Genome-wide Identification, Characterization, and Function Analysis Uncovered Roles of Wheat TaLIM in Responding to Adverse Stress and TaLIM8-4D Function as a Susceptible Gene

Yiting Li

Yangtze University

Xi Liu

Yangtze University

Yongxin Xiao

Yangtze University

Yong Wen

Yangtze University

Xiaowen Han

Yangtze University

Shuo Han

Yangtze University

Lijun Yang

Hubei Academy of Agricultural Sciences

Yongxing Zhu

Yangtze University

Junliang Yin (✉ yinjunliang@yangtzeu.edu.cn)

Yangtze University

Yiqing Liu

Yangtze University

Research Article

Keywords: exon/intron structure, protein motif, GO annotation, qRT-PCR, cell death, subcellular localization

Posted Date: October 12th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-954637/v1>

Abstract

Background: The LIM (Lin-11, Isl-1 and Mec-3 domains) transcription factors play essential roles in regulating plant biological processes, including cytoskeletal organization, development of secondary cell wall, and cell differentiation. Despite their important roles, there lacks a fully understanding of LIMs in wheat.

Results: In this study, 28 TaLIMs (*Triticum aestivum* LIM) were identified and designated as TaLIM1-1A to TaLIM12-7D. Phylogenetic relationship analysis showed that TaLIMs could be divided into two groups (Group I and Group II), and the exon/intron structure analysis suggested that members in the same group have similar gene structure. Chromosome mapping showed that 28 TaLIMs were unevenly distributed on 18 chromosomes. Motif analysis revealed that 28 TaLIMs contained 15 motifs, of which motif1 (LIM-domain) was detected in all TaLIMs. Protein characterization indicated that TaLIMs were hydrophilic proteins and most of them were unstable. They were mainly comprised by α -helix and β -turn and without signal peptides. Duplication, Ka/Ks, and synthesis analyses suggested that polyploidization played fundamental role in the expansion of TaLIM members. Function annotations indicated that all TaLIMs were annotated under GO terms “zinc ion binding” (GO:0008270). The *cis*-regulatory element analysis showed that TaLIMs were rich in elements related to biological/abiotic stress, growth and development, and phytohormone response. RNA-seq data analysis showed certain members of TaLIMs were responsive to biotic and/or abiotic stresses. Such as *TaLIM1-1A*, *TaLIM3-2B*, *TaLIM8-4D*, and *TaLIM10-5D* were significantly induced by heat, drought, NaCl, ABA and *Fusarium graminearum* stresses. The subcellular localization of *TaLIM8-4D* show that it was in the nucleus. Furthermore, the biological function of *TaLIM8-4D* was analyzed and results showed that it could induce weak cell death in *Nicotiana benthamiana* leaf. Besides, overexpression of *TaLIM8-4D* could up-regulate plant *pathogenesis-related* (*PR*) genes, whereas promoted the infection of hemi-biotrophic pathogen, implying that *TaLIM8-4D* could function as susceptible gene in nucleus by up-regulating *PR* genes and inducing cell death to promote the colonization of hemi-biotrophic agent *F. graminearum*.

Conclusion: The systematic identification, characterization, expression profiling, evolutionary, and function analyses provided us the fully understanding to TaLIMs and laid a foundation for the further function study of LIM family members in wheat.

Background

Transcription factors (TFs), also known as *trans*-acting factors, are a kind of proteins that can regulate the expression of target genes [1]. They participate the regulation of plant's response to environmental stresses by directly or indirectly binding to *cis*-elements in the promoter region of stress-induced genes [2]. Typical TFs contain several special structures that determine the transcription and regulation characters of corresponding TFs, such as DNA binding domain, activation domain, oligomerization site and nuclear localization signal [3-5]. LIM (Lin-11, Isl-1 and Mec-3 domains) proteins are widely present in eukaryotes. They are a class of cysteine rich proteins (CRP) and comprise a key transcription factor family [6].

Sequence analysis showed that all LIM proteins contain at least one LIM domain, which is a cysteine- and histidine-rich zinc coordination domain and functionally related to DNA binding during the transcription process [7]. The LIM domain sequence (C-X2-C-X16-23-H-X2-C)-X2-(CX16-23-Z-X2-C) is highly conserved in many species and generally composed of more than 50 amino acid residues [8].

Previously studies showed that LIM proteins have two functions in plants: (1) as a transcription factor, it regulates the expression of genes related to phenylpropanoid biosynthesis pathway, thus involves in plant secondary cell wall development and cell differentiation [9]; (2) As a binding protein of actin, it regulates the cytoskeleton structure of actin, therefore participates in the cytoskeleton organization [10]. Considering their important roles, LIM family members had been studied in a variety of plants, such as *Populus trichocarpa*, *Arabidopsis thaliana*, *Oryza sativa* [11], *Zea mays* [12], *Solanum lycopersicum* [13], and *Chenopodium quinoa willd* [14]. However, there is no comprehensive study of LIM family in wheat.

As a grain plant in Gramineae, wheat is widely grown throughout the world, thereby it becomes one of the most important food crops worldwide [15]. However, the growth of wheat is threatened by many abiotic/biotic stresses [4, 16]. Cultivating stress resistant wheat varieties is an economic and effective way to reduce the lost caused by abiotic/biotic stresses. Since resistant genes are the foundation for breeding stress tolerance wheat cultivars, it is necessary to mine more resistant genes [17]. In this study, the TaLIM (*Triticum aestivum* LIM) family members were systemically identified and characterized by bioinformatics method, and their chromosome distribution, gene structure, motif and *cis*-acting elements, expression patterns, and subcellular localization were further analyzed. Furthermore, the biological function of *TaLIM8-4D* was analyzed. These results can provide a theoretical basis for the further study of LIM gene function in wheat and prepare gene resources for resistance breeding.

