

# Genomic location and expression analysis of expansin gene family reveals the evolutionary and functional significance in *Triticum aestivum*

Nana Li<sup>1</sup> · Yanyan Pu<sup>1</sup> · Yongchao Gong<sup>1</sup> · Yanli Yu<sup>2</sup> · Hanfeng Ding<sup>1</sup>

Received: 4 March 2016 / Accepted: 28 June 2016 / Published online: 9 July 2016  
© The Genetics Society of Korea and Springer-Science and Media 2016

**Abstract** The plant-specific expansin proteins constitute an ancient and major gene family known to have roles in regulating diverse biological processes in plants. Although the functions of many expansin genes have been identified in wheat and other species, little is known about the evolution and genomic locations of the expansin genes in wheat (*Triticum aestivum*). In this study, a total of 87 expansin genes were identified in the wheat genome, including 52 *EXPAs*, 42 *EXPBs* and 4 *EXLAs*. The *EXLB* gene was not found in the wheat genome. Phylogenetic tree and comparative analysis revealed amplification of the *EXPBs* in rice, maize and wheat. The predicted wheat expansins were distributed across 14 of 21 chromosomes with different densities, 3 tightly co-located clusters and 15 paralogous pairs, indicating that tandem duplication and segmental duplication events also played roles in the evolution of expansins in wheat. In addition, the gene structures and conserved protein domains of wheat expansins suggest high levels of conservation within the phylogenetic subgroups. Analysis of a published microarray database showed that most wheat expansin genes exhibit different

expression levels in different tissues and developmental stages. To our knowledge, this is the first report of a genome-wide analysis of the wheat expansin gene family, which should provide valuable information for further elucidating the classification and putative functions of the entire gene family.

**Keywords** Genome-wide analysis · Expansin · Gene family · *Triticum aestivum*

## Abbreviations

BLAST	The Basic Local Alignment Search Tool
HMM	Hidden Markov Model
MEGA	Molecular Evolutionary Genetics Analysis
MUSCEL	Multiple Sequence Comparison by Log-Expectation
NJ	The neighbor-joining
SMART	Simple Modular Architecture Research Tool
GSDS	Gene Structure Display Server
NCBI	National Center for Biotechnology Information
ORF	Open Reading Frame
UTR	Untranslated Regions
Aa	Amino acid
DPBB	Double-psi beta-barrel

**Electronic supplementary material** The online version of this article (doi:10.1007/s13258-016-0446-y) contains supplementary material, which is available to authorized users.

✉ Nana Li  
qyz1127@163.com

✉ Hanfeng Ding  
dinghf2005@163.com

<sup>1</sup> Shandong Center of Crop Germplasm Resources/Shandong Academy of Agricultural Sciences, Jinan 250100, People's Republic of China

<sup>2</sup> Maize Institute, Shandong Academy of Agricultural Sciences, Jinan 250100, People's Republic of China

## Introduction

Expansins comprise a plant-specific superfamily of proteins that characteristically loosen the plant cell wall by weakening the non-covalent bonding of polysaccharides to one another (Cosgrove 2000, 2005; Li et al. 2003; McQueen-Mason and Cosgrove 1994; McQueen-Mason et al. 1992; Sampedro and Cosgrove 2005). Expansin plays

important roles in plant growth and resistance mechanisms, including root growth (Brotman et al. 2008; Guo et al. 2011; Lin et al. 2011), fruit softening (Brummell et al. 1999; Harrison et al. 2001), leaf growth (Cho and Cosgrove 2000; Lu et al. 2013), seed yield (Bae et al. 2014), drought and stress response (Han et al. 2012; Lu et al. 2013). They are also involved in cell enlargement and cell wall changes induced by plant hormones such as gibberellin (Chen et al. 2001), ABA and auxin (Cox et al. 2004), cytokinin (Lee et al. 2008), ethylene (Vreeburg et al. 2005) and brassinosteroids (Park et al. 2010).

Integral expansins range from 250 to 275 amino acids and are composed of three domains: signal peptide, which is presumably removed during processing in the endoplasmic reticulum (Sampedro and Cosgrove 2005); domain 1, which includes a series of conserved cysteines and a His-Phe-Asp (HFD) motif (Cosgrove 2000, 1997); domain 2, which is distantly related to group-2 grass pollen allergens (Cosgrove 2000). Based on phylogenetic and conserved sequence analyses, the expansin superfamily has been classified into four families, EXPA, EXPB, EXLA and EXLB (Kende et al. 2004; Sampedro and Cosgrove 2005).

Wheat (*Triticum aestivum* L.) is one of the most important crops in the world and is relied upon for human food, animal feed and starch ethanol production. In contrast to the extensive research conducted on expansin in both model and crop plants such as *Arabidopsis*, rice, poplar, maize and apple (Hemalatha et al. 2011; Zhang et al. 2014a, b), there are only very limited reports on the characterization of expansin in wheat. Recently, the draft genome sequence of wheat was decoded, which has provided an excellent opportunity for genome-wide analyses of all genes belonging to specific gene families (Mayer 2014; Haider et al. 2013). However, no genome-wide information is currently available on the wheat expansin gene family.

Given the importance of expansin in diverse biological and physiological processes and their potential application for development of transgenic plants, we carried out a systematic analysis of the wheat expansin family in the present study for the first time. Then the chromosome location and gene structure of the putative expansin genes predicted by genome-wide surveys of the wheat genomic sequences were carefully analyzed. Additionally, the putative expansin gene was subjected to phylogenetic analyses with its *Arabidopsis*, rice and maize counterparts. These comparisons enabled the identification of gene orthologs and clusters of orthologous groups that can be subjected to further functional characterization. To our knowledge, this is the first reported genome-wide analysis of the wheat expansin family, which can provide valuable information for understanding the classification and putative functions of *TaEXPs*. These findings open doors to

further improvement of quality and yield in wheat via genetic engineering.

