# Identification and validation of organ-preferential genes and analysis of corresponding upstream tissue-specific promoters in wheat

P.P. SU[1], X. JIN[1], T. SUN[1], L. CHEN[2,3], F. SHI[1], K.X. LI[1], J.L. CHANG[1], G.X. YANG[1]*, and G.Y. HE[1]*

*College of Life Science and Technology, Huazhong University of Science & Technology, Wuhan 430074, China*[1]
*Institute of Food Crops, Hubei Academy of Agricultural Science, Hubei, Wuhan 430064, China*[2]
*Hubei Key Laboratory of Food Crop Germplasm and Genetic Improvement, Hubei, Wuhan 430064, China*[3]

## Abstract

Tissue/organ-specific promoters are important tools in genetic engineering and crop molecular breeding. They are well characterized in dicots, such as *Arabidopsis*, tobacco, and tomato, but not sufficiently in monocots, especially in wheat. In this study, the genes specifically expressed in seven different tissues, including coleoptile, root, leaf, pistil, anther, embryo, and endosperm were identified through analyzing the public transcriptome data from a wheat microarray using the ROKU method. The expression patterns of selected genes were validated by reverse transcription polymerase chain reaction. The results showed that these selected genes were expressed specifically or preferentially in each representative tissue/organ. Moreover, the function of their promoters was verified by transient expression in wheat or stable transformation in *Arabidopsis*. The results showed that these promoters can efficiently and predominantly drive *uidA* (β-glucuronidase) reporter gene expression in different tissues. Due to their tissue-specific nature, these promoters can be used as potential candidates in plant genetic engineering.

*Additional key words*: *Arabidopsis thaliana*, microarray, transgenic plants, *Triticum aestivum*.

## Introduction

Promoters can drive gene expression in different spatial and temporal patterns and play a great role in plant genetic engineering. Among different promoters, constitutive promoters such as *CaMV 35S* from cauliflower mosaic virus (Covey *et al*. 1981), *ZmUbi1* from maize polyubiquitin gene (Christensen *et al*. 1992), *OsAct1* (McElroy *et al*. 1990) and *OsAct2* (He *et al*. 2009) from rice *actin* genes have been widely used. However, continuous overexpression of transgenes driven by these constitutive promoters may not be advantageous because of undesirable side effects and unfavourable phenotypes generated. In some cases, constitutive promoters can even be harmful to the host plants because they can cause extra metabolic burden or toxicity (Hsieh *et al*. 2002, Li *et al*. 2011, Kong *et al*. 2014, Li *et al*. 2015). In addition, the expression of several transgenes controlled by one promoter could be unpredictable in transgenic plants, due to, for example, homology-based transgene silencing (Jorgensen 1992, Butaye *et al*. 2005, Rocha *et al*. 2005, Furtado *et al*. 2009). These problems can be avoided by using tissue-specific promoters which can control gene expression in a desirable tissue at a specific time point. Therefore, more spatiotemporal specific or inducible promoters need to be exploited to drive transgene expression more precisely and efficiently. Until now, many tissue-specific promoters have been identified and isolated (Kato *et al*. 2010, Yang *et al*. 2012, Jeong and Jung 2015, Yan *et al*. 2015, Zhang *et al*. 2016, Ali *et al*. 2017, Xu *et al*. 2018). Studies on these promoters could not only help the understanding of gene regulation but also provide more available promoters for genetic engineering (Hernandez-Garcia and Finer 2014, Cong *et al*. 2009).

Wheat is an important cereal crop in the world, but research on tissue-specific promoters in this plant species is limited compared to the other two cereal crops rice and maize, due to its complex allohexaploid genome and low transformation efficiency (He *et al*. 2015). Previous works demonstrated that *TaPT2* and *TaPT2-1* promoters are exclusively active in roots and their activities are phosphate-dependent in transgenic wheat, *Arabidopsis*, or tobacco (Tittarelli *et al*. 2007, Cui *et al*. 2011), while *TaPSG719* and *TaPSG076* are pollen-specific promoters (Chen *et al*. 2010, 2012). The *TaHMWGlu1D* and *TaHMWGlu1Bx17* promoters have been reported to drive *GUS* specific expression in the endosperm of transgenic wheat (Lamacchia *et al*. 2001, Oszvald *et al*. 2007). To meet the need of efficient wheat genetic engineering and avoid various types of unpredictable expression driven by homologous promoters from other species, more tissue-specific native promoters should be isolated and studied in wheat.

Tissue/organ-specific genes and promoters can be commonly screened by microarray and next-generation sequencing (NGS) technology with high efficiency (Endo *et al*. 2004, Ye *et al*. 2012, Nguyen *et al*. 2016). For instance, the organ-specific genes and promoters in seven tissues were efficiently isolated in rice by using microarrays (Chae *et al*. 2016). This method was also used in a genome-wide analysis of gene expression in wheat but these studies were performed only on one or two target tissues (Crismani *et al*. 2006). However, high-throughput studies of tissue-specific gene expression and promoter prediction in wheat have not been reported and need an in-depth scrutiny. For this aim, publicly available microarray data from seven different wheat tissues were retrieved for high-throughput transcriptome analysis. Preferentially and abundantly expressed genes were identified and candidate genes were chosen for validation. Some promoters were also validated by transient expression in wheat or stable transformation in *Arabidopsis*.

## Materials and methods

**Microarray data analysis:** A public microarray data (*NCBI* GSE12508) (Schreiber *et al*. 2009), which contains 61 127 probe sets representing 55 052 transcripts from seven wheat tissues, namely coleoptile, root, leaf, pistil, anther, embryo, and endosperm, were analyzed. After normalizing the expression intensities by arithmetic RMA normalization method, the microarray data were processed to remove genes with constitutive or low expression. Then tissue-specific genes were selected by the ROKU method with gene entropy $H < 1$ and maximum expression level $> 10$. The ROKU method was used because of its advantages in selecting genes specific to tissues of interest (Kadota *et al*. 2006, Chishima *et al*. 2018). Redundant probes representing the same genes were deleted. The chromosomal distribution of candidate genes was located on genetic maps using *Genome Zipper V5* at *IWGSC* (http://www.wheatgenome.org/) and presented by *MapDraw*.

**Plants and growth conditions:** *Arabidopsis thaliana* L. ecotype Columbia (Col-0) was used. The transgenic *Arabidopsis* seeds were surface sterilized with 75 % (v/v) ethanol for 1 min followed by disinfection with 10 % (v/v) $H_2O_2$ for 8 min. The seeds were rinsed with distilled water 3 - 5 times and placed on Murashige and Skoog (1962; MS) medium containing 0.8 % agar (m/v) and kanamycin (50 mg dm$^{-3}$) and grown at a temperature of 22 °C, a 16-h photo-period, and an irradiance of 150 μmol m$^{-2}$ s$^{-1}$.

The seeds of wheat (*Triticum aestivum* L. cv. Chinese Spring) were germinated in the dark under a temperature of 22 °C. Then the seedlings were transferred in pots containing a mixture of soil and *Vermiculite* (3:1) to the greenhouse (a temperature of 22 °C, an irradiance of 350 μmol m$^{-2}$ s$^{-1}$, a 16-h photoperiod, and an air humidity of 60 %).

**Semi-quantitative RT-PCR and real-time qPCR analysis:** Thirty-six tissue-specific genes were chosen for validation of their expression patterns by reverse transcription (RT)-PCR. For convenience, these genes were named according to their expression patterns shown in microarray data and the amplification order. The coding sequences of these genes were obtained from the *PLEXdb* database based on the accession numbers of their probes (Dash *et al*. 2012) and used for designing specific primers as listed in Table 1 Suppl. *TaActin* (GenBank accession No. AB181991) was used as the internal control to normalize the expressions of tested genes.

