

RESEARCH PAPER

Characterization of the wheat gene encoding a grain-specific lipid transfer protein TdPR61, and promoter activity in wheat, barley and rice

Nataliya Kovalchuk, Jessica Smith, Natalia Bazanova, Tatiana Pyvovarenko, Rohan Singh, Neil Shirley, Ainur Ismagul, Alexander Johnson[†], Andrew S. Milligan[‡], Maria Hrmova, Peter Langridge and Sergiy Lopato[§]

Australian Centre for Plant Functional Genomics, University of Adelaide, Hartley Grove, Urrbrae, South Australia 5064, Australia

* Accession numbers: *TaPR61*, JN400738; *TdPR61*, JN400737; *TaBES1*, JN400739; *TaBES1S*, JN400740.

[†] Present address: School of Botany, The University of Melbourne, VIC 3010, Australia

[‡] Present address: Bio Innovation SA, Level 15, 33 King William Street, SA 5000, Australia

[§] To whom correspondence should be addressed. E-mail: sergiy.lopatov@acpfg.com.au

Received 6 August 2011; Revised 3 November 2011; Accepted 16 November 2011

Abstract

The *TaPR61* gene from bread wheat encodes a lipid transfer protein (LTP) with a hydrophobic signal peptide, predicted to direct the TaPR61 protein to the apoplast. Modelling of TaPR61 revealed the presence of an internal cavity which can accommodate at least two lipid molecules. The full-length gene, including the promoter sequence of a *TaPR61* orthologue, was cloned from a BAC library of *Triticum durum*. Quantitative RT-PCR analysis revealed the presence of *TaPR61* and *TdPR61* mainly in grain. A transcriptional *TdPR61* promoter–GUS fusion was stably transformed into wheat, barley, and rice. The strongest GUS expression in all three plants was found in the endosperm transfer cells, the embryo surrounding region (ESR), and in the embryo. The promoter is strong and has similar but not identical spatial patterns of activity in wheat, barley, and rice. These results suggest that the *TdPR61* promoter will be a useful tool for improving grain quality by manipulating the quality and quantity of nutrient/lipid uptake to the endosperm and embryo. Mapping of regions important for the promoter function using transient expression assays in developing embryos resulted in the identification of two segments important for promoter activation in embryos. The putative *cis*-elements from the distal segment were used as bait in a yeast 1-hybrid (Y1H) screen of a cDNA library prepared from the liquid part of the wheat multinucleate syncytium. A transcription factor isolated in the screen is similar to BES1/BLZ1 from *Arabidopsis*, which is known to be a key transcriptional regulator of the brassinosteroid signalling pathway.

Key words: Barley, BES1 transcription factor, embryo, endosperm transfer cells, lipid transfer protein, promoter, rice, wheat.

Introduction

The embryo develops as a result of fusion of male and female gametes. The mature cereal embryo comprises the embryo axis and scutellum. The embryo axis, in turn, comprises the shoot, mesocotyl, radical, and two sheath-like structures: the coleorrhizae which covers the radicle and the coleoptile which covers the plumule (Evers and Millar, 2002). The embryo is the cereal grain component with the highest concentration of lipids. Barley grain at 25 d after fertilization contains lipids mainly in the scutellum and

nodular region of the embryo; a clear decrease of lipid content occurs in the direction of the coleorrhizae and coleoptiles (Neuberger *et al.*, 2010). The developing embryo relies on nutrition from the endosperm and surrounding maternal tissue (Milligan *et al.*, 2004).

After cellularization, the endosperm of cereals comprises four main tissues: endosperm transfer cells, embryo surrounding region, starchy endosperm, and aleurone (Becraft, 2001; Olsen, 2004). Endosperm transfer cells (ETC) are

involved in solute exchange and are very efficient in the uptake of nutrients from adjacent maternal vascular tissue to the endosperm (Davis *et al.*, 1990; Wang *et al.*, 1994; Thompson *et al.*, 2001; Olsen, 2004). The development of ETC in wheat takes place between 5 d and 10 d after pollination (DAP) (Davis *et al.*, 1990; Olsen, 2004). Many ETC-specific genes encode low molecular weight cysteine-rich proteins with N-terminal hydrophobic signal peptides (Hueros *et al.*, 1995, 1999; Maitz *et al.*, 2000; Serna *et al.*, 2001; Cai *et al.*, 2002; Gutierrez-Marcos *et al.*, 2004; Li *et al.*, 2008; Kovalchuk *et al.*, 2009). The functions of such proteins are not absolutely clear, but there is evidence for a role in the defence of developing grain against mother plant-borne pathogens (Serna *et al.*, 2001) and in cell-to-cell communication and signalling (Gutierrez-Marcos *et al.*, 2004; Gómez *et al.*, 2009). Expression of many maize ETC-specific genes is regulated by a single MYB domain transcription factor ZmMRP-1 (Gomes *et al.*, 2002). Transient overexpression of this factor in plant epidermal cells was demonstrated to be sufficient to transform them into ETC cells (Gómez *et al.*, 2009). Two C₂H₂ zinc finger-containing proteins, ZmMRPI-1 and ZmMPRI-2, were shown to interact physically with the ZmMRP-1 transcription factor and to enhance the activation of ETC-specific genes (Royo *et al.*, 2009).

Cells of the embryo surrounding region (ESR) are compacted because of a dense cytoplasm containing extensive rough endoplasmic reticulum (Schel *et al.*, 1984). Initially, at the beginning of embryo development, the ESR surrounds the whole embryo, but mostly surrounds the suspensor when the embryo reaches its maximum size. Only a few genes are known to be specifically expressed in the ESR. These are *ZmEsr1-3* (Opsahl-Ferstad *et al.*, 1997; Bonello *et al.*, 2000), *ZmAEl* and *ZMAE3* (Magnard *et al.*, 2000), and *ZmESR-6* (Balandin *et al.*, 2005). Identification of these ESR-specific genes confirmed ESR cells as a separate domain of gene expression in the endosperm. However, no transcriptional regulators responsible for ESR-specific expression are reported. The function of the ESR is also not clear. However, it was proposed that it is involved in the production/transport of nutrients and/or transfer of developmental signals from the endosperm to the embryo directly or through the suspensor (Schel *et al.*, 1984; Opsahl-Ferstad *et al.*, 1997; Becraft, 2001; Thompson *et al.*, 2001). It is also speculated to be a possible barrier between the embryo and endosperm for protection of the developing embryo from antibacterial and antifungal activity, because of the ESR-specific expression of the defensin *ZmESR-6* (Balandin *et al.*, 2005).

There are three types of aleurone cells: the aleurone surrounding most of the starchy endosperm; the modified aleurone around the ventral groove which is now referred to as the ETC; and modified aleurone cells around the embryo (Ritchie *et al.*, 2000). The third type has morphological similarities to the aleurone cells around the starchy endosperm; however, histochemical staining suggests these cells have more lipid and protein, but less polysaccharide than the aleurone around the starchy endosperm (Raju and

Walther, 1988). This may reflect their function in lipid transfer to the embryo and/or protection of the embryo from pathogens (Raju and Walther, 1988). Several non-specific lipid transfer proteins (nsLTPs) were found to be expressed in aleurone cells around the embryo and two roles were suggested for these nsLTPs: participation in cuticle development and embryonic vascular bundle formation (Boutrot *et al.*, 2007).

Some LTPs were found to be expressed in several grain tissues. For example, two grain-specific maize genes of unknown function, *ZmEBE-1* and *ZmEBE-2*, were found to be expressed in the embryo sac before pollination and later after pollination in both ETC and ESR cells of the developing endosperm (Magnard *et al.*, 2003). However, no expression of these genes was detected in the embryo or other grain tissues.

In this paper, the cloning and characterization of the gene and promoter of another member of the *TdPR60* (Kovalchuk *et al.*, 2009) clade of nsLTPs are described. The gene, designated *TdPR61*, is an orthologue of *TaPR61*, cDNA of which was isolated from the endosperm of hexaploid wheat and used for cloning of *TdPR61*. Analysis of expression using quantitative RT-PCR (Q-PCR) demonstrated grain-specific expression of both genes. Using transgenic wheat, barley, and rice plants stably transformed with promoter-GUS fusion constructs it was demonstrated that *TdPR61* expression, in contrast to *TdPR60*, is not restricted to ETC but can also be detected in partially cellularized endosperm around the embryo, in embryo surrounding and ETC adjacent parts of aleurone, ESR, scutellum, and radical, and thus marks the whole possible pathway of lipids/nutrients and/or developmental signal delivery from maternal vasculature to the roots of the embryo. Transient GUS expression assays in developing embryos were used for mapping of the DNA segments responsible for promoter activation. Two potential promoter segments responsible for the activity of the promoter were identified and the most distal was mapped further. Two putative *cis*-elements from the distal segment were used in a Y1H screen of a cDNA library prepared from syncytial endosperm. The screen resulted in the isolation of a wheat homologue of BES1/BLZ1, suggesting the possibility of brassinosteroid regulation of *TdPR61* expression.

Materials and methods

Gene cloning and plasmid construction

The cDNA of *TaPR61* was isolated from a cDNA library prepared from the liquid part of the syncytial endosperm of *Triticum aestivum* at 3–6 DAP. A single cDNA of *TaPR61* was identified among about 200 cDNAs randomly selected for sequencing. The grain-specific expression of the gene was demonstrated by Q-PCR and justified the selection of *TaPR61* for further characterization. The full-length cDNA of *TaPR61* was used as a probe to screen a bacterial artificial chromosome (BAC) library prepared from genomic DNA of *Triticum durum* cv. Langdon as described by Kovalchuk *et al.* (2009). Seven BAC clones (#952 M16, #946 M23, #1035 P12, #1076 O7, #1128 M7, #1152 A6, and #1172 E15) were identified; clone #952 M16 was selected for further analysis on the

basis of the strength of the hybridization signals and the results of PCR analysis using primers from the beginning and the end of the *TaPR61* CDS. The cloned gene, designated *TdPR61*, and the 2669 bp promoter sequence were isolated and sequenced as described by Kovalchuk *et al.* (2009).

The 2669 bp long regulatory sequence upstream of the translational start, which includes promoter and 5' UTR of the *TdPR61* gene, was cloned into the pMDC164 vector (Curtis and Grossniklaus, 2003) as described by Kovalchuk *et al.* (2009). The resulting binary vector, designated pTdPR61, contained selectable marker genes conferring hygromycin resistance in plants and kanamycin resistance in bacteria. The construct was introduced into *Agrobacterium tumefaciens* AGL1 strain by electroporation. For wheat transformation, the pTdPR61 construct was linearized using the unique *PmeI* site in the vector sequence.