## Materials and methods

### The identification of expansin genes in wheat

To identify the members of the expansin gene family in *T. aestivum*, two different approaches were performed. First, all of the known *Arabidopsis* and rice expansin gene sequences, which downloaded from the *Arabidopsis* genome TAIR 9.0 release (The Arabidopsis Information Resource, <http://www.arabidopsis.org/>) (Poole 2007) and Rice Genome Annotation Project database (Lamesch et al. 2012), were used as query sequences to perform multiple database searches against the proteome and genome files downloaded from PlantGDB (<http://www.plantgdb.org/>) (Dong et al. 2004). Stand-alone versions of BLASTP and TBLASTN (Basic Local Alignment Search Tool: <http://blast.ncbi.nlm.nih.gov>) (Altschul et al. 1990), which are available from the National Center for Biotechnology Information (NCBI), were used with an e-value cutoff set to  $1e-003$ . All of the protein sequences derived from the collected candidate expansin genes were examined using the domain analysis programs Pfam (Protein family: <http://pfam.sanger.ac.uk/>) (Punta et al. 2012) and Simple Modular Architecture Research Tool (SMART): <http://smart.embl-heidelberg.de/>) (Letunic et al. 2012) with the default cutoff parameters. Second, we analyzed the domains of all of the wheat peptide sequences using a Hidden Markov Model (HMM) (Jeanmougin et al. 1998; Wu et al. 2002) analysis with Pfam searching. Then, we obtained the sequences with the PF03330 and PF01357 Pfam number, which contained typical DPBB\_1 and Pollen\_allerg\_1 domain, from the wheat genome sequences using a Perl-based script. Finally, all of the protein sequences were compared with known expansin sequences using ClustalX (<http://www.clustal.org/>) to verify the sequences were candidate Expansins (Jeanmougin et al. 1998). Cellular roles and gene ontology (GO) categories were predicted using the ProtFun2.2 server (<http://www.cbs.dtu.dk/services/ProtFun/>), which relies solely on the protein sequence as input (Jensen et al. 2002, 2003).

### Chromosomal location and gene structure of the expansin genes in wheat

The chromosomal locations were retrieved from the genome data downloaded from the wheat genome sequences database using a Perl-based program and mapped to the chromosomes using the Circos (Krzywinski et al. 2009), as well as the gene structure of the expansin genes were

generated with the Gene Structure Display Server (GSDS): <http://gsds.cbi.pku.edu.cn/>) (Guo et al. 2007).

### Sequence alignment and phylogenetic analysis

The Expansin sequences were aligned using the ClustalX program with BLOSUM30 as the protein-weight matrix (Jeanmougin et al. 1998). The Multiple Sequence Comparison by Log-Expectation (MUSCLE) program (version 3.52) was also used to perform multiple sequence alignments to confirm the ClustalX results (Edgar 2004). Phylogenetic trees of the expansin protein sequences were constructed using the neighbor-joining (NJ) method of the Molecular Evolutionary Genetics Analysis (MEGA5 program, <http://www.megasoftware.net/>) using the p-distance and complete deletion option parameters (Tamura et al. 2011). The reliability of the obtained trees was tested using a bootstrapping method with 1000 replicates. Phylogenetic and chromosomal location analyses were used to identify duplicated genes. The number of nonsynonymous substitutions per nonsynonymous site (Ka) and synonymous substitutions per synonymous site (Ks) were calculated by DnaSP (Librado and Rozas 2009; Rozas 2009).

### Expression analyses of the expansin genes in microarray

The development and tissue expression data for various tissues/organs and developmental stages were obtained using Genevestigator (<https://www.genevestigator.com/gv/>) with the *T. aestivum* gene chip platforms. Then, the identified expansin-containing gene IDs were used as query sequences to perform searches in the gene chip platform of genevestigator (Grennan 2006).

## Results

### Identification of expansin genes in wheat

To identify genes encoding expansin in the wheat genome, BLASTP searches were performed using the entire wheat peptide sequences of three other well-studied expansins (*Arabidopsis*, rice and maize expansins) as queries. The HMM of the SMART and Pfam tools were then exploited as queries to confirm the putative expansin genes. Finally, 98 typical expansin genes containing full open reading frames (ORF) were identified and manually analyzed using InterProScan and ClustalX program to confirm the presence of conserved domains in expansin, and were used for further analysis as described below. In order to distinguish them from the remaining expansin genes, we provisionally named them as *TaEXPA1*, *TaEXPA2* and *TaEXPB1*, based on their

genome locations (Table S1). The ORF ranged from 573 (*TaEXPB24*) to 1257 bp (*TaEXPB26*) in length (average = 803 bp). The identified *TaEXP* genes encode proteins ranging from 190 (*TaEXPB24*) to 418 (*TaEXPB26*) amino acids (aa) in length (average = 267 aa) (Table S1).

These expansin genes in wheat were larger than those of other reported plants genomes. The EXPB family has more members in wheat compared to rice and maize (Table 1). The wheat EXLs have not expanded, which has occurred in other plants except soybean, and which may lead to functional changes in soybean. Expansion of the EXPB gene family is particularly impressive in grasses (rice, maize and wheat), with 19, 48 and 42 genes in rice, maize and wheat compared to six in *Arabidopsis*, three in poplar and one in apple. The main reason underlying this difference is the existence of a large number of tandem duplications in the rice, maize and wheat genomes. Notably, the components of the cell wall are different in monocots compared to the dicots, and there are more glucuronic arab glycosylation xylan and  $\beta$ -(1 → 3), (1 → 4)-D-glucan in monocots. These polysaccharides may be substrates for the EXPB genes (Sarkar et al. 2009; Yennawar et al. 2006).

### Phylogenetic relationships and comparative analysis of the expansin gene family in wheat

In order to evaluate the evolutionary relationship among the expansin proteins, full-length amino acid sequences of 98 TaEXPs, 35 AtEXPs, 56 OsEXPs and 88 ZmEXPs were subjected to multiple sequence alignment using the MEGA5 program. The multiple sequence alignment file was then used to construct an unrooted phylogenetic tree. As shown in Fig. 1, the phylogenetic tree divided the EXPs into four groups (EXPA, EXPB, EXLA and EXLB), which was defined by the maximum likelihood method (Fig S1). To be consistent with previous studies, we adopted the same name for the 17 subgroups (Li et al. 2002; Lin et al.

**Table 1** Number of expansin genes in 10 plant species

Species	EXPA	EXPB	EXLA	EXLB	Total
<i>Arabidopsis</i>	26	6	3	1	36
Poplar	27	3	2	4	36
Grape	20	4	1	4	29
Apple	34	1	2	4	41
Soybean	49	9	2	15	75
Common bean	25	6	0	5	36
<i>Medicago truncatula</i>	16	1	0	1	18
Rice	34	19	4	1	58
Maize	36	48	4	0	88
<i>Triticum aestivum</i>	52	42	4	0	98

2005; Sampedro and Cosgrove 2005; Zhang et al. 2014a). Meanwhile, 98 *TaEXP*s were used to conduct phylogenetic analyses. Twenty-seven sister pairs and two clusters of paralogous *TaEXP* genes were found and marked by purple shadow in the phylogenetic tree, which have very strong bootstrap support (>90 %) (Fig. 2).