Results And Discussion

Identification and phylogenetic analysis of LIM gene in wheat genome

LIM proteins are widely found in eukaryotes which act as developmental regulators in basic cellular processes such as regulating the transcription or organizing the cytoskeleton [11]. The genome wide identification of LIMs has been conducted in several plant species, such as *Oryza sativa*, *Populus trichocarpa*, *Arabidopsis thaliana* [11], *Zea mays* [12] and *Chenopodium quinoa* [14]. In order to get insights into the function of wheat LIMs (TaLIMs), comprehensive analysis of TaLIMs was performed in the present study. We used two methods to identify the *Triticum aestivum* LIM genes (TaLIMs) from the wheat reference genome (IWGSC v2.1). One is to use the downloaded Pfam domain sequences (PF00412) as seeds to perform BLASTp search against genome. Another one is to use the sequences of the known LIM proteins, including 14 AtLIMs, 9 OsLIMs and 11 ZmLIMs, as the queries to perform the BLASTp search [18]. Results of both methods were combined and redundant sequences were removed. The obtained candidate sequences were further checked by Pfam and InterProScan to excluded sequences that do not contain the domain of LIM. Finally, 28 TaLIMs were identified from wheat genome

(Additional file 1: Table S1). In order to determine the evolutionary relationship between TaLIMs and AtLIMs, OsLIMs, and ZmLIMs, a phylogenetic tree containing 62 LIM protein sequences was constructed by multiple sequence alignment using ClustalW2. Based on the phylogenetic relationship, the 62 LIMs were divided into two groups, Group I containing 36 LIMs and Group II containing 26 LIMs (Fig. 1). Among the 28 TaLIMs, 17 have two LIM conserved domain and belong to Group I; 11 belong to Group II and contain one LIM domain at the N-terminal and another new conserved cysteine-rich domain DA1-like at the C-terminal [19]. Those TaLIMs were named by comprehensive consideration of the evolutionary relationships, the character of conserved domains, and the chromosome position. The name of TaLIMs was shown in Additional file 1: Table S1.

Chromosomal mapping, gene duplication and Ka/Ks analysis of TaLIMs

As showed in Fig. 2, chromosome distribution map of TaLIMs was constructed by using MapInspect according to the physical chromosome location extracted from the GFF3 file. The map showed that the 28 identified TaLIMs were unevenly distributed on 18 chromosomes, whereas no TaLIM was located on chromosome 3. The number of TaLIMs in A, B and D sub-genomes were counted. Results showed that there were 8 TaLIMs in A, 10 TaLIMs in B, and 10 TaLIMs in D sub-genome. Gene copy number variation is the key foundation for the plant genome adaptation to the changing environment since each copy of the gene can evolve independently and diversify its functions [20]. Fragment duplication and tandem duplication are generally considered to be the two main causes of gene copy number variation, which further result in the extension of plant gene family [21]. Therefore, it is necessary to detect the duplication events of *TaLIMs*. Based on the sequence similarity and chromosome position, a total of 26 pairs of fragment duplications were found among 28 *TaLIMs*, and no tandem duplication event was found. The result suggested that fragment duplication is a major contributor to the expansion of LIM family in wheat during evolution. In order to further decipher the evolutionary history of *TaLIMs*, we calculated the Ka/Ks values of 26 pairs of *TaLIM* fragment duplications using TBtools. Generally, the Ka/Ks value (non-synonymous substitution rate/synonymous substitution rate) is widely used to represent the selection pressure and evolutionary speed of gene evolution. Ka/Ks value = 1 represents the neutral selection effect, Ka/Ks value > 1 represents positive selection of evolution acceleration, and Ka/Ks value < 1 represents purifying selection under functional constraints. In this study, the Ka/Ks values of the 26 duplicate pairs were all less than 1 and the average value is 0.11 (Additional file 1: Table S2), indicating that they underwent strong purification selection in the evolutionary process, which may contribute to functional stability.

Protein characterization and homology modeling of TaLIMs

In order to further determine the protein characteristic of TaLIMs, the online tool ExPASy Server10 was used to predict the protein properties of TaLIMs (Additional file 1: Table S1). Results showed that the

amino acid (aa) length of TaLIMs ranged from 196 to 1,279 aa. The instability index showed that the majority of TaLIMs were unstable proteins (instability index > 40). Only Group I member TaLIM11-6A, TaLIM11-6B, TaLIM11-6D, TaLIM1-1A, TaLIM1-1B, TaLIM1-1D, TaLIM10-5A, TaLIM10-5B, and TaLIM10-5D were predicted to be stable protein (instability index < 40). The TaLIMs isoelectric point varied significantly between 4.57 and 8.63, and there were both acidic and basic proteins. The grand average of hydropathicity (GRAVY) of every TaLIMs is less than 0, suggesting that they are all hydrophilic proteins. Since no signal peptides were detected, it was speculated that all TaLIMs were non-secreted proteins. The subcellular localization prediction showed that all 11 TaLIMs in Group I were located in the nucleus. In Group II, 12 TaLIMs were located in the nucleus, while TaLIM11-6A, TaLIM7-4B, TaLIM8-4D, TaLIM10-5B, and TaLIM10-5D were located in both nucleus and golgiosome.

The 3D (three dimension) homology modeling of TaLIMs were completed by SWISS-MODEL and results showed that TaLIMs are mainly comprised by α -helix and β -turn and have considerable simple 3D structures, whereas the proportion of these two kinds of structures are different among TaLIM proteins (Fig. 3). Moreover, the proportion of α -helix (20-38%) was greater than that of β -turn (2.5-14%).