For RNA extraction, root and young leaf tissues were collected from two-week-old seedlings of wheat Chinese Spring. Mature leaf, pistil, and anther were collected at flowering stage. Immature embryo and endosperm were obtained at approximately 15 d after anthesis. All samples were immediately frozen in liquid nitrogen and then stored at -80 °C. Total RNA was extracted from frozen tissues with a plant total RNA extraction kit according to the manufacturer's instructions (*ZOMANBIO*, Beijing, China). The first-strand cDNAs were synthesized using the first-strand cDNA synthesis kit (*Fermentas*, Waltham, USA). Semi-quantitative RT-PCR was used to detect the expression of candidate genes in different wheat organs. The amplification program was as follows: pre-denaturation at 95 °C for 3 min, then 30 cycles of 95 °C for 30 s, 56 - 60 °C for 30 s, and 72 °C for 10 s. Real-time quantitative PCR (qPCR) was performed to determine the gene expression, using the *SYBR Green* kit (*Tiangen*, Beijing, China) and the *CFX Connect*™ *Optics* module real-time system (*Bio-Rad*, Hercules, USA). The qPCR conditions were 95 °C for 15 min followed by 40 cycles of 95 °C for 10 s, 55 - 60 °C for 15 s, and 72 °C for 30 s. The

relative gene expressions were calculated using the $2^{-\triangle\triangle Ct}$ formula (Livak and Schmittgen 2001). Each experiment was repeated three times.

**Promoter cloning and vector construction:** Three tissue-specific genes *TaCol1*, *TaLeaf1*, and *TaRoot2* (GenBank accession numbers: CA675646, CA623997, and BQ294913) were selected for the analyses of their expression patterns in transformation assays. Their upstream sequences were obtained in *Ensembl* database by searching their coding sequences (Kersey *et al.* 2018). The upstream fragments (-1349/-1, -963/-1, -1399/-1, relative to translation start site, TSS) of genes *TaCol1*, *TaLeaf1*, and *TaRoot2* were amplified by PCR and verified by sequencing. Then the promoter fragments were digested and inserted into pBI121 plasmid to replace the *CaMV 35S* promoter. The constructed vectors were named *pCol1::GUS*, *pLeaf1::GUS* and *pRoot2::GUS*, respectively, and they were transferred into *Agrobacterium tumefaciens* strain EHA105 by the freeze-thaw method. Sequences of primers used for amplifying promoters are listed in Table 2 Suppl.

***Agrobacterium*-mediated transformation of *Arabidopsis*:** *Agrobacterium* strain EHA105 separately harbouring *pLeaf1::GUS* and *pRoot2::GUS* vectors, were introduced into *Arabidopsis*, using the *Agrobacterium*-mediated floral-dip transformation method according to the procedure described previously (Clough and Bent 1998). The transgenic seeds were germinated on kanamycin selective medium, then the positive transformants were confirmed by PCR followed by further analysis of $T_1$ - $T_3$ generations.

**Transient expression analysis of the *TaCol1* promoter in wheat tissues:** For *GUS* transient expression analysis, wheat seeds were surface sterilized with 0.1 % (m/v) $HgCl_2$ for 8 min, and then washed with sterile water 5 - 6 times. The seeds were then germinated in sterilized Petri dishes with sterile water in the culture room (22 °C, dark) for 2 - 3 d. Then, wheat seedlings with coleoptiles were collected and infected by 10 $cm^3$ of *Agrobacterium* suspension harbouring the construct *pCol1::GUS* and containing 100 µM acetosyringone. After incubation for approximately 5 - 6 h (25 °C, 70-80 rpm shaking), wheat seedlings were co-cultured in solid MS medium for 2 d (25 °C, dark). Wheat seedlings transformed with or without pBI121 were used as positive or negative controls, respectively.

**Analysis of GUS activity:** Qualitative analysis of β-glucuronidase (GUS) activity was performed as described previously (Jefferson *et al.* 1987). All organs or tissues were immersed in GUS staining solution, which contained 50 mM sodium phosphate buffer (pH 7.0), 0.5 mM each of potassium ferrocyanide and potassium ferricyanide, 0.5 mg $cm^{-3}$ 5-bromo-4-chloro-3-indolyl-β-D-glucuronide (X-Gluc), 0.1 % (v/v) *Triton X-100* and 20 % (v/v) methanol. Carbenicillin (500 mg $dm^{-3}$) was added to the staining solution to avoid non-specific *GUS* expression from residual *Agrobacterium* and endophytes in wheat tissues (Dhadi *et al.* 2012). After incubation in the dark at 37 °C overnight, the chlorophyll was cleared off from the samples by gradient treatment with 70 - 90 % (v/v) ethanol.
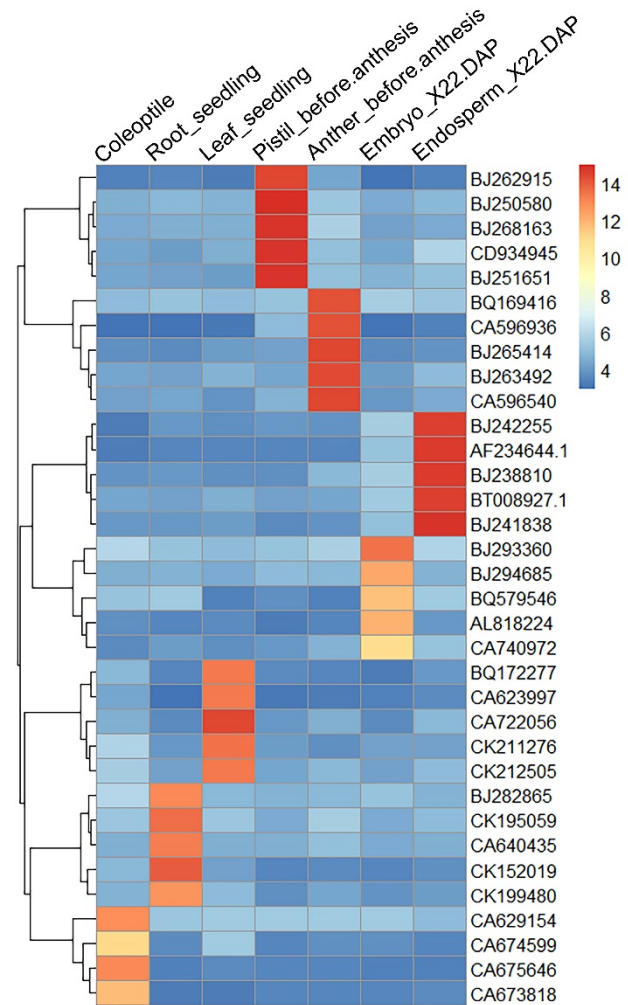


Fig. 1. Expression profile analysis of representative specific genes in seven tissues of wheat using microarray data. The genes presented include 4 candidate specific genes from coleoptile and the top 5 genes highly expressed specifically in tissues including root, leaf, pistil, anther, embryo, and endosperm, respectively. Each row represents the candidate gene, each column represents a tissue. The *colour bar* at the right shows the degree of expression: blue - low expression, red - high expression.

