Phyton-
International Journal of
Experimental Botany

Tech Science Press

DOI: 10.32604/phyton.2022.021406

ARTICLE

Wheat Lysin-Motif-Containing Proteins Characterization and Gene Expression Patterns under Abiotic and Biotic Stress

Minjie Liu*, Na Gao, Yaqiong Zhao, Yingpeng Wu and Zongying Yuan

College of Plant Protection, Shanxi Agricultural University, Taiyuan, 030031, China

*Corresponding Author: Minjie Liu. Email: liuminjie@sxau.com

Received: 12 January 2022 Accepted: 11 April 2022

ABSTRACT

Lysin motif (LysM)-containing proteins (LYPs) are important pattern recognition receptors in plants. However, the evolutionary history and characteristics of LYP genes remain largely unclear in wheat. In this study, 62 LYPs were identified at genome wide in wheat. Based on phylogenetic and domain analysis, wheat LYPs were classified into 6 subgroups (group LysMe, LysMn, LYP, LYK, LysMFbox). Syntenic analysis showed the evolution of LYP genes in wheat. RNA-seq data showed that 22 genes were not expressed at any tissue or stress stimulation period. Some *LYP* and *LYK* genes were tissue- or stage- specific. The majority of *TaLYK5s*, *TaLYK6s*, *TaLYP2s* and *TaLysMns* genes were induced under chitin, flg22 and fungal treatment. qRT-PCR analysis showed that 4 genes were upregulated during *Puccinia triticina* infection with a peak at 18 h post inoculation. Our findings suggested that wheat LYPs may have specific roles in response to fungal infection and provided insights into the function and characteristics of wheat LYP genes.

KEYWORDS

Wheat; lysin motif containing protein; evolution; expression pattern

1 Introduction

Plants are subjected to a wide range of stresses which reduces and limits the productivity of agricultural crops [1,2]. Plants have evolved sophisticated innate immune system to deal with various stimuli [3,4]. Innate immune signaling of plants is initiated by perception of pathogen-associated molecular patterns (PAMPs) by pattern recognition receptors (PRRs) [5–7]. Plant PRRs are either surface-localized receptor kinases or receptor-like proteins containing various ligand-binding ectodomains that perceive PAMPs [8,9].

Lysin motif (LysM), usually about 40 amino acids, is a widely distributed protein motif in prokaryotes and eukaryotes [10]. LysM-containing proteins (LYPs) are important PRRs in plants, which function in the perception of PAMPs and in defense against pathogenic attack [11]. Plant LYPs are also essential molecules for the signaling in root nodule and arbuscular mycorrhizal formation [12]. LYPs have been widely studied in a range of plants, including *Arabidopsis thaliana* (L.) Heynh, *Oryza sativa* L. and so on [13–16]. Rice chitin elicitor receptor kinase (OsCERK1) regulated both chitin-triggered immunity and arbuscular mycorrhizal symbiosis [17]. Rice chitin elicitor binding protein (OsCEBiP) binds chitin oligosaccharides with the extracellular region and forms a complex with OsCERK1 to induce immune signaling [18]. OsLYP4 and OsLYP6 can bind peptidoglycan (PGN) and chitin [19]. In Arabidopsis, LysM receptor-like kinases,

namely LYK1/CERK1 (CHITIN ELICITOR RECEPTOR KINASE 1), LYK4 and LYK5, play a major role in chitin perception and immunity against pathogenic fungi [20]. AtLYK5 is the primary chitin receptor and forming a chitin inducible complex with AtCERK1 to induce plant immunity [21]. LYK4 functions as a LYK5-associated co-receptor or scaffold protein that enhances chitin-induced signaling [22]. AtLYM1 and AtLYM3 (homologs of OsLYP4 and OsLYP6) are required for peptidoglycans sensing in bacteria [23].

Wheat (*Triticum aestivum* L.) is one of the most important crops worldwide [24–26]. Bread wheat is a hexaploid which originated from three diploid ancestors: *Triticum urartu* Tum. (A genome), *Aegilops speltoides* Tasch. (B genome) and *Aegilops tauschii* Coss. (D genome) making the genome more complex [27]. The recent high-quality genome annotation [28] and large scale of RNA-seq datasets [29] provide an opportunity to conduct homologous expression to better understand the expression patterns under a variety of conditions. In this study, we identified and characterized the lysin motif contained proteins (LYPs) family members at genome wide in wheat. We investigated the phylogenetic relationships, chromosomal locations, synteny relationship and expression patterns of by employing bioinformatics and publicly available data. We further investigated the expression pattern of selected genes during wheat leaf rust infection. Taken together, our studies provide a set of LYPs that have potential for further studies in plant immunity and genetic modifications of resistance in wheat.

2 Materials and Methods

2.1 Sequence Detection of LYPs in Wheat Genome

Wheat coding sequence (CDS), and functional annotations of wheat genes were downloaded from IWGSC archive v.1.1 (https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Annotations/v1.1/). Functional annotations were filtered for Protein family database (Pfam) identifiers of the LysM domain (PF01476). The proteins with at least one LysM domain were selected as family member. A total of 65 sequences were selected. Splice variants were excluded and only the longest variant was kept for further analysis. Transmembrane helices were predicted by TMHMM Server v 2.0 [30]. The subcellular localization of these proteins was predicted by TargetP-2.0 Server (http://www.cbs.dtu.dk/services/TargetP/) [31]. The pI and MW were calculated by Expasy's ProtParam (https://web.expasy.org/protparam/). The conserved domain was predicted by PFAM (http://pfam.xfam.org/) [32].

2.2 Phylogenetic Analysis of LYP Genes

Protein sequences of *A. thaliana*, *O. sativa*, *Brachypodium distachyon* (L.) P. Beauv. and *T. aestivum* were aligned by MUSCLE. The phylogenetic tree was constructed by using neighbor-joining and maximum likelihood method with default parameters in Mega 7.0 [33].

2.3 Naming of LYP Genes

We suggest a consistent naming pattern for all LysM genes, considering of their subgroup, phylogenetic relationships, motif contained as well as their subgenome location (A, B or D). Each gene name starts with an abbreviation of *T. aestivum* (*Ta*).

2.4 Chromosomal Localization and Synteny Analysis

The genome annotation information of wheat was used to analyzed the chromosomal localization. The local blast searches of wheat and itself was conducted for considerable pairs of homologous genes. Then TBtools was employed to perform synteny analysis. Ka, Ks, and Ka/Ks values of wheat LYP gene pairs were calculated [34].

2.5 Expression Analysis of LYP Genes

Expression level was downloaded from Wheat Expression Browser (http://www.wheat-expression.com/). We considered a gene expressed when its average expression per treatment was >0.5 tpm in at least one treatment. A heatmap was generated by R studio. The expression data of development was calculated as log2 [(transcript per million)+1] [35]. Genes were clustered according to their expression using K-means. The expression data of biotic and abiotic treatment was normalized by control and calculated as log2 (relative amount of expression). The expression data of biotic treatment contained powdery mildew, stripe rust and wheat head scab.

The relative expression levels of the A, B and D subgenome were analyzed only when the gene triads across the three subgenomes comply with 1:1:1. To standardize the relative expression of each homolog across the triad, we normalized the absolute tpm to the expression ratio of three subgenome. The tenary diagram was plotted by R package ggtern [36].

For qRT-PCR, wheat leaves inoculated with *Puccinia triticina* (race PHJ), the causal agent of wheat leaf rust, were harvested at 0, 6,18,48,168 h post-inoculation (hpi) for detection of the transcript levels of selected wheat *LYP* genes. The infected leaves were ground in liquid nitrogen, and RNA was isolated using MiniBEST Plant RNA Extraction Kit (TaKaRa) following the manufacturer's instructions. qRT-PCR was carried out using Bio-Rad CFX 96, and the $2(^{-\Delta\Delta Ct})$ analysis method was used to determine the relative expression levels of selected genes. Wheat GADPH (GenBank No. AF251217) was used as an internal reference gene [37]. Three independent biological replicates were performed per treatment. Primers were listed in Appendix A.