Gene structure and motifs analysis of TaLIMs

Structure determines function and exon/intron structural diversity usually play key roles in the evolution of gene families, which can provide additional evidence to support phylogenetic grouping [22]. Therefore, we used the GFF3 file of *TaLIMs* to analyze their exon/intron structure and was illustrated by GSDS. Results showed that all *TaLIMs* contained intron. The number of introns ranged from 4 to 12 and the number of exons ranged from 5 to 12. Most of the *TaLIMs* (23) contained complete UTR region, except for *TaLIM5-4A*, *TaLIM5-4B*, and *TaLIM5-4D* that only contained 3'-UTR region, and *TaLIM2-2D* and *TaLIM8-4D* contained no UTR region (Fig. 4b), which may be one of the reasons for the functional differentiation of *TaLIMs*. Protein conserved functional domain analysis showed that all of the 17 members in Group I contain two LIM domains, and all of the 11 members in Group II contain one LIM domain and one DA1-like domain, which is consistent with previous studies (Fig. 4a and c) [23]. Protein motif prediction is an effective method to predict protein function [24]. In order to further study structural diversity and predict the function of TaLIMs proteins, 15 motifs were identified using MEME online software and further annotated using Interproscan5 (Fig. 5b). Detailed information of 15 motifs including specific amino acid sequences of each motif were showed in Additional file 1: Table S4. Meanwhile, WebLogo was used to generate a conservative design logo icon (Fig. 5c). Results showed that TaLIMs closely clustered in phylogenetic tree (Fig. 5a) have similar motif arrangements and positions, suggesting that members in same tree branch may have similar biological function. All the TaLIMs contained motif1 (LIM-domain) and motif6. The function of motif6 is unknown. Group I members specifically contained motif3 (LIM-domain), LIM-domain containing proteins have been shown to play roles in cytoskeletal organization and organ development [9, 10]. Group II members specifically contained motif4, 5, 7, 8, which comprise the DA1-like conserved domain. But the function of DA1-like domain is unknown. These results showed that the motifs between Group I and II members were different, while the members belong

to same Group contain similar motifs, which imply that the functions of different group members might be related to specific motif [25].

Analysis of *cis*-acting element for TaLIMs

Transcription factors are important in the regulation of plant growth and development, hormones metabolisms and stress responses. While *cis*-elements existed in the promoter region play important roles in the transcriptional level of genes in response to multiple stimuli [1]. Therefore, the analysis of promoter *cis*-acting elements is important in the study of genes function [26]. In this study, we identified 48 kinds of *cis*-acting elements in the upstream 1.5 kb region of 28 TaLIMs and these *cis*-acting elements were divided into three groups: 28 *cis*-acting elements belong to “biological/abiotic stress”, 11 belong to “growth and development”, and 9 belong to “phytohormone response”, respectively (Fig.6). The detailed information of these identified *cis*-acting elements including function and location were displayed in Additional file 1: Table S5. Results showed that the promoter sequences of *TaLIMs* have various *cis*-regulatory elements, such as anaerobic-response element (GC-motif and ARE), low temperature responsive element (LTR), *cis*-acting regulatory element involved in light responsiveness (G-box) and MYB binding site involved in drought-inducibility (MBS), which were related to biotic/abiotic stresses.

Abscissic acid (ABA) plays a vital role in plant growth and development as well as in mediating plant response to a wide range of stresses. The abscissic acid response element (ABRE) was identified in most of *TaLIMs* (17). Moreover, the other *cis*-elements involved in osmotic stress, such as MBS and TC-rich repeats, were also observed in TaLIMs promoters. This result suggested that LIM genes in wheat may be regulated by various factors, including drought and ABA, which need to be experimentally proved in further studies.

Transcriptomic profiling revealed various expression patterns of *TaLIMs*

For multigene family, analyses of gene expression patterns often provided useful clues for determining gene function [27]. The transcriptome data sets downloaded from NCBI were divided into three categories according to biotic stress, abiotic stress, and growth and development (Fig. 7, Additional file 1: Table S6). Expression patterns analysis showed that TaLIMs have variable expression patterns. The expression levels of *TaLIM7-4B* and *TaLIM8-4D* were generally high throughout the growth period and under various biotic and abiotic stress conditions. Under various biotic stress including *Zymoseptoria tritici*, powdery mildew, stripe rust, and *Fusarium graminearum*, the expression levels of 17 TaLIMs were highly induced. Whereas the expression levels of *TaLIM1-1A*, *TaLIM1-1B*, *TaLIM1-1D*, *TaLIM2-2A*, *TaLIM2-2B*, *TaLIM2-2D*, *TaLIM11-6A*, *TaLIM11-6B*, and *TaLIM11-6D* were only highly induced under the *Fusarium graminearum* treatment. As shown in Fig. 7, under 17 different growth and development stages, the expression levels of 19 TaLIMs were highly induced. But *TaLIM1-1A*, *TaLIM1-1B*, *TaLIM1-1D*, *TaLIM11-*

6A, *TaLIM11-6B*, and *TaLIM11-6D*, which had high expression only in anther-anthesis stage and stigma & ovary-anthesis stage. *TaLIM2-2A*, *TaLIM2-2B*, and *TaLIM2-2D* were highly expressed only in the anther-anthesis stage. Under abiotic stress, most of the genes were expressed, whereas the expression levels were low (the expression levels showed in Additional file 1: Table S6). *TaLIM2-2A* and *TaLIM11-6D* were not expressed, while *TaLIM2-2B*, *TaLIM2-2D*, *TaLIM11-6A*, and *TaLIM11-6B* were expressed only under certain treatments.