Construction of the *TaPR61* molecular model

A three-dimensional (3D) molecular model of a putative wheat lipid-binding protein (LBP) TaPR61 without a signal peptide was constructed using the Modeller 9v7 program (Sali and Blundell, 1993), as described by Kovalchuk *et al.* (2009). To identify the most suitable template for modelling of TaPR61, searches were performed (Kovalchuk *et al.*, 2009) that identified the LBP defence protein from *Arabidopsis thaliana* (Lascombe *et al.*, 2008), [Protein Data Bank (PDB) ID 2rkn, chain A, designated 2rkn:A], to be the most suitable structural template. The 2rkn:A and TaPR61 sequences were analysed for the positions of secondary structures (Kovalchuk *et al.*, 2009) and for the positions of hydrophobic clusters by hydrophobic cluster analysis (HCA) (Callebaut *et al.*, 1997). The NW-align algorithm (<http://zhanglab.ccmb.med.umich.edu/NW-align/>) was used to align 2rkn:A with TaPR61 and to calculate the positional sequence identity/similarity scores between the two sequences. On the basis of the sequence alignment generated by NW-align and HCA, it was possible to detect the positions of eight paired cysteines that formed a basis for molecular modelling. The aligned 2rkn:A and TaPR61 sequences were used as input parameters to build 3D models on a Linux Red Hat workstation, running a Fedora Linux (release 12) operating system. The final 3D molecular model of TaPR61 was selected from 40 models that showed the lowest values of the 'Modeller Objective Function' and 'Discrete Optimized Potential Energy' analyses. The stereochemical quality and overall G-factors of the 2rkn:A structure and TaPR61 model were calculated with PROCHECK (Laskowski *et al.*, 1993). Z-score values for combined energy profiles were evaluated by Prosa2003 (Sippl, 1993). Spatial superposition of the 2rkn:A structure and TaPR61 model was performed using the 'iterative magic fit' routine in DeepView (Guex and Peitsch, 1997) that performs a sequence-based structural alignment, with minimizing the root-mean-square-deviation (RMSD) value. Here 69 residues excluding insertions and deletions in the C α positions in both architectures were used. The molecular graphics were generated with a PyMol software package (<http://www.pymol.org/>). The molecular model of TaPR61 including the signal peptide was built using I-Tasser (Roy *et al.*, 2010) and evaluated as described above.

Plant transformation and analyses

The construct pTdPR60 was transformed into rice and barley using *Agrobacterium*-mediated transformation and the method developed by Tingay *et al.* (1997) and modified by Matthews *et al.* (2001). Rice (*Oryza sativa* L. ssp. *japonica* cv. Nipponbare) and barley (*Hordeum vulgare* L. cv. Golden Promise) were used as donor plants of embryos. Wheat (*Triticum aestivum* L. cv. Bobwhite) was transformed using biolistic bombardment as described by Kovalchuk *et al.* (2009). Histochemical and histological GUS assays were performed as described by Li *et al.* (2008) using T₀, T₁, and T₂ (only for barley) transgenic plants and T₁, T₂, and T₃ seeds.

Quantitative RT-PCR

RNA isolation and Northern blot hybridization was performed as described elsewhere (Sambrook and Russell, 1989). Q-PCR analysis of the expression of *TaPR61* and *TdPR61* genes in different tissues of untransformed wheat and at different stages of grain development was performed as described by Morran *et al.* (2011).

Promoter analysis

Full-length (2669 bp upstream translational start codon) and 5'-deletions of the *TdPR61* promoter containing full-length 5'-untranslated region were amplified by PCR using AccuPrime™ Pfx DNA polymerase (Invitrogen) from DNA of BAC clone No. 952 M16 as a template. Amplified sequences were cloned into the pENTR-D-TOPO vector (Invitrogen, Mulgrave, Victoria, Australia); the cloned inserts were verified by sequencing and subcloned into the pMDC164 vector. Plasmid constructs were purified using commercially available kits. DNA-gold compositions for biolistic transformation were then produced as follows: 50 µg of gold powder (1.0 µm) in 50% glycerol was mixed with 10 µl of plasmid DNA (1 mg ml⁻¹), 50 µl of 2.5 M CaCl₂, and 20 µl of 0.1 M spermidine. For co-transformation, plasmids were mixed 1:1 v/v. The mixture was vortexed for 2 min, incubated for 30 min at room temperature, briefly centrifuged at 13 000 g, and washed sequentially with 70% and 99.5% ethanol. Finally, the pellet was re-suspended in 60 µl 99.5% ethanol and 6 µl of suspension was used for each shot. The microprojectile bombardment was performed using a biolistic PDS-1000/He particle delivery System (Bio-Rad). Immature embryos (12–15 DAP) were pre-incubated for 4 h on MS2 medium (Murashige and Skoog, 1962) containing 100 g l⁻¹ sucrose. Embryos (10 per plate) were placed in the centre of the plate to form a circle with a diameter of 10 mm, and were bombarded under the following conditions: 900 or 1100 psi, 15 mm distance from a macrocarrier launch point to the stopping screen and 60 mm distance from the stopping screen to the target tissue. The distance between the rupture disc and the launch point of the macrocarrier was 12 mm. After bombardment embryos were incubated in the same medium for 24 h and stained for GUS activity. Two to three repeats (independent bombardment events), with 10 embryos per plate were used for each construct. After overnight GUS activity staining at 37 °C the embryos were examined for GUS colour forming units (CFU) and images were captured using a LEICA DC 300F camera attached to a LEICA MZ FLIII stereomicroscope. The CFU were counted using the images viewed with Microsoft Office Document Imaging.

The yeast 1-hybrid screen

The Y1H screen was performed as described by Pyvovarenko and Lopato (2011). The full-length cDNA of a short form of TaBES1 (TaBES1S) was isolated from a Y2H cDNA library prepared from the endosperm of *Triticum aestivum*, cultivar Chinese Spring at 3–6 DAP. The bait DNA sequences used to screen the cDNA library consisted of three and four tandem repeats, respectively, of the predicted *cis*-elements 61EMB (TCAAACCCACACACGA) and 61END (CGCACGTACAC). A screen of the grain and endosperm cDNA libraries with 61EMB produced no transcription factors (TFs). Of 48 positive clones analysed during the screen of the endosperm library with the 61END element only three positive clones contained inserts encoding TFs. All three clones contained the same cDNA encoding TaBES1S. The full-length cDNA of TaBES1 as well as TaBES1S were cloned using RT-PCR followed by nested PCR; RNA isolated from the developing grain was used as template. Primers for the second PCR round were derived from the beginning and the end of the coding regions. The coding region of TaBES1S and TaBES1 were re-cloned in-frame with the yeast activation domain into the pGADT7 vector and specific

(Kovalchuk *et al.*, 2009). Data about transgenic lines are summarized in Table 1.

Structure and molecular modelling of the TaPR61 protein

Alignment of the protein sequences of TaPR61 and TdPR61, deduced from cDNA and genomic sequence, respectively, shows a difference in two amino acid residues (98% sequence identical and 100% sequence similar) (Fig. 1C).

Analysis of the protein sequences of TaPR61 and TdPR61 using SignalP 3.0 (Bendtsen *et al.*, 2004) suggested the presence of an N-terminal hydrophobic signal peptide of 20 amino acid residues in length, implying secretion of both TaPR61 and TdPR61, most likely to the apoplast. The predicted cleavage site was between amino acid residues 20 and 21, producing a mature protein of 87 amino acid residues with a predicted molecular mass of 9.5 kDa and a pI point of 8.6.

The most appropriate structural template for the TaPR61 protein was found to be the lipid binding protein (LBP) from *A. thaliana*, 2rkn:A (Pasquato *et al.*, 2006). This protein consists of 107 residues and two monoacylated phospholipid molecules [(7R)-4,7-dihydroxy-*N,N,N*-trimethyl-10-oxo-3,5,9-trioxo-4-phosphaheptacosan-1-aminium 4-oxide] that are bound side-by-side in the internal cavity of 2rkn:A. Further, three Zn²⁺ ions are positioned on the surface of 2rkn:A. This template was used for molecular modelling of TaPR61. A total of 77 amino acid residues from 107 residues of TaPR61, excluding the predicted 20-amino-acid-residues long signal peptide and ten COOH-terminal residues that

had no structural counterpart in 2rkn:A, were aligned with 74 residues of the template from a total 77 residues of 2rkn:A, excluding three N-terminal residues. The positional sequence identity and similarity between the 2rkn:A and TaPR61 sequences were 26% and 70%, respectively, and these scores reflected a high complexity of modelling (Sanchez and Sali, 1998). However, analyses of sequences by hydrophobic cluster analysis (HCA) indicated that the positions of eight cysteine residues were highly conserved (Fig. 2A; marked by arrowheads). This indicates that the cysteine residues in TaPR61 could also form disulphide bridges and that the associated parts of the sequences could form α -helical secondary structural elements (Fig. 2A; black lines).

The molecular model of TaPR61 with secondary structural elements superposed on the template structure is shown in Fig. 2B (left panel). This figure also shows the positions of four invariant disulphide bridges (shown as lines in atomic colours). Superposition of TaPR61 on the template crystal structure of 2rkn:A indicated that the two structures matched closely and had a root-mean-square-deviation value of 0.67 Å for the C α backbone positions in 69 structurally equivalent positions (Sippl and Wiederstein, 2008). The positions of two monoacylated phospholipid molecules in the TaPR61 model, which were resolved in the 2rkn:A structure by X-ray crystallography, were also determined (Fig. 2B). Figure 2B shows that the two bound monoacylated phospholipid molecules would accommodate well in the internal central protein cavity of TaPR61. The stereochemical parameters of the template 2rkn:A structure and the modelled TaPR61 were calculated by PROCHECK (Ramachandran *et al.*, 1963; Laskowski *et al.*, 1993), and the calculations showed that 100% of all residues were located in the most favoured, additionally allowed or generously allowed regions of the Ramachandran plot. Therefore, the stereochemistry of the TaPR61 model was correct. The overall average G factor values of 2rkn:A and TaPR61, as measures of normality of main chain bond lengths and bond angles (Ramachandran *et al.*, 1963), were 0.07 and -0.28, respectively. The Z-score values for combined energy profiles were -7.70 and -5.77 for 2rkn:A and TaPR61, respectively. Thus, summarizing, although modelling of TaPR61 represented a so-called 'twilight zone' case, the evaluation data indicated that the TaPR61 model was reliable.

Molecular modelling of a full-length sequence of TaPR61, including its hydrophobic signal peptide, was generated through the I-TASSER Server (Roy *et al.*, 2010). The molecular model shown in Fig. 2B (right panel) indicates that the signal peptide comprising 20 amino acid residues folds into additional α -helix in TaPR61 that has a significant hydrophobic nature, as also suggested by TOPCONS analysis (Bernsel *et al.*, 2009), where SCAMPI, PRODIV, and OCTORPUS algorithms predicted that the first 20 residues could form a membrane-associated α -helix. Hence, it could be concluded that this α -helix could impose a significant hydrophobic character to TaPR61.