3 Results

3.1 Identification and Classification of LYP Genes in Wheat

A total of 65 coding sequences were identified by HMM search using HMM profile (PF01476) in IWGSC wheat genome (Appendix B). The dataset was simplified by retaining 62 coding genes with only one splice variant from each genomic locus for further analysis (Table 1).

In order to classify the LYP genes, phylogenetic analyses were performed and the domains contained in these LYPs were considered simultaneously (Appendix B). Wheat LYP proteins were separated into six major groups: LYP, LYK, LysMn, EMSA, LysMe and LysMFbox. All the proteins have at least one LysM domain. In addition to LysM domain, LYP family contain an another LysM domain. Protein kinase domain (Pfam ID: PF00069) or protein tyrosine and serine/threonine kinase domain (Pfam ID: PF07714) were the characteristic for LYK family. LysMn family was predicted to be intracellular protein. EMSA family members were homologs of OsEMSA1. LysMe family contained proteins with only one LysM domain and have signal peptide. F-box domain (Pfam ID: PF00646) was the characteristic of LysMFbox family.

The characteristics of the wheat LYPs are shown in Table 1. The lengths of LYP protein sequences ranged from 100 to 749 amino acids and the molecular weights were 10.31 to 80.56 kD. Protein isoelectric points (PI) ranged from 4.57 to 9.29. The majority of the LYPs were predicted as secreted protein. The grand average of hydropathy (GRAVY) for LysMe and most of LYP subgroup was positive indicating hydrophobic character, while that of LysMFbox, EMSA, LysMn and most of LYK subgroup as negative indicating hydrophilic character.

3.2 Phylogenetic Analysis of LYPs from Wheat, Arabidopsis, Rice and Brachypodium

To analyze the phylogenetic relationships of LYPs from different species, lysin motif contained protein sequences from *A. thaliana*, *O. sativa*, *B. distachyon* and *T. aestivum* were used to construct a neighbor-joining tree. As shown in Fig. 1A, wheat LYPs shared high homology with that from other species. Regardless of species, LYK was the largest group.

Table 1: Detailed information about 62 lysin motif contained proteins in *Triticum aestivum* (L)

Gene Id	Gene name	Chr[1]	Start[2]	End[3]	strand	Prot (aa)[4]	Exon	TargetP[5]	TMH[6]	pI[7]	MW (kDa)[8]	GRAVY[9]
TraesCS3B02G352100.1	TaLysMe1-3B1	3B	561840722	561841388	+	101	2	SP	1	5.56	10.3659	0.547
TraesCS3B02G351900.1	TaLysMe1-3B2	3B	561592592	561592992	+	101	2	SP	1	5.56	10.33789	0.552
TraesCS3A02G314900.1	TaLysMe2-3A	3A	556340109	556340688	−	103	2	SP	1	6.52	10.64526	0.503
TraesCS3A02G315000.1	TaLysMe3-3A1	3A	556353562	556353963	−	102	2	SP	1	5	10.50808	0.520
TraesCS3D02G316500.1	TaLysMe1-3D	3D	429637289	429637682	+	102	2	SP	1	5.56	10.43901	0.579
TraesCS3A02G314800.1	TaLysMe3-3A2	3A	556260455	556260849	−	101	2	SP	1	5.04	10.49296	0.362
TraesCS3B02G352200.1	TaLysMe3-3B	3B	561889196	561889667	+	102	2	SP	1	5.61	10.66622	0.368
TraesCS3D02G316600.1	TaLysMe3-3D	3D	429647452	429648203	+	102	2	SP	1	5.04	10.60612	0.421
TraesCS3A02G316100.1	TaLysMe5-3A	3A	556683215	556683592	+	125	1	SP	1	6.68	13.07622	0.569
TraesCS3D02G315300.1	TaLysMe5-3D	3D	428993996	428994373	−	125	1	SP	1	6.22	12.95703	0.526
TraesCS3B02G350800.1	TaLysMe5-3B	3B	561038700	561039083	−	100	2	SP	1	6	10.31383	0.461
TraesCS3D02G315500.1	TaLysMe4-3D1	3D	429085177	429085783	−	104	2	SP	1	5.55	10.63323	0.540
TraesCS3D02G315400.1	TaLysMe4-3D2	3D	429043817	429044231	−	101	2	SP	1	6.69	10.33991	0.501
TraesCS3B02G350900.1	TaLysMe4-3B	3B	561044324	561044733	−	101	2	SP	1	6.69	10.32589	0.505
TraesCS3A02G316000.1	TaLysMe4-3A	3A	556678445	556678976	+	101	2	SP	1	6.68	10.31186	0.498
TraesCS4B02G038000.1	TaLysMFbox-4B	4B	27526533	27528622	−	245	2	OTHER	0	8.86	26.88542	−0.342
TraesCS4A02G275700.1	TaLysMFbox-4A	4A	584671905	584675986	+	248	2	OTHER	0	9.29	27.21388	−0.338
TraesCS4D02G035300.1	TaLysMFbox-4D	4D	15927016	15929136	−	245	2	OTHER	0	8.5	26.91644	−0.371
TraesCS5A02G347800.1	TaLYP1-5A	5A	550718922	550721627	+	366	4	SP	0	8.28	37.68609	0.245
TraesCS5B02G348800.1	TaLYP1-5B	5B	530262150	530264999	+	366	4	SP	0	8.28	37.71809	0.230
TraesCS5D02G354000.1	TaLYP1-5D	5D	436257553	436260258	+	370	4	SP	0	8.28	38.05448	0.226
TraesCS4B02G329500.2	TaLYP2-4B	4B	620056284	620059068	+	356	4	SP	0	7.83	37.10154	0.118
TraesCS4D02G326400.2	TaLYP2-4D	4D	486179743	486183283	−	360	4	SP	0	8.09	37.66021	0.159
TraesCS5A02G501100.1	TaLYP2-5A	5A	666740738	666743746	+	418	4	OTHER	0	8.98	43.58167	−0.045
TraesCS5A02G234700.1	TaLYP3-5A	5A	451012317	451015052	+	411	5	SP	0	5.81	41.44488	0.443
TraesCS5B02G233200.1	TaLYP3-5B	5B	411544324	411547275	+	410	5	SP	1	6.04	41.74732	0.427
TraesCS5D02G241600.1	TaLYP3-5D	5D	350592971	350595725	+	406	5	SP	0	5.56	41.3397	0.451
TraesCS7B02G073300.1	TaLYP4-7B	7B	81690952	81694741	+	401	5	SP	1	4.73	39.99348	0.417
TraesCS7D02G169400.1	TaLYP4-7D	7D	120321447	120325549	+	401	5	SP	1	4.57	39.96637	0.419
TraesCS7A02G168500.1	TaLYP4-7A	7A	125262698	125266532	+	401	5	SP	1	4.57	39.92531	0.425
TraesCS6B02G359500.1	TaLYP5-6B1	6B	631200466	631203676	−	420	5	SP	0	5.25	42.36345	0.336
TraesCS6A02G328800.1	TaLYP5-6A	6A	562374700	562378122	−	460	5	luTP	0	5.68	46.78852	0.288
TraesCS6B02G359300.1	TaLYP5-6B2	6B	630846675	630849877	+	420	5	SP	0	5.37	42.37548	0.342
TraesCS6D02G307700.1	TaLYP5-6D	6D	418739676	418742695	−	423	5	SP	0	5.12	42.73099	0.362
TraesCS4B02G353400.1	TaEMSA1-4B	4B	645084113	645084496	−	127	1	OTHER	1	6.27	13.48821	−0.154
TraesCS4D02G347400.1	TaEMSA1-4D	4D	501233795	501234169	−	124	1	SP	1	6.05	13.2852	−0.031
TraesCS1B02G195500.1	TaEMSA2-1B	1B	350630695	350631492	−	116	1	SP	0	5.75	12.27492	−0.028
TraesCS1A02G187700.1	TaEMSA2-1A	1A	338600466	338600819	−	117	1	SP	0	5.48	12.32601	0.015
TraesCS1D02G188900.1	TaEMSA2-1D	1D	261373559	261373912	+	117	1	SP	0	5.75	12.37002	−0.047
TraesCS7B02G486800.1	TaLysMn1-7B	7B	742588952	742591716	+	333	3	OTHER	0	9.02	35.59269	−0.596
TraesCS7D02G549400.1	TaLysMn1-7D	7D	634175263	634178042	+	328	3	OTHER	0	8.35	35.01197	−0.636

