced dosage of *CBL-CIPK* as well: **(1)** the HEB expression seen in several WGD-derived *CIPK*s and *CBL*s; and **(2)** the specificity and/or affinity between interacting CBLs and CIPKs, which may be regulated by the presence or absence of regulatory motifs in CIPK C-terminal, cellular Ca^{2+} concentration and the conformation of CBLs [56,71–73]. The expansion of *TaCIPK* and *TaCBL* families appeared to be different from those seen in Arabidopsis, in which the family expansion has been mainly driven by WGD events (*i.e.*, ζ and ε duplication events) [32]. Similarly, *TaCIPKs* and *AtCIPKs* tend to have more tandemly duplicated genes than those of the *CBL* family. After the expansion of *CIPK* and *CBL* families, Arabidopsis and also differ from hexaploid wheat in the mechanisms for maintaining the balanced dosage between CIPK and CBL. For example, both divergent interaction pattern and divergent expression pattern after duplication could contribute to maintenance of the CIPK-CBL dosage balance in Arabidopsis. Whether divergent interaction pattern might be involved in keeping CIPK-CBL in relatively balanced dosage remain to be investigated. In addition, it is worth noting that some limitations remain in the present study to fully capture the evolutionary strategies of hexaploid wheat for maintaining protein dosage balance. These limitations include: **(1)** The Y2H-based protein–protein interactions reported herein were not comprehensive enough as not all CIPK and CBL proteins were cloned and examined for interactions; **(2)** Other monocot species (such as rice or diploid Triticeae relatives before the wheat allo-hexaploidization) should be used as a comparative species for the protein–protein interaction study; **(3)** Complex alternative splicing of the *CIPK* and/or *CBL* genes might serve as another strategy to affect CIPK-CBL interactions and to maintain balanced dosages between CIPKs and CBLs, which has been indicated in a recent study [67]. Thus, mechanistic studies on how the C-terminal region of CIPK regulates its interaction specificity and affinity with CBLs is of significance not only to the biochemistry of CBL-CIPK module, but also to their co-evolution insights.

Functions of TaCIPK family members

The spatio-temporal expression patterns of *CBL*s and *CIPK*s and specific interactions between members of the two families lay the foundation for specific regulation of the downstream proteins and involvements in numerous biological processes, such as regulation of ion fluxes [15,74], carbohydrate metabolism [75]), sugar transportation [76,77] and ROS scavenging ability [78–80]. Up to now, only a few *TaCIPK*s have been functionally characterized: **(1)** overexpression of *TaCIPK14* improves the resistance to cold and high salinity stress in tobacco [79]; **(2)** overexpression of *TaCIPK24* enhance salt tolerance in Arabidopsis [56]; **(3)** *TaCIPK27* positively regulates drought tolerance through the ABA-dependent pathway [80]; **(4)** TaCIPK5 interact with CBL4 to enhance the resistance to stripe rust fungus [81]. Our expression analysis of *TaCIPK*s/*TaCBL*s coincided with the reported functions of CIPK or CBL members: *TaCIPK14*, *TaCIPK4*, *TaCIPK27* responds to the cold, salt or drought treatment, respectively, (Fig. 2A), while *TaCBL4.1-D2* was up-regulated by the infection of stripe-rust fungus (Fig. 2C). These *TaCIPK*s and *TaCBL*s, together with the functionally studied *TaCIPK17* herein, suggest that the expression profile of *CIPK*s or *CBL*s is a good indicator for the function in the response and regulation of biotic and abiotic stresses, helping to prioritize the *TaCIPK* or *TaCBL* members for functional studies.

Notably, *TaCIPK17*-OE produced distinct phenotypes between tobacco and wheat during the drought and recovery stages. In tobacco plants, the activities of CAT and T-SOD in the *TaCIPK17*-OE lines were comparable to the controls, but significantly higher than those in the controls during the recovery stages. Less cellular membrane damages (MDA, Fig.S7E) and higher POD activity (Fig. S7D) during drought stress may contribute to the drought recovering ability of *TaCIPK17*-OE lines. In contrast, *TaCIPK17*-OE lines of wheat exhibited clearly better physiological status compared to the non-transgenic control, with higher activities of CAT, POD, SOD and lowered MDA content (Fig. 6). These phenotypic differences between transgenic tobacco and wheat plants indicate potentially distinct working mechanisms of *TaCIPK17* in the two species. We infer that TaCIPK17 might interact with different CBL proteins in tobacco and wheat, respectively, and hence downstream targets to produce the different phenotypes. Our results also suggest that eudicot species (such as Arabidopsis and tobacco) may not be quite suitable to accurately uncover the mechanism of *TaCIPK* as in wheat. In future, detailed studies regarding the CBLs interacting with TaCIPK17, and mechanisms to regulate ROS scavenging system are needed.

