Genome-wide identification and expression analysis of YTH domain-containing RNA-binding protein family in common wheat

Jing Sun†, Xiao Min Bie†, Ning Wang, Xian Sheng Zhang and Xin-Qi Gao*

Abstract

Background: N6-Methyladenosine (m6A) is the most widespread RNA modification that plays roles in the regulation of genes and genome stability. YT521-B homology (YTH) domain-containing RNA-binding proteins are important RNA binding proteins that affect the fate of m6A-containing RNA by binding m6A. Little is known about the YTH genes in common wheat (Triticum aestivum L.), one of the most important crops for humans.

Results: A total of 39 TaYTH genes were identified in common wheat, which are comprised of 13 homologous triads, and could be mapped in 18 out of the 21 chromosomes. A phylogenetic analysis revealed that the TaYTHs could be divided into two groups: YT521-B (TaDF) and YTHDC (TaDC). The TaYTHs in the same group share similar motif distributions and domain organizations, which indicates functional similarity between the closely related TaYTHs. The TaDF proteins share only one domain, which is the YTH domain. In contrast, the TaDCs possess three C3H1-type zinc finger repeats at their N-termini in addition to their central YTH domain. In TaDFs, the predicted aromatic cage pocket that binds the methyllysine residue of m6A is composed of tryptophan, tryptophan, and tryptophan (WWW). In contrast, the aromatic cage pocket in the TaDCs is composed of tryptophan, tryptophan, and tyrosine (WWY). In addition to the general aspartic acid or asparagine residue used to form a hydrogen bond with N1 of m6A, histidine might be utilized in some TaDFb proteins. An analysis of the expression using both online RNA-Seq data and quantitative real-time PCR verification revealed that the TaDFa and TaDFb genes are highly expressed in various tissues/organs compared with that of TaDFc and TaDCs. In addition, the expression of the TaYTH genes is changed in response to various abiotic stresses.

Conclusions: In this study, we identified 39 TaYTH genes from common wheat. The phylogenetic structure, chromosome distribution, and patterns of expression of these genes and their protein structures were analyzed. Our results provide a foundation for the functional analysis of TaYTHs in the future.

Keywords: Wheat, YTH domain-containing RNA-binding protein, Expression profiling, Development, Abiotic stress
which m6A modification controls the accessibility of binding sites for RNA binding proteins (Readers) [6]. The recognition of m6A by the readers is important for the metabolism, processing, and folding of RNA, and protein translation. Many m6A reader proteins have been identified, and most contain a YT521-B homologous (YTH) domain with an aromatic cage that specifically recognizes the GG(m6A) C sequence [7]. Five YTH domain-containing RNA-binding proteins in humans were identified, and they can be classified into three categories based on their sequences: YTHDC1 (YTH domain-containing protein 1), YTHDC2 (YTH domain-containing protein 2), and YTHDF (YTH domain-containing family protein) family [8]. YTHDF proteins (DF1, DF2, and DF3) are localized in the cytoplasm and contain a C-terminal YTH domain and a large, low-complexity region. It has been shown that YTHDFs bind all the m6A sites in mRNA. YTHDC1 is nuclear-enriched and contains a YTH domain at its central part and is composed of other multiple functional domains. In contrast, YTHDC2 is a nucleocytoplasmic protein that contains a C-terminal YTH domain and a DEAD-box RNA helicase domain. YTHDC1 binds to certain m6A sites in mRNA. YTHDC2 is a nucleocytoplasmic protein with a C-terminal YTH domain and a DEAD-box RNA helicase domain. YTHDC1 binds to certain m6A sites in mRNA and to noncoding RNAs, whereas YTHDC2 primarily binds to noncoding RNAs [9]. Many YTH genes have been identified in various plant species, including Arabidopsis thaliana, Oryza sativa, Malus domestica, Citrus sinensis and Cucumis sativus [10–13]. Thirteen ECT genes of A. thaliana, homologs of YTHDFs, were identified, some of which function in controlling the timing of the first true leaf formation, proper leaf morphology, and trichome branching [14–17]. ECT2 can interact with m6A-containing RNAs in vivo, indicating that YTH domain is an m6A reading domain in Arabidopsis [15]. AtCPSF30 is a homolog of human HYTDC, which is a cleavage and polyadenylation specificity factor that functions in oxidative and nitrate signaling [18, 19]. In A. thaliana, M. hupehensis and C. sativus, YTH genes are involved in plant responses to pathogens, plant hormones (salicylic acid and abscisic acid), and abiotic stresses (water, temperature, and salinity) [10–13, 20–22]. These studies revealed that YTH domain-containing proteins play important roles in plant development and biotic and abiotic stress responses. Wheat is one of the major cereal crops in the world. To our knowledge, the YTH protein family in common wheat (Triticum aestivum L.) has not yet been analyzed in detail. In this study, we identified 39 genes that encode YTH domain-containing proteins in common wheat and analyzed their phylogenetic relationship and expression in different tissues/organs and in response to various stresses.

