ACCEPTED VERSION

John C. Harris, Pradeep Sornaraj, Mathew Taylor, Natalia Bazanova, Ute Baumann, Ben Lovell, Peter Langridge, Sergiy Lopato, Maria Hrmova

Molecular interaction of the γ-clade homeodomain-leucine zipper class I transcription factors during the wheat response to water deficit

Plant Molecular Biology, 2016; 90(4):435-452

© Springer Science+Business Media Dordrecht 2016

The final publication is available at Springer via http://dx.doi.org/10.1007/s11103-015-0427-6

PERMISSIONS

http://www.springer.com/gp/open-access/authors-rights/self-archiving-policy/2124

Springer is a green publisher, as we allow self-archiving, but most importantly we are fully transparent about your rights.

Publishing in a subscription-based journal

By signing the Copyright Transfer Statement you still retain substantial rights, such as self-archiving:

"Authors may self-archive the author's accepted manuscript of their articles on their own websites. Authors may also deposit this version of the article in any repository, provided it is only made publicly available 12 months after official publication or later. He/ she may not use the publisher's version (the final article), which is posted on SpringerLink and other Springer websites, for the purpose of self-archiving or deposit. Furthermore, the author may only post his/her version provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be provided by inserting the DOI number of the article in the following sentence: "The final publication is available at Springer via http://dx.doi.org/[insert DOI]"."

5 May, 2017

http://hdl.handle.net/2440/97970

Molecular interactions of the γ-Clade Homeodomain-Leucine Zipper Class I transcription factors during the wheat response to water deficit [S]

John C. Harris[1,2], Pradeep Sornaraj[1], Mathew Taylor, Natalia Bazanova, Ute Baumann, Ben Lovell, Peter Langridge, Sergiy Lopato, Maria Hrmova[3]

From the Australian Centre for Plant Functional Genomics, School of Agriculture, Food and Wine, University of Adelaide, Waite Campus, Glen Osmond, South Australia 5064, Australia

[S] This article contains Supporting Figures 1-4 and Supporting Tables 1-6.

[1] These authors contributed equally to this work and are listed in alphabetical order.

[2] Current address: South Australian Research and Development Institute, Waite Campus, Glen Osmond, South Australia 5064, Australia.

[3] To whom correspondence should be addressed: Tel.: +61 8 8313 7160; Fax: +61 8 8313 7102; E-mail: maria.hrmova@adelaide.edu.au.

[4] **Abbreviations**: 3D, three-dimensional; ABA, abscisic acid; CELD, endo-1-4-β-D-glucanase; ChIP, chromatin immuno-precipitation; DMSO, dimethyl sulfoxide; DOPE, discrete optimised protein energy; DPI-ELISA, DNA-protein interaction-enzyme linked immunosorbent assay; EMSA, electrophoretic mobility shift assay; GA, gibberellic acid; GFP, green fluorescent protein; HD-Zip I, homeodomain-leucine zipper class I; LEA, late embryogenesis abundant; LZ, leucine zipper; MOF, modeller objective function; Q-PCR, quantitative PCR; Ta, *Triticum aestivum*; TF, transcription factor; Y1H, yeast-1-hybrid; Y2H, yeast-2-hybrid.

Keywords: Abiotic stress; Homeodomain leucine zipper; Transcription factor networks; Homo- and hetero-dimerisation; DNA binding; Molecular modelling.

GenBank accession numbers: TaHDZipI-3-KT224374, TaHDZipI-4-KT224375, TaHDZipI-5-KT224376