(Continued)

Table 1 (continued)

Gene Id	Gene name	Chr[1]	Start[2]	End[3]	strand	Prot (aa)[4]	Exon	TargetP[5]	TMH[6]	pI[7]	MW (kDa)[8]	GRAVY[9]
TraesCS7A02G560400.1	TaLysMn1-7A	7A	732087221	732090195	+	327	3	OTHER	0	7.75	34.97595	−0.603
TraesCS3B02G588200.1	TaLYK3-3B	3B	814266347	814274743	−	593	4	SP	0	8.23	65.83633	−0.154
TraesCSU02G125700.1	TaLYK3-U	Un	108652929	108662267	+	550	5	SP	0	8.11	61.00381	−0.217
TraesCS5A02G552200.1	TaLYK3-5A	5A	705361926	705376743	−	593	3	SP	0	8.8	65.73226	−0.166
TraesCS3B02G403100.1	TaLYK1-3B	3B	637076174	637079536	+	615	10	SP	2	6.05	68.27747	−0.043
TraesCS3D02G364000.1	TaLYK1-3D	3D	477895788	477898692	+	615	10	SP	2	6.3	67.9702	−0.018
TraesCS3A02G370900.1	TaLYK1-3A	3A	621230384	621233286	+	612	10	SP	1	5.91	67.83787	−0.058
TraesCS6B02G266500.1	TaLYK2-6B	6B	479077207	479083290	−	716	2	OTHER	1	7	76.6752	−0.140
TraesCS6D02G240100.1	TaLYK2-6D	6D	341194264	341198640	+	714	2	OTHER	0	7.87	76.43803	−0.112
TraesCS6A02G258900.1	TaLYK2-6A	6A	481255757	481260121	+	648	2	SP	0	6.12	69.41229	−0.004
TraesCS7D02G056600.1	TaLYK4-7D	7D	30264305	30266686	+	652	1	SP	2	5.53	69.98053	0.087
TraesCS4A02G427200.1	TaLYK4-4A	4A	698169121	698172179	−	749	2	OTHER	0	5.45	80.56417	−0.009
TraesCS7A02G061600.1	TaLYK4-7A	7A	30651156	30653078	+	640	1	SP	1	4.89	68.03837	0.140
TraesCS6A02G168800.1	TaLYK6-6A	6A	175995613	176003026	−	603	2	SP	1	9.05	63.48015	−0.037
TraesCS6D02G157800.1	TaLYK6-6D1	6D	134335551	134337568	−	637	2	SP	3	8.13	67.49872	0.015
TraesCS6D02G158300.1	TaLYK6-6D2	6D	134740207	134742719	−	670	1	SP	3	8.46	70.89443	−0.005
TraesCS6B02G196800.1	TaLYK6-6B	6B	233086725	233089297	−	670	1	SP	3	8.46	71.05764	−0.012
TraesCS6B02G197000.1	TaLYK5-6B	6B	233671844	233674217	−	681	1	SP	1	8.14	71.9735	−0.032
TraesCS6A02G168900.1	TaLYK5-6A	6A	176291828	176294160	−	681	1	SP	1	7.9	71.89042	−0.029
TraesCS6D02G157900.1	TaLYK5-6D1	6D	134586113	134588583	−	681	1	SP	1	8.31	71.70018	−0.013
TraesCS6D02G158400.1	TaLYK5-6D2	6D	135039466	135041947	−	681	1	SP	1	7.88	71.6431	−0.001

Notes: [1]Chromosome location; [2,3]Genomic location; [4]Prot: protein length; aa: amino acid; [5]TargetP: TargetP-2.0 server predicts the presence of N-terminal presequences: signal peptide (SP), thylakoid luminal transit peptide (luTP); [6]TMH: TMHMM 2.0 server predicts the presence of transmembrane helices; [7]pI: isoelectric point; [8]MW: molecular weight; [9]GRAVY: grand average of hydropathy value.

The amount of group numbers was calculated for each species. Rice, Arabidopsis and Brachypodium, despite their phylogenetic distance, have a similar number of LYP genes (15, 11 and 13, respectively) (Figs. 1C–1E). However, the number of LYPs in wheat is as high as 62 (Fig. 1B). The number of LYP genes was 4-fold than those in *A. thaliana*, *O. sativa* and *B. distachyon* (Fig. 1F). One of the main reasons is that wheat is hexaploidy which the number of genes is theoretically three times that of other diploid species. When corrected for ploid level, gene expanding in LysMe, LYK and EMSA subgroup was the main reason (Fig. 1F).

3.3 Chromosomal Locations, Synteny and Evolution Analysis of Wheat LYPs

We found that 62 TaLYPs were localized on chromosomes (1A, 1B, 1D, 3A, 3B, 3D, 4A, 4B, 4D, 5A, 5B, 5D, 6A, 6B, 6D, 7A, 7B, 7D) and the unassembled scaffolds (Un). No gene was located on chromosome 2A, 2B and 2C (Fig. 2A).

In genetics, Ka/Ks represent the ratio between the nonsynonymous substitution rate (Ka) and the synonymous substitution rate (Ks) of two protein-coding genes. The value of Ka/Ks can be used as an indicator of selective pressure on a protein-coding gene. The Ka/Ks ratio was less than one almost in all gene pairs. The EMSA group showed relatively high Ka/Ks ratio, suggesting that these genes evolved at a faster evolutionary rate, which is a feature of new genes (Fig. 2B).