Results
Identification and phylogenetic analysis of the YTH proteins in common wheat
To identify the TaYTH genes in common wheat, we used the hidden Markov model profiles of the YTH domain (PF04146) as queries to search the protein databases (IWGSC RefSeq v1.1) of common wheat (https://www.ebi.ac.uk/Tools/hmmer/search/hmmsearch) that was downloaded as a local protein database of bread wheat. Additionally, the conserved YTH domain was also used to BLASTp the YTH proteins encoded by the wheat genome on the website http://plants.ensembl.org/Multi/Tools/Blast [8]. Additionally, we performed BLASTP analysis against the local wheat protein database (E-value < 1 × 10⁻5) by using the YTH domain-containing protein sequences of Arabidopsis thaliana and Oryza sativa as queries. After the redundant sequences were removed, a total of 39 putative YTH proteins were identified and used for further analysis (Additional file 1). Their encoding genes are 13 sets of homeoalleles with 13 YTH genes in each sub-genome. Additionally, 13, 17, 12, 15, 12, 13 YTH genes in A. thaliana, Glycine max, O. sativa, Zea mays, Brachypodium distachyon, and Hordeum vulgare, respectively, were identified using the same strategy in this study and other study [15]. An unrooted phylogenetic tree of 39 TaYTH genes was constructed based on the YTH domains of the 39 TaYTHs in concert with 82 YTH proteins from A. thaliana, O. sativa, G. max, Z. mays, B. distachyon, and H. vulgare (Additional file 2) to analyze the evolutionary relationships of the TaYTH proteins (Fig. 1a). The phylogenetic analysis revealed that the TaYTH domain proteins were clustered into two groups: YTHDF (TaDF) and YTHDC (TaDC). The YTHDF family contains three subgroups, YTHDFa, YTHDFb, and YTHDFc, which include 12, 9, and 15 members, respectively. The YTHDC family can be divided into two additional subgroups, YTHDCa and YTHDCb. No YTHDCb proteins were identified in common wheat, which is consistent with the previous hypothesis that there are no group-b YTHDCs in monocotyledons [15]. All of the 39 YTH proteins that had been identified were re-named based on their phylogenetic relationship consistent with the YTH proteins that have been identified in other plant species [15]. The subgenome (A, B or D) on which the genes are located is included in the name, and the number before the subgenome indicates the serial number of chromosome. Chromosomal location information of the TaYTH genes was extracted from the GFF3 reference file of the wheat genome, and a chromosomal map was constructed using KnetMiner, an online analytical tool (https://knetminer.rothamsted.ac.uk/Triticum_aestivum/). The map constructed revealed that the TaYTHs are located on all chromosomes with the exception of chromosomes 6A, 6B, and 6D. There are three TaYTH genes on chromosomes 5A, 5B, 5D, 7A, 7B, and 7D, and two on chromosomes 1A, 1B, 1D, 3A, 3B, 3D, 4A, 4B, and 4D. In contrast, only one TaYTH gene is located on each 2A, 2B, and 2D chromosome (Fig. 1b).
Fig. 1 Phylogenetic tree and chromosomal distribution of TaYTH genes. a. Phylogenetic tree of YTH proteins based on the YTH domains of the YTH proteins from wheat, O. sativa, A. thaliana, Z. mays, and G. max. b. Chromosomal distribution of TaYTH genes. All wheat chromosomes are drawn to scale based on their actual lengths.
Gene structure and synteny of the *TaYTH* genes

The exon-intron structure provides insight on the evolution of gene families and additional evidence to support phylogenetic grouping. We then analyzed the exon-intron structure of the *TaYTH* genes based on their evolutionary classification. All the *TaYTH* genes contain introns, and most of the genes contained from five to eight introns (Fig. 2). In contrast, *TaDF3b-7B* contains the largest number of introns (11), including three long ones (> 4 Kb) that are not found in the *YTH* genes of other plant species [11, 12] (Fig. 2). The patterns of distribution of the *TaYTH* exon/introns were similar among the members within each group in the phylogenetic tree. The proportions of intron phases 1, 2, and 0 are 2.55, 23.3, and 74.2%, respectively, of all the *TaYTH* genes (Fig. 2). There are similar patterns of the intron phases of *TaYTH* genes with those of the *YTH* genes in *A. thaliana*, *O. sativa* and other species [11], i.e., the largest proportion of introns was in phase 0, with a few in phase 1 (Fig. 2).

Genomic replication events often result in the expansion of the gene family during the evolution of angiosperms [23]. Three pairs of genes, *TaDF3a-5A* and *TaDF4a-5A*, *TaDF3a-5B* and *TaDF4a-5B*, and *TaDF3a-5D* and *TaDF4a-5D*, are localized in the region within a 200 kb range in chromosomes 5A, 5B, and 5D, respectively (Fig. 1b, Table 1). In addition, these genes share similar exon-intron structure (Fig. 2). These suggest that these *YTH* genes may be produced by tandem gene replication. The synteny diagram of the *YTH* genes between wheat and other plant species (*A. thaliana*; *O. sativa*; *Z. mays*; *Medicago truncatula* and *B. distachyon*) revealed that the *YTH* genes of wheat contained the greatest number of homologous gene pairs with those of *B. distachyon*, followed by rice and maize (Fig. 3). No *YTH* homologous gene pairs were identified between wheat and *Arabidopsis*. Only two homologous gene pairs were identified between wheat and *M. truncatula*. In contrast, 18, 21, 21, and 24 *YTH* gene pairs were identified between barley and wheat, rice and wheat, maize and wheat, *B. distachyon* and wheat, respectively, matching the three expected homologous counterparts in wheat (Fig. 3, Additional file 3). Our results reveals there is a similar copy number of *YTH* genes in different species and similar numbers of *YTH* gene pairs among different species and wheat, which suggest that the *YTH* gene family is conservative and withstand strong selection pressure during plant evolution.