Abstract

The γ-clade of class I homeodomain-leucine zipper (HD-Zip I) transcription factors (TFs) constitute members which play a role in adapting plant growth to conditions of water deficit. Given the importance of wheat (*Triticum aestivum* L.) as a global food crop and the impact of water deficit upon grain yield, we focused on functional aspects of wheat drought responsive HD-Zip I TFs. While the wheat γ-clade HD-Zip I TFs share significant sequence similarities with homologous genes from other plants, the clade-specific features in transcriptional response to abiotic stress were detected. We demonstrate that wheat *TaHDZipI-3*, *TaHDZipI-4*, and *TaHDZipI-5* genes respond differentially to a variety of abiotic stresses, and that proteins encoded by these genes exhibit pronounced differences in oligomerisation, strength of DNA binding, and trans-activation of an artificial promoter. Three-dimensional molecular modelling of the protein-DNA interface was conducted to address the ambiguity at the central nucleotide in the pseudo-palindromic *cis*-element CAATNATTG that is recognised by all three HD-Zip I proteins. The co-expression of these genes in the same plant tissues together with the ability of HD-Zip I TFs of the γ-clade to hetero-dimerise suggests a role in the regulatory mechanisms of HD-Zip I dependent transcription. Our findings highlight the complexity of TF networks involved in plant responses to water deficit. A better understanding of the molecular complexity at the protein level during crop responses to drought will enable adoption of efficient strategies for production of cereal plants with enhanced drought tolerance.

Introduction

Traditional breeding of wheat (*Triticum aestivum* L.) for increased yield under water limiting conditions has been hampered by the complexity of the polygenic plant response and the unpredictability of seasonal conditions. This has made it difficult to observe the hereditability of a selected trait. Given these difficulties, breeders have been selecting for yield improvement under water deficit by phenotyping methods, rather than by genetic association. Hence, an improvement in yield under water limited conditions is often seen under well-watered conditions and so selection has been for yield increase and not “drought tolerance” (Cattivelli et al. 2008). To increase yield specifically under water limited conditions, understanding mechanisms of the cooperative work of protein molecules during the plant drought response will aid efforts, and enable us to better determine gene targets for selection. One of the approaches to reveal the molecular mechanisms of the plant drought response is to study the role of transcription factors (TFs), which act as global regulators of gene expression (Hrmova and Lopato 2014).

It has been reported that members of the γ-clade HD-Zip I TFs can negatively regulate plant growth, suppress lateral root emergence, inhibit leaf and petal senescence, suppress gibberellic acid (GA)

biosynthesis and regulate abscisic acid (ABA) biosynthesis and signalling genes, making them good candidates for study of drought response mechanisms (Ariel et al. 2010; Dezar et al. 2005; Hjellstrom et al. 2003; Lü et al. 2014; Manavella et al. 2006; Olsson et al. 2004; Valdés et al. 2012; Zhang et al. 2012). Reports have shown that members of the HD-Zip I TFs from the γ-clade are regulated by abiotic stress and stress-related treatments such as water deficit, ABA, polyethylene glycols, sodium chloride, and low temperatures-(Ariel et al. 2010; Chang et al. 2014; Hu et al. 2012; Lee and Chun 1998; Zhang et al. 2012; Zhao et al. 2011). Several members of the γ-clade HD-Zip I TFs have also been shown to affect drought tolerance, when their expression levels were modulated using a transgenic approach. Expression of the sunflower γ-clade HD-Zip I TF *HaHB4*, in *Arabidopsis thaliana* resulted in transgenic plants with a delay in drought-induced senescence that correlated with suppression of ethylene response related genes (Cabello et al. 2007; Dezar et al. 2005; Manavella et al. 2006). Transgenic rice seedlings over-expressing *OsHOX22* showed increased susceptibility to terminal drought stress and sodium chloride, stronger responses to ABA treatment, and increased ABA content (Zhang et al. 2012). Supporting these observations, rice T-DNA insertion lines with reduced *OsHOX22* expression showed reductions in ABA content and sensitivity, and increased resistance to terminal drought and sodium chloride stresses (Zhang et al. 2012). Investigation of the mechanisms of function of γ-clade HD-Zip I TFs is crucial to our understanding of the role they play in the plant drought response and may lead to beneficial biotechnological applications in agriculture.

HD-Zip I TFs are proteins that contain DNA binding homeodomain (HD) and leucine zipper (LZ) motifs, which are highly conserved within the family (Mattsson et al. 1992; Ruberti et al. 1991; Schena and Davis 1992). It has been reported that dimerisation is a prerequisite for DNA binding within the HD-Zip I family and through the LZ two HDs are brought within close proximity to interact with the HD-Zip I *cis*-element (Palena et al. 1999; Sessa et al. 1993). These two domains contribute to the predominant mode of gene regulation by these TFs (reviewed in Harris et al. 2011). Defining the mechanism of how HD-Zip I TFs operate and the impact on expression of downstream targets, will greatly enhance our understanding of the role HD-Zip I TFs play in the molecular plant drought response network (Harris et al. 2011; Hrmova and Lopato 2014).