## Results

To identify wheat tissue/organ-specific genes, microarray data (*NCBI* GSE12508) from seven wheat tissues/organs, namely coleoptile, root, leaf, pistil, anther, embryo, and endosperm, were analyzed using the ROKU method. The results showed that genes from different tissues with high expressions were obviously clustered together in the target specific tissue. The heatmap depicting the expression profiles of representative genes from each tissue was shown in Fig. 1. In total, 604 probes displayed specific expression patterns (Fig. 1 and Table 3 Suppl.). Among them, 235 probes covering over one-third of the whole probes were found in anthers, whereas 157 probes accounting for one-fourth of the total number of probes were detected in endosperms. Besides, 78, 86, and 37 probes were found in roots, leaves, and pistils, respectively. Only 4 and 7 probes were detected in coleoptiles and embryos, respectively. These results indicate that the number of genes preferentially expressed in anther or endosperm tissues was greater than the numbers found in the other five tissues.
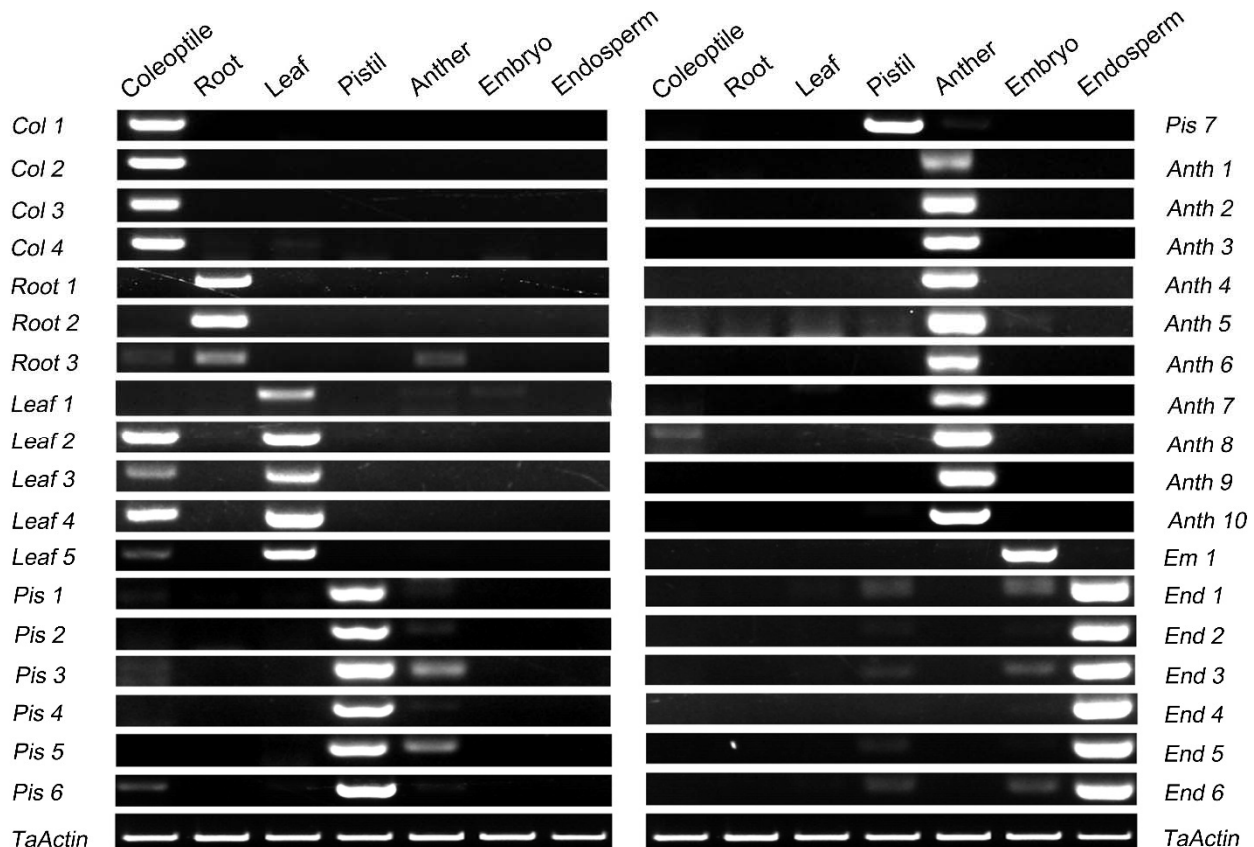


Fig. 2. Semi-quantitative RT-PCR analysis of selected genes in seven different tissues of wheat: coleoptile, root, leaf, pistil, anther, embryo, and endosperm. *TaActin* was used as an internal reference control.

In addition, the loci of 330 genes predicted from tissue-specific probes were identified and mapped to the wheat genetic linkage map. These genes were unevenly located on different chromosomes. Some gene loci representing the same tissue were identified to be clustered in the same sites on the chromosome. For example, BJ249627, BJ252359, BQ169477, BQ169653, and BJ254190 representing the anther tissue were found clustered on chromosome 1B, whereas BQ804326, BQ805009, CK205651, BJ239761, BJ237853, BQ804999, and CA713323, endosperm-specific gene loci, were clustered on chromosome 7D (Fig. 1 Suppl.). Interestingly, the sequences of some genes were found on multiple chromosomal sites in the genome. For instance, CB412210 (TaAffx.128473.1.A1_s_at) was found on chromosome 1D and 3D while CK215309 (Ta.345.1.S1_at) was found on chromosome 4B and 4D. The duplication may contribute to high gene expressions, which may be caused by the tandem and segmental duplication events that always evolve in sub-functionalization or neo-functionalization under selective environmental pressures (Wei *et al*. 2014, Chen *et al*. 2016). The probes and their corresponding loci on the chromosome are shown in Table 4 Suppl.

For validation of gene expression patterns, thirty-six predicted organ-preferential genes were selected from coleoptile (4), root (3), leaf (5), pistil (7), anther (10), embryo (1), and endosperm (6) microarray data and tested by semi-quantitative RT-PCR. Some genes showed very high specific expression in target tissues. For example, three (*Col1, 2, 3*) of the four genes displayed high expression only in coleoptiles. Two (*Root1, 2*) of the three genes showed exclusive expression in roots. *Em1* was confirmed to be exclusively expressed in embryos. Eight (*Anth1, 2, 3, 4, 6, 7, 9, 10*) of ten genes were found only in