Table 1. Information about transgenic plants transformed with pTdPR61 construct

Number of transgenic lines	Wheat	Barley	Rice
T ₀ lines selected on Hg	56	11	32
T ₀ lines confirmed by PCR	43	11	No data
T ₀ lines tested for GUS activity	43	11	32
T ₀ lines with strong expression	9	4	21
T ₀ lines with weak expression	1	4	5
T ₀ lines with no expression	30	3	4
Sterile T ₀ lines	3	–	2
T ₀ lines with different pattern of expression	–	–	–
T ₁ and T ₂ lines tested for GUS activity	–	L1-3; L2-3; L11.1-3	–
T ₁ plants with strong expression	–	L1-1; L2-1; L11.1-3	–
T ₁ plants with weak expression	–	L1-1; L2-1; L11.1-0	–
T ₁ plants with no expression	–	L1-1; L2-1; L11.1-0	–

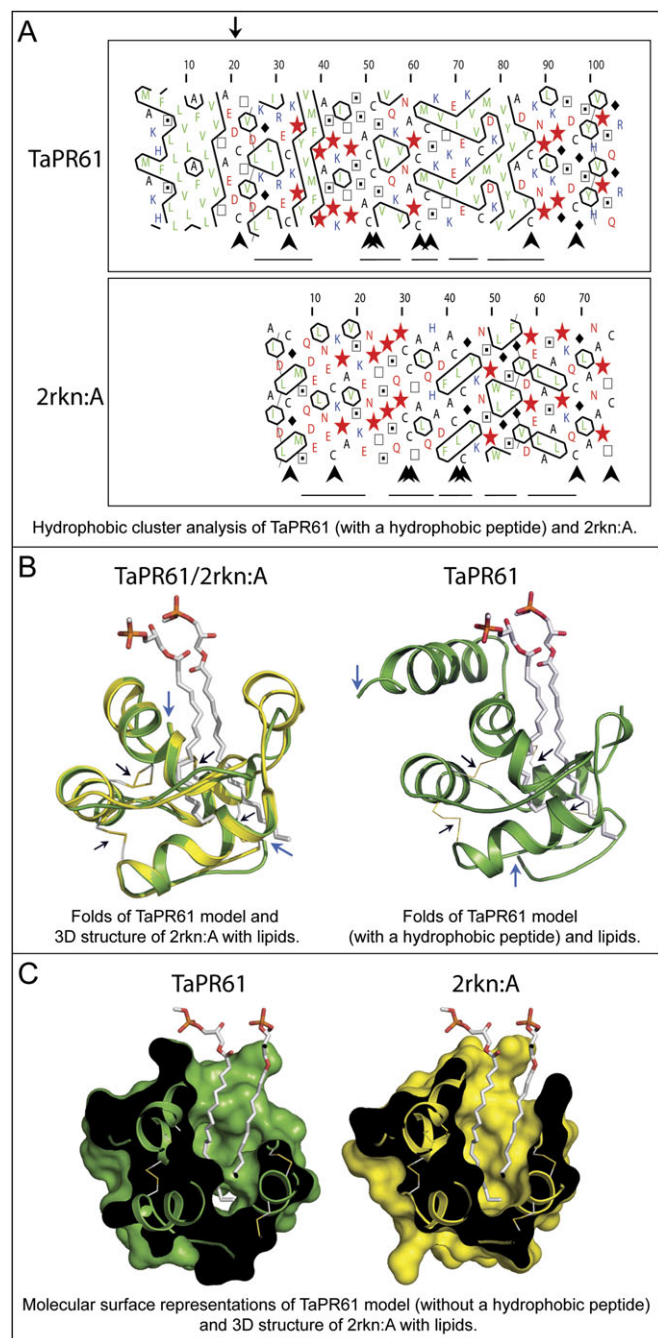


Fig. 2. Molecular model of the TaPR61 lipid-binding protein. (A) Hydrophobic cluster analysis of TaPR61 (with a hydrophobic signal sequence) and 2rkn:A. Positions of four paired conserved cysteine disulfide bridges (arrowheads) and the positions of α -helices (lines) are shown. (B) Left panel, Superposition of TaPR61 (green) on the template structure of 2rkn:A (yellow) indicating the dispositions of secondary structural elements. The positions of the two monoacylated phospholipid molecules in the protein cavities are illustrated. Right panel, molecular fold of TaPR61 (with a hydrophobic signal peptide) containing two monoacylated phospholipids. The NH₂- and COOH-termini are indicated by upper and lower blue arrows, respectively. The positions of four invariant disulfide bridges, shown as lines in atomic colours are indicated by black arrows. (C) Cross-sectional views through molecular folds and surfaces of TaPR61 (green) and 2rkn:A (yellow) show the dispositions of

Spatial and temporal activity of the TdPR61 promoter in transgenic wheat

From 43 wheat transgenic lines confirmed by PCR, nine had strong and one weak GUS expression. GUS expression was not detected in another 30 transgenic lines. The data about transgenic lines are summarized in Table 1.

Analysis of the activity of the *TdPR61* promoter in transgenic wheat plants using a promoter–GUS fusion construct demonstrated good spatial and temporal correlation between the activity of the *TdPR61* promoter and expression of *PR61* genes from *T. aestivum* and *T. durum*.

At 9 DAP, *TdPR61* promoter activity was detected in both the embryo and endosperm; in the endosperm it was observed in aleurone and starchy endosperm close to the grain poles (Fig. 3, A1–A3, B1–B4). The strongest expression was seen in and around the embryo, in ESR and in the upper cell layers. At 10–14 DAP the activity became stronger in aleurone, ETC, and the starchy endosperm over ETC, however, it disappeared inside the embryo and was observed only in a fraction of the outer cell layer of embryo adjacent to the ESR (Fig. 3, A6–A8, B6–B12). At 21 DAP the GUS expression was detected only in ETC, the ESR, and coleorhizae, marking a possible pathway of lipids/nutrients and/or developmental signal delivery from the main maternal vascular bundle to roots of the embryo (Fig. 3, B13–B18). In embryos, isolated from 9 DAP transgenic grains, GUS activity was observed along the embryonic axes; however, in contrast to barley it was not found in the scutellum (Fig. 3, A4, A5). In 27 DAP embryos the activity of the promoter was observed in the coleorhizae only in the region closer to the scutellum (Fig. 3, A10). Longitudinal sections of embryos revealed strong activity of the promoter in the outer cell layers of the coleorhizae (Fig. 3, B17, B18). After 30 DAP promoter activity was observed only in ETC and was not detectable in the embryo (Fig. 3, B19, B20).

Spatial and temporal activity of the TdPR61 promoter in transgenic barley

Eleven transgenic barley plants were generated; four of them had strong and four had weak GUS activity. The pattern of GUS expression was the same in all eight lines. No GUS activity was detected in three transgenic lines (Table 1). Similarly to wheat, the activity of the promoter was first detected in cellularizing endosperm around the embryo at 5 DAP (Fig. 4A, L). At 8 DAP it was observed in ETC, close to the embryo part of the aleurone and starchy endosperm, ESR, and the strongest expression was around and in the embryo (Fig. 4B, G). Later, the activity of the promoter started gradually to decrease in the endosperm, but in the embryo it was localized in the scutellum (as two cylinders directed from the endosperm to points of connection of the scutellum and embryonic axis) and further in the

secondary structure elements (illustrated as cartoons), morphology of internal cavities and positions of disulphide bridges (as lines coloured in atomic colours).

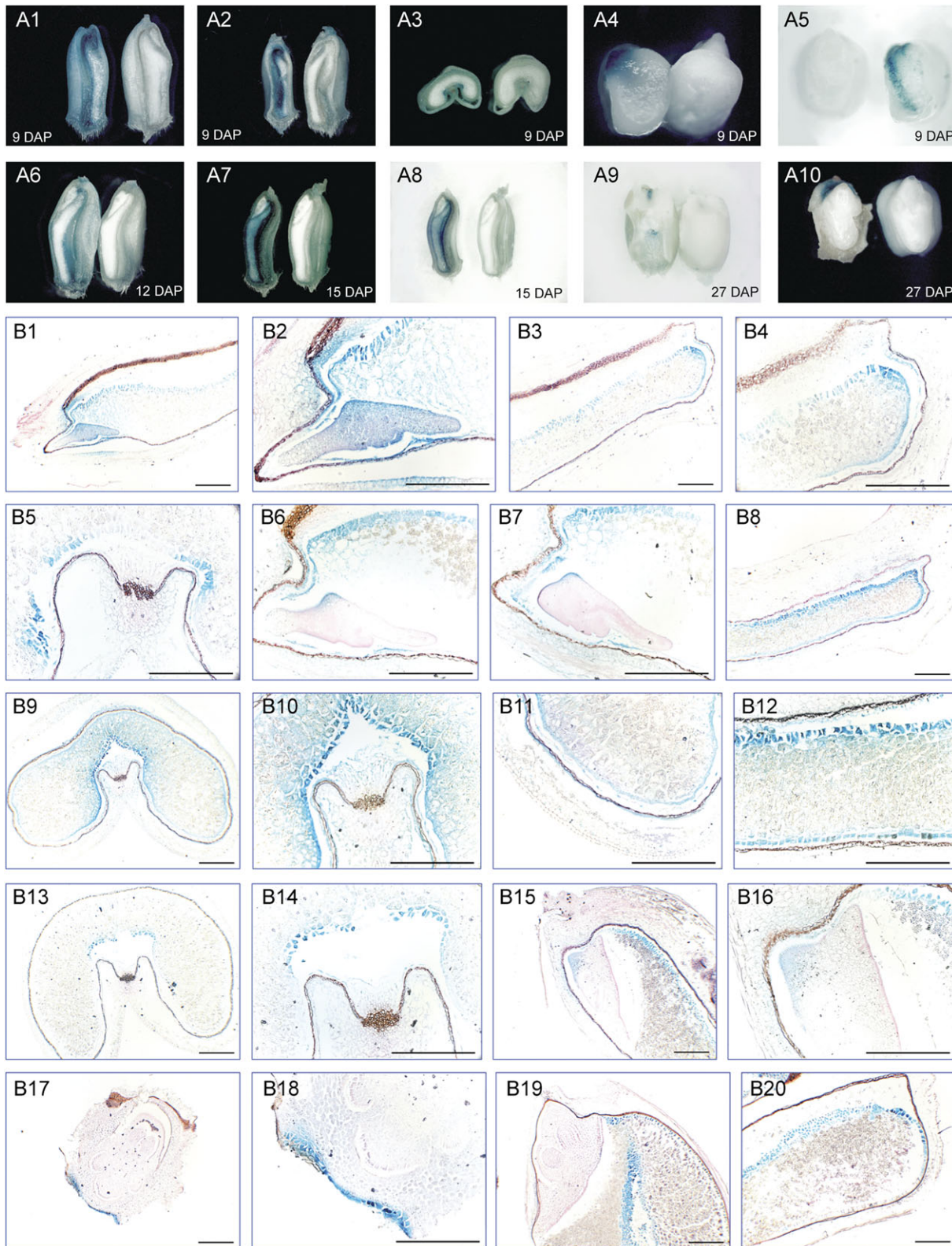


Fig. 3. Spatial and temporal GUS expression in wheat directed by the *TdPR61* promoter. GUS activity in grain tissues (A1–A3, A6–A8) and in isolated embryos (A4, A5, A9, A10) at 9 DAP (A1–A5), 12 DAP (A6), 15 DAP (A7, A8), and 27 DAP (A9, A10). Control grain and embryos at the same stage of development are shown on the right hand side of each panel. Histochemical GUS assay counterstained with Safranin O in 10 μ m thick longitudinal sections (B1–B4, B6–B8, B12, B17–B20) and cross-sections (B5, B9–B11, B13, B14) of grain (B1–B16, B19, B20) and isolated embryos (B17, B18) at 9 DAP (B1–B5), 10 DAP (B6–B8), 14 DAP (B9–B12), 21 DAP (B13–B18), and 30 DAP (B19, B20). Bars: 0.2 mm.