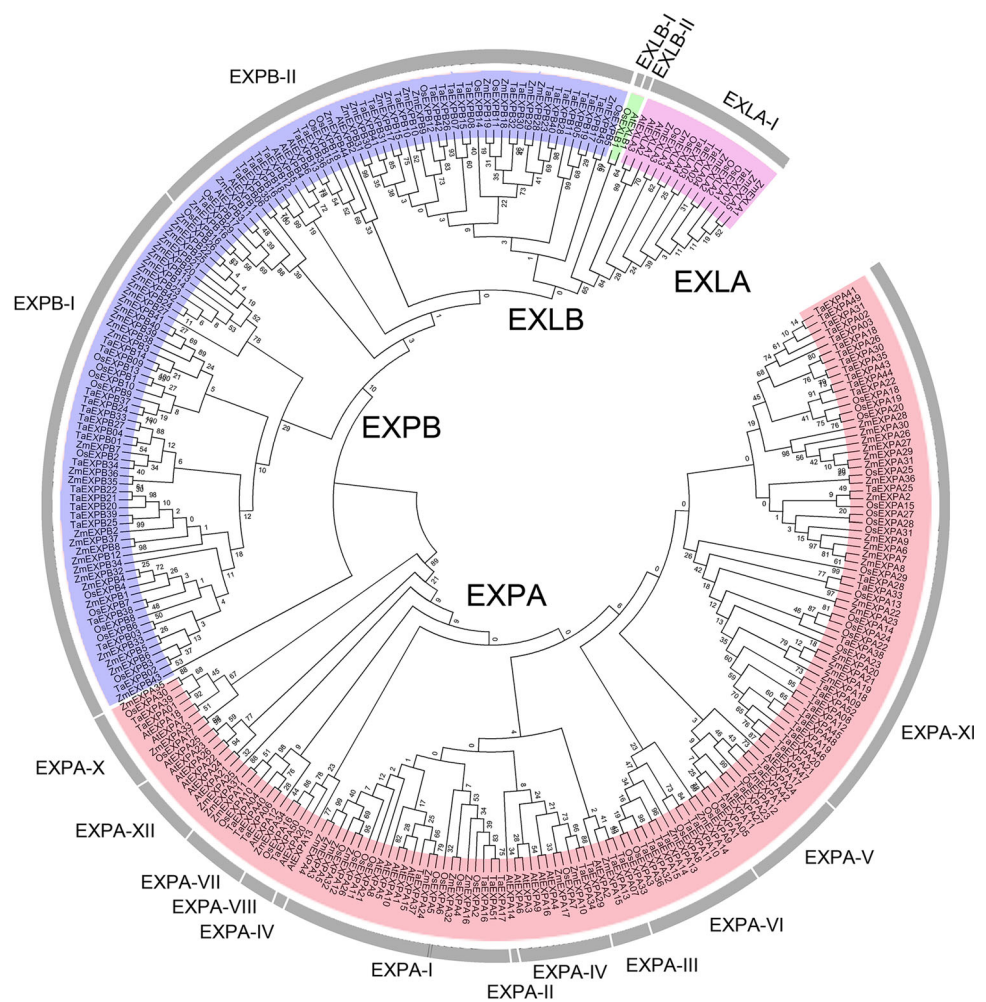
Structural analyses are supposed to provide valuable information concerning duplication events when interpreting phylogenetic relationships within gene families. In the *TaEXP* gene family, the number of exons varied from 1 (*TaEXPB24* and *TaEXPB37*) to 5 (*TaEXPB26* and *TaEXPB38*) (Fig. 2). Additionally, most members within the same subfamily shared a similar exon structure and length. In the EXLA subfamily, three expansins had 3 exons and *TaEXLA4* had 4 exons. Most *TaEXPAs* had 2, 3 or 4 exons, 24 *TaEXPBs* had 4 exons, 41 *TaEXPBs* had 3 exons, and 29 genes had 2 exons. Furthermore, by comparing the gene structures of *Arabidopsis* and rice, we found that they shared similar expansin sequences and intron/exon structures, which indicated that all the *TaEXP*

**Fig. 2** Phylogenetic relationships, gene structure and protein structure of *TaEXP* genes. The following panels are shown, from left to right. The phylogenetic tree was constructed based on a complete protein sequence alignment of *TaEXP*s by the neighbor-joining method with bootstrapping analysis (1000 replicates). The subgroups are marked by colorful background. The numbers next to the branches indicate the bootstrap values that support the adjacent node. Ten sister pairs of paralogous expansin genes are found and indicated by yellow shadow, which had very strong bootstrap support (>90 %). The gray boxes, blue boxes and black lines in the gene structure diagram represent exons, UTR and introns, respectively. The scale bar represents 1000 bp. Protein structure domain were obtained from PFAM and SMART, the red box in the protein structure diagram represents the signal peptide; the pink and the green boxes represent a DPBB domain and Pollen allergen domain, respectively