anthers, which was in corroboration with the computational microarray expression data (Fig. 2). It is worth noting that some genes were detected as preferential in one tissue but with slight expression in other tissues. For example, all seven genes preferentially expressed in pistils were also slightly expressed in anthers and/or coleoptiles. Similarly, the five genes showing preferentially high expression in endosperms were slightly expressed in pistils and embryos except for gene *End4*. Nevertheless, the four leaf-preferential genes (*Leaf2, 3, 4, 5*) were also found in coleoptiles with high expression levels, especially *Leaf2*
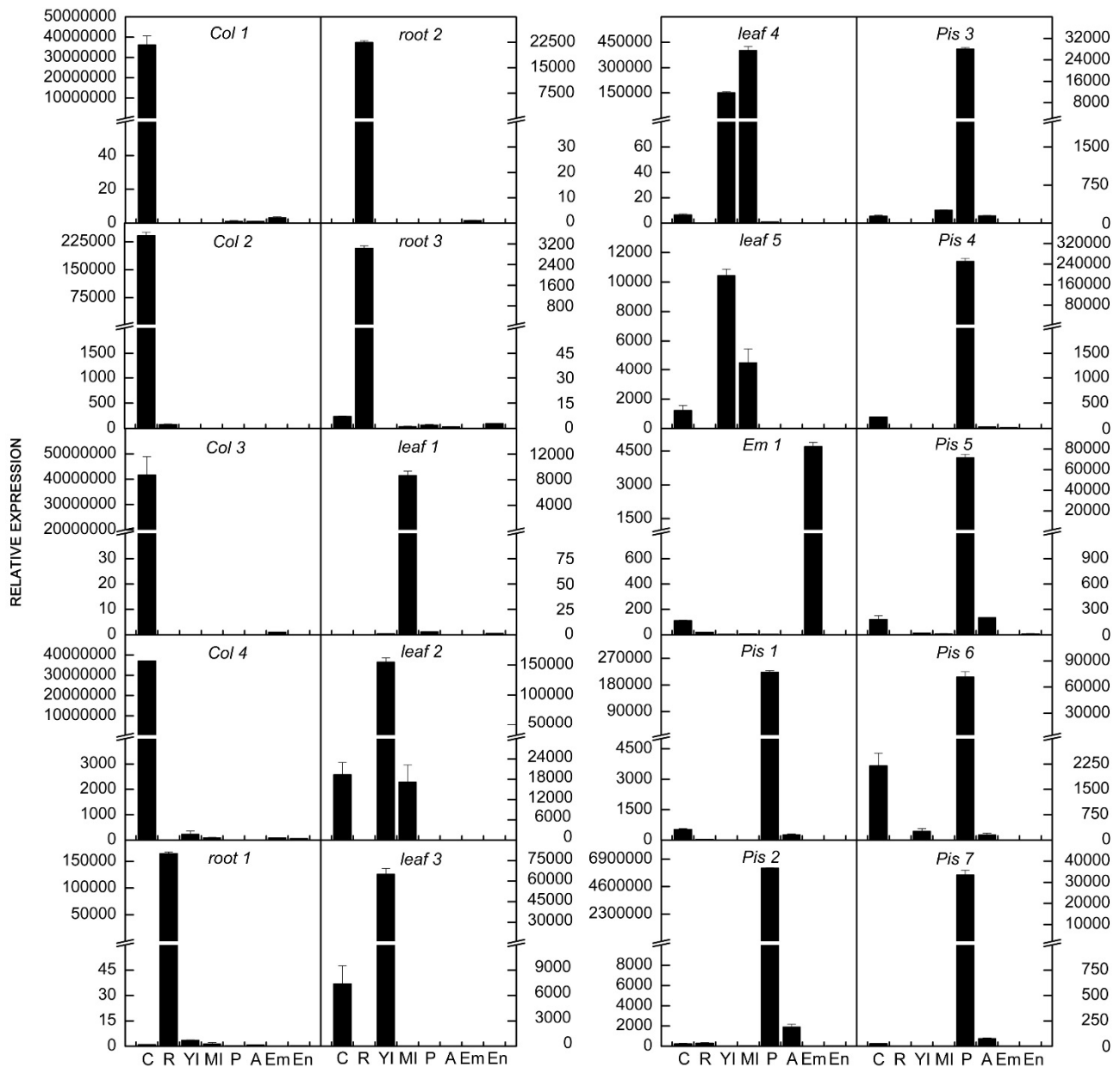


Fig. 3. Relative expression analysis of selected genes in different tissues/organs of wheat by real-time qPCR method. The tissues used are: coleoptile (C), root (R), young leaf (Yl), mature leaf (Ml), pistil (P), anther (A), embryo (Em), and endosperm (En). The coleoptiles were sampled 2 d after seed germination, the roots and young leaves in two-week-old seedlings, the mature leaf, pistil, anther, embryo, and endosperm in mature plants. Means ± SDs of three independent replicates.

and *Leaf4* genes (Fig. 2).

The spatial and temporal expression patterns of thirty-six genes were further analyzed by real-time qPCR method. All the genes were tested with highly specific transcript abundances in the target tissues although their expressions varied significantly among each other (Figs. 3 and 4). Indeed, the expressions of *Col1-4*, *Root1*, *2*, *Leaf1*, *Anth1-10* obtained from qPCR showed high consistency with the semi-quantitative RT-PCR results. As an example, *anth10* transcript was not detected in leaf tissue but showed specifically high transcript abundance in anther tissue, which was more than 130 thousand-fold higher compared with the basic expression level detected in other tested tissues (Fig. 4). The expression patterns of other genes obtained from qPCR experiment were also consistent, to some extent, with the semi-quantitative RT-PCR results. For example, two genes (*Leaf2, 3*) showed high expressions in coleoptiles and leaves in RT-PCR experiment (Fig. 2). Real-time qPCR results also showed a certain degree of expression of these genes in

coleoptile, but the expressions were lower (about 1/9) compared to those in the leaf tissue. The results of real-time qPCR and RT-PCR conformed with the high-throughput microarray data to some degree.

As coleoptile is a tissue of monocots, especially the gramineous plants, the activity of *TaCol1* promoter was verified in wheat coleoptile by transient expression assay. The vector *pCol1::GUS* was constructed by fusing *TaCol1* promoter and the *GUS* reporter gene (Fig. 5), and was introduced into germinated wheat seeds by *Agrobacterium*-mediated transformation. GUS histochemical analysis showed high GUS expression in wheat coleoptiles (Fig. 6*B*). On the contrary, no GUS expression was detected in coleoptiles of the negative controls (Fig. 6*A,C*), while strong GUS staining was observed in roots and coleoptiles of the positive control (Fig. 6*D*). The above-mentioned results indicated that the *TaCol1* promoter can regulate *GUS* gene preferential expression in wheat coleoptiles. It also indicates that it is feasible to isolate new promoters by searching microarray data in wheat.
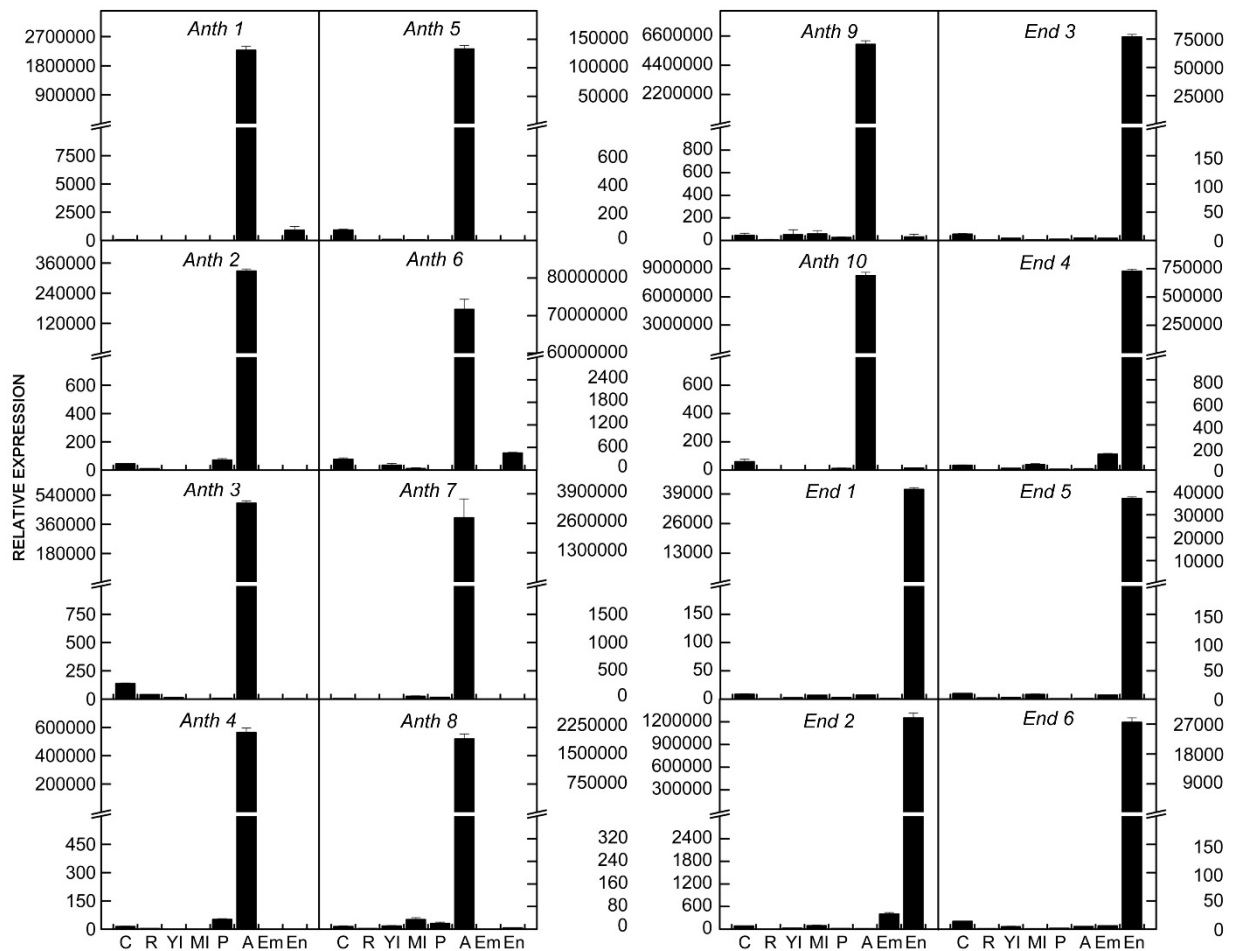