Protein features of *TaYTHs*

The lengths of *TaYTHs* ranged from 588 to 767 amino acid residues. The molecular weight (MW) and
isoelectric point (pI) of the TaYTH proteins varied from 6.39 to 8.21 kDa and from 5.4 to 8.54, respectively (Table 1). The subcellular localization of TaYTH proteins was predicted by the online tool PSORT (https://www.psort.org/). TaDF1b-7D, TaDF2c-3A, TaDF2c-3B, TaDF2c-3D, and TaDF4c-1B were predicted to be localized in the cytoplasm, while other TaYTHs were localized in nucleus (Table 1).
**Fig. 3** Synteny analysis of YTH genes between wheat and other plant species, *A. thaliana; O. sativa; Z. mays; M. truncatula* and *B. distachyon*. Gray lines in the background indicate the collinear blocks within the genomes of wheat and other plants, and the blue lines indicate the syntenic YTH gene pairs.
Sequence alignments revealed that the TaYTH proteins only share 30.85% identity, and the highly conserved regions are located in the YTH domains with 69.57% identity (Additional file 4), which is similar to the YTH domain-containing proteins of other species [6]. In most of the TaDF proteins, the YTH domain is the only recognizable module at their C-terminus (Fig. 4a) that is similar with those of other species [11]. In contrast, in addition to the YTH domain at their middle parts, all three TaDC proteins harbor two additional domains, an N-terminal YTH1 superfamily domain and a C-terminal PRK12678 superfamily domain. The YTH1 superfamily domain is composed of three repeated C3H1-type zinc fingers [13] (Fig. 4b). All these TaDC proteins were classified as type-a YTHDC (Fig. 1). The type-a YTHDCs in other plants also possess the YTH1 superfamily domain at their N-terminus, but their C-terminus harbors different types of domains (Additional files 5 and 6). In contrast, the type-b YTHDCs in dicotyledonous plants do not contain additional domains besides the YTH domain, which is like that of the YTHDC1s in animals (Additional files 5 and 6). To further analyze the structural diversity in the TaYTH proteins, the conserved motifs in TaYTHs were identified using the online tool MEME (http://meme-suite.org/tools/meme) (Additional file 7). All the TaYTH proteins contained motifs 2, 4, and 6. In the TaDFs, the YTH domain is composed of three motifs (2, 4, and 1) and followed by motif 3 at their C-termini. In contrast, the YTH domain of TaDCs consists of three motifs (2, 4, and 6) and is followed by an additional motif 6 at their C-termini (Fig. 5c). Upstream of the YTH domain, motifs 5, 6, and 8 were shared by subgroup TaDFa; in contrast, they are comprised of motif 5, 6, 7 and 8 in subgroup TaDFb and motif 6 and/or 8 in subgroup TaDFc (Fig. 4c).

An aromatic cage pocket in the YTH domain of YTHDF and YTHDC is utilized to bind the methyllysine residue of m6A [6, 8]. This positively-charged pocket is formed by the side chains of tryptophan (W411), tryptophan (W465), and tryptophan (W470) in human YTHDF1; and tryptophan (W377), tryptophan (W428),...
and leucine (L439) in human YTHDC1. In this study, we found that the aromatic cage in the TaDFs is composed of tryptophan, tryptophan, and tryptophan (WWW), but the cage in TaDCs is composed of tryptophan, tryptophan, and tyrosine (WWY) (Fig. 5a and b). In addition to the methylation-dependent interactions, m6A also forms base-specific hydrogen bonds with the YTH proteins [7]. A comparison of the binding affinity between YTHDF1 and YTHDC1 with m6A revealed that the asparagine (N367) in YTHDC1 forms a hydrogen bond with N1 of m6A that is stronger than the corresponding aspartic acid (D401) residue with N1 of m6A in YTHDF1. In this study, we found that all the TaDFa, six TaDFb and three TaDFc proteins used aspartic acid to form a hydrogen bond with N1 of m6A; however, aspartic acid has been replaced by asparagine in six TaDFb, three TaDFc, and three TaDC proteins (Fig. 5a and b). Additionally, histidine is used in the corresponding position in three TaDFb proteins (Fig. 5a and b), and this mode was not found in other YTH proteins identified in plant species [15]. The N-termini of the TaDF proteins possess low-complexity regions containing Y/P/Q-rich regions (Fig. 5c, Additional file 8); however, the Y/P/Q-rich regions in TaDCs were located between the zinc finger repeat (YTH1 superfamily domain) and the YTH domain (Fig. 5c).