The current putative HD-Zip I *cis*-element sequences were initially determined by *in vitro* reiterations of protein binding to randomised oligonucleotides and PCR amplifications, which returned consensus sequences containing 9 bp pseudo-palindromes with variations at the central nucleotide, CAAT(A/T)ATTG or CAAT(C/G)ATTG (Sessa et al. 1993). Due to the palindromic nature of this *cis*-element and the proximity of two HDs in the dimer complex, it is believed that protein-DNA contact is a result of each HD interacting with distinct yet overlapping *cis*-elements (Tron et al. 2005). It has also been proposed that only one HD makes contact with the central nucleotide (Tron et al. 2005) and that the responsible Arg55 was conserved (Palena et al. 2001; Sessa et al. 1993). However, attempts by other

groups to confirm the binding properties of γ-clade homologues have presented difficulties. It was reported that AtHB7 and AtHB12 did not bind this *cis*-element, when assessed using electrophoretic mobility shift assays (EMSA, Johannesson et al. 2001) yet, *in vivo* transient assays revealed they could specifically activate a promoter with six repeats of CAATTATTG (Henriksson et al. 2005). In contrast, HaHB4 from sunflower was able to bind the assumed HD-Zip I *cis*-element *in vitro* in an EMSA (Palena et al. 1999). The only conclusive study to date confirming the *in planta* regulation of promoter activity through the HD-Zip I *cis*-element was demonstrated with MtHB1 from *Medicago truncatula* (Ariel et al. 2010). It was shown using a hairy root system and chromatin immuno-precipitation (ChIP)-PCR that the HD-Zip I specific *cis*-element (CAATAATTG) was necessary to maintain typical spatial gene expression of *LOB-Binding Domain 1*. In addition, it was confirmed by EMSA that MtHB1 could bind the same *cis*-element *in vitro* (Ariel et al. 2010).

Opposing the proposed conservation of the CAAT(A/T)ATTG HD-Zip I family *cis*-element, the *Arabidopsis* δ-clade HD-Zip I TF AtHB21 has been shown to activate the *COX5b-1* promoter requiring two repeats of ATCATT in the *cis*-element ATCATTGTATCATT (Comelli et al. 2012). The presented *cis*-element does bear resemblance to the assumed HD-Zip I *cis*-element CAAT(A/T)ATTG at two different points around the ATCATT repeats, <u>ATCATTG</u>T<u>ATCATT</u> (similarities underlined). AtHB7 and AtHB12 have also recently been shown to bind the promoter regions of the genes *PYL5* and *PYL8*, PP2C protein phosphatase genes, and *SnRK2* genes of the ABA signalling pathway (Valdés et al. 2012). This promoter interaction was determined by ChIP that looked selectively for enrichment of genes involved in ABA regulation and results were supported by further analysis of transgenic plants. The HD-Zip I responsive *cis*-element was found in the 1kb long promoter regions of only two of the 11 identified genes (Valdés et al. 2012). It has also been shown using EMSA and a yeast 1-hybrid (Y1H) assay that the rose RhHB1 protein can interact with a HD-Zip I like *cis*-element, AATATTATT, found in the upstream region of the gibberellin *20-oxidase 1* (*GA20ox1*) gene (Lü et al. 2014). Further, it was demonstrated that this element was necessary for suppression of *GA20ox1* expression by RhHB1 in *Arabidopsis* protoplasts (Lü et al. 2014). A truncated LeHB1 protein from the Tomato HD-Zip I β_1-clade was shown to bind in EMSA upstream regions of the *Lycopersicon esculentum ACC oxidase 1* gene at HD-Zip I like *cis*-elements, CAAT(A/T)ATTG and AATA(A)TATT (Lin et al. 2008). And finally, AtHB5 was demonstrated to suppress *BODENLOS* (*BDL*) expression through interaction with a 36 bp fragment of the *BDL* gene promoter lacking the putative pseudo-palindromic *cis*-element, yet harbouring an apparent half site CAATT (De Smet et al. 2013). Therefore, while it is considered that two different *cis*-elements exist for HD-Zip I TFs, the nature of the *cis*-elements and the conservation of the regulon within the HD-Zip I family remains somewhat ambiguous.