Fig. 4. Relative expression analysis of selected genes in different tissues/organs of wheat by real-time qPCR method. The tissues used are: coleoptile (C), root (R), young leaf (Yl), mature leaf (Ml), pistil (P), anther (A), embryo (Em), and endosperm (En). The coleoptiles were sampled 2 d after seed germination, the roots and young leaves in two-week-old seedlings, the mature leaf, pistil, anther, embryo, and endosperm in mature plants. Means ± SDs of three independent replicates.

Next, we investigated the spatial and temporal regulation patterns of gene promoters during the life cycle of stable transgenic plants. Based on the restriction enzyme sites and the quality of designed primers, 963 and 1399 bp fragments of *TaLeaf1* and *TaRoot2* gene promoters were amplified by PCR, respectively, using wheat genomic DNA as template. To confirm the activity of *TaLeaf1* and *TaRoot2* promoters, vectors *pLeaf1::GUS* and *pRoot2::GUS* were constructed (Fig. 5) and transformed into *Arabidopsis* by floral-dip transformation method. For each construct, three to five independent transformants were subjected to GUS staining. GUS activity was found in all tissues including root, leaf, inflorescence, and silique of transformants harboring pBI121. On the contrary, no

GUS expression was detected in wild-type plants (Fig. 7). In transformants containing *pLeaf1::GUS*, strong blue staining was found in young leaves. High GUS expression was also observed in mature leaves. In addition, slight GUS staining was detected in siliques and inflorescences. No GUS activity was visible in young roots and mature roots (Fig. 7*F-J*). The results indicated that the *TaLeaf1* promoter can drive *GUS* gene expression in a leaf-preferential pattern. In transgenic lines harbouring *pRoot2::GUS*, GUS signals were only detected in young roots and mature roots, but not in mature leaves, inflorescences or siliques (Fig. 7*K-O*). This result indicates that the *TaRoot2* is a root-preferential promoter in *Arabidopsis*.
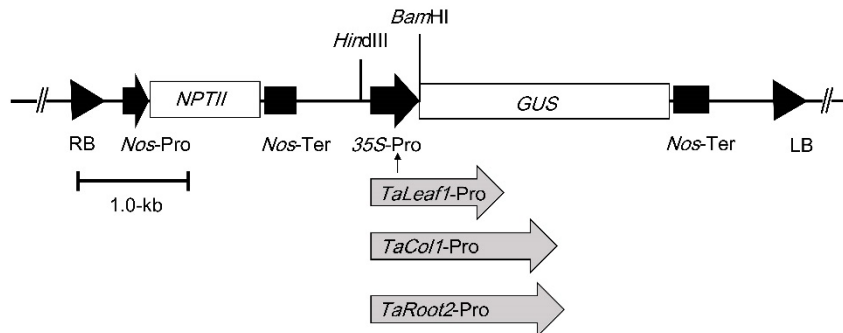


Fig. 5. Schematic representation of vector construction of the transformation plasmids. *Promoter::GUS* expression vectors. The *35S* promoter of pBI121 was replaced by the *TaCol1*, *TaLeaf,1* or *TaRoot2* promoter fragment. *Nos*-Ter, *Agrobacterium tumefaciens* nopaline synthase (*Nos*) terminator.
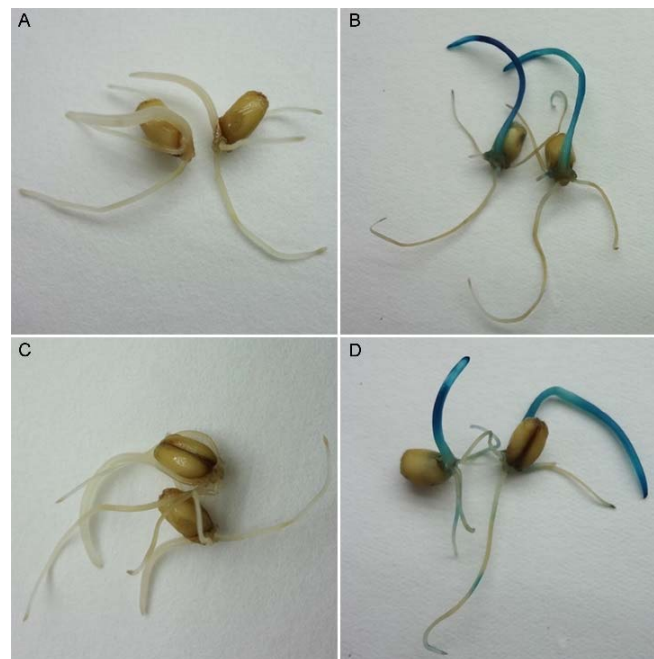


Fig. 6. Transient expression of *pCol1::GUS* in germinating wheat seeds. *A* - Wild-type with MS medium incubation as a negative control. *B* - Germinating seeds infected with EHA105 harboring *pCol1::GUS*. *C* - Germinating seeds infected with EHA105 as a negative control. *D* - Germinating seeds infected with EHA105 harboring the pBI121 expression vector as a positive control.