Conclusion

In summary, we present an example for the proof-of-concept that the candidates genes deserving detailed functional and genetic experiments can be efficiently mined out by integrating the duplication analysis, expression profiling and/or information regarding protein–protein or protein-DNA interactions. By analyzing the CIPK family across several Triticeae species, we highlight the new function of *TaCIPK17-A2* in the response and regulation of drought stress, which are validated in both tobacco and wheat with the transgenic approach. Our analysis also provides evolutionary insights into the dosage balance maintenance between *TaCBL*s and *TaCIPK*s. In future studies, it will be promising to apply the integrative gene duplication and genome-wide analysis (iGG analysis) to other grasses or plant species with complex duplication events and polyploidization history (such as the Saccharum and Brassica species) [82,83]. Additionally, the power of iGG analysis may be fully unleashed by incorporating more omics information, such as epigenomics, epitranscriptomics and comparative network analyses across multiple species [84].

Compliance with Ethics requirement

This article does not contain any studies with human or animal subjects.

CRediT authorship contribution statement

Ya'nan Wu: Methodology, Validation, Investigation, Data curation, Writing – original draft. **Jialu Feng:** Methodology, Validation, Investigation, Data curation, Writing – review & editing. **Qian Zhang:** Methodology, Validation, Investigation, Data curation, Writing – review & editing. **Yaqiong Wang:** Investigation, Writing – review & editing. **Yanbin Guan:** Investigation, Writing – review & editing. **Ruibin Wang:** Investigation, Writing – review & editing. **Fu Shi:** Investigation, Writing – review & editing. **Fang Zeng:** Investigation, Writing – review & editing. **Yuesheng Wang:** Investigation, Writing – review & editing, Resources, Supervision, Writing – review & editing. **Mingjie Chen:** Resources, Supervision, Writing – review & editing. **Junli Chang:** Resources, Supervision, Writing – review & editing. **Guangyuan He:** Conceptualization, Resources, Supervision, Project administration, Funding acquisition, Writing – original draft, Writing – review & editing. **Guangxiao Yang:** Conceptualization, Resources, Supervision, Project administration, Funding acquisition, Writing – original draft, Writing – review & editing. **Yin Li:** Conceptualization, Resources, Supervision, Project administration, Funding acquisition, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Dr. Kexiu Li (College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, China) for her assistance in wheat transformation. We thank Zhenwu He for his contribution in the plant nursery. We acknowledge the Core Facility of Life Sciences, HUST, for providing service of the MultiSkan microplate reader.

This work was supported by the National Genetically Modified New Varieties of Major Projects of China (2016ZX08010004-004), the National Natural Science Foundation of China (Nos. 31771418, 31570261 and 32272126) the Fundamental Research Funds for Central Universities, HUST (2021XXJS070, 3004170157) and Wuhan Knowledge Innovation Project (2022020801010073).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jare.2023.09.005.