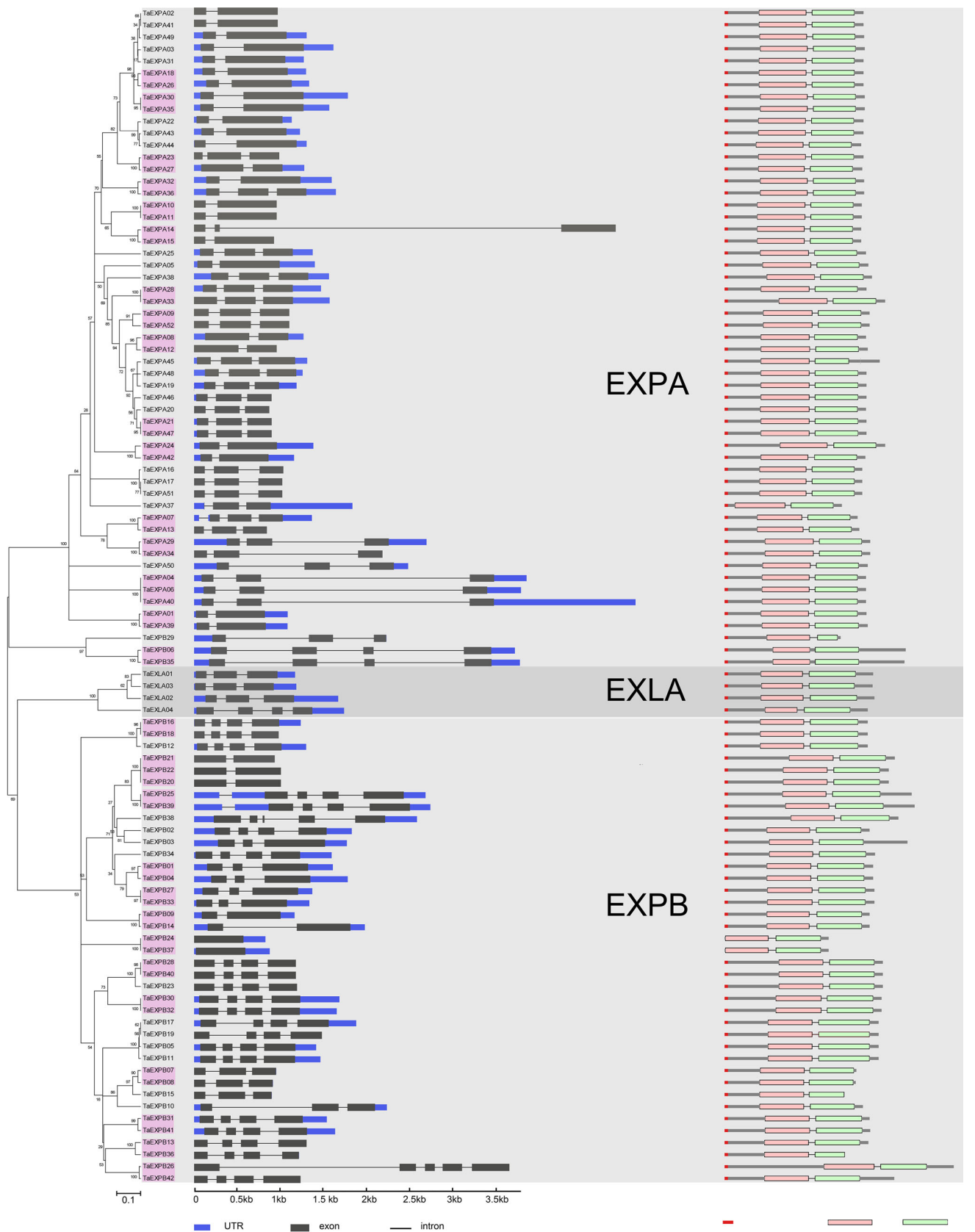
genes belonged to the expansin gene superfamily. Thus, our findings reveal that the gene structure of expansins is relatively simple.

Next, the protein structure of each expansin gene was analyzed using the SignalP, SMART and Pfam databases

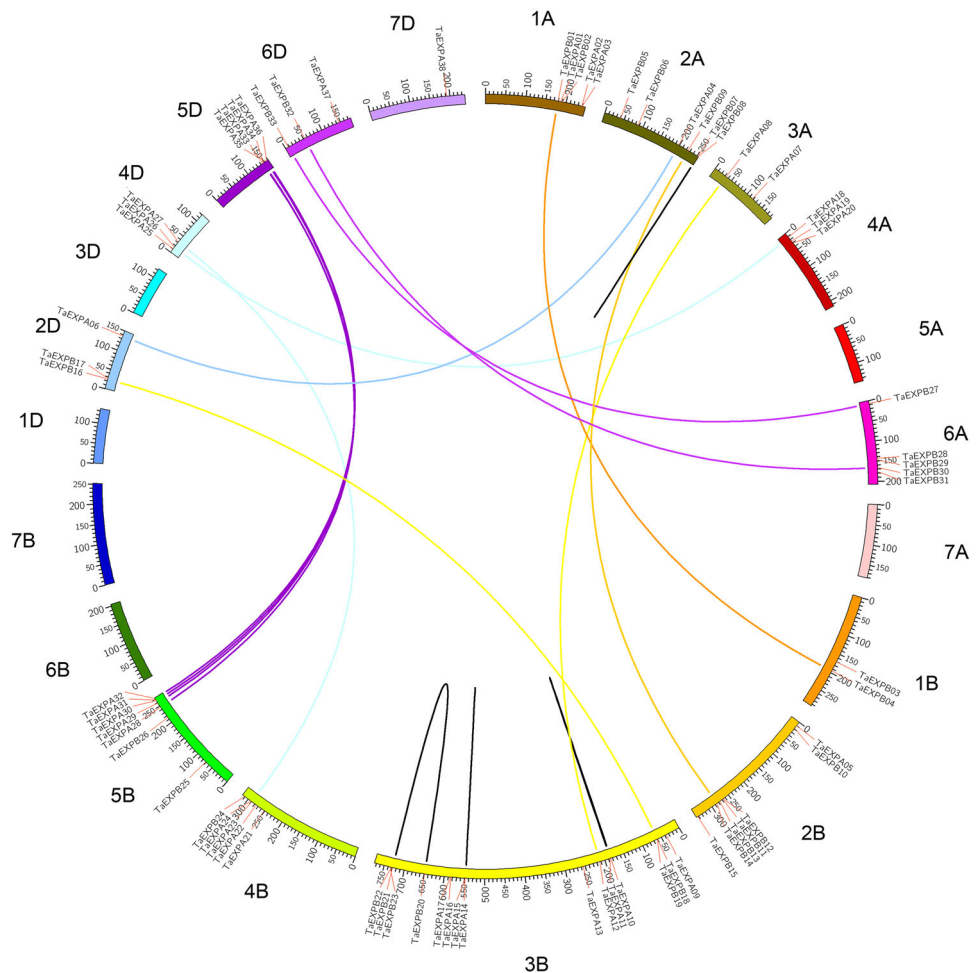
**Fig. 1** Phylogenetic relationship of *Arabidopsis*, rice, maize and wheat expansin genes. The phylogenetic tree was constructed based on a complete protein sequence alignment of expansins in *Arabidopsis*, rice, maize and wheat by the neighbor-joining method with bootstrapping analysis (1000 replicates). The subgroups are marked by colorful background







**Fig. 3** Positions of expansin gene family members on the wheat chromosomes. Sister paralogous pairs are indicated by color line