To fully understand the role of the HD-Zip I TF network in the plant molecular drought response, a comprehensive understanding of the dimerisation role within the family will need to be achieved.

Currently, little is known of the protein interaction profile of HD-Zip I TFs and the transcriptional regulation properties conferred to either homo- or hetero-dimers. Differences could arise in the strength of *cis*-element binding as demonstrated *in vitro* with OsHOX4 and OsHOX5 (of the ζ and β_1 clades respectively) which bind DNA more strongly as hetero-dimers than as homo-dimers (Meijer et al. 2000). It has also been seen within the basic leucine zipper family that hetero-dimerisation regulates the specificity of *cis*-element interaction (Marco Llorca et al. 2014), a property that may extend to the HD-Zip I family.

In the current work we show that the γ-clade HD-Zip I TFs of wheat (*Triticum aestivum,* Ta) *TaHDZipI-3*, *TaHDZipI-4* and *TaHDZipI-5* are differentially expressed under abiotic stress and show further differences in wheat cultivars that differ in their yield responses under water deficit. We also demonstrate differences within the γ-clade for preferential dimerisation partners, *cis*-element binding properties, and trans-activation function. Three-dimensional (3D) modelling of the protein-DNA interface has addressed ambiguity at the central nucleotide of the pseudo-palindromic *cis*-element. Our findings on γ-clade HD-Zip I TFs highlight the complexity of wheat TF networks under water deficit.

Materials and methods

Gene isolation

Nested PCR was used to isolate a wheat homologue of the rice γ-clade *OsHOX6* from wheat heat and drought stress cDNA libraries (Kovalchuk et al. 2009) using primers derived from an expressed sequence tag contig which was subsequently named *TaHDZipI-3*. As HD-Zip TFs are known to hetero-dimerise within each class, *TaHDZipI-3* was then used in yeast-2-hybrid screens (Supporting Fig. 1) to isolate dimerisation partners from the wheat cDNA library as described (Eini et al. 2013). More specifically, TaHDZipI-4 and TaHDZipI-5 were identified in the screen using the wheat cDNA library with a full-length sequence of TaHDZipI-3 as bait (Supporting Fig. 1a). Further, homo-dimerisation of HD-Zip I TFs in the presence of specific *cis*-elements was demonstrated in a yeast-1-hybrid assay (Supporting Fig. 1b). Through this approach partial cDNAs for two more γ-clade HD-Zip I TFs, designated *TaHDZipI-4* and *TaHDZipI-5* were isolated. These sequences were complemented with ESTs from NCBI databases to obtain full length cDNAs of *TaHDZipI-4* and *TaHDZipI*-5, which were then cloned by nested PCR.

In silico **sequence identification, polypeptide alignment, and phylogenetic analysis**

TBLASTN searches, using γ-clade HD-Zip I translated protein sequences from various species, were performed to identify full length genomic sequences of homologues from multiple genome database servers accessed from: arabidopsis.org, brachypodium.org, rice.plantbiology.msu.edu, phytozome.net,

maizegdb.org. Secondary TBLASTN searches were also performed using the Genbank non-redundant nucleotide and EST databases at NCBI to confirm correct annotation of the genomic sequences. Alignments were performed using MAFFT version 6 and the L-INS-i algorithm (Katoh et al. 2002). Phylogenetic reconstruction was performed with the Neighbour Joining algorithm and the boot strap method with 1000 replications in the MEGA5 program (Tamura et al. 2007). Tree images were exported to PowerPoint for further processing.