Evolutionary conserved TaLIMs were family member expanded by polyploidization

Previous researches showed that the evolutionary history of common wheat is quite complex, and its ancestral origin is influenced by many factors. But it is commonly believed that wild diploid wheat (*Triticum urartu*) crossed with the B genome ancestor that is the closest relative of goat grass (*Aegilops speltoides*) to produce wild emmer wheat (*Triticum dicoccoides*) [28-30]. In order to further understand the evolutionary relationship and homology of the LIM genes among allohexaploid common wheat (AABBDD) and ancestral species, we used BLASTp to identify homologous genes from *T. urartu* (AA), *Aegilops tauschii* (DD) and *T. dicoccoides* (AABB), which resulted in the detection of 8, 10, and 20 homologous genes, respectively (Fig. 8a). Therefore, *T. aestivum* inherited more LIM genes from *T. dicoccoides* than *T. urartu* and *Ae. tauschii*. Furthermore, we speculate that the expansion of *TaLIMs* may derived from ancestors and expanded by polyploidization after the crossing between ancestral species. To further confirm this speculation, we identified the homologous gene pairs between *T. aestivum* and ancestral species, resulting in the detection of 4, 11, and 36 homologous pairs between *T. aestivum* and *T. urartu* (Ta vs Tu), *T. aestivum* and *Ae. tauschii* (Ta vs Ae), and *T. aestivum* and *T. dicoccoide* (Ta vs Td), respectively (Fig. 8b). Among the 51 homologous gene pairs, the Ka/Ks ratios for 49 pairs were less than 1 (Fig. 8c), implying that most duplicated gene pairs underwent purification selection pressure, which helps their function to be maintained and reflects that they do not have much divergence during evolution. To further validate the conservation of LIMs, we identified homologous genes in the Gramineae family and calculate their Ka/Ks. The result showed that 41 pairs of homologous genes among wheat, maize, and rice were identified. The Ka/Ks analysis showed that ratios for homologous pairs were less than 1 and the average value is 0.14 (Fig. 8d, Additional file 1: Table S7), implying that the evolution of LIM genes were conserved in Gramineae family species. The above results indicate that LIM genes were not only conserved in the evolution of wheat, but also highly conserved in grass plants.

Expression analyses of TaLIMs

In order to further understand the role of TaLIM genes in response to biological/abiotic stress, the log2-transformed TPM values of the RNA-seq transcriptomic dataset were mined (Fig. 7). Furthermore, several highly expressed (*TaLIM8-4D*) and/or differentially expressed (*TaLIM1-1A*, *TaLIM3-2B* and *TaLIM10-5D*)

LIM genes were selected to analysis their expression patterns in responding to *Fusarium graminearum*, heat, drought, ABA, and NaCl stresses using qRT-PCR (Fig. 9).

In the pollen, under heat treatment, the expression levels of *TaLIM1-1A* and *TaLIM3-2B* were down-regulated, with the expression levels being the lowest after 48 h. The expression levels of *TaLIM8-4D* and *TaLIM10-5D* were up-regulated after 12 h, 24 h, 48 h and 72 h of treatment. The expression levels of *TaLIM10-5D* were up-regulated after 24 h, 48 h and 72 h of treatment, but the expression level was decreased after 12 h of treatment. Under drought treatment, the expression levels of *TaLIM1-1A*, *TaLIM3-2B* and *TaLIM10-5D* in the pollen were down-regulated. The expression levels of *TaLIM8-4D* were up-regulated after 12 h, 24 h and 72 h of treatment, but after 48 h of treatment, the expression level was down-regulated compared with control.

In the leaf, under ABA treatment, the expression levels of *TaLIM1-1A* were up-regulated after 2 h of treatment, but down-regulated after 6 h, 12 h, 24 h and 48 h of treatment. The expression levels of *TaLIM8-4D* and *TaLIM10-5D* were down-regulated after 2 h, 6 h, 24 h and 48 h of treatment, but up-regulated after 12 h of treatment. The expression levels of *TaLIM3-2B* were up-regulated after 2 h, 6 h, 12 h and 48 h of treatment, but down-regulated after 24 h of treatment.

In the leaf, under NaCl treatment, the expression levels of *TaLIM1-1A* were up-regulated after 2 h of treatment, but down-regulated after 6 h, 12 h, 24 h, and 48 h of treatment. The expression levels of *TaLIM3-2B* were up-regulated after 2 h and 24 h of treatment, but down-regulated after 6 h, 12 h and 48 h of treatment. The expression levels of *TaLIM8-4D* were down-regulated after 2 h, 6 h and 48 h of NaCl treatment, but up-regulated after 12 h and 24 h of treatment. The expression levels of *TaLIM10-5D* were down-regulated after 2 h, 6 h and 24 h of treatment, but up-regulated after 12 h and 48 h of treatment.

In the coleoptiles, under the treatment of *Fusarium graminearum*, the expression levels of *TaLIM1-1A*, *TaLIM3-2B*, *TaLIM8-4D* and *TaLIM10-5D* were down-regulated after 1 d, 2 d and 3 d of treatment, but up-regulated after 5 d of treatment. Furthermore, the expression levels of *TaLIM1-1A* and *TaLIM3-2B* were lowest after 2 d of treatment.