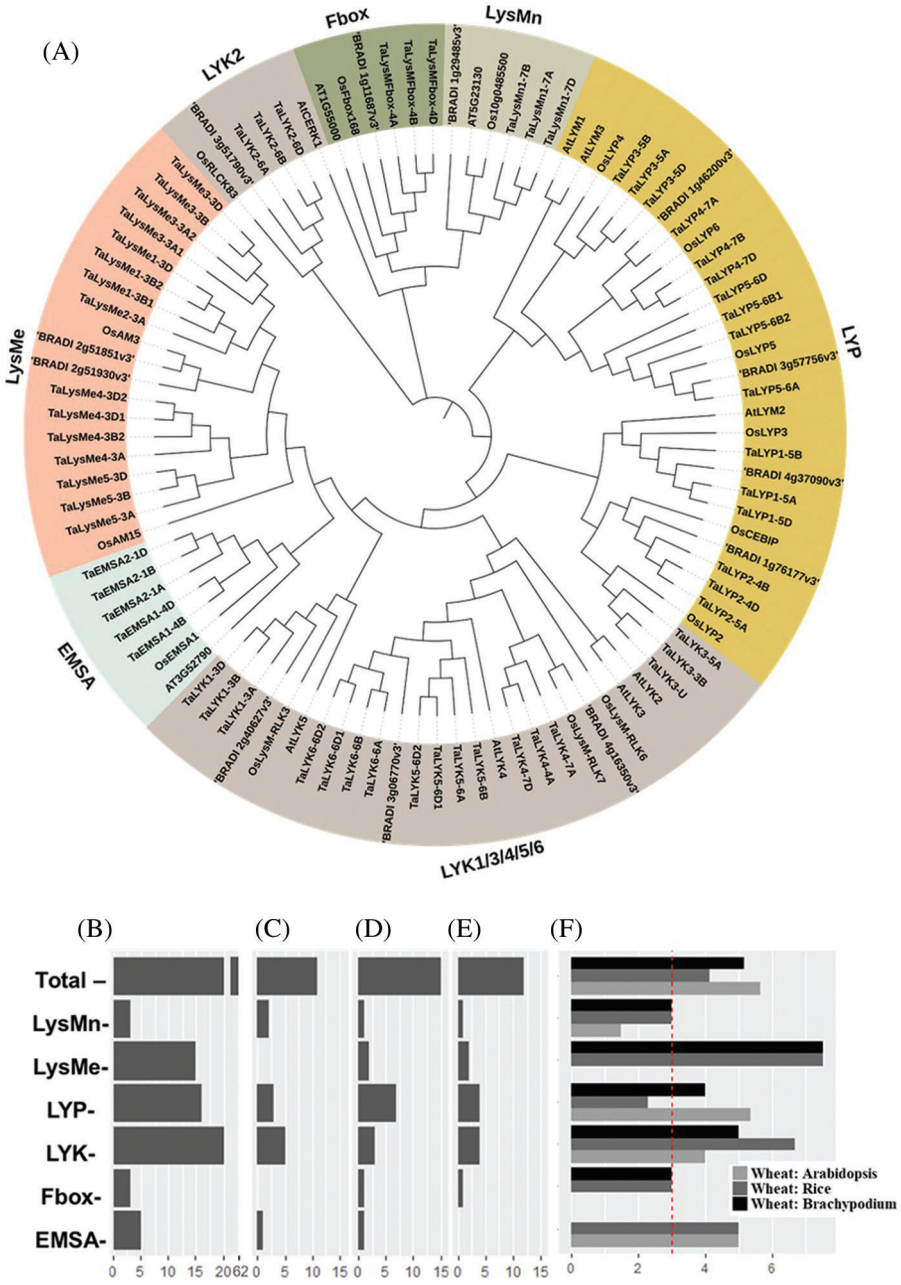

Figure 1: Phylogenetic analysis and number of LYP proteins from wheat, Arabidopsis, rice and Brachypodium. (A) A phylogenetic tree of LYPs from wheat, rice, Arabidopsis and Brachypodium was formed via mega 7.0 with the neighbor-joining method and was displayed in iTOL. Different background colors indicate the different groups of the LYP proteins. (B–E) The number of LYP genes was identified in each group in (B) wheat, (C) Arabidopsis, (D) rice and (E) Brachypodium. (F) The ratio of LYP genes in each group was shown in wheat: Arabidopsis (light grey), wheat: Rice (dark grey) and wheat: Brachypodium (black). The expected ratio (3:1) was indicated in a red dashed line

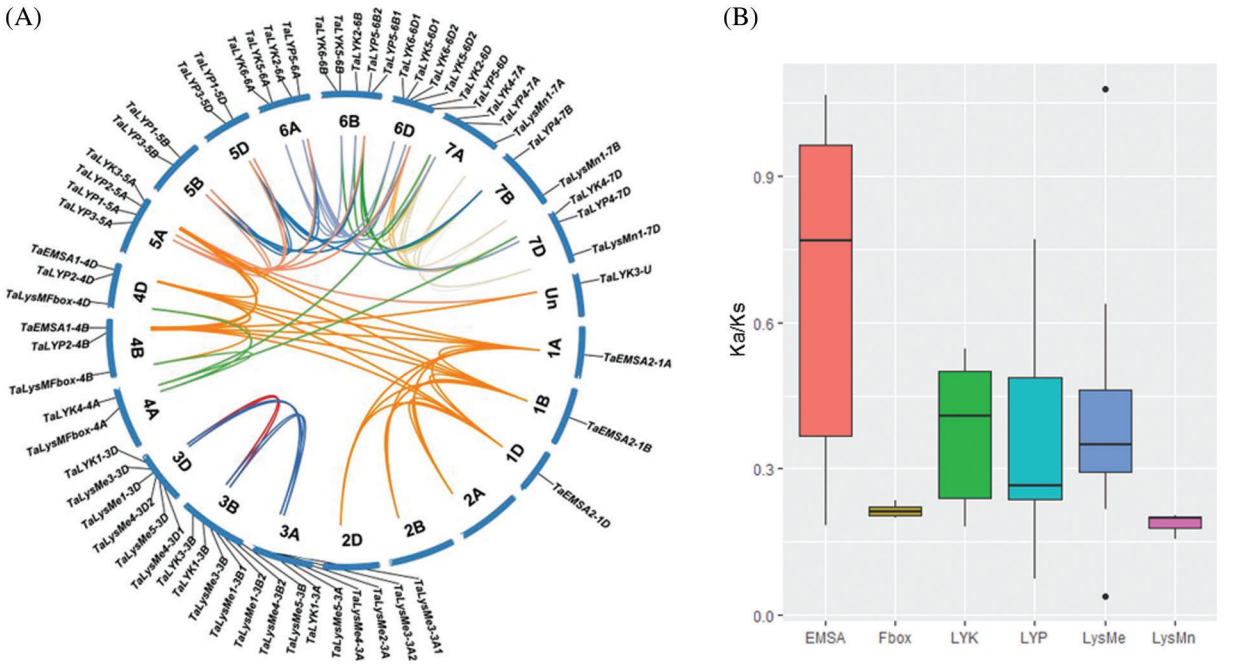

Figure 2: (A) Chromosomal localization and syntenic relationships of LYP genes in wheat. LYP genes are mapped on different chromosomes. Homologous genes are linked by lines. (B) The ratio of nonsynonymous to synonymous substitutions (Ka/Ks) of wheat LYPs in each group

8.39% of LYP genes triads across the three subgenomes were comply with 1:1:1. The percentage of LYP genes with homolog-specific duplication is much higher than in all wheat genes (32.26% *vs*. 5.76%) [38]. Loss of one homolog is less pronounced in LYP genes (3.23% *vs*. 13.22%) (Table 2).

Table 2: Groups of homologous *LYPs* in wheat

Homologous group (A:B:D)	All wheat genes (%)[1]	Gene number of LYPs	Composition of genes (%)
1:1:1	35.8	30	48.39
n:1:1/1:n:1/1:1:n[2]	5.76	20	32.26
0:1:1/1:0:1/1:1:0	13.22	2	3.23
Other ratio	8	9	14.52
Orphans/singletons	37.22	1	1.61
	100	62	100

Notes: [1]According to IWGSC (2018); [2]$n > 1$.

3.4 The Expression Patterns of LYP Gene during Wheat Development

36 genes (58%) of the 62 genes were expressed in at least one developmental stage (based on the >0.5 TPM criteria). The remaining 26 genes which tpm <0.5 were considered not expressed. The *LysMe* clade has been expanded during wheat evolution. Many of the genes from this clade were not expressed or expressed on a very low level. However, the expression of *LysMn* and *LysMFbox* genes showed no significant changes during development. The expression peak of the *EMSA* subclade genes appeared in

root at the reproductive stage. *TaLYP5* and *TaLYP1* clusters were expressed at seedling and vegetative period, but not at reproductive period. *TaLYK1* genes and *TaLYK4* genes were not expressed at any developmental period. *TaLYK3* were highly expressed on leaves at reproductive period and spike (Fig. 3A).

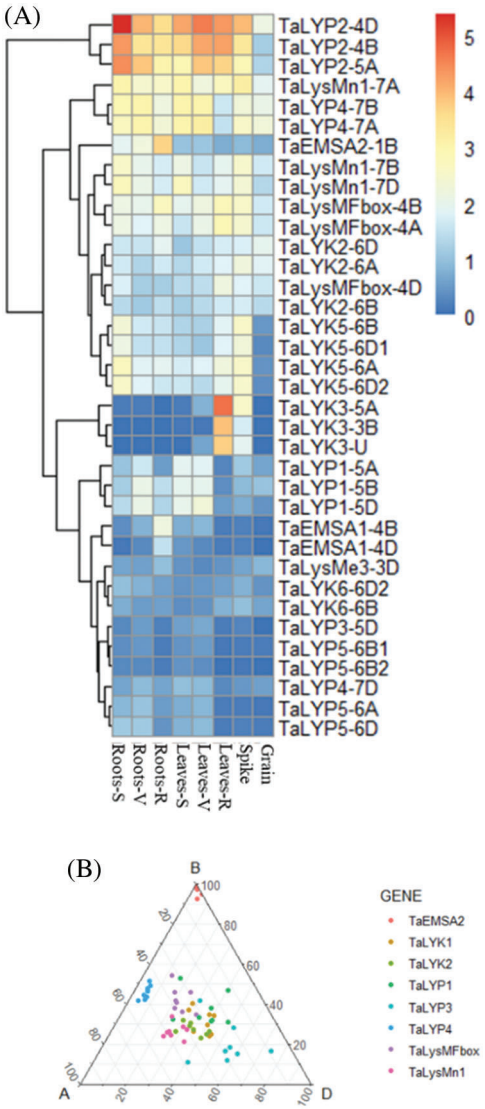

Figure 3: The expression patterns of LYP gene during wheat development. (A) Genes were clustered according to their expression using K-means. R: reproductive stage V: vegetative stage S: seedling stage (B) The Ternary plot showing relative expression abundance for all 1:1:1 wheat LYPs triad during wheat development. Colors in different circles represent subgroups

We also analyzed the expression pattern of each triad. *TaEMSA2* was expressed only from B subgenome. *TaLYP4* was B suppressed. *TaLysMFbox* was D-suppressed in root at reproductive period. *TaLYP1* was A-suppressed in root at reproductive period. *TaLYP3* was B-suppressed in root at vegetative period and in spike at reproductive period. *TaLYP3* was D-dominant in leaves/shoots at reproductive period (Fig. 3B).