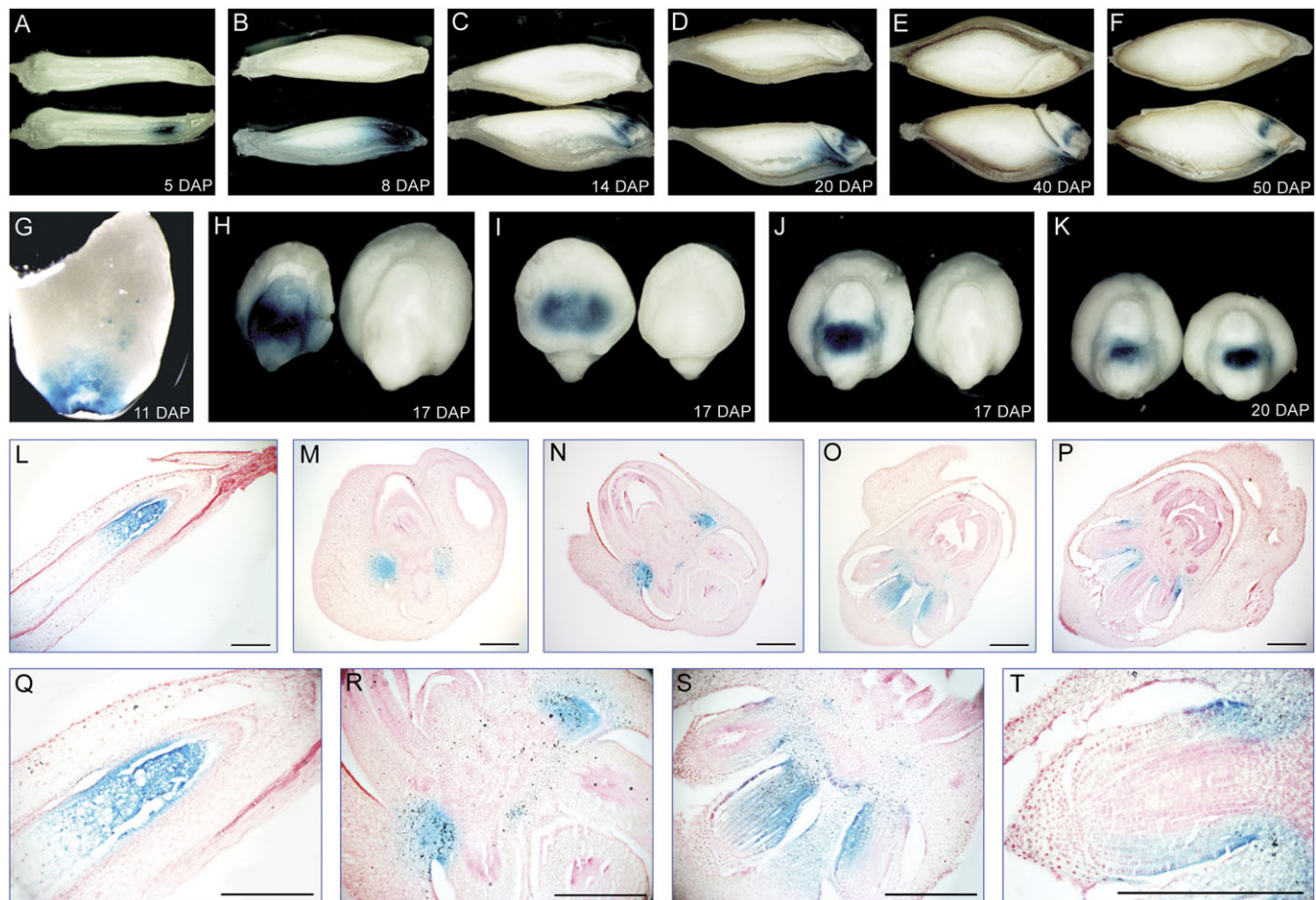


Fig. 4. Spatial and temporal GUS expression in barley directed by the *TdPR61* promoter. GUS activity in grain tissues (A–F), in endosperm after embryo isolation (G), and in isolated embryos (H–K) at 5 DAP (A), 8 DAP (B), 11 DAP (G), 14 DAP (C), 17 DAP (H–J), 20 DAP (D, K), 40 DAP (E), and 50 DAP (F). Control grain and embryos at the same stage of development are shown on the upper part of panels B–F and right hand side of panels H–J. Histochemical GUS assay counterstained with Safranin O in 10 μ m thick longitudinal sections of grain at 5 DAP (L, Q) and isolated embryo at 20 DAP (M–P, R–T). Bars: 0.2 mm.

coleorhizae (Fig. 4C–F, H–T). The activity of the promoter could be detected until at least 50 DAP. No GUS activity was detected in leaves, stems, and roots of mature plants, roots and shoots of seedlings, and in germinating seed.

Spatial and temporal activity of the TdPR61 promoter in transgenic rice

Thirty-two independent transgenic lines were selected using hygromycin resistance. The presence of the transgene was confirmed by PCR in all lines using primers specific for the GUS gene. Strong GUS activity, driven by the *TdPR61* promoter, was detected in developing grains of 21 lines, five lines showed very weak GUS activity, no activity was detected in four lines, and two lines were sterile and were not analysed (Table 1). A pattern of *TdPR61* promoter activity similar to wheat was found in transgenic rice plants, except that the activity of the promoter was less specific and much stronger than in transgenic wheat. GUS activity was not detected before 7 DAP and at 9–10 DAP could be seen mainly in the embryo and around the embryo, as well as in

the opposite pole of the grain (Fig. 5, A2, B1–B9). In one of three tested transgenic lines it was detected in the vascular tissues of the lemma at 8 DAP (Fig. 5, A1). Between 8 and 21 DAP, strong GUS activity was detected everywhere in the embryo; in the embryo-surrounding part of aleurone and ESR; and, similarly to wheat, GUS activity was observed in ETC and part of a single cell layer aleurone along the whole main vascular bundle (Fig. 5, A3–A11, A21, B1–B15). However, in contrast to wheat, where GUS activity disappeared from the embryo earlier than from ETC, in rice, GUS activity initially slowly diminished in ETC and after 26 DAP promoter activity was detected only in the embryo (Fig. 5, A11–A20, A22–A25). Strong GUS activity could still be seen in the embryo and in some lines in the ESR and close to the embryo part of the ETC at 69 DAP. At this point, in lines with weaker transgene expression GUS activity was detected in the scutellum, but not in the embryonic axes (Fig. 5, A18, A19). No analyses were performed at later stages of grain development. No GUS activity was detected in leaves, stems, roots, florets and anthers of transgenic plants.

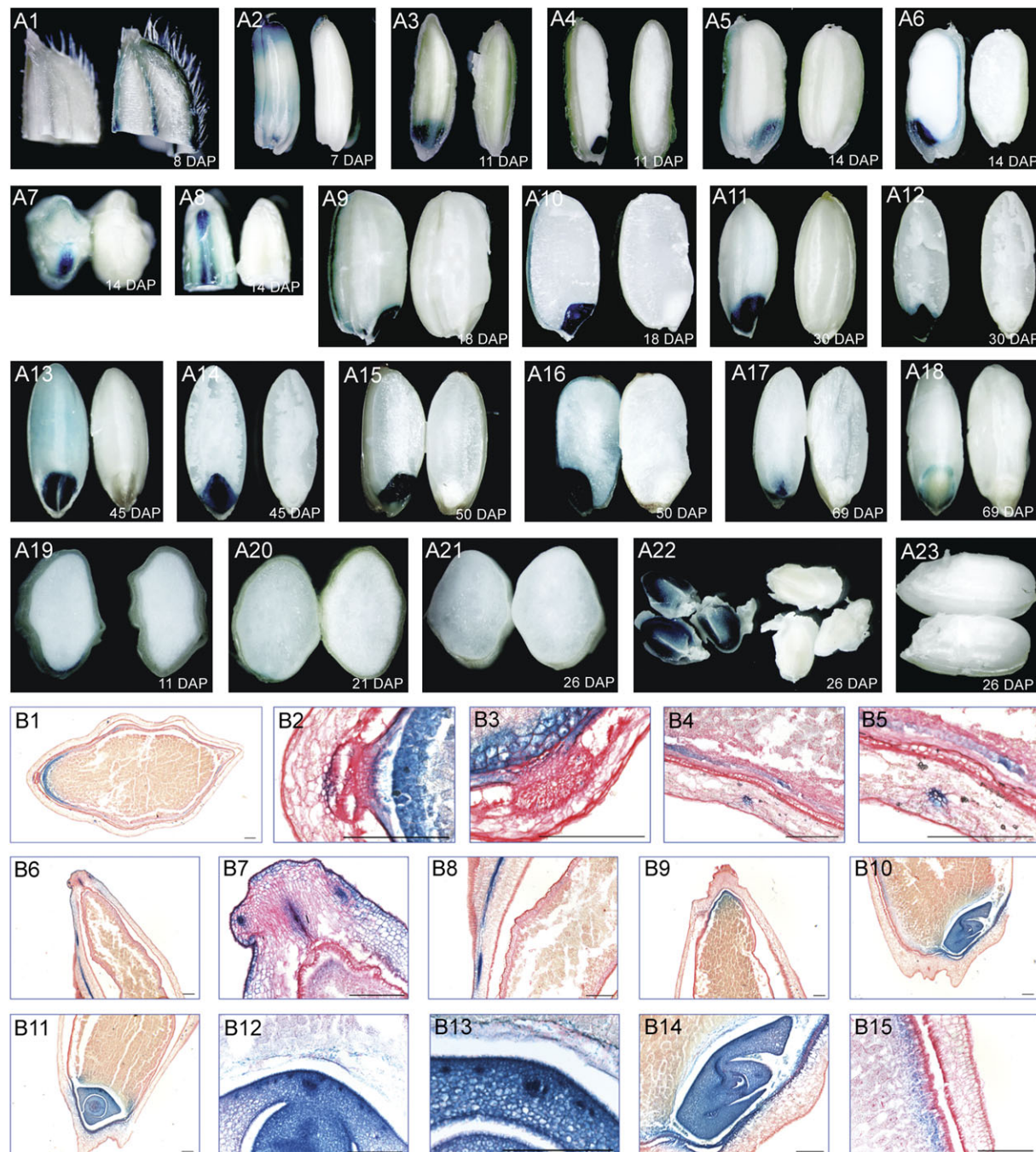


Fig. 5. Spatial and temporal GUS expression in rice directed by the *TdPR61* promoter. GUS activity in vascular tissues of lemma (A1), tissues of hand-cut grain (A2–A26), isolated embryos (A22), and endosperm after embryo isolation (A23) at 7 DAP (A2), 8 DAP (A1), 11 DAP (A3, A4, A19), 14 DAP (A5–A8), 18 DAP (A9, A10), 21 DAP (A20), 26 DAP (A21–A23), 30 DAP (A11, A12), 45 DAP (A13, A14), 50 DAP (A15, A16), and 69 DAP (A17, A18). Control grain and embryos at the same stage of development are shown on the left hand side of panel A1, right hand side of panels A2–A22, and the upper part of the panel A23. Histochemical GUS assay counterstained with Safranin O in 10 µm thick cross-sections (B1–B5, B12, B13) and longitudinal sections (B6–B11, B14, B15) of rice grain at 8 DAP (B1–B7) and 10 DAP (B9–B15). Bars: 0.2 mm.