Plant material, growth, treatments, and quantitative PCR

Conditions used for plant treatment with ABA, low temperatures, and water deficit were described earlier by Kovalchuk et al. (2013). For the cyclic drought experiment three south Australian wheat cultivars (drought-tolerant Excalibur and RAC875, and the drought-sensitive Kukri) were grown in a controlled environment room under a 12 hour day light cycle (700 μmol m2s-1). Plants were grown in watertight bags containing 6 kg of dried Roseworthy soil/Waikerie sand (50/50%, w/w) containing basal nutrients as described by Izanloo et al. (2008). Temperatures were set at 16°C day/5°C night for the first four weeks, 17°C day/6°C night for the next four weeks, then 22°C day/10°C night for the remainder of the experiment. Relative humidity was maintained at 55-68% during the day and at approximately 80% during the night cycle. Plants were watered to field capacity daily with tap water (determined by weight) until the emergence of the first flag leaf (Zadok stage=37). Drought treatment was applied by gradually reducing the quantity of water added each day until the drought sensitive Kukri showed wilting. Plants were re-watered to field capacity (weight) and again left to dry without watering until rewatering again at the wilting point (Supporting Fig. 2). Plant water status was monitored by measuring leaf relative water content of the second leaf of all cultivars over the life of the experiment as described by Izanloo et al. (2008).

RNA preparation

For gene expression analysis RNA was extracted from leaves of plants treated with ABA, or subjected to cold or water deficit; cDNA was synthesised, and quantitative PCR (Q-PCR) performed as described by Kovalchuk et al. (2013). Flag leaf samples of plants exposed to cyclic drought were collected at five different stages of the drought stress from five biological replicates (indicated by x in Supporting Fig. 2), snap frozen in liquid nitrogen and stored at -80°C until processing. Samples were ground to a fine powder using an IKA A11 Basic Grinding Mill and total RNA was extracted from 500 mg of ground tissue using the standard TRIZOL method (Life technologies, Australia). 100 μg of total RNA from each sample was DNAse I treated and purified on QIAgen RNeasy Plant Mini Kit spin columns (Qiagen, Germany) following the manufacturer's instructions. RNA concentrations were quantified using a ND-1000

spectrophotometer (NanoDrop Technologies, USA) and the integrity of the RNA was determined on an Agilent BioAnalyzer 2010 (Agilent Technologies, USA). To observe changes in gene expression under cyclic drought, cDNA was synthesised in nuclease free PCR plates using SuperScript® III Reverse Transcriptase (Invitrogen, USA). Several reactions/sample were prepared to yield a total of more than 40 μg of cDNA per flag leaf sample. After synthesis, RNA was hydrolysed with sodium hydroxide and cDNA was subsequently purified with the MiniElute PCR Purification Kit (Qiagen, Germany), following the manufacturer's instructions. Q-PCR was performed as described by Kovalchuk et al. (2013) using the primers listed in Supporting Table 6.

Artificial promoter activation

Artificial promoter activation was analysed using a *Triticum monococcum* cell culture based transient assay using particle bombardment as described (32). The TOPO cloning and the Gateway Technology system (Invitrogen, Carlsbad, CA, USA) were used for generation of the artificial promoter:GUS and Ubi:TF plasmid constructs: *Ubi*:*TaHDZipI-3*, *Ubi*:*TaHDZipI-4* and *Ubi*:*TaHDZipI-5*. Sequences were shuttled from the pENTR-D-TOPO (Invitrogen, Carlsbad, CA, USA) entry vector to the final destination gateway vector pMDC164 or pUbi (Curtis and Grossniklaus 2003). For the artificial promoter:reporter construct forward and reverse oligonucleotides were designed to amplify the 251 nucleotide region upstream of the *TdCor410b* transcription start site, the forward primer also incorporated an artificial *cis*-element in tandem named HDZ2 (CAATAATTGCAATAATTG). This 251 nucleotide region of the *TdCor410b* promoter was previously determined as being the minimal allowable region which enabled basal transcription without stress activation (Eini et al. 2013).