3.5 Some Genes were Highly Expressed in Response to Abiotic and Biotic Response

We also analyzed the expression pattern of wheat *LYPs* to various biotic and abiotic stress (Fig. 4). Like the expression pattern during development, the *LysMe* subgroup were nearly not expressed upon various treatment. The expression of *TaLysMFbox* family was not changed significantly in response to various stress. For the *LysMn* family, genes were upregulated slightly when treated by chitin, flg22 or abiotic stress.

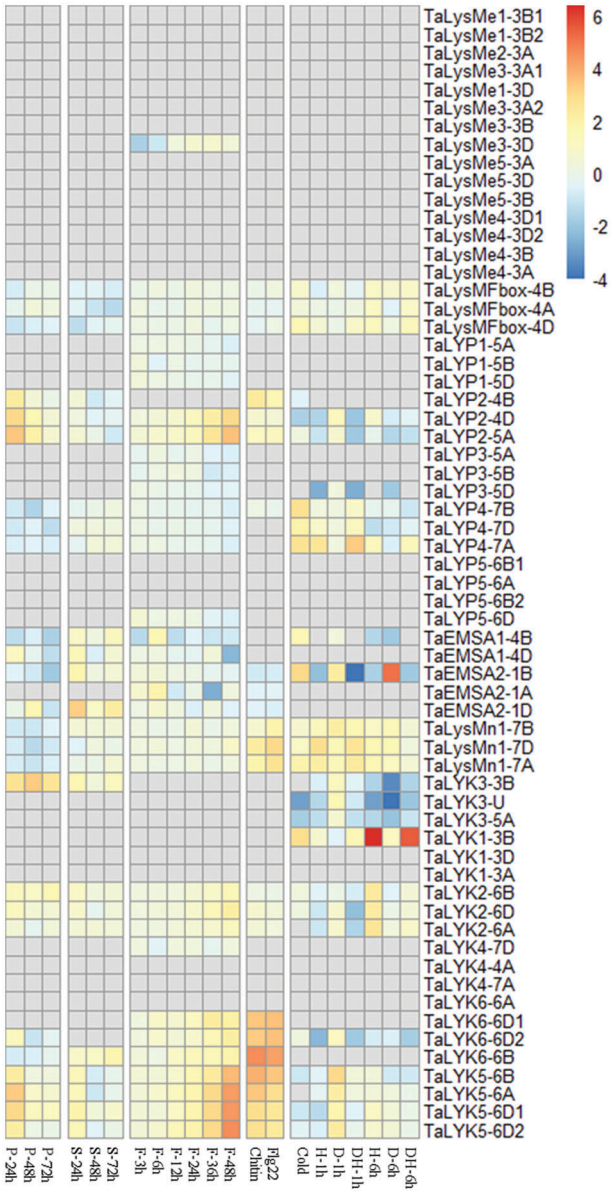

Figure 4: Expression of LYP genes in response to wheat powdery mildew ('P'), stripe rust ('S'), fusarium head blight ('F'), chitin, flg22, cold, drought ('D'), heat ('H') and combined drought and heat stress ('DH'). Grey squares: genes not expressed

Except for the genes not expressed, most of the wheat LYK family members were upregulated in response to biotic and abiotic stress. Interestingly, the expression of *TaLYK5s* and *TaLYK6s* (except

TaLYK6-6A) increased dramatically when treated by chitin and flg22, which may infer that they were important for PAMPs recognition. They were also slightly increased upon drought stress. *TaLYK2s* were induced by fungal infection and PAMPs treatment. *TaLYK1-3B* was dramatically increased when treated by heat or heat and drought after 6 h (Fig. 4).

The expression levels of *TaLYP2* clusters were upregulated when inoculated with fungal disease or PAMPs. In this study, *TaLYP3* and *TaLYP4* clusters were not significantly changed under biotic stress. However, *TaLYP4* gene cluster were induced by cold stress, drought, and heat treatment. Meanwhile, *TaLYP3* cluster were depressed when treated by drought and heat stress.

We further used qRT-PCR analysis to investigate the expression of selected genes in response to wheat leaf rust. The selected four genes were upregulated during wheat leaf rust infection with similar expression patterns. The expression of *TaLysMFbox* (Fig. 5A) and *TaLysMn* (Fig. 5B) increased steadily from 6 to 18 h, fell markedly at 48 h and increased slightly at 168 h. *TaLYK5* showed remarkably upregulated during the infection with a peak at 18 h (Fig. 5C). In particular, *TaLYK6* was upregulated 300 to 400-fold at 6 and 18 h upon wheat leaf rust infection (Fig. 5D).

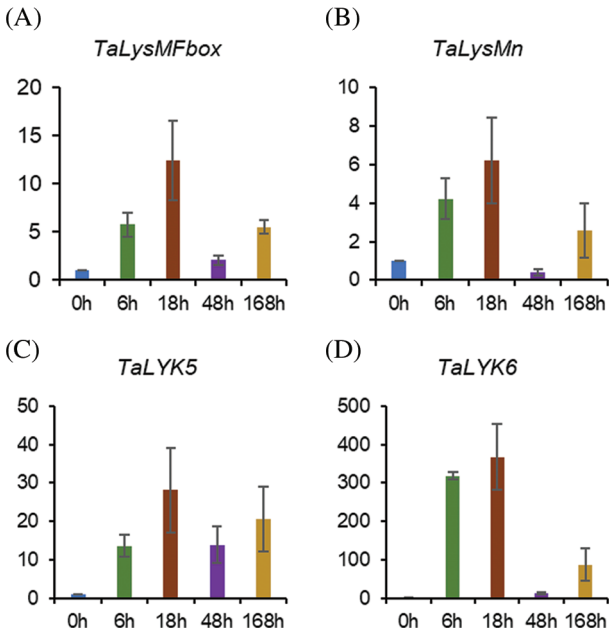

Figure 5: Expression pattern of 4 LYP genes in response to wheat leaf rust

4 Discussion

LYP genes have been widely studied in a range of plants, including monocots and dicots [13]. Previous studies have shown that LysM domain containing proteins are involved in plant-microbe interactions, glycine metabolism, embryo sac development and other biological processes [39,40]. In this study, sixty-two LysM domain containing proteins were identified in wheat genome. The number of LYP genes was 4-fold than those in *A. thaliana*, rice and *B. distachyon*, exceeding the expectations of hexaploidy. This is similar to the Rosaceae species, in which 13 to 21 LYP genes were identified [41]. We found that in wheat genome, genes were expanding in LysMe, LYK and EMSA subgroup, indicating that these groups may play diverse roles in the adaptive evolution to environmental stresses. The genes distributed not equally among the chromosomes, ranging from zero to seven. The chromosomes 3 and 6 contained more genes than expected from the chromosome lengths. This is mainly a result of expanding of LysMe gene cluster in

chromosome 3 and LYK gene cluster in chromosome 6D. Intriguingly, we found *TaLYK5* and *TaLYK6* were tandem duplications in the three homologous chromosomes, as well as *TaLysMe4* and *TaLysMe5*. Similar tandem duplications exist in other wheat genes. For example, five wheat *CYP81D* genes was located within a genomic region associated with the salinity response [42]. We speculate that the evolved variation in copy number provided wheat redundancy function to adapt to the environment.