GO annotation and subcellular localization of TaLIMs

TaLIMs functions were predicted by Gene Ontology (GO) annotation. As shown in Fig. 10a (detailed information listed in Additional file 1: Table S8), all TaLIMs (28) were annotated under GO terms “zinc ion binding” (GO:0008270). Half of TaLIMs (14) were annotated under “actin filament binding” (GO:0051015) and “actin filament bundle assembly” (GO:0051017). The results implying TaLIMs may be an actin binding protein that are involved in various developmental activities and cellular processes in plants as suggested by previous studies [31]. Moreover, some triplets of TaLIMs have the same function annotation; such as *TaLIM10-5A*, *TaLIM10-5B* and *TaLIM10-5D* were annotated under “sequence-specific DNA binding transcription factor activity”, “plasma membrane”, “regulation of transcription, DNA-templated” and “membrane fusion”. *TaLIM5-4A*, *TaLIM5-4B* and *TaLIM5-4D* were annotated under

“phloem development” and “root development”. Consistently with the function annotation, the subcellular localization prediction showed that 28 TaLIM proteins were localization in the nucleus, while TaLIM11-6A, TaLIM7-4B, TaLIM8-4D, TaLIM10-5B and TaLIM10-5D were predicted to be located in the nucleus and golgiosome. In order to verify subcellular localization prediction and reveal the possible functions of TaLIMs, we constructed the GFP fusion protein expression vector pART-27:TaLIM8-4D:GFP to do the subcellular analysis. Compared to the control (Fig.10b), TaLIM8-4D is localization in the nucleus (Fig. 10c). These results suggest that TaLIM8-4D may play a biological role in the nucleus.

TaLIM8-4D promoted the infection of pathogen

To preliminary test the biological function and reveal whether TaLIM8-4D participates in the pathogenesis process of plant to pathogen, this gene was cloned and transiently expressed in *Nicotiana benthamiana* using Agrobacterium-mediated expression followed by a hemi-biotrophic pathogen *Phytophthora infestans* infection. As shown in Fig. 11a and b, the result suggested TaLIM8-4D promoted the colonization of *P. infestans*, implying TaLIM8-4D function as a susceptible gene. Meanwhile, weak cell death was observed in *N. benthamiana* leaf after 5 days transient expression, indicating that TaLIM8-4D is a cell death inducer (leave cell death as shown in Fig. 11c).

Pathogenesis-related (PR) genes were up-regulated by TaLIM8-4D

TaLIM8-4D was transiently overexpressed in *N. benthamiana* leaves to explore the influence of TaLIM8-4D on the expression of *pathogenesis-related (PR)* genes. The expression of three PTI related genes (*NbWRKY7*, *NbWRKY8* and *NbACRE31*), one reactive oxygen related gene (*NbRbohA*), three SA response genes (*NbPR1a*, *NbPR1b* and *NbPR2*), and three JA-dependent immunity genes (*NbPR3*, *NbPR4*, *NbLOX*) were analyzed. As shown in Fig. 12, *PR* genes, such as *NbPR1b*, *NbACRE31*, *NbWRKY7* and *NbRbohA*, were up-regulated at 1 to 4 d in *TaLIM8-4D* overexpression plant, but down-regulated at 5 d. The expression of *NbPR3*, *NbPR4*, *NbPR2* and *NbLOX* were up-regulated at specific time point. For example, *NbLOX* and *NbPR2* were down-regulated at 1 d, but up-regulated at 2 to 5 d. *NbPR4* was up-regulated at 1 to 3 d, but down-regulated at 4 to 5 d; *NbPR3* was down-regulated at 1d and 5 d, but up-regulated at 2 to 4 d. These results demonstrate the *TaLIM8-4D* genes could trigger the expression of *PR* genes to activate the hypersensitive response (Fig. 11c), which might be the reason why the larger disease lesion was observed in *TaLIM8-4D* transgenic plant after pathogen infection as shown in Fig. 11a and b. Integrating considering the results of qRT-PCR and overexpression of *TaLIM8-4D*, it was speculated that *TaLIM8-4D* was function as susceptible gene in nucleus by up-regulating *PR* genes and inducing cell death to promote the colonization of hemi-biotrophic agent *Fusarium graminearum*. However, this function mechanism needs to be further confirmed.

Conclusions

In conclusion, 28 TaLIM genes were identified in wheat and were grouped into two groups. Group I members contain two LIM domains and Group II members contains one LIM and one DA1-like domains. They are unevenly distributed on 18 chromosomes. TaLIMs are hydrophilic proteins and most of them are unstable. They are mainly comprised by α -helix and β -turn and without signal peptides. Polyploidization played fundamental role in the expansion of TaLIM members, because: (1) TaLIMs showed fragment duplication, whereas no tandem duplication; (2) they were under strong purification selection and evolutionary conserved among ancestral species, even among Gramineae family; (3) identical LIM orthologues were found in similar genomic regions among the A, B, and D sub-genomes and their corresponding ancestral species. TaLIMs are functionally annotated to GO term “zinc ion binding” and predicted to be localization in the nucleus. TaLIMs showed variable expression patterns and certain members are responsive to biotic and/or abiotic stresses, and growth and development conditions. Such as *TaLIM1-1A*, *TaLIM3-2B*, *TaLIM8-4D*, and *TaLIM10-5D* were significantly induced by heat, drought, NaCl, ABA and *Fusarium graminearum* stresses. Furthermore, the biological function of TaLIM8-4D were systemically analyzed and results revealed that TaLIM8-4D showed nucleus localization and is induced by hemi-biotrophic agent *F. graminearum*. It is a weak cell death inducer and can up-regulate the expression of *PR* genes, whereas promote the infection of hemi-biotrophic pathogen. Summarily, TaLIM8-4D could function as susceptible gene in nucleus through up-regulating *PR* genes and inducing cell death to promote the colonization of hemi-biotrophic *F. graminearum*.