Analysis of the *TdPR61* promoter

Computer analysis of the *TdPR61* promoter using PLACE software and a database of plant *cis*-acting regulatory DNA elements (Higo *et al.*, 1999) revealed a large number of elements potentially responsible for endosperm- and

embryo-specific expression of *TdPR61*. Among them are AACA and ACGT motifs found in the *GLUB-1* gene from rice (Wu *et al.*, 2000). Together with the GCN4 motif they were sufficient to confer endosperm-specific expression. There were also a large number of E-box elements, CANNTG, which is responsible for the transcription of

seed-specific promoters (Stalberg *et al.*, 1996). It was also reported that key transcriptional regulators of the brassinosteroid pathway, BES1 and BIM1, can synergistically interact with E-boxes on the *SAUR-AC1* promoter (Yin *et al.*, 2005). The ACACNNG element responsible for embryo-specific expression was repeated in the *TdPR61* promoter several times. This element can bind bZIP transcription factors, which activate late embryogenesis (Kim *et al.*, 1997; Finkelstein and Lynch, 2000). Another endosperm-specific element is the GCAAC repeat, GCAACGCAAC, which was found in the promoter of the beta-Zein gene from maize (So and Larkins, 1991). Two different soybean embryo factor (SEF) binding sites, AACCCA for SEF3 and RTTTTTR for SEF4, from the beta-conglycinin promoter, responsible for embryo-specific expression (Lessard *et al.*, 1991), were also found in the *TdPR61* promoter. Besides embryo- and endosperm-specific elements, the *TdPR61* promoter contains four GCC-boxes found in many pathogen-responsive and wounding-inducible genes and responsible for jasmonate-responsive gene expression (Brown *et al.*, 2003). However, no induction of the *TdPR61* promoter activation by mechanical wounding has been detected (data not shown).

The full-length promoter (D0) and four truncated versions of the full-length promoter, D1, D2, D3, and D4 (Fig. 6), were cloned in the pMDC164 binary vector upstream of a *GUS* gene. The resultant constructs were used for biolistic bombardment. Bombardment of suspension cell cultures of *Triticum monococcum* L. and whole developing grains of *Triticum aestivum* cv. Bobwhite gave no GUS activity (data not shown). However, bombardment of embryos isolated from wheat grain at 12–15 DAP demonstrated a strong activity of the full-length promoter. The locations of at least two putative regulatory elements can be deduced from the transient expression assay data: the promoter fragment between D3 and D4 deletions contains a *cis*-element which is sufficient for the weak promoter activation; another fragment between D0 and D1 contains a binding site for another putative activator/enhancer of promoter activity, which is responsible for the additional (about 5-fold) increase of promoter activity in the embryo. Attempts at further mapping of the proximal *cis*-activating element(s) were not successful, because the relatively weak activation of respective promoter deletions did not permit accurate measurements. A second set of truncated *TdPR61* promoters was produced with the aim of more accurate mapping of the distal *cis*-element(s), responsible for strong promoter activation. These truncated promoters were designated D0.1, D0.2, D0.3, D0.4, and D0.5 (Fig. 6). Because the activity of the D0.5 promoter was roughly the same as the activity of the full-length promoter D0, it was concluded that a putative distal *cis*-element is situated in the 145 bp long promoter fragment between D0.5 and D1. Analysis of the sequence of this fragment using PLACE software (Higo *et al.*, 1999) revealed three potential *cis*-elements, which were earlier found to be responsible for embryo- and endosperm-specific gene expression. Two of the predicted embryo-specific elements

A

```

>Full length (D0)
1  CGGTAAATG TCAGGATATG CTGTCTATCT GCTCTGGCAT CGTTTGGCTT
51  TAGTTACTGC CTTCCTGAGC ATGCACAGAA AACACAATAG CAAATTAGTG
    >D0.1
101  TCAGTTTCTC AGATGACAGA TTTAGTACAA GATGCTTACT TGCACATGGG
    E-box
151  AACGTTCTTC ATCTACGAAG AACACAGAGTA CTAAGATTTT GACCAATGTA
    >D0.2
201  TAGCGGAGAC GTTCTACTGA ACATGCGCAG GCACGAAGCT ACCTGTCACC
251  GTCCGATAGA GTGGTGATGC AGCATGCTTG AAGTGATGAT TGATGTGACG
    >D0.3
301  GCCTAACGTG GCCTCTGTAT GGCCTGCGCG CCGTAGTAAT CTCGAATTTA
    >D0.4
351  GTTTTGTCAA ATATTACGCG TGCGCTTGAG CTTCGGAAAA ACAAGCAATG
401  TGTGAACATG AGCTGGGCGC TTAACAGATA CAAACCAAGG GGTAAATCCA
451  AAGCACCCGT CTTCTGACCA ACCAAGATCC CGACATATAT AACACAGCCG
    >D0.5 SEF3 bZIP
501  ATCCCAATGC AACCCACAC ACGAACCTAC AAGAGCCAGC CACCCACCCA
    E-box
    End. -spec.
551  CTATTACAC AGCAACAGCT ACCGCACTCC CAACGTACAC GCAAGAACTG
    >D1
601  AAACCTGGTG CCATGCGCAG GCTCTTTGCT GTGTGCTTGG TTCTGCTGCG
    GCC-box
651  CTTCAACATG GCGCTGGCGG CAGATATGGC GCGCATGATG GCGCGATGG
    GCC-box
701  CGGCAGATAT GCGCGCGGAT ATGGCCCGGA TGCGCGCGCC GCGCGCGGAC
    E-box
751  GCGGCGGACT GCAACACGGA CTTGCGAGGC CTGGTTGCCA ACTGCCAGAA
801  CTACATCATG TTCCCGGCGG AGCCGAAGAT AACGCCGCTG CCGCGCTGCT
851  GCGCGCTGAT CCAGAGGGCG GACATGCGCT GCTCTGGGCG CAAGGTGACC
901  CCGCGCATGT AGAAGGTGCT TCGCATGGAT TCTGCTGGTG TTCTGCGCCA
    E-box
951  AATACCTCAA GAGGCGCTTG CAGCTTGCTG CCACTGCGGG AAGTAAGTGT
    E-box
1001  TAGACCATAT GCTGTGTTAT ATCATTTTTC AATTTTCCAA ATATATGTTG
1051  GATCAACTAA AATTTTCTGA AAGTTGGACA AAGGGAAATG CCAATATAT
1101  GTTTTATATG GTTATCTTTT TTCTGTGTGT TTTCGAGCCT ACCCGTCTCC
1151  CTGTGTCGTT CGTGTAGTGC AGCATATCTT CTGATTTGAT GATCACAAG
    E-box
1201  TCTGCAGATG CTATTTCAGT TTTTCAACAT TATTGTACTG GTGCTTGGA
1251  CTTGCATATT TGCTGAGAGA TAAGTTGCTG GATTATCACC TAGATAAATC
    E-box
1301  CAGATATCTT ATGATAAGTA TAAATAAAGT TACGGGCAAA TGTATGATTG
    E-box
1351  CCAGTTAAAA AAAGCAGTTG GCTATGTAGC TGTACTTGCG TTTCTTTAAA
1401  GGGCAGCTTA TTTTGTGGGC GTTTGTTTTT TCGAAGAGGA TCAATCTAT
1451  TATAAAGACT CATGAGAAGT ACAAAATACC GGAATATATA TGAAATTTAC
    E-box GCC-box
1501  ATCGAGTCC ATGGACCGCA GAACGACCAT TGCGCGCCCA AACAGGTGTA
1551  CCGATGCGGC GATGTCGTGC CTCCTATACC GAAGCGGGCC TGACCTTGTC
    E-box
1601  GATGACAGCT GGGAAAGCTG AGTGCATGTA CCCTAAAGA CCAGCGCCCC
1651  GAAGCCACAG TCGTCGCTCT TGAATCCTTG AATAGACCTG AAGAATCTGA
    End. -spec.
1701  TACCAATAT TCGCGTTGCG TACACAGCAG AAGAAACGAA TCTCACTACC
    SEF4
1751  CCGAGGAGTC AAAAATACAA ATATGTACTC CACTGTAGAA GTAATGTTG
    >D3
1801  GACTCTAGTA AGTGCCACAC ATAAGATACG AGCAGATGCC AACTTAGAAC
1851  ATTTTATATG GTAAGTTTAG TTATGCGAGG ATGGCAATTT TAGTTGGCAA
1901  GCATGGCATC TTTGTGCTTC AGTTTATTTT ACCTTGGAAA CTTGCTGTGT
1951  GTCACTGGAA TTTGTCTATC TAGCGGGACT GGAATTGCCA TCGAATAATG
    E-box
2001  TCGACTGGCC GGAGTGCCAC GCACTGACGC TGTGTCACTT ATCATTAGG
2051  TAAATATTTG TGAATATGGA TTGATGCTTC AATACATATC ATAATAATG
2101  GTCTCATAT ATCTCTACAA ATATATACTA AACATATAAT CAGTGAATA
    E-box
2151  TATGCTGCAA ATGCCTTGST CTGCATTATC AGTTTGCATC AATATGATAC
2201  ATCATACATG TATCCAACTT GGTGTGCTAT GGATGTATTC ATTAATAGC
    >D4
2251  TGCTTTTGTG GAGATGGGTC AGTACAACAA ACATAAGTGC TCACATTTGA
2301  CTATAAGACA AACAGATGGG AGGATTGAAA TTTCTTAGAA AACATTATAT
    End. -spec.
2351  TTTGGAACAT ATATGCAACG CAACAAGAAA ACAGAGGTGT AGACACGTGT
    E-box
2401  CTACATAGGA ACTAAGTGTG GACACTTGTC AATGTAACAA ACGCATGTGC
    End. -spec.
2451  GTACATCTGT CTACATGAGA AACAAAAGTC GTGACACTTG TCAATCAGCA
2501  CCACTTAGCC CTGTATGATA ATAAATATT TCCCTGGGCT GATCTATTC
2551  CCATCATGCG TCTGATATA ACCCAAGAT CCAACAGAGG ACACCCACAC
2601  CAAAACACCA CAACCTTAGA TTTACATCTT AAGAGAGAGG TAGCCACCGC
    Rev Primer< M
2651  ATTACTACTG GTAGTAGTTA TG

```

B

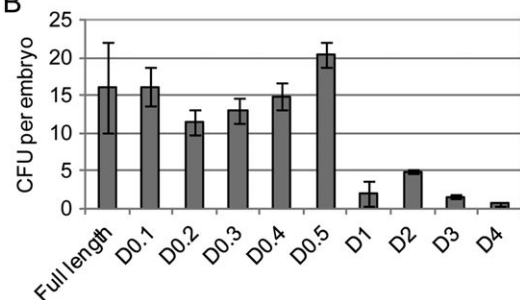


Fig. 6. Identification of sequences responsible for the activation of the *TdPR61* promoter. (A) The *TdPR61* promoter sequence. Deletions used in the mapping of functional *cis*-elements are indicated, the first base pair of each deletion is underlined; elements used as baits in the Y1H screen are in grey boxes. The predicted endosperm and embryo specific elements are underlined and the names of elements are indicated over underlined sequences. The putative TATA-box is in bold. (B) Results of the bombardment of developing wheat embryos with constructs containing different promoter deletions cloned upstream of the reporter gene.