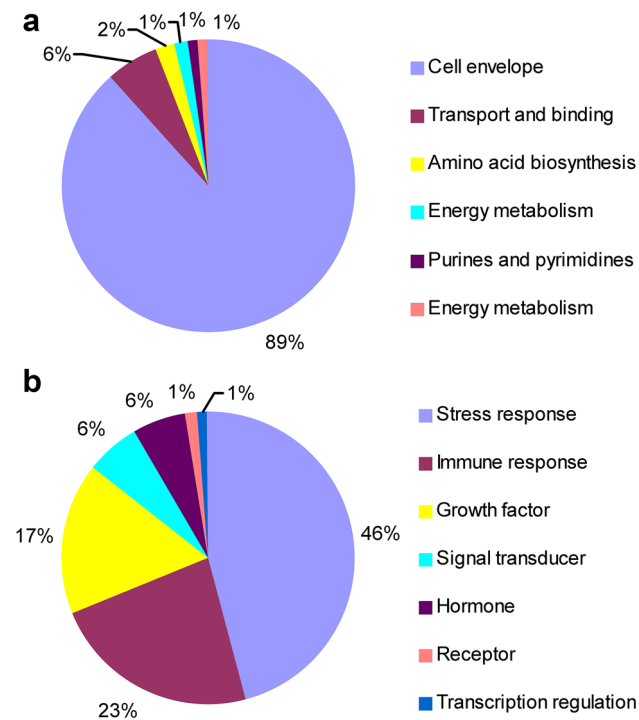


(Fig. 2). All wheat expansins contained a conserved region that showed a double-psi beta-barrel (DPBB) fold, which is an enzymatic domain, and pollen allergen in the N-terminus (Fig. 2). The conserved intron/exon structure and protein domain in each subfamily supports their close evolutionary relationship and the classification introduced here. Expansin proteins from different families share only 20–40 % identity with each other. The degree of conservation is highest in domain 1 (Sampedro and Cosgrove 2005). Expansin domain 1 has distant homology to GH45 proteins. Proteins from this family have been crystallized and their mechanism of action has been determined (Sampedro and Cosgrove 2005). On the basis of hydrophobic cluster analysis, this structural motif is also known to be present in expansins (Sampedro and Cosgrove 2005). EXLA and EXLB proteins lack the HFD motif, which suggests that their action may differ from that of others, but the functional implications of these differences are currently unknown. Furthermore, by comparing their structure with those of other plants, we found that all expansins share a similar protein conserved domain and intron/exon structure.

### Chromosomal locations of expansins in the wheat genome

The physical locations of expansins on wheat chromosomes are depicted in Fig. 3. Seventy-one *TaEXPs* were found in 14 of the 21 wheat chromosomes. The A, B and D genomes had 21, 36 and 14 expansins, respectively. Chromosome 3B contained 15 expansins, which was the largest; five clusters containing 11 genes were tightly co-located in chromosome 3B, which suggested that tandem duplication events played roles in the evolution of expansin in chromosome 3B. Then, chromosomes 2B and 5B were found to have seven genes, respectively, while chromosome 7D had only 1 *TaEXP*.

Next, we searched for evidence of past gene duplication events (tandem and segmental) to elucidate the mechanism underlying the expansion of the *TaEXP* gene family. As shown in Fig. 3, 15 segmental duplication events of *TaEXPs* (20 genes) were found in the wheat genome. All of these segmentally duplicated genes were also found to be paralogs in our phylogenetic analysis, as shown in Fig. 2. These results indicated that segmental duplications played



**Fig. 4** Functional categorization analysis of wheat expansins using ProtFun. **a** Functional category, **b** Gene ontology category

important roles in expansin gene family expansion in the wheat genome. Moreover, three clusters were tightly collocated in the wheat genome, which suggested that tandem duplication events also played important roles in the evolution of the expansin gene family in wheat. To explore different selection constraints on the duplicated *ZmEXP* genes, we calculated  $K_s$  and  $K_a/K_s$  ratios for each duplicated pair. A  $K_a/K_s$  ratio less than one generally indicates accelerated evolution with positive selection, a ratio equal to one corresponds to neutral selection, and a ratio less than one indicates negative or purifying selection. The  $K_a/K_s$  ratios of all duplicated pairs were less than one, implying strong purifying selection (Table S2). These results suggest that the functions of the duplicated genes have not diverged drastically over the course of evolution following the duplication events.

### GO analysis

To predict the biological functions of TaEXP, cellular roles and GO category annotations were explored using the ProtFun server. The majority of genes fell into the cell envelope functional category and GO categories of stress response and immune response (Fig. 4), indicating their possible roles in these processes. In particular, 89 % of TaEXPs were classified into the cell envelope functional category, and more than 46 % of TaEXPs belonged to the stress response category.

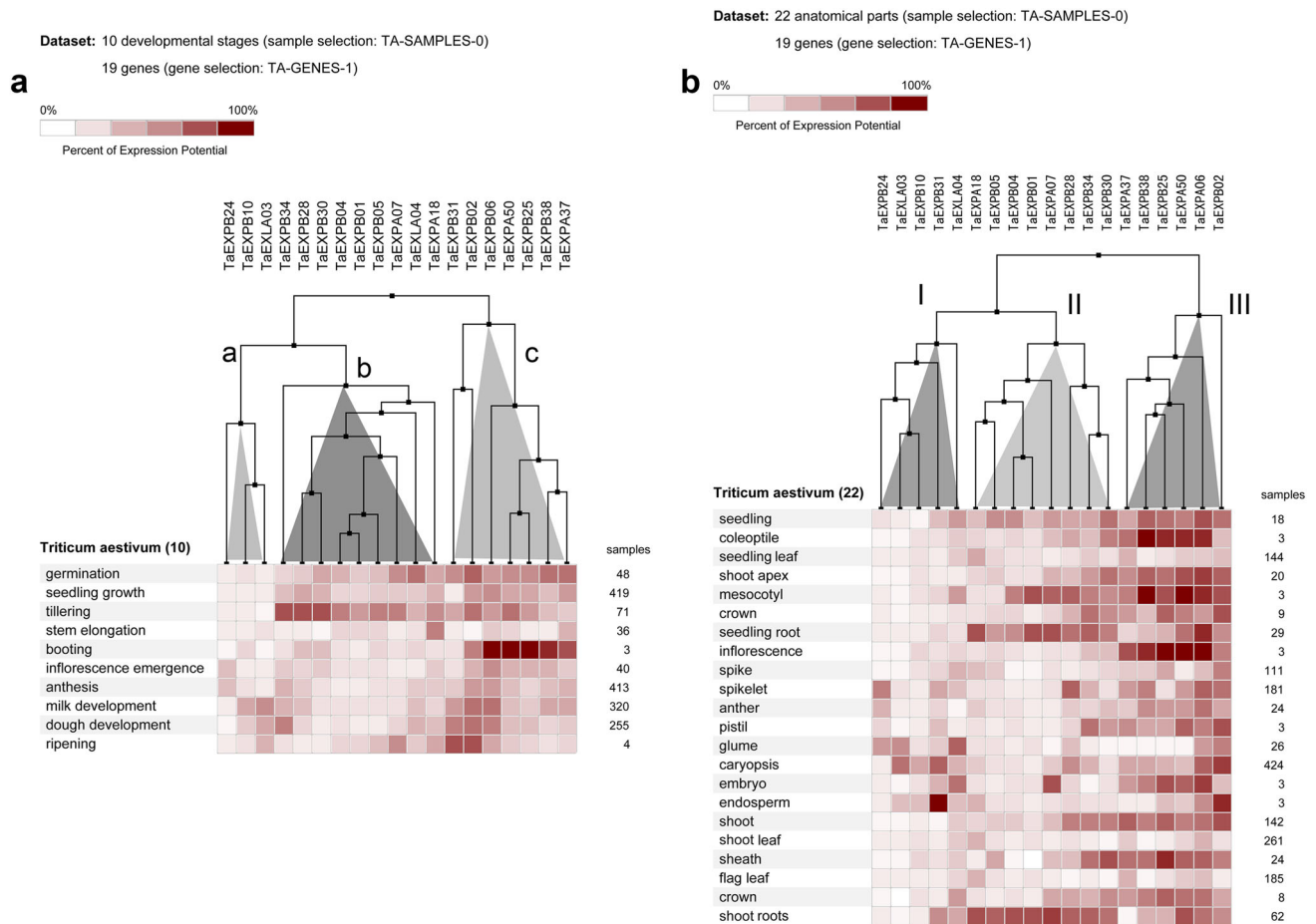
### Determination of TaEXP expression pattern by microarray analysis