Cloning and construction of pTaHDZipI-CELD and pTaHDZipI-BIOTIN plasmids and protein purification

Fusion plasmids pTaHDZipI-CELD and pTaHDZipI-BIOTIN were derived from the pTacDBP-CELD (Xue 2005) and pTacDBP-BIOTIN (Xue et al. 2006) fusion vectors. Three fusion pTaHDZipI-CELD plasmids were constructed by translational fusion of *TaHDZipI-3*, *TaHDZipI-4* or *TaHDZipI-5* coding sequences to the N-terminus of the 6xHis tagged endo-1-4-β-D-glucanase (CELD), which acts as a reporter by hydrolysing 4-nitrophenyl β-D-cellobioside. Three fusion pTaHDZipI-BIOTIN plasmids were constructed by translational fusion of the *TaHDZipI-3*, *TaHDZipI-4* or *TaHDZipI-5* coding sequences to the N-terminus of a synthetic biotinylation (BIOTIN) module of 72 residues, derived from the *Klebsiella pneumoniae* oxaloacetate decarboxylase. PCR amplifications of wheat HD-Zip I coding sequences were performed using primers containing *NheI* and *BamH1* adaptors in forward and reverse primers, respectively. The amplified DNA fragment was ligated in-frame to the N-terminus of either a CELD

protein or a biotinylation peptide. Protein expression and purification methods are found in the Supporting Information.

Protein-protein interaction assays

Assays were performed as described by Xue (2005) with the following modifications. Equimolar concentration of pairs of TaHDZipI-CELD and TaHDZipI-BIOTIN fusion proteins (15 μL) were mixed with a dissociation buffer (120 μL) containing 250 mM HEPES-NaOH pH 7.0, 1.25 M KCl, 2.5 mM EDTA, 1% (v/v) Triton X-100 and 2 mM DTT, and incubated at room temperature for 30 minutes with gentle shaking. Association buffer (1.2 mL) containing 10 mM HEPES-NaOH pH 7.0, 20 mM KCl, 5 mM $MgCl_2$ and 1 mM DTT was added to the mixture and incubated at room temperature for 60 min with gentle shaking. 10 μL of a pre-equilibrated *Strep*-Tactin resin (GE Healthcare, Sweden) was added to the mixture and incubated at ambient temperature for 1 hour with gentle shaking. Unbound proteins were removed by centrifugation (1000xg, 2 min, 4˚C). The supernatant was discarded and the pellet was washed thrice in a wash buffer containing 25 mM HEPES-NaOH pH 7.0, 120 mM KCl, 2 mM $MgCl_2$ and 0.01% (v/v) Nonidet P40. After centrifugation, the pellet was subjected to a CELD activity assay. The CELD activity of the TaHDZipI-CELD fusion proteins bound to either the biotin-fusion proteins or to biotinylated oligonucleotides was assayed by incubation in 200 μL of the CELD substrate solution (pH 5.5) containing 100 mM sodium acetate and 1 mM 4-nitrophenyl β-D-cellobioside at 37º C for 1 to 18 hours. The reaction was stopped by an equal volume of 2 M sodium carbonate and absorbance was measured at 405 nm.

DNA-binding assays using the CELD-fusion method

Biotinylated oligonucleotides (Geneworks, Australia) of known concentrations were incubated with pre-equilibrated *Strep*-Tactin resin and any unbound oligonucleotides were washed away with a *Strep*-Tactin wash buffer. Each TaHDZipI-CELD fusion protein (30 μL) was subjected to dissociation and association as described in the protein-protein interaction assay and incubated with each of *Strep*-Tactin bound biotinylated-oligonucleotides, HDZ1 3x(5'-CAATCATTG-3'), HDZ2 3x(5'-CAATAATTG-3'), HDZ-mutant 3x(5'-CAGTTACTG -3') or a GCC box 3x(5'-AGCCGCC-3'), for one hour at ambient temperature with gentle shaking. Unbound proteins were removed by centrifugation (1000xg, 2 min, 4˚C), supernatant was discarded and the pellet washed thrice in a wash buffer (25 mM HEPES-NaOH pH 7.0, 120 mM KCl, 2 mM $MgCl_2$ and 0.01% (v/v) Nonidet P40). After washing, the pellet was subjected to CELD activity assays as described above.