Meanwhile, we analyzed the expression pattern of LYP genes during wheat development. Previous study showed that OsEMSA1 appeared in QTL for panicle, seeds, and sterility but not in any other QTL for morphological/physiological traits or QTL for tolerance/resistance [43]. Another study showed that OsEMSA1 involved in embryo sac development in rice [44]. In wheat, the expression peak of the EMSA subclade genes appeared in root at the reproductive stage. In addition, EMSA group showed relatively high Ka/Ks ratio, suggesting that these genes evolved at a faster evolutionary rate, which is a feature of new genes.

To verify whether wheat LYPs were involved in the biotic and abiotic reaction, we analyzed the expression pattern by using the RNA-seq data from Wheat Expression Browser. Except for the genes not expressed, most of the wheat LYK family members were upregulated in response to biotic and abiotic stress. Different from other LYK family members which contain the protein kinase domain (Pfam ID: PF00069), wheat LYK5 and LYK6 proteins have the protein tyrosine and serine/threonine kinase domain (Pfam ID: PF07714). Wheat LYK5 and LYK6 were resided in the same clade with AtLYK4 and BdLYK4. No rice homologs were found in this clade. Meanwhile, wheat LYK5 and LYK6 genes derived from tandem duplications and were important contributors to the expansion of LYK gene family. Previous reports have shown that Arabidopsis AtLYK4 is important for chitin recognition during fungal infection [22,45,46]. Interestingly, the expression of TaLYK5s and TaLYK6s (except TaLYK6-6A) increased dramatically when treated by chitin and flg22, which may infer that they were important for PAMPs recognition. Another LYK family member TaLYK2, homologs of AtCERK1 in wheat, were induced by fungal infection and PAMPs treatment. Previous research suggested that AtCERK1 is a chitin co-receptor and mediates chitin-induced signaling through homodimerization and phosphorylation [20,47]. AtCERK1 can interact with AtLYK5 and forms a chitin-induced complex to induce plant immunity. Recent studies have shown that heterologous expression of the *Haynaldia villosa* lysin-motif contained receptor CERK1-V in wheat increases resistance to powdery mildew, yellow rust, and *Fusarium* head blight [48]. TaLYK2 may have similar functions in defense signaling.

In this study, TaLYP2 was induced by both fungal infection and PAMPs triggered treatment. TaLYP2 was rice OsCEBiP homolog in wheat. Similar to OsCEBiP, TaLYP2 were predicted to be secreted proteins with no transmembrane helices. OsCEBiP binds chitin oligosaccharides with the extracellular region and forms a complex with OsCERK1 to induce immune signaling [18,49]. TaLYP2 may bind chitin oligosaccharides with the extracellular region and interacting with other transmembrane proteins to activate downstream defense signaling pathways. TaLYP3 and TaLYP4 were phylogenetically related to rice OsLYP4/OsLYP6 and Arabidopsis AtLYM1/AtLYM3, respectively. Previously, AtLYM1 and AtLYM3 were identified as PGN but not chitin receptors [23]. However, OsLYP4 and OsLYP6 have dual function sensing both PGN and fungal chitin [19]. In this study, *TaLYP3* and *TaLYP4* clusters were not significantly changed under biotic stress. However, *TaLYP4* gene cluster were induced by cold stress, drought, and heat treatment. Meanwhile, *TaLYP3* cluster were depressed when treated by drought and heat stress. Whether TaLYP3 and TaLYP4 participate in biotic or abiotic defense should be further studied.

In summary, our studies provide the phylogeny and diversification of LYPs in wheat, including the evolutionary relationship, synteny analyze and the expression patterns. All of the 62 TaLYPs were divided into 6 subgroups. The LysMe and LYK subgroup were expanded during evolution. The

expression of some *LYP* and *LYK* genes were tissue- or stage- specific. Most of the wheat *LYK*s and *LYPs* were upregulated in response to biotic and abiotic stress. qRT-PCR analysis showed that 4 *LYP* genes were upregulated during *Puccinia triticina* infection. This study will serve as a foundation for further elucidation of the function of LYPs in wheat and other plants.

Authorship: The authors confirm contribution to the paper as follows: study conception and design: Liu MJ, Yuan ZY; data collection: Gao N, Wu YP; analysis and interpretation of results: Liu MJ, Gao N, Zhao YQ; draft manuscript preparation: Liu MJ, Zhao YQ. All authors reviewed the results and approved the final version of the manuscript.

Funding Statement: This work is supported by National Natural Science Foundation of China (Grant No. 31801693) and National Natural Fund Cultivation Project of Shanxi Academy of Agricultural Sciences (Grant No. YGJPY1902).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Al-Ashkar, I., Alderfasi, A., Ben Romdhane, W., Seleiman, M. F., El-Said, R. A. et al. (2020). Morphological and genetic diversity within salt tolerance detection in eighteen wheat genotypes. *Plants, 9(3),* 287. DOI 10.3390/plants9030287.
2. Ding, Z., Ali, E. F., Elmahdy, A. M., Ragab, K. E., Seleiman, M. F. et al. (2021). Modeling the combined impacts of deficit irrigation, rising temperature and compost application on wheat yield and water productivity. *Agricultural Water Management, 244,* 106626. DOI 10.1016/j.agwat.2020.106626.
3. Dangl, J. L., Horvath, D. M., Staskawicz, B. J. (2013). Pivoting the plant immune system from dissection to deployment. *Science, 341(6147),* 746–751. DOI 10.1126/science.1236011.
4. Jones, J. D. G., Dangl, J. L. (2006). The plant immune system. *Nature, 444(7117),* 323–329.
5. Boller, T., Felix, G. (2009). A renaissance of elicitors: Perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review of Plant Biology, 60,* 379–406.
6. Schwessinger, B., Ronald, P. C. (2012). Plant innate immunity: perception of conserved microbial signatures. *Annual Review of Plant Biology, 63(1),* 451–482. DOI 10.1146/annurev-arplant-042811-105518.
7. Dodds, P. N., Rathjen, J. P. (2010). Plant immunity: Towards an integrated view of plant-pathogen interactions. *Nature Reviews Genetics, 11(8),* 539–548. DOI 10.1038/nrg2812.
8. Zipfel, C. (2014). Plant pattern-recognition receptors. *Trends in Immunology, 35(7),* 345–351. DOI 10.1016/j.it.2014.05.004.
9. Boutrot, F., Zipfel, C. (2017). Function, discovery, and exploitation of plant pattern recognition receptors for broad-spectrum disease resistance. *Annual Review of Phytopathology, 55(1),* 257–286. DOI 10.1146/annurev-phyto-080614-120106.
10. Buist, G., Steen, A., Kok, J., Kuipers, O. P. (2008). LysM, a widely distributed protein motif for binding to (peptido) glycans. *Molecular Microbiology, 68(4),* 838–847. DOI 10.1111/j.1365-2958.2008.06211.x.
11. Desaki, Y., Miyata, K., Suzuki, M., Shibuya, N., Kaku, H. (2018). Plant immunity and symbiosis signaling mediated by LysM receptors. *Innate Immunity, 24(2),* 92–100.
12. Zhang, X. W., Dong, W. T., Sun, J. H., Feng, F., Deng, Y. W. et al. (2015). The receptor kinase CERK1 has dual functions in symbiosis and immunity signalling. *Plant Journal, 81(2),* 258–267.
13. Tombuloglu, G., Tombuloglu, H., Cevik, E., Sabit, H. (2019). Genome-wide identification of *Lysin-Motif Receptor-Like Kinase (LysM-RLK)* gene family in *Brachypodium distachyon* and docking analysis of chitin/LYK binding. *Physiological and Molecular Plant Pathology, 106,* 217–225.
14. Zhou, Z., Tian, Y., Cong, P. H., Zhu, Y. M. (2018). Functional characterization of an apple (*Malus* x *domestica*) LysM domain receptor encoding gene for its role in defense response. *Plant Science, 269,* 56–65.