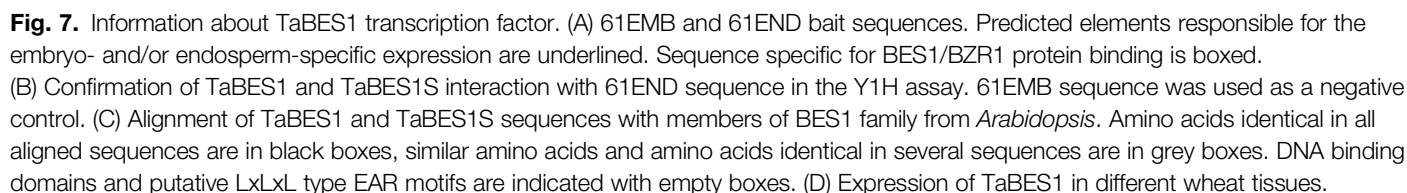
were adjacent to one another (AACCCACACACG) and had a single base overlap, which is shown in bold. The first predicted embryo-specific element (AACCCA) shows homology to the Soybean Embryo Factor 3 (SEF3) binding site from the soybean consensus sequence for the 5' upstream region of the beta-conglycinin (7S globulin) gene (AACCCA-27 bp-AACCCA) (Allen *et al.*, 1989). The second predicted embryo-specific element (ACACACG) shows homology to the binding site of bZIP transcription factors DPBF-1 and 2 factors (*Dc3* promoter-binding factor-1 and 2) from carrot (Kim *et al.*, 1997). *Dc3* expression was embryo-specific and could also be induced by ABA. The predicted endosperm-specific element (CACGTAC) is situated relatively close to the embryo-specific elements. This element shows homology to an element which was originally found in the *GluB-1* gene in rice and which is required for endosperm-specific expression and conserved in the 5'-flanking region of glutelin genes (Washida *et al.*, 1999). Sequences of these predicted elements, shown with two base pairs flanking each end, CAAACCCACACACGAA and CGCACGTACAC, were repeated three and four times, respectively, and designated 61EMB and 61END; they were used as baits to screen cDNA libraries prepared from whole grain and the liquid fraction of developing endosperm, respectively (Fig. 7A).

Screens with the full-length D0.5-D1 fragment or with the PR61EMB bait sequence did not identify any transcription factor. However, a screen with the PR61END bait sequence resulted in several independent clones encoding protein with a DNA binding domain similar to that of the BES1/BZR1 family of TFs (Mora-Garcia *et al.*, 2004; He *et al.*, 2005; Vert *et al.*, 2005). Analysis of the sequence of the cloned cDNA revealed that it is identical to part of the sequence of cDNA encoding the BES1/BZR1-like protein from wheat, designated TaBES1. The cDNA isolated in the Y1H screen contained a long deletion which led to a change of reading frame and, as a result, encoded a truncated form of TaBES1 protein, designated TaBES1S, which includes only the N-terminal nuclear localization signal and DNA binding domain. The full-length cDNA of TaBES1 was isolated using primers from the beginning and the end of the TaBES1S cDNA. Both TaBES1S and TaBES1 were shown to interact with 61END, but not with 61EMB (negative control), in the Y1H assay (Fig. 7B). It was reported that a close homologue of AtBES1, AtBRZ1, binds to the promoter of the BR biosynthetic gene *CPD* via a brassinosteroid response element (BRRE) with the sequence CGTG(T/C)G and, as a result, suppresses expression of this gene (He *et al.*, 2005). The 61END bait sequence and hence the *TdPR61* promoter also contained the BRRE, CGTGCG. Alignment of the TaBES1 and TaBES1S protein sequences to sequences of members of the BES1/BZR1 family from *Arabidopsis* revealed the presence of the full-length DNA binding domain in the TaBES1S, however, most other conserved motifs including the potential C-terminal LxLxL-type of the ethylene-responsive element binding factor-associated amphiphilic repression (EAR) motif (Kagale *et al.*, 2010) were absent in the short form of the

protein (Fig. 7C). Analysis of the expression of *TaBES1* using Q-PCR and primers common for both *TaBES1* and *TaBES1S* cDNAs showed low levels of expression of these genes/gene forms in all tested tissues with the maximum expression in the pistil shortly before fertilization and in the leaf (Fig. 7D). Both genes/gene forms were found by RT-PCR using RNA from the endosperm fraction of grain at 3–6 DAP (data not shown).

Discussion

nsLTP PR60, previously characterized from wheat, was specifically expressed in the ETC layer and in several layers of starchy endosperm cells over the ETC. In this paper, a homologous nsLTP, designated PR61, is described that is expressed in several endosperm and embryo tissues and which could be involved in the transport of nutrients/lipids or the transfer of developmental signals from maternal vascular tissues to embryo roots during grain development. The cDNA encoding TaPR61 was isolated from a cDNA library prepared from the liquid fraction of wheat syncytial/cellularizing endosperm at 3–6 DAP. It encodes another nsLTP, which belongs to a subfamily of nsLTPs that includes several already characterized ETC-specific proteins: HvEND1, ZmBETL-9, OsPR602, TaPR60, and TdPR60 (Doan *et al.*, 1996; Gruis *et al.*, 2006; Li *et al.*, 2008; Gómez *et al.*, 2009; Kovalchuk *et al.*, 2009). Plant nsLTPs are low molecular mass, soluble proteins that were originally proposed to facilitate the transfer of fatty acids, phospholipids, glycolipids, and steroids between membranes so they can perform these functions *in vitro* (Yeats and Rose, 2008). Despite their designation, their role in intracellular lipid transport is considered to be unlikely, as they occur extracellularly. A number of other biological roles, including antimicrobial defence, signalling, cuticle deposition, and cell wall loosening have been proposed, although conclusive evidence for their functional role is lacking, as the aforementioned functions are not well correlated with their *in vitro* activity or structure (Yeats and Rose, 2008). Molecular modelling is a feasible *in silico* approach that complements biological data and which can shed light on structural and functional properties of proteins. In the current structural proteomics era, new 3D structures and reliable molecular models are being generated on a daily basis, and 3D molecular models are becoming integral components of biological research (Hrmova and Fincher, 2009). To this end, a molecular model of the wheat TaPR61 protein was built using a lipid binding protein (LBP) from *A. thaliana* (PDB accession number 2rkn:A) as a template. The template was identified by a multitude of prediction tools, as described previously by Kovalchuk *et al.* (2009). This 77-residue long protein folds into a canonical 'all alpha protein' class that is categorized in the SCOP protein classification system (Andreeva *et al.*, 2007). Strictly speaking, 2rkn:A and similar proteins belong to the 'bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin' superfamily of proteins, and to the 'plant lipid-transfer and



To envisage if TaPR61 could enclose lipid molecules and thus function as a lipid carrier, two monoacylated phospholipids were modelled in the central cavity of basic TaPR61. These lipid molecules were also contained in the hydrophobic cavity of its template structure that is formed in the central region of the α -helical bundle. The volume of this cavity in TaPR61 is approximately 952 Å³, compared with the

volume of the cavity of 2rkn:A, which equals 1 069 Å³. The side chains of the highly hydrophobic Val, Leu, Ile, and Met and the hydrophobic Ala and Cys residues (Monera *et al.*, 1995) lined the cavity, in addition to a few charged Lys and Glu residues. It is of note that side-chains of cognate residues were rotated to the sides of the model cavity, whereas a large hollow space could be created that internalizes potentially several, but at least two lipid molecules. In the cavity of TaPR61, five residues Leu26, Met63, Met77, Val80, and Val81 are predicted to make hydrophobic interactions of less than 3.2 Å with the first

lipid molecule on one side of the cavity, while Ile30 and Lys34 are predicted to contact the second lipid molecule on the opposite side of the cavity, also at separations of less than 3.2 Å. Similar types of interactions have been noted in the 3D structure of LBP defence protein from *A. thaliana*, 2rkn:A (Lascombe *et al.*, 2008). The cavity enclosing lipids could be flexible and the volume of the cavity could be increased upon encasing a variety of lipids, including bulky sterols (Charvolin *et al.*, 1999; Yeats and Rose, 2008). In summary, molecular modelling suggested that TaPR61 could be involved in carrying lipids, as TaPR61 contains a large plastic cavity lined by hydrophobic residues. Hence, one putative function of TaPR61 could be in lipid transport and membrane metabolism, as TaPR61 could carry its lipid-like cargo to or through membranes during transport, metabolite processing or secretion. The other functional alternatives for TaPR61 include direct participation in antimicrobial defence signalling, cuticle formation, and cell wall re-modelling (Yeats and Rose, 2008). Nonetheless, all of the latter functions are also linked to the ability of TaPR61 to carry lipid-like molecules.

There is significant structural similarity between the TaPR60 (Kovalchuk *et al.*, 2009) and TaPR61 models (this work) that allows the two proteins to be classified in the same 'bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin' superfamily of proteins. This classification reflects close sequence identity and similarity scores between TaPR60 and TaPR61, which are 50% and 88%, respectively. Nevertheless, the differences in the primary structures of TaPR60 and TaPR61 allowed the selection of two similar, yet not identical, structural templates for molecular modelling, which were 2alg:A for TaPR60 (Pasquato *et al.*, 2006) and 2rkn:A for TaPR61 (Lascombe *et al.*, 2008), respectively. It is remarkable that the volumes of the cavities of 2alg:A and 2rkn:A were similar, nevertheless, the positions of lipid-like (2alg:A) and lipid molecules (2rkn:A) that were enclosed in the cavities differed. This observation indicates that both proteins could bind a variety of lipid molecules in many different positions and thus these proteins could rather be designated as non-specific lipid-binding (transfer) proteins (Charvolin *et al.*, 1999).