To investigate the expression profiles of expansin during wheat development, we analyzed the expression of the expansin genes at various developmental stages using genevestigator. A total of 19 genes were found to have different expression levels across 10 stages (germination, seedling growth, tillering, stem elongation, booting, inflorescence emergence, anthesis, milk development, dough development and ripening). All 19 genes were detected in more than one stage (Fig. 5a). Most genes were expressed in all developmental stages, such as *TaEXPB1*, *TaEXPB4*, *TaEXPB5*, *TaEXLA3*, and *TaEXLA4*. Based on hierarchical clustering, the expression patterns of the wheat expansin genes were divided into three groups, marked a, b and c. Group a genes exhibited low expression value at all ten stages, 3 of 9 group b genes showed higher expression at tillering, and the expression values of all 7 group c genes at booting stage were found to be higher than those of others. Overall, the broad expression patterns of expansins suggest that they play important roles in wheat development.

Next, we investigated wheat expression patterns in a variety of tissues. Heatmap representation of the expression profiles of 19 expansin genes in 22 wheat tissues is shown in Fig. 5b. The expansin gene expression patterns could be divided into three groups, namely group I, II and III, based on the results of hierarchical clustering analysis. Most group I expansin genes were found to have higher expression in caryopsis, embryo and endosperm. Most group II genes had higher expression levels in rhizome roots. Group III genes had higher expression in seedling, coleoptile, seedling leaf, shoot apex, mesocotyl, crown, seedling root and inflorescence than other developmental stages, which suggested that group III genes may play important roles in seedling development.

### Discussion

Expansins play important roles in diverse processes, including developmental programs, defense and abiotic stress responses (Sampedro and Cosgrove 2005). Besides, the presence of expansin genes in land plants ranging from mosses to eudicots makes them interesting candidates for studying the evolution of plant development (Carey and Cosgrove 2007; Dal Santo et al. 2013; Zhu et al. 2014). A previous genome-wide analysis showed that expansins are widely present in plants, concluding mosses (Carey and Cosgrove 2007), legumes (Zhu et al. 2014), as well as cruciferous (Sampedro et al. 2006), rosaceae (Zhang et al.



**Fig. 5** The expression profiles of *TaEXP* genes in development stages and tissues. **a** The expression profile of *TaEXP* genes in 10 stages of development. The dark and light colour shadings represent relative high or low expression levels, respectively, of the *TaEXP* genes in ten stages of development (germination, seedling growth, tillering, stem elongation, booting, inflorescence emergence, anthesis, milk development, dough development and ripening). **b** The

expression profile of *TaEXP* genes in 22 different tissues. The dark and light colour shadings represent relative high or low expression levels, respectively, of the *TaEXP* genes in 22 tissues, including seedling, coleoptile, seedling leaf, shoot apex, mesocotyl, crown, seedling root, inflorescence, spike, spikelet, anther, pistil, glume, caryopsis, embryo, endosperm, shoot, shoot leaf, sheath, flag leaf, crown and rhizome roots

2014a), grassy (Sampedro et al. 2006) plants. However, genomic data on wheat expansins is not available to date.

In this study, we first conducted genome-wide analysis and identified 98 expansins containing full ORFs in wheat, which constitutes the largest expansin family among the well-studied species (Krishnamurthy et al. 2015; Zhang et al. 2014a; Zhu et al. 2014). By comparing across different plant species, we found the large expansion of EXPB subfamily in rice and wheat, which may due to the difference in substrates in the cell walls of monocots and dicots. EXPB may interact with glucuronic arab glycosylation xylan and  $\beta$ -(1  $\rightarrow$  3), (1  $\rightarrow$  4)-D-glucan, which are abundant in monocots (Sarkar et al. 2009; Yennawar et al. 2006). What factors drove expansion of the EXPB subfamily during plant evolution? In wheat, 3 tightly co-located clusters were found to contain 6 *EXPBs*, and 15 sister gene pairs were found in *EXPBs*. So, segmental

duplications and tandem duplication events may play roles in the evolution of *EXPBs* in wheat.

Generally, orthologous genes and genes from the same subfamily hold similar functions. In *Arabidopsis*, root hair-specific *AtEXPA7* and *AtEXPA18* genes are necessary for root hair elongation (Cho and Cosgrove 2002; Lin et al. 2011). *TaEXPA33* and *TaEXPA35* are phylogenetically related to *AtEXPA7* as shown in Fig. 1, suggesting that they may share similar expression patterns and functions. *AtEXPA1*, another well-studied *expansin*, can accelerate stomatal opening by decreasing the volumetric elastic modulus (Zhang et al. 2011), so the homologous genes *TaEXPA15* and *TaEXPA24* may play similar roles in wheat. Likewise, *ZmEXLAs* may have similar function in abiotic stress as the homologous gene *AtEXLA2* in *Arabidopsis* (Abuqamar et al. 2013). Moreover, *TaEXPB4* and *TaEXPB32* may play roles in



seedling growth based on the known role of *OsEXP4* in rice (Choi et al. 2003).

Gene expression data revealed the expression patterns of 19 expansins during 10 different wheat development stages and 22 tissues. Results of our expression pattern analysis suggest that expansin genes may play different roles during development and in different tissues. Furthermore, the similarities in expression patterns of most sister genes indicate that functional redundancy is widespread in expansins. GmEXPB2 affects soybean nodulation by modifying the root architecture and promoting nodule formation and development (Li et al. 2015). HvEXPB7, which was revealed by studying the root hair transcriptome of Tibetan wild barley, improves root hair growth under drought stress (He et al. 2015). Overexpression of CIEIPA1 and CIEIPA2 affect growth and development in transgenic tobacco and increase the amount of cellulose in stem cell walls (Wang et al. 2011). DzEXP1 and DzEXP2 are involved in fruit development (Palapol et al. 2015). Thus, the expansins play very important roles in development of plants, including roots, stems, and fruits.