References

[1] Vergauwen D, De Smet I. From early farmers to Norman Borlaug-the making of modern wheat. Curr Biol 2017;27(17):R858–62.
[2] . Science 2018;361(6403):eaar7191.
[3] Wang X, Yan X, Hu Y, Qin L, Wang D, Jia J, et al. A recent burst of gene duplications in Triticeae. Plant Commun 2021;3(2):100268.
[4] Zhao X, Fu X, Yin C, Lu F. Wheat speciation and adaptation: perspectives from reticulate evolution. aBIOTECH 2021;2(4):386–402.
[5] Krasileva KV, Vasquez-Gross HA, Howell T, Bailey P, Paraiso F, Clissold L, et al. Uncovering hidden variation in polyploid wheat. PNAS 2017;114(6):E913–21.
[6] Wang D, Li Y, Wang H, Xu Y, Yang Y, Zhou Y, et al. Boosting wheat functional genomics via indexed EMS mutant library of KN9204. Plant Commun 2022;100593.
[7] Pont C, Leroy T, Seidel M, Tondelli A, Duchemin W, Armisen D, et al. Tracing the ancestry of modern bread wheats. Nat Genet 2019;51(5):905–11.
[8] Pang Y, Liu C, Wang D, St Amand P, Bernardo A, Li W, et al. High-Resolution Genome-wide association study identifies genomic regions and candidate genes for important agronomic traits in wheat. Mol Plant 2020;13(9):1311–27.
[9] Chen Y, Song W, Xie X, Wang Z, Guan P, Peng H, et al. A collinearity-incorporating homology inference strategy for connecting emerging assemblies in the Triticeae Tribe as a pilot practice in the plant pangenomic Era. Mol Plant 2020;13(12):1694–708.
[10] Richardson T, Thistleton J, Higgins TJ, Howitt C, Ayliffe M. Efficient Agrobacterium transformation of elite wheat germplasm without selection. Plant Cell Tiss Organ Cult 2014;119(3):647–59.
[11] Ishida Y, Tsunashima M, Hiei Y, Komari T. Wheat (*Triticum aestivum L.*) transformation using immature embryos. Methods Mol Bio 2015;1223:189–98.
[12] Wang K, Liu H, Du L, Ye X. Generation of marker-free transgenic hexaploid wheat via an Agrobacterium-mediated co-transformation strategy in commercial Chinese wheat varieties. Plant Biotechnol J 2017;15(5):614–23.
[13] Wang K, Shi L, Liang X, Zhao P, Wang W, Liu J, et al. The gene TaWOX5 overcomes genotype dependency in wheat genetic transformation. Nat Plants 2022;8(2):110–7.
[14] Wang Y, Zeng J, Su P, Zhao H, Li L, Xie X, et al. An established protocol for generating transgenic wheat for wheat functional genomics via particle bombardment. Front Plant Sci 2022;13:979540.
[15] Tang R, Wang C, Li K, Luan S. The CBL-CIPK calcium signaling network: Unified paradigm from 20 years of discoveries. Trends Plant Sci 2020;25(6):604–17.
[16] Dong Q, Wallrad L, Almutairi BO, Kudla J. Ca^{2+} signaling in plant responses to abiotic stresses. J Integr Plant Biol 2022;64(2):287–300.
[17] Kolukisaoglu U, Weinl S, Blazevic D, Batistic O, Kudla J. Calcium sensors and their interacting protein kinases: genomics of the Arabidopsis and rice CBL-CIPK signaling networks. Plant Physiol 2004;134(1):43–58.
[18] Ho CH, Lin SH, Hu H, Tsay YF. CHL1 functions as a nitrate sensor in plants. Cell 2009;138(6):1184–94.
[19] Maierhofer T, Diekmann M, Offenborn JN, Lind C, Bauer H, Hashimoto K, et al. Site- and kinase-specific phosphorylation-mediated activation of SLAC1, a guard cell anion channel stimulated by abscisic acid. Sci Signal 2014;7(342):ra86..
[20] Brandt B, Munemasa S, Wang C, Nguyen D, Yong T, Yang P, et al. Calcium specificity signaling mechanisms in abscisic acid signal transduction in Arabidopsis guard cells. Elife 2015;4:e03599.
[21] Förster S, Schmidt LK, Kopic E, Anschütz U, Huang S, Schlücking K, et al. Wounding-induced stomatal closure requires jasmonate-mediated activation of GORK K^+ channels by a Ca^{2+} sensor-kinase CBL1-CIPK5 complex. Dev Cell 2019;48(1):87–99.
[22] Xiang Y, Huang Y, Xiong L. Characterization of stress-responsive CIPK genes in rice for stress tolerance improvement. Plant Physiol 2007;144(3):1416–28.
[23] Avni R, Nave M, Barad O, Baruch K, Twardziok SO, Gundlach H, et al. Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. Science 2017;357(6346):93–7.