The gene and promoter of the *TaPR61* orthologue from durum wheat were cloned as described for *TdPR60* (Kovalchuk *et al.*, 2009). The activity of the *TdPR61* promoter was studied using promoter-GUS constructs in three agriculturally important species: wheat, barley, and rice. It was earlier demonstrated for several grain-specific promoters (Li *et al.*, 2008; Kovalchuk *et al.*, 2009, 2010) that activity of the same tissue-specific promoter can be significantly different in these three species. This was observed for the *TdPR61* promoter as well. Furthermore, although the *TdPR61* promoter like other members of this subfamily (aligned in Fig. 1C) is active in ETC, the GUS expression driven by this promoter was observed in a broader range of grain tissues than expression driven by the *TaPR60* and *OsPR602* promoters. Besides ETC, the activity of the *TdPR61* promoter was also detected in other parts of the endosperm, i.e. aleurone, starchy endosperm,

and ESR. The promoter was also active in the embryo in all three tested plant species; however, the spatial pattern of activity and the strength of GUS expression in the embryos were different.

Attempts to identify promoter fragments responsible for *TdPR61* gene activation using biolistic bombardment of developing embryos resulted in the identification of two promoter fragments important for promoter activation in the embryo. However, because of the relatively weak activation of the promoter through the potential *cis*-element(s) of the proximal fragment, further attempts to narrow down important sequences on the proximal fragment gave unconvincing and non-repeatable results (data not shown). The attempt to use the whole proximal D2–D3 fragment (Fig. 6A) as bait in the Y1H screen produced many false positives but revealed no TFs. A second round of the distal D0–D1 fragment mapping resulted in the identification of a shorter promoter fragment D0.5–D1 (Fig. 6B). Nevertheless, use of the whole D0.5–D1 fragment as bait in the Y1H screen did not identify any transcription factor. Three embryo- and endosperm-specific *cis*-elements were predicted in the D0.5–D1 fragment of the *TdPR61* promoter by PLACE software. They were used as baits to screen cDNA libraries prepared from the whole wheat grain at 0–6 DAP and from the liquid fraction of endosperm at 3–6 DAP. The screen of the endosperm library with PR61END as bait resulted in the isolation of cDNA which encoded a protein containing a DNA binding domain similar to DNA binding domains of BES1/BZR1 TFs from *Arabidopsis* (Yin *et al.*, 2002, 2005; He *et al.*, 2005). Both BES1 and BZR1 were demonstrated to function as transcription repressors in *Arabidopsis* (Mora-Garcia *et al.*, 2004; He *et al.*, 2005; Vert *et al.*, 2005) and as being involved in different aspects of brassinosteroid (BRs) regulation of growth processes, including vascular differentiation, flower and fruit development, root growth, seedling photomorphogenesis, and shade avoidance (Nemhauser *et al.*, 2003). The deduced protein, designated TaBES1S, was a short form of the wheat homologue of AtBES1, designated TaBES1. Expecting that the *TaBES1S* cDNA is a product of alternative splicing of *TaBES1*, the full-length Coding Domain Sequence (CDS) of the wheat *TaBES1* gene was isolated from the grain cDNA using primers from the beginning and the end of the short cDNA. Comparison of isolated *TaBES1* and *TaBES1S* cDNA sequences revealed a long deletion in the coding region of the *TaBES1S* sequence isolated in the Y1H screen. This deletion changed a reading frame leading to the formation of a premature stop codon and, consequently, to the production of a truncated protein containing a full-length DNA binding domain (Fig. 7C). However, all other conserved motifs of BES1/BZL1 proteins, including the C-terminal EAR motif (Kagale *et al.*, 2010), were not presented in the truncated TaBES1S protein. The genomic clone, which was isolated, using primers derived from the *TaBES1S* cDNA, represented the gene of a *TaBES1* homoeologue. This gene had 98% sequence identity at the nucleotide level, to *TaBES1* in the coding region, contained

no deletions, and encoded a protein with 98% amino acid sequence identity to TaBES1. No other homoeologues or homologues of *TaBES1* were isolated by PCR (data not shown). It was found that the deletion in *TaBES1S* cDNA was not a result of alternative splicing or premature polyadenylation inside an intron sequence, because a single intron of the homoeologue to the *TaBES1* gene is situated in a different position and is correctly spliced in both short and long cDNAs of *TaBES1* (data not shown). Because the presence of a short variant of cDNA in the cDNA pool from grain was confirmed (data not shown), it is clear that the short cDNA is not an aberrant product generated in the process of the cDNA library preparation. Yet, the origin of deletion in *TaBES1S* cDNA remains unclear, although, potentially, TaBES1S can be a product of the mutated allele. It was demonstrated in the Y1H assay that both full-length and truncated proteins could bind specific bait sequences equally well in the Y1H assay, suggesting that the DNA binding domain of the truncated protein is intact and functionally independent from the rest of the protein sequence (Fig. 7B). However, the protein–promoter interactions remained to be confirmed by (i) spatial analysis of *TaBES1* expression using plants transformed with the promoter–GUS fusion construct or, alternatively with *in situ* hybridization to show coincident sites of expression with the putative target gene of *TaBES1*; (ii) analysis of the *PR61* transcript levels in transgenic plants and/or mutants with up- or down-regulated *TaBES1*; (iii) demonstration of interaction between the TaBES1 and *PR61* promoters either *in vitro* or in a transient expression system. As the TaBES1S contains no EAR repression domain, it may potentially act as a tissue-specific competitor of the full-length form of the TdBES1 repressor in particular cells/tissues, and thus act as a molecule that decreases/abolishes formation of the repression protein complex and permits as yet unknown transcription factors to activate the *TdPR61* promoter. Unfortunately, at this stage of our studies these suggestions can be considered as preliminary speculations, which have to be confirmed in the future by more extensive research.

Acknowledgements

We thank Margaret Pallotta and Anita Lapina for technical assistance with BAC DNA isolation and promoter cloning, and Ursula Langridge, Robin Hosking, and Lorraine Carruthers for assistance with growing plants in the greenhouse. This work was supported by the Australian Research Council, the Grains Research and Development Corporation and the Government of South Australia.

References

- Allen RD, Bernier F, Lessard PA, Beachy RN. 1989. Nuclear factors interact with a soybean β -conglycinin enhancer. *The Plant Cell* **1**, 623–631.
- Andreeva A, Prlic A, Hubbard TJP, Murzin AG. 2007. SISYPHUS: structural alignments for proteins with non-trivial relationships. *Nucleic Acids Research* **35**, D253–D259.
- Balandin M, Royo J, Gomez E, Muniz L, Molina A, Hueros G. 2005. A protective role for the embryo surrounding region of the maize endosperm, as evidenced by the characterisation of *ZmESR-6*, a defensin gene specifically expressed in this region. *Plant Molecular Biology* **58**, 269–282.
- Becraft PW. 2001. Cell fate specification in the cereal endosperm. *Seminars in Cell and Developmental Biology* **12**, 387–394.
- Bendtsen JD, Nielsen H, von Heijne G, Brunak S. 2004. Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology* **340**, 783–795.
- Bernsel A, Viklund H, Hennerdal A, Elofsson A. 2009. TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acids Research* **37**, W465–W468.
- Bonello JF, Opsahl-Ferstad HG, Perez P, Dumas C, Rogowsky PM. 2000. *Esr* genes show different levels of expression in the same region of maize endosperm. *Gene* **246**, 219–227.
- Boutrot F, Meynard D, Guiderdoni E, Joudrier P, Gautier MF. 2007. The *Triticum aestivum* non-specific lipid transfer protein (*TaLtp*) gene family: comparative promoter activity of six *TaLtp* genes in transgenic rice. *Planta* **225**, 843–862.
- Brown RL, Kazan K, McGrath KC, Maclean DJ, Manners JM. 2003. A role for the GCC-box in jasmonate-mediated activation of the PDF1.2 gene of *Arabidopsis*. *Plant Physiology* **132**, 1020–1032.
- Cai G, Faleri C, Del Casino C, Hueros G, Thompson RD, Cresti M. 2002. Subcellular localisation of BETL-1, -2, and -4 in *Zea mays* L. endosperm. *Sexual Plant Reproduction* **15**, 85–98.
- Callebaut I, Labesse G, Durand P, Poupon A, Canard L, Chomilier J, Henrissat B, Mornon JP. 1997. Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. *Cellular and Molecular Life Sciences* **53**, 621–645.
- Charvolin D, Douliez JP, Marion D, Cohen-Addad C, Pebay-Peyroula E. 1999. The crystal structure of a wheat nonspecific lipid transfer protein (ns-LTP1) complexed with two molecules of phospholipid at 2.1 angstrom resolution. *European Journal of Biochemistry* **264**, 562–568.
- Curtis MD, Grossniklaus U. 2003. A gateway cloning vector set for high-throughput functional analysis of genes *in planta*. *Plant Physiology* **133**, 462–469.
- Davis RW, Smith JD, Cobb BG. 1990. A light and electron microscope investigation of the transfer cell region of maize caryopsis. *Canadian Journal of Botany* **68**, 471–479.
- Doan DNP, Linnestad C, Olsen OA. 1996. Isolation of molecular markers from the barley endosperm coenocyte and the surrounding nucellus cell layers. *Plant Molecular Biology* **31**, 877–886.
- Evers T, Millar S. 2002. Cereal grain structure and development: some implications for quality. *Journal of Cereal Science* **36**, 261–284.
- Finkelstein RR, Lynch TJ. 2000. The *Arabidopsis* abscisic acid response gene *ABI5* encodes a basic leucine zipper transcription factor. *The Plant Cell* **12**, 599–609.
- Gómez E, Royo J, Guo Y, Thompson R, Hueros G. 2002. Establishment of cereal endosperm expression domains: identification and properties of a maize transfer cell-specific transcription factor, ZmMRP-1. *The Plant Cell* **14**, 599–610.