Construction of three-dimensional (3D) models of HDs of TaHDZipI-3 in complex with HDZ1, HDZ2 and HDZ1mutant cis-elements

To evaluate DNA binding at a molecular level, 3D molecular models of dimeric HDs of TaHDZipI-3 in complex with HDZ1 (5'-CAATCATTGC-3'/3'-GCAATGATTG-5'), HDZ2 (5'-CAATAATTG-3'/3-CAATTATTG-5') and HDZmutant (5'-CAGTTACTG-3'/3'-CAATTATTG-5') *cis*-elements were constructed. The HD sequence of TaHDZipI-3 (Fig. 7) was analysed with the PHYRE2 fold recognition server (Kelley and Sternberg 2009) to determine a suitable template, which was that of even-skipped HDs from *Drosophila melanogaster* in complex with the AT rich *cis*-element 5'-TAATTGAATT-3'/3'-AATTCAATTA-5' (Hirsch and Aggarwal 1995; PDB accession 1JGG). Alignments of the TaHDZipI-3 HD with 1JGG:A were carried out using ProMals3D (Pei et al. 2008), T-coffee (Notredame et al. 2000), MUSCLE (Edgar 2004) and Clustal W 2.1 (Larkin et al. 2007) algorithms. The alignment using ProMals3D returned the most accurate alignment (Fig. 7a) with a sequence identity and similarity of 30% and 75%, respectively (EMBOSS Needle global alignment; Rice et al. 2000). The aligned template and target sequences were used as input parameters to generate 3D models of the TaHDZipI-3 HD in complex with HDZ1, HDZ2, and HDZmutant DNA using Modeller 9v8 (Sali and Blundell 1993). As DNA was bound in 1JGG through water-mediated contacts (Hirsch and Aggarwal 1995), these water molecules were retained in the template structure. Nucleotide variations required in the DNA template from 1JGG were introduced using Coot (Emsley and Cowtan 2004) and the resultant double-stranded DNA molecules were subjected to energy minimisation using Yasara (Krieger et al. 2002). The final models were evaluated using Discrete Optimised Protein Energy (DOPE; Shen and Sali 2006), Modeller Objective Function (MOF; Sali and Blundell 1993), Ramachandran statistics and G-factor (Laskowski et al. 1993) and ProSa 2003 z-score (Sippl 1993) parameters that are summarised in Supporting Table 4.

Results

Confirmation of γ-clade origins

Phylogenetic analysis of protein sequences was performed as the first step towards assessing the relationship of *TaHDZipI-3*, *TaHDZipI-4*, and *TaHDZipI-5* TFs within the HD-Zip I family. The *Arabidopsis* HD-Zip I protein sequences were included as dicot representatives and monocot sequences were obtained from previous works (Ruberti et al. 1991, Mattsson et al. 1992, Sippl 1993) or from genome sequence databases. Our analysis confirmed that the HD-Zip I γ-clade holds true across monocots and dicots (Fig. 1a). Although, the relationship of *TaHDZipI-3* to *TaHDZipI-4* and *TaHDZipI-5* appeared distant, the evolutionary distance was comparable with that of the *Arabidopsis* γ-clade. This suggested that the monocot γ-clade has diversified during evolution and is composed of two different sub-clades that

we propose naming as γ1-monocot and γ2-monocot, the remaining dicot sequences are considered as γ-dicot (Fig. 1a).

However, intron/exon structures are known to be conserved within HD-Zip I clades and would further support a common origin. To investigate this, the genomic sequences of putative γ-clade HD-Zip I sequences of other monocots and dicots were analysed. Confirming this conservation of gene structure, a single intron is found in γ-clade coding sequences in the nucleotide region encoding the region between the 5th and 6th Leu residues of the LZ motif, which is conserved across all analysed genomes (Fig. 1b and *cf.* Supporting Table 1). These results confirmed that isolated wheat HD-Zip I sequences are members of the HD-Zip I γ-clade and indicated that they have arisen from a common ancestor.

Abiotic stress responses of *TaHDZipI-3, TaHDZipI-4 and TaHDZipI-5* transcript levels