Expression analysis of the TaYTH genes in different tissues/organs and under abiotic stresses

To analyze the patterns of TaYTH expression, we downloaded the RNA-Seq data of wheat from the Wheat Expression Browser website (http://www.wheat-expression.com/) [24]. The expression data of 19 tissues/organs (grain, radicle, embryo proper, coleoptile, endosperm, seedling, root, root apical meristem, stem, shoot, shoot apical meristem, leaf, flag leaf, spike, spikelet, glume, lemma, microspore, stigma, and ovary) at different development stages and abiotic/biotic stress conditions were included in our analysis. The results revealed that all the TaDFa and TaDFb genes were highly expressed in the 19 tissues/organs with the exception of TaDF3bs (Fig. 6a). In contrast, the levels of expression of TaDFc...
and TaDCas were lower in various tissues/organisms with the exception of TaDF3cs and TaDF5cs (Fig. 6a). Most of the TaYTH genes exhibited obvious changes of expression in response to abiotic stresses, such as phosphorus starvation, cold stress, and heat stress (Fig. 6b). In addition, the expression of TaDF3b-7B and TaDF4c-IA are low in all the tissues/organs and stress conditions analyzed (Fig. 6a and b). TaFda and TaFdb genes showed highly expressed under biotic stresses with the exception of TaDF3b-7B (Additional file 9). In contrast, TaDFc and TaDca are low expressed except that TaDFsc is high expression under biotic stresses (Additional file 9).

To confirm the tissue-expression of the TaYTH genes in common wheat, we performed qRT-PCR analyses in five different tissues/organisms of winter wheat cultivar Jimai 22, including roots, stems, leaves at vegetative development, and spikes and flowers at reproductive development. A total of 13 genes (TaDF1a-4D, TaDF2a-5B, TaDF3a-5D, TaDF4a-5B, TaDF1b-7A, TaDF2b-1D, TaDF3b-7A, TaDF1c-2B, TaDF2c-3D, TaDF3c-4A, TaDF4c-1D, TaDF5c-3A, and TaDC1a-7A) and one of each set of homeoalleles were included in our analysis. TaDF1a-4D exhibited the highest level of expression in five tissues among the 13 TaYTH genes (Fig. 6c), which is consistent with the results based on the RNA-Seq data (Fig. 6a). Our results revealed that 11 out of 13 TaYTHs were expressed at relatively higher levels in leaves with the exception of TaDF5c-3A, and TaDC1a-7A (Fig. 6c). In addition, the levels of expression of TaDF3a-5D, TaDF1b-7A, TaDF3c-3A and TaDC1a-7A in the stems and TaDF3a-5D, TaDF4a-5B, TaDF1b-7A, TaDF2c-3D, and TaDF3c-4A in the flowers were at relatively higher levels. In contrast, only TaDF3b-7A and TaDF3c-4A exhibited a higher level of expression in the roots and spikes, respectively (Fig. 6c).

Discussion

In our study, a total of 39 putative YTH domain-containing protein family genes were identified in common wheat. Yue et al. identified 41 m6A reader proteins in wheat [25], and 38 of them contained the YTH domain. We identified an additional wheat gene, TaDC1a-7D (TraesCS7D02G450100.1), which encodes a protein containing YTH domain at its middle part, and was designated member of the wheat YTHDC group (Table 1). In the study by Yue et al. (2019) [25], TaCPSF30–2 (TraesCS6A01G138900.1) and TaCPSF30–5 (TraesCS6B01G167200.1) are designated YTHDC proteins. However, no YTH domain was identified in these two proteins. TaCPSF30–2 and TaCPSF30–5 were localized in chromosomes 6A and 6B [25]. However, in our results, no YTH gene localized in any form of chromosomes 6 (A, B, and D) was identified. These discrepancies might have resulted from the different annotation of version 1 and 1.1 of the wheat genome (IWGSC RefSeq Annotations v1.0 and v1.1).

A chromosomal distribution map revealed that the TaYTH genes are dispersed across all the chromosomes of all three wheat sub-genomes with the exception of chromosome 6. A similar pattern of dispersion was also found in the YTH genes in A. thaliana O. sativa and M. domestica, but it is different from those in C. sativus, in which the CsYTH genes are located on three out of seven of the C. sativus chromosomes [13].