Methods

Identification of LIM gene family members from *Triticum aestivum* L. genome

The wheat reference genome (IWGSC v2.1) was downloaded from Wheat Whole Genome Database (<https://wheat-urgi.versailles.inra.fr/Seq-Repository/Assemblies>). The Hidden Markov Model (HMM) of LIM domain (PF00412) was downloaded from Pfam (v32.0, <http://pfam.xfam.org/>) and used as query sequences to perform the first round BLASTp against wheat reference protein sequences (e-value < 10^{-10}). Meanwhile, 34 reported LIMs protein sequences, including 14 LIMs from *Arabidopsis thaliana* (AtLIMs, TAIR10, <http://www.arabidopsis.org/index.jsp>), 9 LIMs from *Oryza sativa* (OsLIMs, RGAP, <http://rice.plantbiology.msu.edu/index.shtml>) [11], and 11 LIMs from *Zea mays* L. (ZmLIMs, maizeGDB, <https://www.maizegdb.org/>) [12], were collected and used as a query sequences to perform the second round BLASTp. After merging two rounds of BLASTp searching results and deleting redundant sequences, unique ones were further validated through Pfam (v32.0, <http://pfam.xfam.org/>) and InterProScan (v71.0, <http://www.ebi.ac.uk/InterProScan>) to determine whether sequences contain PF00412 domain and exclude sequences that do not contain LIM domain.

Phylogenetic analysis of wheat LIMs (TaLIMs)

Multiple sequence alignments (MSA) of the identified LIM protein sequences were performed using the ClustalW2 program through the Neighbor-Joining (NJ) method with 1000 replicated bootstraps [32]. Phylogenetic tree, containing TaLIMs, AtLIMs, OsLIMs, and ZmLIMs, was further illustrated using ITOL tools (Interactive Tree of Life, <http://itol.embl.de>) to infer the phylogenetic relationships of TaLIMs and assisting the nomenclature of them.

Chromosomal mapping, gene duplication and Ka/Ks analysis

Detailed information of corresponding TaLIMs were extracted from GFF3 file of reference genome. The software MapInspect was used to mark the start and end positions of TaLIMs to each chromosome [33]. Gene duplicates are divided into tandem duplicates and fragment duplicates. Tandem repeated events were identified using the following assessment criteria: (1) length of the aligned sequence > 80% of the length of each sequence; (2) identity > 80%; (3) threshold $\leq 10^{-10}$; (4) only one duplication can be admitted when genes are linked closely; and (5) intergenic distance is less than 25 kb. If genes satisfied criteria (1), (2) and (3) and located on different chromosomes, they were judged as segmental duplications [18]. The software TBtools was used to calculate the non-synonymous substitution rate (Ka) and synonymous substitution rate (Ks) of TaLIM duplication pairs.

Gene structure, motif and *cis*-element analysis of TaLIMs

According to GFF3 annotation, the gene exon/intron structures of TaLIMs was drawn by TBtools. TaLIMs protein structures were analyzed by Pfam and results were used to build protein functional domains (Biological Sequence, <http://ibs.biocuckoo.org/>) [34]. The online website MEME (Version 5.3.1, <http://meme-suite.org/tools/meme>) was used to identify the conserved motifs from TaLIMs according to the following conditions. Each sequence could contain any number of non-overlapping motifs, the number of different motifs was 15, and the width of motifs ranged from 6 to 50 aa (amino acid). The output results were explored by TBtools to draw the motif structures of TaLIMs. Function of these motifs were further predicted using the online tool in Pfam. The upstream promoter sequences (1-1,500 bp) were extracted manually from genome sequence and uploaded into PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) to identify the *cis*-elements in promoter regions. Then R package pheatmap was used to sort out and display the analysis results [35].

Protein characterization and homology modeling of TaLIMs

The protein basic characteristics of the TaLIMs, including the length of the protein, molecular weight (MW), isoelectric point (pI), grand average of hydropathicity (GRAVY), were analyzed using online website ExPASy Server10 (<https://prosite.expasy.org/PS50011>). Signal peptide was predicted with SignalP 5.0 (<http://www.cbs.dtu.dk/services/SignalP/>) and subcellular localization was predicted with Plant-mPLOC

(http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/?tdsourcetag=s_pcqq_aiomsg). The three-dimensional homology modeling of TaLIMs were performed using online server SWISS-MODEL (<https://www.swissmodel.expasy.org/>).

Expression pattern profiling of TaLIMs

Wheat transcriptome sequencing data were downloaded from the NCBI Short Read Archive (SRA) database and compared to the wheat reference genome using Hisat2 [36]. Then, the expression levels of *TaLIMs* genes were calculated and normalized by Cufflinks (represented by TPM, Transcripts Per Kilobase of exon model per Million mapped reads) [37]. $\text{Log}_2(\text{TPM}+1)$ values were used to draw heatmap by R package “pheatmap” to show the expression pattern of *TaLIMs* gene under different conditions [38].

Inter-species evolution analysis of TaLIMs

The homologous of LIM genes among *Triticum urartu*, *Triticum dicoccoides* and *Aegilops tauschii* were identified by reciprocal BLASTp analysis [24]. The cutoff values (e-value < 10^{-10} , identity > 80%) were used to ensure the reliability of homologous. The R package “circlize” was used to indicate the chromosomal relationship between them.