[24] Luo M, Gu Y, Puiu D, Wang H, Twardziok SO, Deal KR, et al. Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. Nature 2017;551(7681):498–502.
[25] Zhao G, Zou C, Li K, Wang K, Li T, Gao L, et al. The *Aegilops tauschii* genome reveals multiple impacts of transposons. Nat Plants 2017;3(12):946–55.
[26] Ling H, Ma B, Shi X, Liu H, Dong L, Sun H, et al. Genome sequence of the progenitor of wheat a subgenome *Triticum urartu*. Nature 2018;557(7705):424–8.
[27] Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H, Spannagl M, et al. Durum wheat genome highlights past domestication signatures and future improvement targets. Nat Genet 2019;51(5):885–95.
[28] Walkowiak S, Gao L, Monat C, Haberer G, Kassa MT, Brinton J, et al. Multiple wheat genomes reveal global variation in modern breeding. Nature 2020;588(7837):277–83.
[29] Li L, Zhang Z, Wang Z, Li N, Sha Y, Wang X, et al. Genome sequences of the five sitopsis species of *Aegilops* and the origin of polyploid wheat B-subgenome. Mol Plant 2022;15(3):488–503.
[30] Birchler JA, Newton KJ. Modulation of protein levels in chromosomal dosage series of maize: the biochemical basis of aneuploid syndromes. Genetics 1981;99(2):247–66.
[31] Birchler JA, Riddle NC, Auger DL, Veitia RA. Dosage balance in gene regulation: biological implications. Trends Genet 2005;21(4):219–26.
[32] Zhang X, Li X, Zhao R, Zhou Y, Jiao Y. Evolutionary strategies drive a balance of the interacting gene products for the CBL and CIPK gene families. New Phytol 2020;226(5):1506–16.
[33] Beckmann L, Edel KH, Batistic O, Kudla J. A calcium sensor-protein kinase signaling module diversified in plants and is retained in all lineages of Bikonta species. Sci Rep 2016;6:31645.
[34] Zhu K, Chen F, Liu J, Chen X, Hewezi T, Cheng Z. Evolution of an intron-poor cluster of the CIPK gene family and expression in response to drought stress in soybean. Sci Rep 2016;6:28225.
[35] Zhang Y, Zhou X, Liu S, Yu A, Yang C, Chen X, et al. Identification and functional analysis of tomato CIPK gene family. Int J Mol Sci 2019;21(1):110.
[36] Wang L, Feng X, Yao L, Ding C, Lei L, Hao X, et al. Characterization of CBL-CIPK signaling complexes and their involvement in cold response in tea plant. Plant Physiol Biochem 2020;154:195–203.
[37] Du W, Yang J, Ma L, Su Q, Pang Y. Identification and characterization of abiotic stress responsive CBL-CIPK family genes in Medicago. Int J Mol Sci 2021;22(9):4634.
[38] Wang S, Li Q. Genome-wide identification of the Salvia miltiorrhiza SmCIPK gene family and revealing the salt resistance characteristic of SmCIPK13. Int J Mol Sci 2022;23(12):6861.
[39] Chen X, Gu Z, Xin D, Hao L, Liu C, Huang J, et al. Identification and characterization of putative CIPK genes in maize. J Genet Genomics 2011;38(2):77–87.
[40] Gong HK, Moon S, Jung KH. A systematic view of the rice calcineurin B-like protein interacting protein kinase family. Genes Genom 2014;37(1):55–68.
[41] Kanwar P, Sanyal SK, Tokas I, Yadav AK, Pandey A, Kapoor S, et al. Comprehensive structural, interaction and expression analysis of CBL and CIPK complement during abiotic stresses and development in rice. Cell Calcium 2014;56(2):81–95.
[42] Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, et al. Ensembl 2022. Nucleic Acids Res 2022;50(D1):D988–95.
[43] Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 2011;40(Database issue):D1178–86.
[44] Ramírez-González RH, Borrill P, Lang D, Harrington SA, Brinton J, Venturini L, et al. The transcriptional landscape of polyploid wheat. Science 2018;361(6403):eaar6089.
[45] Paterson AH, Bowers JE, Chapman BA. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. PNAS 2004;101(26):9903–8.
[46] Tang H, Bowers JE, Wang X, Paterson AH. Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. PNAS 2010;107(1):472–7.
[47] Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol 2018;35(6):1547–9.
[48] Bailey TL, Johnson J, Grant CE, Noble WS. The MEME suite. Nucleic Acids Res 2015;43(W1):W39–49.
[49] Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant 2020;13(8):1194–202.