It has been reported that a conserved mechanism was utilized by the YTH domain proteins YTHDF and YTHDC to recognize m6A, an aromatic cage that recognizes the methyl moiety of m6A [6, 8]. The YTHDF family of proteins has a WWW cage, while the YTHDCs have a WWL-type cage in humans. In this study, we found that the aromatic cage in TaDFs is comprised of WW, but the cage in TaDCs is comprised of WWY. This rule is employed by the YTHDC proteins from other plant species. These analyses indicated that the plant YTHDCs are different from the animal YTHDCs in m6A binding. In animals, two types of YTHDCs were identified, YTHDC1 and YTHDC2. In YTHDC1, the YTH domain is located in the central part of protein, while the YTH domain is located at the C-terminus of YTHDC2. In plant species, two types of YTHDC were also identified. However, the YTH domains in plant YTHDCs were all localized in the central part of proteins, which is similar to that of the animal YTHDC1s. This suggests that the YTHDCs in plants could share similar molecular mechanisms with those of the animal YTHDC1s. In humans, YTHDC1 mediates pre-mRNA splicing by interacting with splicing factor [26]. Pre-mRNA splicing is involved in the regulation of plant development, sexual reproduction, and abiotic stress responses [27–35]. Thus, the functions of plant YTHDCs in pre-mRNA splicing merit further study. Additionally, the N-termini of plant YTHDFs do not possess a zinc finger domain like those found in the plant YTHDCs. This is consistent with the suggestion that, aside from their conserved YTH domain, YTHDCs are unrelated to YTHDFs based on their amino acid sequences [6]. Animal YTHDF proteins harbor YTH domain at their C-
termini and a low-complexity region at their N-termini. In this low-complexity region, several P/Q/N-rich patches were found in animal YTHDF proteins. P/Q/N-rich patches also were identified in TaYTHs in this study. A P/Q/N-rich region is suggested to function in the regulation of the stability of m6A-containing mRNAs by localizing the YTH domain-containing proteins to the RNA decay sites \[36--38\]. m6A forms base-specific hydrogen bonds with YTH proteins that enhance these interactions \[7\]. In human YTHDF1 and YTHDC1, asparagine or aspartic acid is used to form the hydrogen bond with N1 of m6A, although asparagine interacts more strongly with N3 of m6A than does aspartic acid. However, in some wheat YTH proteins, histidine might be utilized for the interaction with N3 of m6A. Asparagine and aspartic acid are neutral amino acids, while histidine is basic. Thus, the binding of m6A with YTH proteins in wheat might have substantial differences with those in other plant and animal species \[6, 15\]. Human YTHDC2 is a nucleocytoplasmic protein that contains a DEAD-box RNA helicase domain in addition to a C-terminal YTH domain. The DEAD-box RNA helicase domain-containing proteins function in unwinding double stranded RNA \[39\]. In fact, the YTHDC2 proteins play roles in unwinding 5′-UTR and increasing the translational efficiency of the target gene \[40\]. However, no DEAD-box RNA helicase domain-containing YTH proteins have been identified in the plant kingdom. Thus, whether YTH-involved unwinding is required for m6A recognition by the readers remains unknown.

Wheat is an important crop around the world, which provides more than 20% of the protein and caloric intake of humans. The yield of wheat is closely related with the development and architecture of spikes. Many signaling pathways have been identified that are involved in the regulation of spike development and architecture in wheat \[41\], during which the regulation of gene expression and hormone signaling play critical roles. In our study, a set of YTH genes were expressed at higher levels during the development of spikes and flowers of wheat. Thus, these highly expressed YTH genes in the wheat spikes might function in the development and architecture of spikes. The selection and creation of wheat germplasm with tolerance to multiple abiotic and biotic stresses is a key challenge for wheat breeding \[42--56\]. Our results and those of previous studies reveal that the expression of YTH gene is involved in the responses to various stresses and in plant development \[10, 13, 15, 17, 21\]. Therefore, the function of wheat YTH genes in the response to abiotic stresses merits further study, and these YTH genes might be valuable for wheat genetic improvement.

Conclusions
In this study, 39 TaYTH genes were identified in common wheat \((Triticum aestivum \ L.)\). A phylogenetic analysis revealed that the TaYTHs could be divided into two groups, YTHDF (TaDF) and YTHDC (TaDC). The chromosomal location and exon-intron structure of the genes, and the motif distribution and domain organization of the proteins were analyzed. In addition, the aromatic cage pockets of TaYTH proteins were predicted. The aromatic cage pocket of TaDFs is composed of tryptophan, tryptophan, and tryptophan (WWW). In contrast, this pocket is composed of tryptophan, tryptophan, and tyrosine (WWY) in TaDC proteins. In addition to the general aspartic acid or asparagine residues used to form a hydrogen bond with N1 of m6A, some TaDFb proteins might utilize histidine in this role. The results of the analysis of expression revealed that the TaDFa and TaDFb genes were highly expressed in various tissues/organs. In addition, the expression of TaYTH genes is changed in response to various abiotic stresses. Our results provide the foundation for the functional analysis of TaYTHs in the future.

Methods
Identification of the YTH genes in wheat
The YTH protein family in wheat was identified as described by Yan et al. \[57\]. The conserved sequence of the YTH domain was used to make a Hidden Markov model and used in BLAST homology searches. The reference genome sequence and annotations (GFF files) of wheat were downloaded from the wheat genome Database (IWGSC RefSeq v2.0) at Ensemble Plants \(\text{http://plants.ensembl.org/Triticum_aestivum/}\). The conserved YTH domain was also used to BLASTp the YTH proteins encoded by the wheat genome on the website \(\text{http://plants.ensembl.org/Multi/Tools/Blast}\) \[8\]. Additionally, BLASTp analysis against the local wheat protein database (E-value \(< 1 \times 10^{-5}\)) was performed by using the \textit{Arabidopsis} and rice YTH domain-containing protein sequences as queries (Additional file 2).

Phylogenetic analysis, gene structure, and chromosomal locations
ClustalX2.1 was used for multiple sequence alignments with default parameters. A maximum likelihood phylogenetic tree was constructed using MEGA 6.0 (https://www.megasoftware.net/) with 1000 bootstrap replicates. The phylogenetic tree was visualized by FigTree (http://tree.bio.ed.ac.uk/software/). TBtools software \[58\] and the online analytical tool KnetMiner (https://knetminer.rothamsted.ac.uk/Triticum_aestivum/) were used to map the locations of HTY genes in chromosomes and construct the map of exon-intron structure.