In summary, we present here a complete genomic analysis of the expansin family in wheat along with a wide range of expression data. Our findings should be helpful in laying the foundation for functional characterization of the expansin gene family and further gaining an understanding of the structure–function relationship between these family members. Additionally, our study provides comprehensive information and novel insights into the evolution and divergence of expansin genes in plants. Taken together, these data may aid in elucidating the molecular basis of many important aspects of wheat physiology, such as root development and other processes.

**Acknowledgments** This work was supported by Youth Scientific Research Foundation of Shandong Academy of Agricultural Sciences (Grant No. 2014QNM09 to Nana Li), Natural Science Foundation of Shandong Province (Grant No. ZR2014YL017 to Nana Li) and Shandong Province Science and Technology Development Program (Grant No. 2012G0021031 to Nana Li).

#### Compliance with ethical standards

**Conflict of interest** Nana Li, Yanyan Pu, Yongchao Gong, Yanli Yu and Hanfeng Ding declares no conflict of interest.

**Ethical approval** The article does not contain any studies with human participants or animals performed by any of the authors.

#### References

Abuqamar S, Ajeb S, Sham A, Enan MR, Iratni R (2013) A mutation in the expansin-like A2 gene enhances resistance to necrotrophic fungi and hypersensitivity to abiotic stress in *Arabidopsis thaliana*. *Mol Plant Pathol* 14:813–827

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Bae JM, Kwak MS, Noh SA, Oh MJ, Kim YS, Shin JS (2014) Overexpression of sweetpotato expansin cDNA (IbEXP1) increases seed yield in *Arabidopsis*. *Transgenic Res* 23:657–667
- Brotman Y, Briff E, Viterbo A, Chet I (2008) Role of swollenin, an expansin-like protein from *Trichoderma*, in plant root colonization. *Plant Physiol* 147:779–789
- Brummell DA, Harpster MH, Civeello PM, Palys JM, Bennett AB, Dunsmuir P (1999) Modification of expansin protein abundance in tomato fruit alters softening and cell wall polymer metabolism during ripening. *Plant Cell* 11:2203–2216
- Carey RE, Cosgrove DJ (2007) Portrait of the expansin superfamily in *Physcomitrella patens*: comparisons with angiosperm expansins. *Ann Bot* 99:1131–1141
- Chen F, Dahal P, Bradford KJ (2001) Two tomato expansin genes show divergent expression and localization in embryos during seed development and germination. *Plant Physiol* 127:928–936
- Cho HT, Cosgrove DJ (2000) Altered expression of expansin modulates leaf growth and pedicel abscission in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 97:9783–9788
- Cho HT, Cosgrove DJ (2002) Regulation of root hair initiation and expansin gene expression in *Arabidopsis*. *Plant Cell* 14:3237–3253
- Choi D, Lee Y, Cho HT, Kende H (2003) Regulation of expansin gene expression affects growth and development in transgenic rice plants. *Plant Cell* 15:1386–1398
- Cosgrove DJ (1997) Relaxation in a high-stress environment: the molecular bases of extensible cell walls and cell enlargement. *Plant Cell* 9:1031–1041
- Cosgrove DJ (2000) Loosening of plant cell walls by expansins. *Nature* 407:321–326
- Cosgrove DJ (2005) Growth of the plant cell wall. *Nat Rev Mol Cell Biol* 6:850–861
- Cox MC, Benschop JJ, Vreeburg RA, Wagemaker CA, Moritz T, Peeters AJ, Voeselek LA (2004) The roles of ethylene, auxin, abscisic acid, and gibberellin in the hyponastic growth of submerged *Rumex palustris* petioles. *Plant Physiol* 136:2948–2960
- Dal Santo S, Vannozzi A, Tornielli GB, Fasoli M, Venturini L, Pezzotti M, Zenoni S (2013) Genome-wide analysis of the expansin gene superfamily reveals grapevine-specific structural and functional characteristics. *PLoS One* 8:e62206
- Dong Q, Schlueter SD, Brendel V (2004) PlantGDB, plant genome database and analysis tools. *Nucleic Acids Res* 32:354–359
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
- Grennan AK (2006) Genevestigator facilitating web-based gene-expression analysis. *Plant Physiol* 141:1164–1166
- Guo AY, Zhu QH, Chen X, Luo JC (2007) GSDBS: a gene structure display server. *Yi Chuan* 29:1023–1026
- Guo W, Zhao J, Li X, Qin L, Yan X, Liao H (2011) A soybean beta-expansin gene GmEXPB2 intrinsically involved in root system architecture responses to abiotic stresses. *Plant J* 66:541–552
- Haider N (2013) The origin of the B-genome of bread wheat (*Triticum aestivum* L.). *Genetika* 49:303–314
- Han Y, Li A, Li F, Zhao M, Wang W (2012) Characterization of a wheat (*Triticum aestivum* L.) expansin gene, TaEXPB23, involved in the abiotic stress response and phytohormone regulation. *Plant Physiol Biochem* 54:49–58
- Harrison EP, McQueen-Mason SJ, Manning K (2001) Expression of six expansin genes in relation to extension activity in developing strawberry fruit. *J Exp Bot* 52:1437–1446
- He X, Zeng J, Cao F, Ahmed IM, Zhang G, Vincze E, Wu F (2015) HvEXPB7, a novel beta-expansin gene revealed by the root hair