## Discussion

Numerous developmental and physiological processes, such as organelle organization, transport, and energy consumption, are all constant in the entire life cycle of plants. This requires constitutive expression of a considerable number of genes in organs. Tissues/organs are distinguished by gene expression patterns, which implies that distinct tissue-specific genes play important roles in controlling tissue/organ morphogenesis and the maintenance of relevant functions. Isolation and functional analysis of these genes and promoters shed light on understanding the regulatory mechanism of plant development. Microarray and NGS technologies have been proven to be effective tools on screening such genes and promoters (Sharma *et al.* 2012, Liu *et al.* 2014). To search tissue/organ-specific genes and promoters in wheat, a public microarray data was analyzed and 604 probes which represent putative tissue-specific genes were obtained according to the selection criteria set for computational predictions of gene-specificity expression (Fig. 1). As expected, in the list of candidate anther-specific genes which matched to probes, we found that *TaPSG076* (Ta.10341.1.S1_at, AY451239.1) and
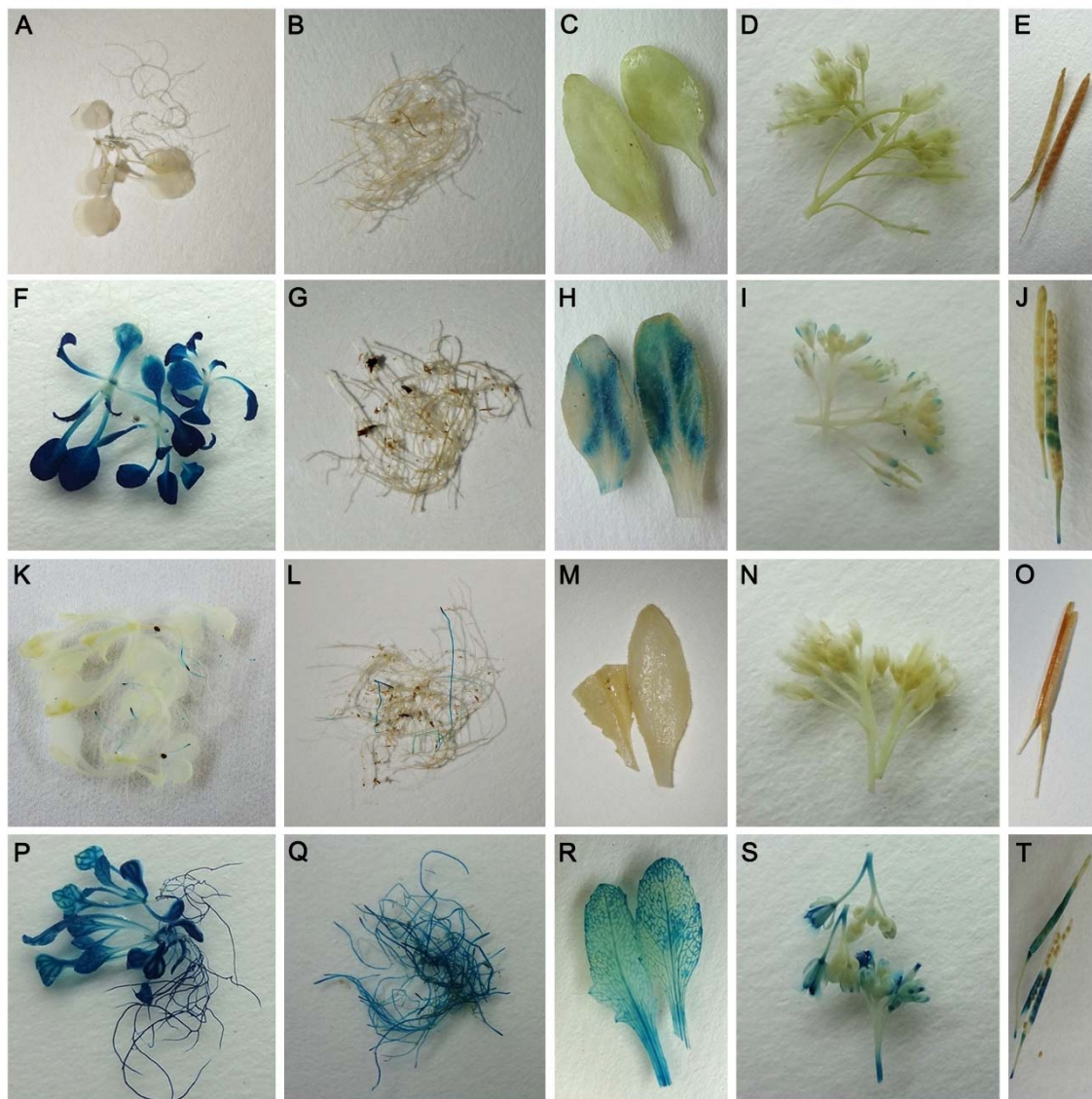


Fig. 7. GUS staining in different organs of transgenic *A. thaliana* harboring *pLeaf1::GUS* or *pRoot2::GUS* compared with WT and positive control of pBI121. The organ samples stained by GUS staining solution include whole seedlings, mature roots, mature leaves, inflorescences, and siliques. *A* to *E* - Different tissues/organs from wild-type plant. *F* to *J* - Different tissues/organs from *Arabidopsis* transformed with *pLeaf1::GUS* vector. *K* to *O* - Different tissues/organs from *Arabidopsis* transformed with *pRoot2::GUS* vector. *P* to *T* - Different tissues/organs from *Arabidopsis* transformed with PBI121 vector as a positive control.

*TaPSG719* (Ta.30672.1.S1_at, AY451238.1) genes were also screened out successfully by the ROKU method employed in this research. These two genes have been confirmed as pollen-specific expression genes (Jin and Bian 2004a,b). Our previous studies showed that *TaPSG076* and *TaPSG719* are pollen-specific promoters in transgenic tobacco (Chen *et al.* 2010, 2012). The expression of *TaPSG076* was higher than that of *TaPSG719* based on the results from the microarray data analysis (Table 3 Suppl.). The expression of probes representing *TaPSG076* and *TaPSG719* ranked 70th and 138th among the candidate anther-specific probe sets (235 probes in total) (Table 3 Suppl.). This was in agreement with the results from the RT-PCR analysis of *TaPSG076* and *TaPSG719* genes and the activity of promoters demonstrated in transgenic tobacco (Jin and Bian 2004a,b, Chen *et al.* 2010, 2012).

To verify the results of microarray data analysis, thirty-six putative tissue-specific genes were selected for experimental confirmation. The results of RT-PCR and real-time qPCR showed that most of these genes are tissue/organ-specific or preferentially expressed genes, which confirmed the microarray analysis data (Figs. 2, 3, 4). However, some differences were found between the results from microarray analysis and those from the RT-PCR. For example, *Leaf2* exhibited relatively high expression in leaves and coleoptiles based on the results of RT-PCR and qPCR (Figs. 2, 3). In contrast, the expression of this gene in coleoptiles was lower than that in leaves but comparable with the expressions in roots, pistils, anthers, embryos, and endosperms according to the microarray analysis (Ta.20830.1.A1_at in Table 3 Suppl.). The differences observed may be caused by the following reasons: different sensitivity and algorithm exist between the two analysis methods (Fan *et al.* 2011, Chae *et al.* 2016), the samples were not collected precisely at the same time, or the highly homologous genes are not easy to be distinguished.

To get tissue-specific promoters, the upstream sequences of three tissue-specific genes were verified in transient or stable transformation systems. The results showed that *TaCol1* promoter was able to effectively drive *GUS* reporter gene expression (Fig. 6). The *TaRoot2* presented root-specific regulation pattern in *Arabidopsis*, which suggested that *trans*-acting factors and *cis*-acting elements of this promoter related to root-specific regulation were conserved in dicot and monocot systems (Fig. 7). Gene expression analysis indicated that *TaLeaf1* was exclusively expressed in wheat leaves (Figs. 1, 2, 3) while *GUS* expression driven by the *TaLeaf1* promoter was high in leaves and low in siliques and inflorescences of transgenic plants. This implied that *TaLeaf1* is a leaf-preferential promoter in *Arabidopsis* (Fig. 7). We deduced that this discrepancy in the expression pattern of *TaLeaf1* promoter in *Arabidopsis* may be caused by ectopic expression due to different plant structure and gene regulation systems in heterologous species of monocot wheat and dicot *Arabidopsis*. Similar results were observed in previous reports (Lü *et al.* 2007, Huang *et al.* 2016). Similarly, interspecies differences in the expression of genes regulated by *CaMV 35S* and *nos* promoters also occurred in transgenic *Arabidopsis* and tobacco plants (Wilkinson *et al.* 1997). Thus further studies should be performed in wheat to validate the regulatory pattern of *TaLeaf1* in a homologous system. These novel organ-preferential promoters have a potential application in transgenic engineering.