Amino acid sequence analysis
Online tool PSORT (https://www.psort.org/) was used to predict the subcellular localization of the TaYTH
proteins. The molecular masses and isoelectric points of the TaYTH proteins were predicted using the web tool ExPASy (https://web.expasy.org/compute_pi/). The domains of TaYTH proteins were analyzed using the web tools CDD (https://www.ncbi.nlm.nih.gov/) and ExPASy (https://prosite.expasy.org/). Finally, the domain graphs were visualized using TBtools [48]. The motif compositions of TaYTHs were identified by MEME (http://meme-suite.org/tools/meme) and exhibited using TBtools software. Multiple protein sequence alignments were performed using DNAMAN8. Sequence logos of the conserved motifs of proteins were created using Weblogo (http://weblogo.berkeley.edu/logo.cgi).

Digital gene expression pattern analysis
The transcript per million values of expression of TaYTH genes in 19 tissues/organs (grain, radicle, embryo proper, coleoptile, endosperm, seedling, root, root apical meristem, stem, shoot, shoot apical meristem, leaf, flag leave, spike, spikelet, glume, lemma, microspore, stigma and ovary) and in response to various stresses were downloaded from the Wheat Expression Browser (www.wheat-expression.com) [24]. Heat maps that showed the relative expression levels were constructed using TBtools.

Plant materials and growth conditions
The common wheat used in this study is Jimai22, a winter wheat cultivar that is widely planted in North China. The seeds of Jimai22 were provide by Dr. Yu Xiu Dong (Shandong Agricultural University, China). The wheat seedlings were grown in a greenhouse at 22 °C with a 16 h light/8 h dark cycle at Shandong Agricultural University. The roots, stems, and leaves were sampled from 7-day seedlings. To sample the spikes and flowers, Jimai22 seeds were first vernalized in soil for 30 days at 4 °C. The vernalized seedlings were then transferred into greenhouses. The spikes and flowers were sampled from the plants grown for 30 days and 40 days in a greenhouse, respectively.

Expression of the TaYTH genes analyzed by qRT-PCR
Total RNA was extracted using the TRIzol reagent (Invitrogen, Beijing, China) as described by Zhao et al. [59]. The total RNA was reverse transcribed using a PrimeScript RT Reagent Kit (TakaRa, Dalian, China) according to the manufacturer’s instructions. The first strand cDNAs for qRT-PCR were synthesized using Synthesis SuperMix (TransGen, Beijing, China). The qRT-PCR amplifications were performed as described by Zhao et al. [60]. Three biological replicates and three technical replicates were utilized. The Actin gene was used as an internal control. The sequences of the primers used in this study are listed in Additional file 10.

Supplementary information
Supplementary information accompanies this paper at https://doi.org/10.1186/s12870-020-02505-1.

Additional file 1. Sequences of TaYTHs.
Additional file 2. Sequences of YTHs used for the construction of phylogenetic tree.
Additional file 3. Gene pairs between M. truncatula and wheat, barley and wheat, rice and wheat, barley and wheat, B. distachyon and wheat.
Additional file 4. Sequence alignments among the identified TaYTH proteins.
Additional file 5. Organization and distribution of the conserved domain in YTHDC1s in animals and YTHDCs in plants.
Additional file 6. Sequences of animal YTHDC1s and plant YTHDCs.
Additional file 7. Conserved motifs of TaYTHs.
Additional file 8. Y/P/Q-rich region in TaYTHs.
Additional file 9. Heat map showing the expression of TaYTH genes in response to biotic stress.
Additional file 10. Primers used in this study.

Abbreviations
MW: Molecular weight; m6A: N6-Methyladenosine; pI: Isoelectric point; YTH: YT521-B homology

Acknowledgements
We would like to thank the reviewers for their comments and helpful suggestions.

Authors’ contributions
XQG, XSZ conceived and designed the research. JS, XMB, XQG performed the gene expression analysis. JS, NW performed the bioinformatics analysis. JS, XMB, XQG drafted the manuscript. All authors have read and approved the manuscript.

Funding
This work is supported by National Natural Science Foundation of China (91935302) focusing on the genetic improvement of yield traits in wheat, and National Transgenic Science and Technology Program (2019ZX08010–003) focusing on the genetic improvement of yield traits in wheat. The funding agencies were not involved in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials
All data generated or analyzed in this study are included in this published article and its Additional files. The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 27 February 2020 Accepted: 18 June 2020
Published online: 23 June 2020

References
1. Yang Y, Hsu PJ, Chen Y-S, Yang Y-G. Dynamic transcriptomic m6A decoration: writers, erasers, readers and functions in RNA metabolism. Cell Res. 2018;28(6):616–24.
8. Liao S, Sun H, Xu C, YTH domain: a family of N(6)-methyladenosine (m(6)a) binding proteins. Trends Cell Biol. 2018;28(2):113–27.

9. Xu C, Wang X, Liu K, Roundtree IA, Tempel W, Li Y, Lu Z, He C, Min J. XTH domain messenger RNA adenosine methylation and interacts with a homolog of a sex-specific splicing factor. Plant Cell. 2008;20(5):1278–88.

11. Ouyang Z, Duan H, Mi L, Hu W, Chen J, Li X, Zhong B. Genome-wide identification and expression analysis of their responsiveness to senescence and abiotic stresses. Gene. 2014;538(2):292–9.

14. Scutenaire J, Deragon J-M, Jean V, Benhamed M, Raynaud C, Favory J-J, Brodersen P. An m6A-YTH module controls developmental timing and regulates the nitrate transceptor gene NRT1.1. New Phytol. 2015;207(3):817–33.