- transcriptome of tibetan wild barley, improves root hair growth under drought stress. *J Exp Bot* 66:7405–7419
- Hemalatha N, Rajesh MK, Narayanan NK (2011) Genome-wide analysis and identification of genes related to expansin gene family in indica rice. *Int J Bioinform Res Appl* 7:162–167
- Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple sequence alignment with clustal X. *Trends Biochem Sci* 23:403–405
- Jensen LJ, Gupta R, Blom N, Devos D, Tamames J, Kesmir C, Nielsen H, Staerfeldt HH, Rapacki K, Workman C, Andersen CA, Knudsen S, Krogh A, Valencia A, Brunak S (2002) Prediction of human protein function from post-translational modifications and localization features. *J Mol Biol* 319:1257–1265
- Jensen LJ, Gupta R, Staerfeldt HH, Brunak S (2003) Prediction of human protein function according to gene ontology categories. *Bioinformatics* 19:635–642
- Kende H, Bradford K, Brummell D, Cho HT, Cosgrove D, Fleming A, Gehring C, Lee Y, McQueen-Mason S, Rose J, Voeselek LA (2004) Nomenclature for members of the expansin superfamily of genes and proteins. *Plant Mol Biol* 55:311–314
- Krishnamurthy P, Hong JK, Kim JA, Jeong MJ, Lee YH, Lee SI (2015) Genome-wide analysis of the expansin gene superfamily reveals *Brassica rapa*-specific evolutionary dynamics upon whole genome triplication. *Mol Genet Genom* 290:521–530
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circo: an information aesthetic for comparative genomics. *Genome Res* 19:1639–1645
- Lamesch P, Berardini TZ, Li D, Swarbrick D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, Karthikeyan AS, Lee CH, Nelson WD, Ploetz L, Singh S, Wensel A, Huala E (2012) The *Arabidopsis* information resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res* 40:1202–1210
- Lee A, Giordano W, Hirsch AM (2008) Cytokinin induces expansin gene expression in melilotus alba deswild-type and the non-nodulating, non-mycorrhizal (NodMyc) mutant masym3. *Plant Signal Behav* 3:218–223
- Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* 40:302–305
- Li Y, Darley CP, Ongaro V, Fleming A, Schipper O, Baldauf SL, McQueen-Mason SJ (2002) Plant expansins are a complex multigene family with an ancient evolutionary origin. *Plant Physiol* 128:854–864
- Li Y, Jones L, McQueen-Mason S (2003) Expansins and cell growth. *Curr Opin Plant Biol* 6:603–610
- Li X, Zhao J, Tan Z, Zeng R, Liao H (2015) GmEXPB2, a cell wall beta-expansin, affects soybean nodulation through modifying root architecture and promoting nodule formation and development. *Plant Physiol* 169:2640–2653
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452
- Lin Z, Ni Z, Zhang Y, Yao Y, Wu H, Sun Q (2005) Isolation and characterization of 18 genes encoding  $\alpha$ - and  $\beta$ -expansins in wheat (*Triticum aestivum* L.). *Mol Genet Genom* 274:548–556
- Lin C, Choi HS, Cho HT (2011) Root hair-specific EXPANSIN A7 is required for root hair elongation in *Arabidopsis*. *Mol Cells* 31:393–397
- Lu P, Kang M, Jiang X, Dai F, Gao J, Zhang C (2013) *RhEXPA4*, a rose expansin gene, modulates leaf growth and confers drought and salt tolerance to *Arabidopsis*. *Planta* 237:1547–1559
- Mayer KFX (2014) A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* 345:1251788
- McQueen-Mason S, Cosgrove DJ (1994) Disruption of hydrogen bonding between plant cell wall polymers by proteins that induce wall extension. *Proc Natl Acad Sci USA* 91:6574–6578
- McQueen-Mason S, Durachko DM, Cosgrove DJ (1992) Two endogenous proteins that induce cell wall extension in plants. *Plant Cell* 4:1425–1433
- Palapol Y, Kunyamee S, Thongkhum M, Ketsa S, Ferguson IB, van Doorn WG (2015) Expression of expansin genes in the pulp and the dehiscence zone of ripening durian (*Durio zibethinus*) fruit. *J Plant Physiol* 182:33–39
- Park CH, Kim TW, Son SH, Hwang JY, Lee SC, Chang SC, Kim SH, Kim SW, Kim SK (2010) Brassinosteroids control *AtEXPA5* gene expression in *Arabidopsis thaliana*. *Phytochemistry* 71:380–387
- Poole RL (2007) The TAIR database. *Methods Mol Biol* 406:179–212
- Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer EL, Eddy SR, Bateman A, Finn RD (2012) The Pfam protein families database. *Nucleic Acids Res* 40:290–301
- Rozas J (2009) DNA sequence polymorphism analysis using DnaSP. *Methods Mol Biol* 537:337–350
- Sampedro J, Cosgrove DJ (2005) The expansin superfamily. *Genome Biol* 6:242
- Sampedro J, Carey RE, Cosgrove DJ (2006) Genome histories clarify evolution of the expansin superfamily: new insights from the poplar genome and pine ESTs. *J Plant Res* 119:11–21
- Sarkar P, Bosneaga E, Auer M (2009) Plant cell walls throughout evolution: towards a molecular understanding of their design principles. *J Exp Bot* 60:3615–3635
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739
- Vreeburg RA, Benschop JJ, Peeters AJ, Colmer TD, Ammerlaan AH, Staal M, Elzenga TM, Staals RH, Darley CP, McQueen-Mason SJ, Voeselek LA (2005) Ethylene regulates fast apoplastic acidification and expansin a transcription during submergence-induced petiole elongation in rumex palustris. *Plant J* 43:597–610
- Wang G, Gao Y, Wang J, Yang L, Song R, Li X, Shi J (2011) Overexpression of two cambium-abundant Chinese fir (*Cunninghamia lanceolata*) alpha-expansin genes ClEXPA1 and ClEXPA2 affect growth and development in transgenic tobacco and increase the amount of cellulose in stem cell walls. *Plant Biotechnol J* 9:486–502
- Wu X, Song C, Wang B, Cheng J (2002) Hidden markov model used in protein sequence analysis. *J Biomed Eng* 19:455–458
- Yennawar NH, Li LC, Dudzinski DM, Tabuchi A, Cosgrove DJ (2006) Crystal structure and activities of EXPB1 (*Zea m* 1), a  $\beta$ -expansin and group-1 pollen allergen from maize. *Proc Natl Acad Sci USA* 103:14664–14671
- Zhang XQ, Wei PC, Xiong YM, Yang Y, Chen J, Wang XC (2011) Overexpression of the *Arabidopsis* alpha-expansin gene *AtEXPA1* accelerates stomatal opening by decreasing the volumetric elastic modulus. *Plant Cell Rep* 30:27–36
- Zhang S, Xu R, Gao Z, Chen C, Jiang Z, Shu H (2014a) A genome-wide analysis of the expansin genes in malus x domestica. *Mol Genet Genom* 289:225–236
- Zhang W, Yan H, Chen W, Liu J, Jiang C, Jiang H, Zhu S, Cheng B (2014b) Genome-wide identification and characterization of maize expansin genes expressed in endosperm. *Mol Genet Genom* 289:1061–1074
- Zhu Y, Wu N, Song W, Yin G, Qin Y, Yan Y, Hu Y (2014) Soybean (*Glycine max*) expansin gene superfamily origins: segmental and tandem duplication events followed by divergent selection among subfamilies. *BMC Plant Biol* 14:93–111