In summary, a set of tissue/organ-specific genes in seven tissues of wheat were successfully screened out by analyzing microarray data. Among them, thirty-six genes were selected and confirmed as tissue-specific/preferential genes by transcriptional experiments, and three promoters were verified as tissue/organ-preferential promoters. All these results indicated that it is feasible to obtain tissue-specific/preferential genes and promoters through analyzing public microarray data in an efficient and accurate way. Isolation and identification of these genes and promoters can not only provide resources for plant transgenic engineering but also help us to understand regulatory mechanisms of different tissues in wheat development.

## References

Ali, S., Chandrashekar, N., Rawat, S., Nayanakantha, N.M.C., Mir, Z.A., Manoharan, A., Sultana, M., Grover, A.: Isolation and molecular characterization of pathogenesis related *PR2* gene and its promoter from *Brassica juncea*. - Biol. Plant. 1-11, 2017.

Butaye, K.M.J., Cammue, B.P.A., Delauré, S.L., Bolle, M.F.C.D.: Approaches to minimize variation of transgene expression in plants. - Mol. Breed. **16**: 79-91, 2005.

Chae, S., Kim, J.S., Jun, K.M., Pahk, Y.M., Kim, M.J., Lee, S.B., Park, H.M., Lee, T.H., Nahm, B.H., Kim, Y.K.: Analysis of representative organ-specific genes and promoters of rice using a 3' ORF-oriented long oligomer microarray. - J. Plant Biol. **59**: 579-593, 2016.

Chen, L., Miao, Y.J., Wang, C., Su, P.P., Li, T.H., Wang, R., Hao, X.L., Yang, G.X., He, G.Y., Gao, C.B.: Characterization of a novel pollen-specific promoter from wheat (*Triticum Aestivum* L.). - Plant mol. Biol. Rep. **30**: 1426-1432, 2012.

Chen, L., Tu, Z., Hussain, J., Cong, L., Yan, Y., Jin, L., Yang, G., He, G.: Isolation and heterologous transformation analysis of a pollen-specific promoter from wheat (*Triticum aestivum* L.). - Mol. Biol. Rep. **37**: 737-744, 2010.

Chen, L.H., Han, J.P., Deng, X.M., Tan, S.L., Li, L.L., Li, L., Zhou, J.F., Peng, H., Yang, G.X., He, G.Y., Zhang, W.X.: Expansion and stress responses of AP2/EREBP superfamily in *Brachypodium distachyon*. - Sci. Rep. **6**: 21623, 2016.

Chishima, T., Iwakiri, J., Hamada, M.: Identification of transposable elements contributing to tissue-specific expression of long non-coding RNAs. - Genes 9: 23, 2018.

Christensen, A.H., Sharrock, R.A., Quail, P.H.: Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. - Plant mol. Biol. 18: 675-689, 1992.

Clough, S.J., Bent, A.F.: Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. - Plant J. 16: 735, 1998.

Cong, L., Wang, C., Chen, L., Liu, H.J., Yang, G.X., He, G.Y.: Expression of phytoene synthase1 and carotene desaturase crtI genes result in an increase in the total carotenoids content in transgenic elite wheat (Triticum aestivum L.). - J. Agr. Food Chem. 57: 8652, 2009.

Covey, S.N., Lomonossoff, G.P., Hull, R.: Characterisation of cauliflower mosaic virus DNA sequences which encode major polyadenylated transcripts. - Nucl. Acids Res. 9: 6735-6747, 1981.

Crismani, W., Baumann, U., Sutton, T., Shirley, N., Webster, T., Spangenberg, G., Langridge, P., Able, J.A.: Microarray expression analysis of meiosis and microsporogenesis in hexaploid bread wheat. - BMC Genomics 7: 267, 2006.

Cui, X., Zhang, Y., Zhao, F., Guo, C., Juntao, G.U., Wenjing, L.U., Xiaojuan, L.I., Xiao, K.: Molecular characterization and expression analysis of phosphate transporter gene TaPT2-1 in wheat (Triticum aestivum L.). - Front. Agr. China 5: 274-283, 2011.

Dash, S., Hemert, J.V., Hong, L., Wise, R.P., Dickerson, J.A.: PLEXdb: gene expression resources for plants and plant pathogens. - Nucl. Acids Res. 40: D1194, 2012.

Dhadi, S.R., Deshpande, A., Ramakrishna, W.: A novel non-wounding transient expression assay for cereals mediated by Agrobacterium tumefaciens. - Plant mol. Biol. Rep. 30: 36-45, 2012.

Endo, M., Tsuchiya, T., Saito, H., Matsubara, H., Hakozaki, H., Masuko, H., Kamada, M., Higashitani, A., Takahashi, H., Fukuda, H.: Identification and molecular characterization of novel anther-specific genes in Oryza sativa L. by using cDNA microarray. - Genes Genet. Syst. 79: 213, 2004.

Fan, X.H., Shao, L., Fang, H., Tong, W.D., Cheng, Y.Y.: Cross-platform comparison of microarray-based multiple-class prediction. - PLoS ONE 6: e16067, 2011.

Furtado, A., Henry, R.J., Pellegrineschi, A.: Analysis of promoters in transgenic barley and wheat. - Plant Biotechnol. J. 7: 240, 2009.

He, C., Lin, Z., Mcelroy, D., Wu, R.: Identification of a rice actin2 gene regulatory region for high-level expression of transgenes in monocots. - Plant Biotechnol. J. 7: 227, 2009.

He, Y., Wang, Q., Zeng, J., Sun, T., Yang, G.X., He, G.Y.: Current status and trends of wheat genetic transformation studies in China. - J.integr. Agr. 14: 438-452, 2015.

Hernandez-Garcia, C.M., Finer, J.J.: Identification and validation of promoters and cis-acting regulatory elements. - Plant Sci. 217-218: 109-119, 2014.

Hsieh, T.H., Lee, J.T., Yang, P.T., Chiu, L.H., Charng, Y.Y., Wang, Y.C., Chan, M.T.: Heterology expression of the Arabidopsis C-Repeat/Dehydration Response Element Binding Factor 1 gene confers elevated tolerance to chilling and oxidative stresses in transgenic tomato. - Plant Physiol. 129: 1086, 2002.

Huang, W.F., Huang, P.L., Do, Y.Y.: Characterization of promoter activity of the ethylene receptor gene OgERS1 from Oncidesa in transgenic Arabidopsis. - Biol. Plant. 60: 261-268, 2016.

Jefferson, R.A., Kavanagh, T.A., Bevan, M.W.: GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. - EMBO J. 6: 3901, 1987.

Jeong, H.J., Jung, K.H..: Rice tissue-specific promoters and condition-dependent promoters for effective translational application. - J. integr. Plant Biol. 57: 913-924, 2015.

Jin, Y.F., Bian, T.F.: Isolation and expression of a wheat pollen-specific gene with long leader sequence. - Acta bot. sin. 46: 1347-1356, 2004a.

Jin, Y.F., Bian, T.F.: Isolation and partial characterization of a novel pollen-specific cDNA with multiple polyadenylation sites from wheat. - Acta biochem. biophys. sin. 36: 467, 2004b.

Jorgensen, R.: Silencing of plant genes by homologous transgenes. - Agbiotech News Inform. 4: 265-273, 992.