15. Wang N, Yue Z, Liang D, Ma F. Genome-wide identification of members in the YTH domain-containing RNA-binding protein family in Citrus sinensis. J Am Soc Hortic Sci. 2019;144(2):79–91.

16. Wang N, Yue Z, Liang D, Ma F. Genome-wide identification of members in the YTH-domain-containing RNA-binding protein family in Citrus sinensis. J Am Soc Hortic Sci. 2019;144(2):79–91.

17. Wang N, Guo T, Wang P, Sun X, Shao Y, Liang B, Gong X, Li Y, Zhao Y, Guo Y, Zhang G, Zhao S. m6A binding proteins. Trends Cell Biol. 2018;28(2):113–27.

18. Scutenaire J, Deragon J-M, Jean V, Benhamed M, Raynaud C, Favory J-J, Brodersen P. An m6A-YTH module controls developmental timing and regulates the nitrate transceptor gene NRT1.1. New Phytol. 2015;207(3):817–33.

19. Scutenaire J, Deragon J-M, Jean V, Benhamed M, Raynaud C, Favory J-J, Brodersen P. An m6A-YTH module controls developmental timing and regulates the nitrate transceptor gene NRT1.1. New Phytol. 2015;207(3):817–33.

20. Wang N, Guo T, Sun X, Jia X, Wang P, Shao Y, Liang B, Gong X, Ma F. Functions of two Mdxu hupesheni (Pamp.) Rehd. YTPs (MYTH1P and MYTH2P) in biotic- and abiotic-stress responses. Plant Sci. 2017;261:18–27.

21. Wang N, Guo T, Wang P, Sun X, Shao Y, Jia X, Liang B, Gong X, Ma F. MYTH1P and MYTH2P from apple confer tolerance to multiple abiotic stresses in Arabidopsis thaliana. Front Plant Sci. 2017;8:1367.

22. Guo T, Wang N, Xue Y, Guan Q, van Nocker S, Liu C, Ma F. Overexpression of the RNA binding protein MYTH1P in transgenic apple enhances drought tolerance and WUE by improving ABA level under drought condition. Plant Cell Physiol. 2017;58(12):2313–23.

23. Adams KL, Wendel JF. Polyploidy and genome evolution in plants.Curr Opin Plant Biol. 2005;8(2):135–41.

24. Ramírez-González R, Borrill P, Lang D, Harrington S, Branton J, Venturini L, Davey M, Jacobs J, Van Ex F, Pasha A. The transcriptional landscape of polyploid wheat. Science. 2018;361(6403):eaar5889.

25. Yue H, Nie X, Yan Z, Weinig S. N6-methyladenosine regulatory machinery in plants: composition, function and evolution. Plant Biotechnol J. 2019;17(1):194–208.

26. Xiao W, Adhikari S, Dахal U, Chen Y-S, Hao Y-J, Sun B-B, Sun H-Y, Li A, Ping X-L, Lai W-Y. Nuclear m6a reader YTHDC1 regulates mRNA splicing. Mol Biol Cell. 2017;28(14):2071–89.

27. Xiong F, Ren J, Ju Y, Wang YY, Kong L, Otegui MS, Wang XL. AtUBD13 affects pre-mRNA splicing and is essential for embryo development in Arabidopsis. Plant J. 2019;98(4):714–26.

28. Xiong F, Ren J, Ju Y, Wang YY, Lu CC, Kong L, Otegui MS, Wang XL. AtUBD65s functions in abscisic acid mediated flowering via regulating the precursor messenger RNA splicing of AB5 and PLC in Arabidopsis. New Phytol. 2019;229(1):277–92.

29. Wang Y-Y, Xiong F, Ren QP, Wang X-L. Regulation of flowering time by alternative splicing: the role of the U2 auxiliary factor. J Exp Bot. 2020;71(3):751–8.

30. Thatcher SR, Zhou W, Leonard A, Wang B-B, Beatty M, Zastrow-Hayes G, Zhao X, Baumgarten A, Li B. Genome-wide analysis of alternative splicing in Zea mays: landscape and genetic regulation. Plant Cell. 2014;26(9):3472–87.

31. Qin Z, Wu J, Geng S, Feng N, Chen F, Kong X, Song G, Chen K, Li A, Mao L. Regulation of FT splicing by an endogenous cue in temperate grasses. Nat Commun. 2017;8(1):1–12.

32. Zhu F, Chen MX, Ye NH, Shi L, Ma KL, Yang JF, Cao YY, Zhang Y, Yoshida T, Fernie AR. Proteogenomic analysis reveals alternative splicing and translation as part of the aspic acid response in Arabidopsis seedlings. Plant J. 2017;91(3):518–33.

33. Chen MX, Zhu F-Y, Wang F-Z, Yu N-H, Gao B, Chen X, Zhao S-S, Fan T, Cao Y-Y. Alternative splicing and translation play important roles in hypocotyl germination in rice. J Exp Bot. 2019;70(3):817–33.

34. Ren RC, Wang LL, Zhang Z, Zhao YI, Wu J, Wei YM, Zhang XS, Zhao XY. DEK43 is a P-type pentatricopeptide repeat (PPR) protein responsible for the Cis-splicing of nad4 in maize mitochondria. J Integr Plant Biol. 2019. https://doi.org/10.1111/jipb.12843.