Growth and stress treatment of wheat seedlings

Common hexaploid wheat cultivar “Yangmai 20”, which is moderate susceptible to powdery mildew and fusarium head blight, was surface-sterilized with 1% H_2O_2 , rinsed thoroughly with distilled water, and germinated in Petri dishes above three layers of filter paper saturated with water at 25°C for 2 days [39]. Then seedlings were transferred and cultured in 1/4 strength Hoagland solution and increased to 1/2 after 3 days [39]. *Fusarium graminearum* (PH-1) was routinely cultured on potato dextrose agar (PDA) plate and then cultured in liquid mung bean medium (60 g/L) with 25°C and 150 rpm to produce spores. When coleoptiles reached 0.5 cm, seedlings were soaked in spore suspension (5×10^5 spores/ml) for 1 min [40], then were wrapped in a piece of moist paper towel and cultured at 25°C with relative humidity 65%. Seedlings coleoptiles were sampled at 1, 2, 3 and 5 d after inoculating. When seedlings grow to one leaf and one heart stage, 150 mM NaCl was used to treat plants. The pH of the nutrient solution was adjusted to 6.0 every 2 days using 1 M KOH or 0.2 M H_2SO_4 . Growth conditions are 16 h/8 h (day/night) photoperiod at 25/20°C. Leaves were collected at 2, 6, 12, 24 and 48 h after treatment. For ABA (abscisic acid) treatment, 1/2 strength Hoagland solution containing 100 μM ABA were used to culture the wheat seedlings under conditions as mentioned above [41]. After 2, 6, 12, 24 and 48 h, leaves samples were harvested respectively. For heat and drought treatment, germinated seeds were planted in a mix of peat moss and vermiculite (V/V = 1:1) and cultured under 25°C day/23°C night with 16 h light/8 h dark photoperiod until flowering stage. Then one group of plants were placed in a light incubator under heat at 16 h light/8 h dark photoperiod. Another group was irrigated with water containing 20% PEG6000 to

induce drought treatment [42]. The pollen tissues were harvested after 12, 24, 48, and 72 h treatments, respectively. All samples were frozen immediately with liquid nitrogen and saved at -80°C until use. Each sample includes three biological replications.

RNA extraction and Quantitative-real time PCR analysis (qRT-PCR)

Total RNA was extracted using the TRizol reagent (GenStar, Beijing, China) and digested with DNaseI (TaKaRa, Dalian, China) to remove DNA. RNA was reverse-transcribed to cDNA using RevertAid Reverse Transcriptase (Vazyme, Nanjing, China). Then the cDNA was diluted to 100 ng/μl with enzyme-free water. The qRT-PCR was carried out on a CFX 96 Real-Time PCR system (Bio Rad, Hercules, CA, USA) using SYBR Green Master Mix (Vazyme, Nanjing, China) following manufacturer's instruction [43]. Gene-specific primers were designed using software Primer Premier 5 and were listed in Additional file 1: Table S9. The 20 μL qRT-PCR reaction system consisted of 10 μL of 2 × SYBR Premix Extaq, forward and reverse primers 0.4 μL each, and 2 μL diluted cDNA and 7.2 μL ddH₂O. The protocol was carried out as following: pre-denaturation at 95°C for 30 s (step 1), denaturation at 95°C for 5 s (step 2), and primer annealing/extension and collection of fluorescence signal at 60°C for 30 s (step 3). The next 40 loops started in step 2. Three technical replicates were used for each cDNA. Relative expression levels were determined using the $2^{-\Delta\Delta C_t}$ method [44]. ADP-ribosylation factor Ta2291 (Forward: GCTCTCCAACAACATTGCCAAC, Reverse: GCTTCTGCCTGTACATACGC), whose expression was stable under various conditions, was used as an internal reference gene for qRT-PCR analysis [45].

Subcellular localization analysis and functional annotation of TaLIMs

Full-length primers for *TaLIM8-4D* were designed using Primer Premier 5 (Forward: TTTGGAGAGGACACGCTCGAGATGTTTCAGCGGGACGCAG, Reverse: ATAATCCATGAATTCCTCGAGGGAGGATTTCAGCAGCCGC). Then, fragment was amplified by Phanta HS Master Mix (Vazyme, Nanjing, China) and inserted into the *Xho*I (NEB, Beijing, China) digested vector pART27:GFP by using ClonExpress II One Step Cloning Kit (Vazyme, Nanjing, China). Fusion construct was transformed into *Agrobacterium tumefaciens* strain GV3101 and the pART27:GFP transformation was used as negative control. GV3101 carrying vectors were injected into *Nicotiana benthamiana* leaves [45]. Three days later, leaves fluorescence was visualized through confocal microscopy (TCS SP8, Germany) with 488 nm exciting light wavelength. Gene Ontology database (GO) was used to perform functional annotation of TaLIMs [46].

Transient agro-infiltration assays

The pART-27:TaLIM8-4D:GFP fusion vectors were transient transformed into *N. benthamiana* plants by *A. tumefaciens* strain GV3101, meanwhile the pART27:GFP transformation was used as negative control. The agroinfiltration leaves were collected after 1, 2, 3, 4, and 5 d treatment and quickly frozen in liquid nitrogen and stored at -80°C for RNA extraction. Then the relative expression levels of *pathogenesis-related* (*PR*) genes were calculated with the $2^{-\Delta\Delta Ct}$ method and the primer sequences of these genes were shown in Additional file 1: Table S10. Furthermore, the phenotype of pART-27:TaLIM8-4D:GFP was assessed in response to pathogen attack (*Phytophthora infestans* strain “88069”). *P. infestans* was cultured at 18°C in darkness for 2 weeks on RSA (rye sucrose agar) [47]. The sporangia suspension was prepared with 5 mL distilled water, and the concentration of the sporangia suspension was adjusted to 4×10^4 sporangia/mL under the microscope. The spore suspension was placed in a refrigerator at 4°C for 2 hours to release zoospores. Agro-infiltration experiments were performed on leaves of 6- to 8-week-old *N. benthamiana*. After 24 h, infiltrated *N. benthamiana* leaves were detached and transferred into moist sealed plastic trays and inoculated with 10 μ l *P. infestans* zoospore suspension (100 zoospores/ μ l) at the infiltration sites. The inoculated leaves were incubated at 18°C in darkness. Photographs were taken under ultraviolet light and lesion diameters of the inoculated leaves were recorded 4 days after inoculated. Plants were cultivated in cultivation room with temperature of 22–25°C and under 16/8 h day/night photoperiod. Assay consisted of three biological replications, each biological replication contained ten technical duplications [48].