15. Zhang, L., Yuan, L. B., Staehelin, C., Li, Y., Ruan, J. X. et al. (2019). The lysin motif-containing receptor-like kinase 1 protein of banana is required for perception of pathogenic and symbiotic signals. *New Phytologist, 223(3),* 1530–1546.
16. Xu, J., Wang, G. L., Wang, J., Li, Y. Q., Tian, L. L. et al. (2017). The lysin motif-containing proteins, Lyp1, Lyk7 and LysMe3, play important roles in chitin perception and defense against *Verticillium dahliae* in cotton. *BMC Plant Biology, 17(1),* 1–18.
17. Wang, C., Wang, G., Zhang, C., Zhu, P. K., Dai, H. L. et al. (2017). OsCERK1-mediatedchitin perception and immune signaling requires receptor-like cytoplasmic kinase 185 to activate an MAPK cascade in rice. *Molecular Plant, 10(4),* 619–633.
18. Shimizu, T., Nakano, T., Takamizawa, D., Desaki, Y., Ishii-Minami, N. et al. (2010). Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. *Plant Journal, 64(2),* 204–214.
19. Liu, B., Li, J. F., Ao, Y., Qu, J. W., Li, Z. Q. et al. (2012). Lysin motif-containing proteins LYP4 and LYP6 play dual roles in peptidoglycan and chitin perception in rice innate immunity. *Plant Cell, 24(8),* 3406–3419.
20. Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K. et al. (2007). CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis*. *Proceedings of the National Academy of Sciences, 104(49),* 19613–19618.
21. Cao, Y., Liang, Y., Tanaka, K., Nguyen, C. T., Jedrzejczak, R. P. et al. (2014). The kinase LYK5 is a major chitin receptor in Arabidopsis and forms a chitin-induced complex with related kinase CERK1. *Elife, 3,* e03766.
22. Xue, D. X., Li, C. L., Xie, Z. P., Staehelin, C. (2019). LYK4 is a component of a tripartite chitin receptor complex in *Arabidopsis thaliana*. *Journal of Experimental Botany, 70(19),* 5507–5516.
23. Willmann, R., Lajunen, H. M., Erbs, G., Newman, M. A., Kolb, D. et al. (2011). Arabidopsis lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proceedings of the National Academy of Sciences, 108(49),* 19824–19829.
24. Lobell, D. B., Schlenker, W., Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. *Science, 333(6042),* 616–620.
25. Seleiman, M. F. (2019). Use of plant nutrients in improving abiotic stress tolerance in wheat. In: Hasanuzzaman, M., Nahar, K., Hossain, M. (Eds.), *Wheat production in changing environments*, pp. 481–495. Singapore: Springer.
26. Seleiman, M. F., Abdel-Aal, S. M., Ibrahim, M. E., Monneveux, P. (2010). Variation of yield, milling, technological and rheological characteristics in some Egyptian bread wheat (*Triticum aestivum* L.) cultivars. *Emirates Journal of Food and Agriculture, 22(2),* 84–90.
27. Marcussen, T., Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M. et al. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. *Science, 345(6194),* 1250092. DOI 10.1126/science.1250092.
28. Alaux, M., Rogers, J., Letellier, T., Flores, R., Alfama, F. et al. (2018). Linking the international wheat genome sequencing consortium bread wheat reference genome sequence to wheat genetic and phenomic data. *Genome Biology, 19(1),* 111. DOI 10.1186/s13059-018-1491-4.
29. Ramírez-González, R., Borrill, P., Lang, D., Harrington, S. A., Brinton, J. et al. (2018). The transcriptional landscape of polyploid wheat. *Science, 361(6403),* 662.
30. Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology, 305(3),* 567–580. DOI 10.1006/jmbi.2000.4315.
31. Almagro Armenteros, J. J., Salvatore, M., Emanuelsson, O., Winther, O., Nielsen, H. et al. (2019). Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance, 2(5),* e201900429. DOI 10.26508/lsa.201900429.
32. El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A. et al. (2019). The Pfam protein families database in 2019. *Nucleic Acids Research, 47,* D427–D432. DOI 10.1093/nar/gky995.
33. Kumar, S., Stecher, G., Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology Evolution, 33(7),* 1870–1874. DOI 10.1093/molbev/msw054.

34. Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Xia, R. (2020). TBtools: An Integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant, 13(8),* 1194–1202. DOI 10.1016/j.molp.2020.06.009.
35. Borrill, P., Ramirez-Gonzalez, R., Uauy, C. (2016). expVIP: A customizable RNA-seq data analysis and visualization platform. *Plant Physiology, 170(4),* 2172–2186. DOI 10.1104/pp.15.01667.
36. Hamilton, N. E., Ferry, M. (2018). ggtern: Ternary diagrams using ggplot2. *Journal of Statistical Software, 87,* 1–17. DOI 10.18637/jss.v087.c03.
37. Zhang, Y., Geng, H., Cui, Z., Wang, H., Liu, D. (2021). Functional analysis of wheat NAC transcription factor, TaNAC069, in regulating resistance of wheat to leaf rust fungus. *Frontiers in Plant Science, 12,* 604797. DOI 10.3389/fpls.2021.604797.
38. International Wheat Genome Sequencing Consortium (IWGSC) (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science, 361(6403),* eaar7191.
39. Guo, J., Gong, B. Q., Li, J. F. (2021). Arabidopsis lysin motif/F-box-containing protein InLYP1 fine-tunes glycine metabolism by degrading glycine decarboxylase GLDP2. *The Plant Journal, 106(2),* 394–408. DOI 10.1111/tpj.15171.
40. Wang, P., Zhou, L., Jamieson, P., Zhang, L., Zhao, Z. et al. (2020). The cotton all-associated kinase GhWAK7A mediates responses to fungal wilt pathogens by complexing with the chitin sensory receptors. *The Plant Cell, 32(12),* 3978–4001. DOI 10.1105/tpc.19.00950.
41. Chen, Q., Li, Q., Qiao, X., Yin, H., Zhang, S. (2020). Genome-wide identification of lysin motif containing protein family genes in eight Rosaceae species, and expression analysis in response to pathogenic fungus *Botryosphaeria dothidea* in Chinese white pear. *BMC Genomics, 21(1),* 1–20. DOI 10.1186/s12864-020-07032-9.
42. Wang, M., Yuan, J., Qin, L., Shi, W., Xia, G. et al. (2020). TaCYP81D5, one member in a wheat cytochrome P450 gene cluster, confers salinity tolerance via reactive oxygen species scavenging. *Plant Biotechnology Journal, 18(3),* 791–804.
43. Tsaneva, M., de Schutter, K., Verstraeten, B., van Damme, E. J. M. (2019). Lectin sequence distribution in QTLs from rice (*Oryza sativa*) suggest a role in morphological traits and stress responses. *International Journal of Molecular Sciences, 20(2),* 437.
44. Zhu, Q., Zhang, X. L., Nadir, S., DongChen, W. H., Guo, X. Q. et al. (2017). A LysM domain-containing gene OsEMSA1 involved in embryo sac development in rice (*Oryza sativa* L.). *Frontiers in Plant Science, 8,* 1596.
45. Wan, J. R., Tanaka, K., Zhang, X. C., Son, G. H., Brechenmacher, L. et al. (2012). LYK4, a lysin motif receptor-like kinase, is important for chitin signaling and plant innate immunity in Arabidopsis. *Plant Physiology, 160(1),* 396–406.
46. Cheval, C., Samwald, S., Johnston, M. G., de Keijzer, J., Breakspear, A. et al. (2020). Chitin perception in plasmodesmata characterizes submembrane immune-signaling specificity in plants. *Proceedings of the National Academy of Sciences, 117(17),* 9621–9629.
47. Yamada, K., Yamaguchi, K., Shirakawa, T., Nakagami, H., Mine, A. et al. (2016). The *Arabidopsis* CERK1-associated kinase PBL27 connects chitin perception to MAPK activation. *The EMBO Journal, 35(22),* 2468–2483.
48. Fan, A. Q., Wei, L. Y., Zhang, X., Liu, J., Sun, L. et al. (2022). Heterologous expression of the *Haynaldia villosa* pattern-recognition receptor CERK1-V in wheat increases resistance to three fungal diseases. *The Crop Journal,* DOI 10.1016/j.cj.2022.02.005.
49. Gong, B. Q., Xue, J., Zhang, N. N., Xu, L. H., Yao, X. R. et al. (2017). Rice chitin receptor OsCEBiP is not a transmembrane protein but targets the plasma membrane via a GPI anchor. *Molecular Plant, 10(5),* 767–770.