[50] Chen Y, Song W, Xie X, Wang Z, Guan P, Peng H, et al. A collinearity-incorporating homology inference strategy for connecting emerging assemblies in the Triticeae Tribe as a pilot practice in the plant pangenomic era. Mol Plant 2020;13(10):1694–708.
[51] . Science 2014;345(6194):1251788.
[52] Zhang H, Yang Y, Wang C, Liu M, Li H, Fu Y, et al. Large-scale transcriptome comparison reveals distinct gene activations in wheat responding to stripe rust and powdery mildew. BMC Genom 2014;15(1):898.
[53] Li Q, Zheng Q, Shen W, Cram D, Fowler DB, Wei Y, et al. Understanding the biochemical basis of temperature-induced lipid pathway adjustments in plants. Plant Cell 2015;27(1):86–103.
[54] Liu Z, Xin M, Qin J, Peng H, Ni Z, Yao Y, et al. Temporal transcriptome profiling reveals expression partitioning of homeologous genes contributing to heat and drought acclimation in wheat (*Triticum aestivum L.*). BMC Plant Biol 2015;15:152.
[55] Ma S, Wang M, Wu J, Guo W, Chen Y, Li G, et al. WheatOmics: A platform combining multiple omics data to accelerate functional genomics studies in wheat. Mol Plant 2021;14(12):1965–8.
[56] Sun T, Wang Y, Wang M, Li T, Zhou Y, Wang T, et al. Identification and comprehensive analyses of the CBL and CIPK gene families in wheat (*Triticum aestivum L.*). BMC Plant Biol 2015;15:269.
[57] Wang C, Zeng J, Li Y, Hu W, Chen L, Miao Y, et al. Enrichment of provitamin A content in wheat (*Triticum aestivum L.*) by introduction of the bacterial carotenoid biosynthetic genes CrtB and CrtI. J Exp Bot 2014;65(9):2545–56.
[58] Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SC, Fraley RT. A simple and general method for transferring genes into plants. Science 1985;227:1229–31.
[59] Qiu D, Hu W, Zhou Y, Xiao J, Hu R, Wei Q, et al. TaASR1-D confers abiotic stress resistance by affecting ROS accumulation and ABA signaling in transgenic wheat. Plant Biotechnol J 2021;19(8):1588–601.
[60] Luo Q, Wei Q, Wang R, Zhang Y, Zhang F, He Y, et al. Ectopic expression of BdCIPK31 confers enhanced low-temperature tolerance in transgenic tobacco plants. Acta Biochim Biophys Sin (*Shanghai*) 2018;50(2):199–208.
[61] Wang D, Pan Y, Zhao X, Zhu L, Fu B, Li Z. Genome-wide temporal-spatial gene expression profiling of drought responsiveness in rice. BMC Genomics 2011;12:149.
[62] Huang L, Zhang F, Zhang F, Wang W, Zhou Y, Fu B, et al. Comparative transcriptome sequencing of tolerant rice introgression line and its parents in response to drought stress. BMC Genomics 2014;15(1):1026.
[63] Chung PJ, Jung H, Jeong DH, Ha S, Choi YD, Kim JK. Transcriptome profiling of drought responsive noncoding RNAs and their target genes in rice. BMC Genomics 2016;17:563.
[64] Beena R, Kirubakaran S, Nithya N, Manickavelu A, Sah RP, Abida PS, et al. Association mapping of drought tolerance and agronomic traits in rice (*Oryza sativa L.*) landraces. BMC Plant Biol 2021;21(1):484.
[65] Salse J, Bolot S, Throude M, Jouffe V, Piegu B, Quraishi UM, et al. Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution. Plant Cell 2008;20(1):11–24.
[66] Jiao Y, Li J, Tang H, Paterson AH. Integrated syntenic and phylogenomic analyses reveal an ancient genome duplication in monocots. Plant Cell 2014;26(7):2792–802.
[67] Liu X, Wang X, Yang C, Wang G, Fan B, Shang T, et al. Genome-wide identification of TaCIPK gene family members in wheat and their roles in host response to *Blumeria graminis* f. sp. *tritici* infection. Intl J Biol Macromol 2023;248:125691.