Abbreviations

TFs: Transcription factors; LIM: Lin-11, Isl-1 and Mec-3 domains; CRP: cysteine rich proteins; TaLIM: *Triticum aestivum* LIM; GRAVY: grand average of hydropathicity; 3D: three dimension; PEG: Polyethylene glycol; ABA: Absciscic acid; NaCl: Sodium chloride; qRT-PCR: Quantitative real time PCR; GO: Gene Ontology; *PR*: *pathogenesis-related*; PTI: triggered immunity; MSA: Multiple sequence alignments; NJ: Neighbor-Joining; ITOL: Interactive Tree of Life; Ka: non-synonymous substitution rate; Ks: synonymous substitution rate; aa: amino acid; bp: base pair; MW: molecular weight; pI: isoelectric point; SRA: Short Read Archive; TPM: Transcripts Per Kilobase of exon model per Million mapped reads; PDA: potato dextrose agar; RSA: rye sucrose agar.

Declarations

Acknowledgements

We would like to thank Yiqin He and Wendi Huang for help in bioinformatic work, Xingping Zhang for assisting in the confocal microscopy studies.

Authors' contributions

JLY and YTL conceived the study. YTL, YXZ, YQL and JLY designed the experiments and wrote the manuscript. YTL, YXX and XL performed the plant growth and sampling. YTL, SH and LJY carried out the qRT-PCR analysis. YTL, XWH and YW contributed to the subcellular localization and the infection of *Phytophthora Infestans*. All authors reviewed and approved the final manuscript.

Funding

This work was partially supported by the Open Project Program of Key Laboratory of Integrated Pest Management of Crops in Central China, Ministry of Agriculture/Hubei Key Laboratory of Crop Diseases, Insect Pests and Weeds Control (2019ZTSJJ8).

Availability of data and materials

All datasets supporting the conclusions of this article are included within the article (and its Additional files). The genome data and sequences and expression profiles of *TaLIMs* genes used in the current study are available in the Wheat Whole Genome Database (<https://wheat-urgi.versailles.inra.fr/Seq-Repository/Assemblies>). The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

The seed materials of the cultivars ‘Yangmai 20’ and ‘*Nicotiana benthamiana*’ used in this study were commercially available and collected from market. Plant materials used in the analysis are maintained in accordance with the institutional guidelines of Yangtze University, Hubei, China. This article did not contain any studies with human participants or animals and did not involve any endangered or protected species.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

References are not available with this version

Figures

Figure 1

Phylogenetic tree of *Triticum aestivum*, *Zea mays*, *Oryza sativa*, and *Arabidopsis thaliana* LIMs.

Figure 2

Chromosomal distribution of TaLIMs. Chr represents the chromosome. The ruler on the left represents the physical map distance among genes (Mbp). The start and end information of 28 TaLIMs on chromosomes was shown in Additional file 1: Table S3.

Figure 3

Predicted 3D models of TaLIMs. Models were generated by using online website SWISS-MODEL.

Figure 5

Conserved motifs of 28 TaLIMs. (a) Phylogenetic tree of TaLIMs. (b) Motif analysis of TaLIMs. Motifs were represented in different colors. (c) Conserved amino acid sequences of 15 motifs in TaLIMs.

Figure 6

Statistical analysis of cis-acting elements in the promoter of TaLIMs. The color and number of the grid represent the number of cis-acting elements in corresponding TaLIMs.

Figure 9

Expression patterns of four TaLIM genes under various treatments. The treatments included heat, drought, 150 mM NaCl, 100 μ M ABA, and *Fusarium graminearum*. The Y-axis represents the relative expression levels and the X axis shows the time points of stresses treatment. * indicate significant differences in comparison with the control at $p < 0.05$, respectively. ** indicate significant differences in comparison with the control at $p < 0.01$, respectively.

Figure 10

GO annotation and the subcellular localization of TaLIM8-4D. (a) GO (Gene Ontology) annotation of TaLIMs. * indicates TaLIM8-4D was annotated under the GO terms. (b) Subcellular localization of GFP. (c) Subcellular localization of TaLIM8-4D. Photographs are taken using confocal microscopy. Scale bar: 25 μ m.

Figure 11

Function analysis of TaLIM8-4D. (a) Overexpression of TaLIM8-4D promoted *Phytophthora infestans* colonization. Photographs were taken under ultraviolet headlights. The white circle represents the area of the lesion. (b) lesion diameter of control (GFP) and TaLIM8-4D. ** indicates significant differences at the level of 0.01. Scale bar: 1 cm. (c) TaLIM8-4D induced weak cell death.

Figure 12

Relative expression level of pathogenesis-related (PR) genes in transient overexpression of TaLIM8-4D plants. Each bar represents the average of three replicates, and error bar represents standard deviation (SD).

Supplementary Files

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

- [Additionalfile1.xlsx](#)