Appendix A. Primers used in this study

Primer	Sequence (5′ to 3′)
TaLysMFbox-qRT-F	AGGCAAAGCAAACGGATTCTC
TaLysMFbox-qRT-R	TTGTCGCTGCTCCCACCTAA
TaLysMn1-qRT-F	CGCTGTCCGACGAGTTCTA
TaLysMn1-qRT-R	CCGTACTTGATGGCGATGC
TaLYK6-qRT-F	CACTCTGCTAATCCCGCTCAA
TaLYK6-qRT-R	GCAAGAACACCGACACCAACA
TaLYK5-qRT-F	TACCTCCTCAACACCACCC
TaLYK5-qRT-R	CGACGAGTTTGCGGCTAT
TaGADPH-F	CTGCATCATACGATGACATC
TaGADPH-R	TGTCACCGACAAAGTCAGTG

Appendix B. List of all wheat sequences identified by PFAM domain

Gene-ID	Name	splice variant	number of splice variants	Pfam-IDs1	Location	Pfam-IDs2	Location	Pfam-IDs3	Location
TraesCS4B02G353400.1	TaEMSA1-4B	1	1	PF01476	71–113				
TraesCS4D02G347400.1	TaEMSA1-4D	1	1	PF01476	69–111				
TraesCS1A02G187700.1	TaEMSA2-1A	1	1	PF01476	66–108				
TraesCS1B02G195500.1	TaEMSA2-1B	1	1	PF01476	65—107				
TraesCS1D02G188900.1	TaEMSA2-1D	1	1	PF01476	66–108				
TraesCS3A02G370900.1	TaLYK1-3A	1	1	PF01476	162–206	PF00069	308–583		
TraesCS3B02G403100.1	TaLYK1-3B	1	1	PF01476	165–209	PF00069	311–586		
TraesCS3D02G364000.1	TaLYK1-3D	1	1	PF01476	165–209	PF00069	311–586		
TraesCS6A02G258900.1	TaLYK2-6A	1	1	PF01476	185–220	PF00069	368–621		
TraesCS6B02G266500.1	TaLYK2-6B	1	1	PF01476	251–286	PF00069	434–689		
TraesCS6D02G240100.1	TaLYK2-6D	1	1	PF01476	252–287	PF00069	432–687		
TraesCS3B02G588200.1	TaLYK3-3B	1	1	PF01476	151–195	PF00069	282–554		
TraesCS5A02G552200.1	TaLYK3-5A	1	1	PF01476	150–194	PF00069	282–554		
TraesCSU02G125700.1	TaLYK3-U	1	1	PF01476	151–195	PF00069	239–511		
TraesCS4A02G427200.1	TaLYK4-4A	1	1	PF01476	211–257	PF01476	284–321	PF00069	441–725
TraesCS7A02G061600.1	TaLYK4-7A	1	1	PF01476	193–233	PF00069	352–630		
TraesCS7D02G056600.1	TaLYK4-7D	1	1	PF01476	210–247	PF00069	363–580		
TraesCS6A02G168900.1	TaLYK5-6A	1	1	PF01476	205–248	PF07714	404–656		
TraesCS6B02G197000.1	TaLYK5-6B	1	1	PF01476	205–248	PF07714	403–656		
TraesCS6D02G157900.1	TaLYK5-6D1	1	1	PF01476	205–248	PF07714	384–656		
TraesCS6D02G158400.1	TaLYK5-6D2	1	1	PF01476	205–248	PF07714	394–656		
TraesCS6A02G168800.1	TaLYK6-6A	1	1	PF01476	124–167	PF07714	303–510		
TraesCS6B02G196800.1	TaLYK6-6B	1	1	PF01476	198–241	PF07714	394–645		
TraesCS6D02G157800.1	TaLYK6-6D1	1	1	PF01476	200–243	PF07714	373–612		
TraesCS6D02G158300.1	TaLYK6-6D2	1	1	PF01476	198–241	PF07714	397–645		
TraesCS5A02G347800.1	TaLYP1-5A	1	1	PF01476	110–156	PF01476	174–217		

(Continued)

Appendix B. List of all wheat sequences identified by PFAM domain (continued)

Gene-ID	Name	splice variant	number of splice variants	Pfam-IDs1	Location	Pfam-IDs2	Location	Pfam-IDs3	Location
TraesCS5B02G348800.1	TaLYP1-5B	1	1	PF01476	110–156	PF01476	174–217		
TraesCS5D02G354000.1	TaLYP1-5D	1	1	PF01476	114–160	PF01476	178–221		
TraesCS4B02G329500.2	TaLYP2-4B	3	2	PF01476	109–155	PF01476	174–217		
TraesCS4D02G326400.2	TaLYP2-4D	2	2	PF01476	115–161	PF01476	180–223		
TraesCS5A02G501100.1	TaLYP2-5A	1	1	PF01476	174–220	PF01476	238–281		
TraesCS5A02G234700.1	TaLYP3-5A	1	1	PF01476	116–164	PF01476	184–226		
TraesCS5B02G233200.1	TaLYP3-5B	1	1	PF01476	116–163	PF01476	183–225		
TraesCS5D02G241600.1	TaLYP3-5D	1	1	PF01476	110–159	PF01476	179–221		
TraesCS7A02G168500.1	TaLYP4-7A	1	1	PF01476	109–158	PF01476	178–220		
TraesCS7B02G073300.1	TaLYP4-7B	1	1	PF01476	109–158	PF01476	178–220		
TraesCS7D02G169400.1	TaLYP4-7D	1	1	PF01476	109–158	PF01476	178–220		
TraesCS6A02G328800.1	TaLYP5-6A	1	1	PF01476	216–258				
TraesCS6B02G359500.1	TaLYP5-6B1	1	1	PF01476	176–218				
TraesCS6B02G359300.1	TaLYP5-6B2	1	1	PF01476	176–218				
TraesCS6D02G307700.1	TaLYP5-6D	1	1	PF01476	179–221				
TraesCS3B02G352100.1	TaLysMe1-3B1	1	1	PF01476	55–97				
TraesCS3B02G351900.1	TaLysMe1-3B2	1	1	PF01476	55–97				
TraesCS3D02G316500.1	TaLysMe1-3D	1	1	PF01476	56–98				
TraesCS3A02G314900.1	TaLysMe2-3A	1	1	PF01476	57–99				
TraesCS3A02G315000.1	TaLysMe3-3A1	1	1	PF01476	56–98				
TraesCS3A02G314800.1	TaLysMe3-3A2	1	1	PF01476	55–97				
TraesCS3B02G352200.1	TaLysMe3-3B	1	1	PF01476	56–98				
TraesCS3D02G316600.1	TaLysMe3-3D	1	1	PF01476	56–98				
TraesCS3A02G316000.1	TaLysMe4-3A	1	1	PF01476	55–97				
TraesCS3B02G350900.1	TaLysMe4-3B	1	1	PF01476	55–97				
TraesCS3D02G315500.1	TaLysMe4-3D1	1	1	PF01476	58–100				
TraesCS3D02G315400.1	TaLysMe4-3D2	1	1	PF01476	55–97				
TraesCS3A02G316100.1	TaLysMe5-3A	1	1	PF01476	79–120				
TraesCS3B02G350800.1	TaLysMe5-3B	1	1	PF01476	54–96				
TraesCS3D02G315300.1	TaLysMe5-3D	1	1	PF01476	79–121				
TraesCS4A02G275700.1	TaLysMFbox-4A	1	1	PF01476	106–149	PF00646	29–71		
TraesCS4B02G038000.1	TaLysMFbox-4B	1	1	PF01476	103–146	PF00646	29–68		
TraesCS4D02G035300.1	TaLysMFbox-4D	1	1	PF01476	103–146	PF00646	27–68		
TraesCS7A02G560400.1	TaLysMn1-7A	1	1	PF01476	51–94				
TraesCS7B02G486800.1	TaLysMn1-7B	1	1	PF01476	55–98				
TraesCS7D02G549400.1	TaLysMn1-7D	1	1	PF01476	51–94				
TraesCS4B02G329500.1		3	1	PF01476	109–155	PF01476	174–217		
TraesCS4B02G329500.3		3	3	PF01476	109–155	PF01476	174–217		
TraesCS4D02G326400.1		1	1	PF01476	115–161	PF01476	180–223		