- Gómez E, Royo J, Muñiz LM, Sellam O, Paul W, Gerentes D, Barrero C, López M, Perez P, Hueros G.** 2009. The maize transcription factor myb-related protein-1 is a key regulator of the differentiation of transfer cells. *The Plant Cell* **21**, 2022–2035.
- Gruis D, Guo HN, Selinger D, Tian Q, Olsen OA.** 2006. Surface position, not signaling from surrounding maternal tissues, specifies aleurone epidermal cell fate in maize. *Plant Physiology* **141**, 898–909.
- Guex N, Peitsch MC.** 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723.
- Gutierrez-Marcos JF, Costa LM, Biderre-Petit C, Khbaya B, O'Sullivan DM, Wormald M, Perez P, Dickinson HG.** 2004. *Maternally expressed gene1* is a novel maize endosperm transfer cell-specific gene with a maternal parent-of-origin pattern of expression. *The Plant Cell* **16**, 1288–1301.
- He JX, Gendron JM, Sun Y, Gampala SSL, Gendron N, Sun CQ, Wang ZY.** 2005. BZR1 is a transcriptional repressor with dual roles in brassinosteroid homeostasis and growth responses. *Science* **307**, 1634–1638.
- Higo K, Ugawa Y, Iwamoto M, Korenaga T.** 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research* **27**, 297–300.
- Hrmova M, Fincher GB.** 2009. Functional genomics and structural biology in the definition of gene function. In: Somers DJ, Langridge P, Gustafson JP, eds. *Plant genomics: methods and protocols*. Heidelberg: Springer, 199–227.
- Hueros C, Gomez E, Cheikh N, Edwards J, Weldon M, Salamini F, Thompson RD.** 1999. Identification of a promoter sequence from the BETL1 gene cluster able to confer transfer-cell-specific expression in transgenic maize. *Plant Physiology* **121**, 1143–1152.
- Hueros G, Varotto S, Salamini F, Thompson RD.** 1995. Molecular characterization of BET1, a gene expressed in the endosperm transfer cells of maize. *The Plant Cell* **7**, 747–757.
- Kagale S, Links MG, Rozwadowski K.** 2010. Genome-wide analysis of ethylene-responsive element binding factor-associated amphiphilic repression motif-containing transcriptional regulators in Arabidopsis. *Plant Physiology* **152**, 1109–1134.
- Kim SY, Chung HJ, Thomas TL.** 1997. Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the *Dc3* promoter using a modified yeast one-hybrid system. *The Plant Journal* **11**, 1237–1251.
- Kovalchuk N, Li M, Wittek F, et al.** 2010. Defensin promoters as potential tools for engineering disease resistance in cereal grains. *Plant Biotechnology Journal* **8**, 47–64.
- Kovalchuk N, Smith J, Pallotta M, et al.** 2009. Characterization of the wheat endosperm transfer cell-specific protein TaPR60. *Plant Molecular Biology* **71**, 81–98.
- Lascombe MB, Bakan B, Buhot N, Marion D, Blein JP, Larue V, Lamb C, Prange T.** 2008. The structure of 'defective in induced resistance' protein of *Arabidopsis thaliana*, DIR1, reveals a new type of lipid transfer protein. *Protein Science* **17**, 1522–1530.
- Laskowski RA, Macarthur MW, Moss DS, Thornton JM.** 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **26**, 283–291.
- Lessard PA, Allen RD, Bernier F, Crispino JD, Fujiwara T, Beachy RN.** 1991. Multiple nuclear factors interact with upstream sequences of differentially regulated beta-conglycinin genes. *Plant Molecular Biology* **16**, 397–413.
- Li M, Singh R, Bazanova N, Milligan AS, Shirley N, Langridge P, Lopato S.** 2008. Spatial and temporal expression of endosperm transfer cell-specific promoters in transgenic rice and barley. *Plant Biotechnology Journal* **6**, 465–476.
- Magnard JL, Le Deunff E, Domenech J, Rogowsky PM, Testillano PS, Rougier M, Risueno MC, Vergne P, Dumas C.** 2000. Genes normally expressed in the endosperm are expressed at early stages of microspore embryogenesis in maize. *Plant Molecular Biology* **44**, 559–574.
- Magnard JL, Lehouque G, Massonneau AE, Frangne N, Heckel T, Gutierrez-Marcos JF, Perez P, Dumas C, Rogowsky PM.** 2003. ZmEBE genes show a novel, continuous expression pattern in the central cell before fertilization and in specific domains of the resulting endosperm after fertilization. *Plant Molecular Biology* **53**, 821–836.
- Maitz M, Santandrea G, Zhang ZY, Lal S, Hannah LC, Salamini F, Thompson RD.** 2000. *rgf1*, a mutation reducing grain filling in maize through effects on basal endosperm and pedicel development. *The Plant Journal* **23**, 29–42.
- Matthews PR, Wang MB, Waterhouse PM, Thornton S, Fieg SJ, Gubler F, Jacobsen JV.** 2001. Marker gene elimination from transgenic barley, using co-transformation with adjacent 'twin T-DNAs' on a standard *Agrobacterium* transformation vector. *Molecular Breeding* **7**, 195–202.
- Milligan AS, Lopato S, Langridge P.** 2004. Functional genomics of seed development in cereals. In: Gupta PK, Varshney RK, eds. *Cereal genomics*. Heidelberg: Springer, 447–481.
- Monera OD, Sereda TJ, Zhou NE, Kay CM, Hodges RS.** 1995. Relationship of sidechain hydrophobicity and alpha-helical propensity on the stability of the single-stranded amphipathic alpha-helix. *Journal of Peptide Science* **1**, 319–329.
- Mora-Garcia S, Vert G, Yin YH, Cano-Delgado A, Cheong H, Chory J.** 2004. Nuclear protein phosphatases with Kelch-repeat domains modulate the response to brassinosteroids in Arabidopsis. *Genes and Development* **18**, 448–460.
- Morran S, Eini O, Pyvovarenko T, Parent B, Singh R, Ismagul A, Eliby S, Shirley N, Langridge P, Lopato S.** 2011. Improvement of stress tolerance of wheat and barley by modulation of expression of DREB/CBF factors. *Plant Biotechnology Journal* **9**, 230–249.
- Murashige T, Skoog F.** 1962. A revised medium for rapid growth and bioassay with tobacco tissue culture. *Physiologia Plantarum* **15**, 473–497.
- Nemhauser JL, Maloof JN, Chory J.** 2003. Building integrated models of plant growth and development. *Plant Physiology* **132**, 436–439.
- Neuberger T, Rolletschek H, Webb A, Borisjuk L.** 2010. Non-invasive mapping of lipids in plant tissue using magnetic resonance imaging. *Methods in Molecular Biology* **579**, 485–496.
- Olsen OA.** 2004. Nuclear endosperm development in cereals and *Arabidopsis thaliana*. *The Plant Cell* **16**, S214–S227.

- Opsahl-Ferstad HG, LeDeunff E, Dumas C, Rogowsky PM.** 1997. *ZmEsr*, novel endosperm-specific gene expressed in a restricted region around the maize embryo. *The Plant Journal* **12**, 235–246.
- Pasquato N, Berni R, Folli C, Folloni S, Cianci M, Pantano S, Helliwell JR, Zanotti G.** 2006. Crystal structure of peach Pru p 3, the prototypic member of the family of plant non-specific lipid transfer protein pan-allergens. *Journal of Molecular Biology* **356**, 684–694.
- Pyvovarenko T, Lopato S.** 2011. Isolation of plant transcription factors using a yeast one-hybrid system. *Methods in Molecular Biology* **754**, 45–66.
- Raju MVS, Walther A.** 1988. Heterogeneity and behavior of aleurone cells in the caryopsis of wild oats (*Avena fatua*). *Flora* **180**, 417–427.
- Ramachandran GN, Ramakrishnan C, Sasisekharan V.** 1963. Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* **7**, 95–99.
- Ritchie S, Swanson SJ, Gilroy S.** 2000. Physiology of the aleurone layer and starchy endosperm during grain development and early seedling growth: new insights from cell and molecular biology. *Seed Science Research* **10**, 193–212.
- Roy A, Kucukural A, Zhang Y.** 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* **5**, 725–738.
- Royo J, Gomez E, Barrero C, Muniz LM, Sanz Y, Hueros G.** 2009. Transcriptional activation of the maize endosperm transfer cell-specific gene *BETL1* by *ZmMRP-1* is enhanced by two C2H2 zinc finger-containing proteins. *Planta* **230**, 807–818.
- Sali A, Blundell TL.** 1993. Cooperative protein modeling by satisfaction of spatial restraints. *Journal of Molecular Biology* **234**, 779–815.
- Sambrook J, Russell D.** 2001. *Molecular cloning: a laboratory manual*, 3rd edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Sanchez R, Sali A.** 1998. Large-scale protein structure modeling of the *Saccharomyces cerevisiae* genome. *Proceedings of the National Academy of Sciences, USA* **95**, 13597–13602.
- Serna A, Maitz M, O'Connell T, et al.** 2001. Maize endosperm secretes a novel antifungal protein into adjacent maternal tissue. *The Plant Journal* **25**, 687–698.
- Schel JHN, Kieft H, van Lammeren AAM.** 1984. Interactions between embryo and endosperm during early developmental stages of maize caryopses (*Zea mays*). *Canadian Journal of Botany* **62**, 2842–2853.
- Sippl MJ.** 1993. Recognition of errors in 3-dimensional structures of proteins. *Proteins: Structure, Function and Genetics* **17**, 355–362.
- Sippl MJ, Wiederstein M.** 2008. A note on difficult structure alignment problems. *Bioinformatics* **24**, 426–427.
- So JS, Larkins BA.** 1991. Binding of an endosperm-specific nuclear protein to a maize beta-zein gene correlates with zein transcriptional activity. *Plant Molecular Biology* **17**, 309–319.
- Stalberg K, Ellerstrom M, Ezcurra I, Ablov S, Rask L.** 1996. Disruption of an overlapping E-box/ABRE motif abolished high transcription of the *napA* storage-protein promoter in transgenic *Brassica napus* seeds. *Planta* **199**, 515–519.
- Thompson RD, Hueros G, Becker HA, Maitz M.** 2001. Development and functions of seed transfer cells. *Plant Science* **160**, 775–783.
- Tingay S, McElroy D, Kalla R, Firs S, Wang MB, Thornton S, Brettell R.** 1997. *Agrobacterium tumefaciens*-mediated barley transformation. *The Plant Journal* **11**, 1369–1376.
- Vert G, Nemhauser JL, Geldner N, Hong FX, Chory J.** 2005. Molecular mechanisms of steroid hormone signaling in plants. *Annual Review of Plant Cell and Developmental Biology* **21**, 177–201.
- Wang HL, Offler CE, Patrick JW, Ugalde TD.** 1994. The cellular pathway of photosynthate transfer in the developing wheat grain. 1. Delineation of a potential transfer pathway using fluorescent dyes. *Plant, Cell and Environment* **17**, 257–266.
- Washida H, Wu CY, Suzuki A, Yamanouchi U, Akihama T, Harada K, Takaiwa F.** 1999. Identification of *cis*-regulatory elements required for endosperm expression of the rice storage protein glutelin gene GluB-1. *Plant Molecular Biology* **40**, 1–12.
- Wu CY, Washida H, Onodera Y, Harada K, Takaiwa F.** 2000. Quantitative nature of the Prolamin-box, ACGT and AACAA motifs in a rice glutelin gene promoter: minimal *cis*-element requirements for endosperm-specific gene expression. *The Plant Journal* **23**, 415–421.
- Yeats TH, Rose JKC.** 2008. The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). *Protein Science* **17**, 191–198.
- Yin YH, Vafeados D, Tao Y, Yoshida S, Asami T, Chory J.** 2005. A new class of transcription factors mediates brassinosteroid-regulated gene expression in *Arabidopsis*. *Cell* **120**, 249–259.
- Yin YH, Wang ZY, Mora-Garcia S, Li JM, Yoshida S, Asami T, Chory J.** 2002. BES1 accumulates in the nucleus in response to brassinosteroids to regulate gene expression and promote stem elongation. *Cell* **109**, 181–191.