35. Laloum T, Martin G, Duque P. Alternative splicing control of abiotic stress responses. Trends Plant Sci. 2018;23(2):140–50.

36. Fu Y, Dominiessi D, Rechavi G, He C. Gene expression regulation mediated through reversible m6A RNA methylation. Nat Rev Genet. 2014;15(5):293.

37. Yang Y, Sun B-F, Xiao W, Yang X, Sun H-Y, Zhao Y-L, Yang Y-G. Dynamic m6a modification and its emerging regulatory role in mRNA splicing. Sci Bull. 2015;60(1):1–32.

38. Wang X, Lu Z, Gornik A, Hong C, Yue Y, Han D, Fu Y, Parisien M, Dai Q, Jia G. N6-methyladenosine-dependent regulation of messenger RNA stability. Nature. 2014;505(7481):117.

39. Zhu DZ, Zhao XF, Liu CZ, Ma FF, Wang F, Gao X-Q, Zhang X. Interaction between RNA helicase ROOT INITIATIVE DECEPTIVE 1 and GAMETOPTYCHIC FACTOR 1 is involved in female gametophyte development in Arabidopsis. J Exp Bot. 2016;67(19):958–65.

40. Tanabe A, Tanikawa K, Tsuchimoto M, Takai K, Ikeda H, Konno J, Jorige T, Maeda H, Kutomi G, Okita K. RNA helicase YTHDC2 promotes cancer metastasis via the enhancement of the efficiency by which HIF-1α mRNA is translated. Cancer Lett. 2016;376(1):34–42.

41. Gao X-Q, Wang N, Wang X-L, Zhang X. Architecture of wheat yield: insights from rice. Trends Plant Sci. 2019;24(9):802–9.

42. Gou J-Y, Li K, Wu K, Wang X, Lin H, Cantu D, Uauy C, Delbon-Alonso A, MIDORIKAWA T, Inoue K. Wheat stripe rust resistance protein WKS1 reduces α-metastasis via the enhancement of the efficiency by which HIF-1α mRNA is translated. Cancer Lett. 2016;376(1):34–42.
48. Zhang C, Huang L, Zhang H, et al. An ancestral NB-LRR with duplicated 3’UTRs confers stripe rust resistance in wheat and barley. Nat Commun. 2019;10(1):4023.
49. Chovancek E, Zivcak M, Botyanszka L, et al. Transient heat waves may affect the photosynthetic capacity of susceptible wheat genotypes due to insufficient photosystem I photoprotection. Plants (Basel). 2019;8(8):282.
50. Zheng M, Chen J, Shi Y, et al. Manipulation of lignin metabolism by plant densities and its relationship with lodging resistance in wheat. Sci Rep. 2017;7:41805.
51. Zhang G, Zhang M, Zhao Z, Ren Y, Li Q, Wang W. Wheat TaPUB1 modulates plant drought stress resistance by improving antioxidant capability. Sci Rep. 2017;7(1):7549.
52. Liu S, Sehgal SK, Lin M, Li J, Trick HN, Gill BS, Bai G. Independent mis-splicing mutations in TaPHS1 causing loss of preharvest sprouting (PHS) resistance during wheat domestication. New Phytol. 2015;208(3):928–35.
53. Yu G, Hou W, Du X, Wang L, Wu H, Zhao L, Kong L, Wang H. Identification of wheat non-specific lipid transfer proteins involved in chilling tolerance. Plant Cell Rep. 2014;33(10):1757–66.
54. Tian F, Gong J, Zhang J, Zhang M, Wang G, Li A, Wang W. Enhanced stability of thylakoid membrane proteins and antioxidant competence contribute to drought stress resistance in the tasg1 wheat stay-green mutant. J Exp Bot. 2013;64(6):1509–20.
55. Biswas DK, Xu H, Li YG, Ma BL, Jiang GM. Modification of photosynthesis and growth responses to elevated CO2 by ozone in two cultivars of winter wheat with different years of release. J Exp Bot. 2013;64(6):1485–96.
56. Bie X, Wang K, She M, Du L, Zhang S, Li J, Gao X, Lin Z, Ye X. Combinational transformation of three wheat genes encoding fructan biosynthesis enzymes confers increased fructan content and tolerance to abiotic stresses in tobacco. Plant Cell Rep. 2012;31(12):2229–38.
57. Yan J, Su P, Wei Z, Nevo E, Kong L. Genome-wide identification, classification, evolutionary analysis and gene expression patterns of the protein kinase gene family in wheat and Aegilops tauschii. Plant Mol Biol. 2017;95(3):327–42.
58. Chen C, Xia R, Chen H, He Y. TBtools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. BioRxiv. 2018;289660.
59. Zhao XY, Hong P, Wu JY, Chen XB, Ye XG, Pan YY, Wang J, Zhang XS. The tae-miR408-mediated control of TaTOC1 genes transcription is required for the regulation of heading time in wheat. Plant Physiol. 2016;170(3):1578–94.
60. Zhao B, Wu TT, Ma SS, Jiang DJ, Bie XM, Sui N, Zhang XS, Wang F. TaD27-B gene controls the tiller number in hexaploid wheat. Plant Biotechnol J. 2020;18(2):513–25.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.