Kadota, K., Ye, J., Nakai, Y., Terada, T., Shimizu, K.: ROKU: a novel method for identification of tissue-specific genes. - BMC Bioinformatics 7: 294, 2006.

Kato, H., Xie, G., Sato, Y., Imai, R.: Isolation of anther-specific gene promoters suitable for transgene expression in rice. - Plant mol. Biol. Rep. 28: 381-387, 2010.

Kersey, P.J., Allen, J.E., Allot, A., Barba, M., Boddu, S., Bolt, B.J., Carvalho-Silva, D., Christensen, M., Davis, P., Grabmueller, C., Kumar, N., Liu, Z., Maurel, T., Moore, B., McDowall, M.D., Maheswari, U., Naamati, G., Newman, V., Ong, C.K., Paulini, M., Pedro, H., Perry, E., Russell, M., Sparrow, H., Tapanari, E., Taylor, K., Vullo, A., Williams, G., Zadissia, A., Olson, A., Stein, J., Wei, S., Tello-Ruiz, M., Ware, D., Luciani, A., Potter, S., Finn, R.D., Urban, M., Hammond-Kosack, K.E., Bolser, D.M., De Silva, N., Howe, K.L., Langridge, N., Maslen, G., Staines, D.M., Yates, A.: Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species. - Nucl. Acids Res. 46: D802-D808, 2018.

Kong, K., Ntui, V.O., Makabe, S.: Transgenic tobacco and tomato plants expressing Wasabi defensin genes driven by root-specific LjNRT2 and AtNRT2.1 promoters confer resistance against Fusarium oxysporum. - Plant Biotechnol. 31: 89-96, 2014.

Lamacchia, C., Shewry, P.R., Di, F.N., Forsyth, J.L., Harris, N., Lazzeri, P.A., Napier, J.A., Halford, N.G., Barcelo, P.: Endosperm-specific activity of a storage protein gene promoter in transgenic wheat seed. - J. exp. Bot. 52: 243, 2001.

Li, A.X., Han, Y.Y., Wang, X., Chen, Y.H., Zhao, M.R., Zhou, S.M., Wang, W.: Root-specific expression of wheat expansin gene TaEXPB23 enhances root growth and water stress tolerance in tobacco. - Environ. exp. Bot. 110: 73-84, 2015.

Li, F., Xing, S., Guo, Q., Zhao, M., Zhang, J., Gao, Q., Wang, G., Wang, W.: Drought tolerance through over-expression of the expansin gene TaEXPB23 in transgenic tobacco. - J. Plant Physiol. 168: 960-966, 2011.

Liu, X.Q., Tian, J., Zhou, X.J., Chen, R.M., Wang, L., Zhang, C.Y., Zhao, J., Fan, Y.L.: Identification and characterization of promoters specifically and strongly expressed in maize embryos. - Plant Biotechnol. J. 12: 1286-1296, 2014.

Livak, K.J., Schmittgen, T.D.: Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. - Methods 25: 402-408, 2001.

Lü, S., Gu, H., Yuan, X., Wang, X., Wu, A.M., Qu, L., Liu, J.Y.:

The GUS reporter-aided analysis of the promoter activities of a rice metallothionein gene reveals different regulatory regions responsible for tissue-specific and inducible expression in transgenic *Arabidopsis*. - Transgenic Res. **16**: 177-191, 2007.

McElroy, D., Zhang, W., Cao, J., Wu, R.: Isolation of an efficient actin promoter for use in rice transformation. - Plant Cell **2**: 163, 1990.

Murashige, T., Skoog, F.: A revised medium for rapid growth and bioassays with tobacco tissue cultures. - Physiol. Plant. **15**: 473-497, 1962.

Nguyen, T.D., Moon, S., Nguyen, V.N., Gho, Y., Chandran, A.K., Soh, M.S., Song, J.T., An, G., Oh, S.A., Park, S.K.: Genome-wide identification and analysis of rice genes preferentially expressed in pollen at an early developmental stage. - Plant mol. Biol. **92**: 1-18, 2016.

Oszvald, M., Gardonyi, M., Tamas, C., Takacs, I., Jenes, B., Tamas, L.: Development and characterization of a chimaeric tissue-specific promoter in wheat and rice endosperm. - *In Vitro* Cell Dev. Plant **44**: 1-7, 2007.

Rocha, P.S., Sheikh, M., Melchiorre, R., Fagard, M., Boutet, S., Loach, R., Moffatt, B., Wagner, C., Vaucheret, H., Furner, I.: The *Arabidopsis HOMOLOGY-DEPENDENT GENE SILENCING1* gene codes for an S-adenosyl-L-homocysteine hydrolase required for DNA methylation-dependent gene silencing. - Plant Cell **17**: 404-417, 2005.

Schreiber, A.W., Sutton, T., Caldo, R.A., Kalashyan, E., Lovell, B., Mayo, G., Muehlbauer, G.J., Druka, A., Waugh, R., Wise, R.P., Langridge P., Baumann U.: Comparative transcriptomics in the *Triticeae*. - BMC Genomics **10**: 285, 2009.

Sharma, R., Agarwal, P., Ray, S., Deveshwar, P., Sharma, P., Sharma, N., Nijhawan, A., Jain, M., Singh, A.K., Singh, V.P.: Expression dynamics of metabolic and regulatory components across stages of panicle and seed development

in *indica* rice. - Funct. integr. Genomics **12**: 229, 2012.

Tittarelli, A., Milla, L., Vargas, F., Morales, A., Neupert, C., Meisel, L., Salvog, H., Peñaloza, E., Muñoz, G., Corcuera, L.: Isolation and comparative analysis of the wheat *TaPT2* promoter: identification *in silico* of new putative regulatory motifs conserved between monocots and dicots. - J. exp. Bot. **58**: 2573, 2007.

Wei, B., Zhang, R.Z., Guo, J.J., Liu, D.M., Li, A.L., Fan, R.C., Mao, L., Zhang, X.Q.: Genome-wide analysis of the MADS-box gene family in *Brachypodium distachyon*. - PLoS One **9**: e84781, 2014.

Wilkinson, J.E., Twell, D., Lindsey, K.: Activities of *CaMV 35S* and *nos* promoters in pollen: implications for field release of transgenic plants. - J. exp. Bot. **48**: 265-275, 1997.

Xu, W., Liu, W., Ye, R., Mazarei, M., Huang, D., Zhang, X., Stewart, C.N.: A *profilin* gene promoter from switchgrass (*Panicum virgatum* L.) directs strong and specific transgene expression to vascular bundles in rice. - Plant Cell Rep. 1-11, 2018.

Yan, S., Wang, Z., Yuan, L., Wei, L., Feng, W., Lin, X., Zheng, M.: Functional architecture of two exclusively late stage pollen-specific promoters in rice (*Oryza sativa* L.). - Plant mol. Biol. **88**: 415-428, 2015.

Yang, L., Li, T., Zhang, S.C., Gao, G.L., Yang, C.W.: Characterization of the *GLP13* gene promoter in *Arabidopsis thaliana*. - Biol. Plant. **57**: 231-237, 2012.

Ye, R., Zhou, F., Lin, Y.: Two novel positive *cis*-regulatory elements involved in green tissue-specific promoter activity in rice (*Oryza sativa* L ssp.). - Plant Cell Rep. **31**: 1159-1172, 2012.

Zhang, C., Pan, S., Chen, H., Cai, T., Zhuang, C., Ye, D., Zhuang, Y., Zeng, Y., Chen, S., Zhuang, W.: Characterization of *NtREL1*, a novel root-specific gene from tobacco, and upstream promoter activity analysis in homologous and heterologous hosts. - Plant Cell Rep. **35**: 757-769, 2016.