[68] Freeling M. Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. Annu Rev Plant Biol 2009;60:433–53.
[69] Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, et al. Ancestral polyploidy in seed plants and angiosperms. Nature 2011;473(7345):97–100.
[70] Dong N, Chen L, Ahmad S, Cai Y, Duan Y, Li X, et al. Genome-wide analysis and functional characterization of pyruvate kinase (PK) gene family modulating rice yield and quality. Intl J Mol Sci 2022;23:15357.
[71] Ohta M, Guo Y, Halfter U, Zhu J. A novel domain in the protein kinase SOS2 mediates interaction with the protein phosphatase 2C ABI2. PNAS 2003;100(20):11771–6.
[72] Sanchez-Barrena MJ, Fujii H, Angulo I, Martinez-Ripoll M, Zhu JK, Albert A. The structure of the C-terminal domain of the protein kinase AtSOS2 bound to the calcium sensor AtSOS3. Mol Cell 2007;26(3):427–35.
[73] Hashimoto K, Eckert C, Anschütz U, Scholz M, Held K, Waadt R, et al. Phosphorylation of calcineurin B-like (CBL) calcium sensor proteins by their CBL-interacting protein kinases (CIPKs) is required for full activity of CBL-CIPK complexes toward their target proteins. J Biol Chem 2012;287(11):7956–68.
[74] Rodenas R, Ragel P, Nieves-Cordones M, Martinez-Martinez A, Amo J, Lara A, et al. Insights into the mechanisms of transport and regulation of the Arabidopsis high-affinity K^+ transporter HAK5. Plant Physiol 2021;185(4):1860–74.
[75] Lee KW, Chen P, Lu C, Chen S, Ho TH, Yu S. Coordinated responses to oxygen and sugar deficiency allow rice seedlings to tolerate flooding. Sci Signal 2009;2(91):ra61.
[76] Ma Q, Sun M, Kang H, Lu J, You C, Hao Y. A CIPK protein kinase targets sucrose transporter at MdSUT2.2 Ser^{254} for phosphorylation to enhance salt tolerance. Plant Cell Environ 2019;42(3):918–30.
[77] Ma Q, Sun M, Lu J, Kang H, You C, Hao Y. An apple sucrose transporter MdSUT2.2 is a phosphorylation target for protein kinase MdCIPK22 in response to drought. Plant Biotechnol J 2019;17(3):625–37.
[78] Deng X, Hu W, Wei S, Zhang F, Han J, Chen L, et al. TaCIPK29, a CBL-interacting protein kinase gene from wheat, confers salt stress tolerance in transgenic tobacco. PLoS One 2013;8(7):e69881.
[79] Deng X, Zhou S, Hu W, Feng J, Zhang F, Chen L, et al. Ectopic expression of wheat TaCIPK14, encoding a calcineurin B-like protein-interacting protein kinase, confers salinity and cold tolerance in tobacco. Physiol Plant 2013;149(3):367–77.
[80] Wang Y, Li T, John SJ, Chen M, Chang J, Yang G, et al. A CBL-interacting protein kinase TaCIPK27 confers drought tolerance and exogenous ABA sensitivity in transgenic Arabidopsis. Plant Physiol Biochem 2018;123:103–13.
[81] Liu P, Duan Y, Liu C, Xue Q, Guo J, Qi T, et al. The calcium sensor TaCBL4 and its interacting protein TaCIPK5 are required for wheat resistance to stripe rust fungus. J Exp Bot 2018;69(18):4443–57.
[82] Zhang J, Zhang X, Tang H, Zhang Q, Hua X, Ma X, et al. Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. Nat Genetics 2018;50(11):1565–73.
[83] Song X, Wei Y, Xiao D, Gong K, Sun P, Ren Y, et al. Brassica carinata genome characterization clarifies U's triangle model of evolution and polyploidy in Brassica. Plant Physiol 2021;186(1):388–406.
[84] Miao Z, Zhang T, Xie B, Qi Y, Ma C. Evolutionary implications of the RNA N6-methyladenosine methylome in plants. Mol Biol Evol 2022;39(1):msab299.