Expression of the HD-Zip I γ-clade transcripts is known to be induced by abiotic stresses and ABA in many species, which suggests a role for these genes in the plant drought response (Agalou et al. 2008; Ariel et al. 2010; Gago et al. 2002; Lee and Chun 1998; Liu et al. 2013; Soderman et al. 1996). To confirm that the expression of the three isolated wheat γ-clade HD-Zip I genes is responsive to abiotic stresses and ABA, the drought tolerant wheat cultivar RAC875 was subjected to ABA treatment, cold and water deficit (Fig. 2) and gene expression analysis was performed by Q-PCR. Fig. 2a shows that *TaHDZipI-4* and *TaHDZipI-5* were strongly induced in leaves 4 hours after exposure to 200 µM ABA, whereas *TaHDZipI-3* expression levels remained unchanged. Exposure to 4°C for 48 hours led to induction of *TaHDZipI-4* and *TaHDZipI-5* which displayed diurnal patterns with expression peaking during the evening (Fig. 2b). Samples were taken one hour after the light period and one hour before the dark period cycles commenced and so expression shows correlation with the known diurnal patterns of ABA accumulation (Fujita et al. 2011). When making comparisons of relative levels in the ABA and cold experiments, *TaHDZipI-5* was induced much more strongly by cold treatment than *TaHDZipI-4*. By contrast, *TaHDZipI-3* showed a decrease in transcript levels during cold conditions which recovered during the return to 18°C. Pot grown RAC875 subjected to a prolonged water deficit showed increasing levels of expression of all three HD-Zip I genes as soil water potential decreased from day 8 to day 25 (Fig. 2c). The dramatic reduction of expression seen at day 31 in response to re-watering, also confirmed a water deficit response in expression of HD-Zip I genes.

Varietal differences exist in the expression characteristics of wheat γ-clade HD-Zip I TFs under cyclic drought

The three wheat cultivars Excalibur, Kukri, and RAC875 are suited to the South Australian Mediterranean climate and differ in yield responses under drought (Fleury et al. 2010; Izanloo et al. 2008). However, all

three show distinct phenotypes and with regards to yield, RAC875 and Excalibur are considered drought tolerant whereas Kukri is deemed to be drought sensitive (Izanloo et al. 2008). Many physiological responses to water deficit were monitored previously and Supporting Table 2 summarises the relative response of each cultivar (Izanloo et al. 2008). Regarding these three cultivars as potential sources of variation to observe differences in plant molecular drought responses, expression of the wheat γ-clade HD-Zip I TF genes was analysed from an experiment designed to mimic the cyclic drought imposed by Izanloo *et al*. (2008). Fig. 3a shows that the transcriptional responses of *TaHDZipI-3*, *TaHDZipI-4*, and *TaHDZipI-5* under cyclic drought conditions were different in the three wheat cultivars.

While *TaHDZipI-3* expression was weakly induced in all three wheat cultivars when compared with the drought inducible expression of *TaHDZipI-4* and *TaHDZipI-5*, increased expression under water deficit was apparent compared with controls (Fig. 3a). Kukri displayed much higher *TaHDZipI-3* basal expression under well-watered conditions than RAC875 and Excalibur. Also, a stronger drought response was shown for *TaHDZipI-3* expression in Kukri than RAC875, and Excalibur (Fig. 3a).

By comparison with *TaHDZipI-3*, expression of both *TaHDZipI-4* and *TaHDZipI-5* was dramatically up regulated in all three cultivars, whereas *TaHDZipI-5* was most strongly expressed. Comparison of the absolute levels of the wheat HD-Zip I γ-clade genes in Excalibur showed that the genes were expressed at much lower levels than in Kukri or RAC875.

As an indication of the general drought response, the expression levels of a *dehydrin* gene (*TaWZY2*, Accession EU395844) and a *Late Embryogenesis Abundant (LEA)* gene (*TaWcor410*, Accession JN681186) were also investigated (Fig. 3b). This analysis suggested that the different responses shown by the γ-clade HD-Zip I TF genes were not solely accountable for observed cultivar differences but were part of the general, yet distinct, drought response of each cultivar.

Regulation of synthetic promoters by wheat HD-Zip I γ-clade TFs

HD-Zip I proteins are reported to act as TFs that interact with variants of the CAATNATTG *cis*-element, although as discussed, the evidence for a universal *cis*-element is not strong. To determine if TaHDZipI-3, TaHDZipI-4 and TaHDZipI-5 can activate transcription *in vivo*, using an artificial (synthetic) promoter with the CAAT(A/T)ATTG *cis*-element twice in tandem, transient expression assays in *T. monococcum* cell cultures were performed (Fig. 4). TaHDZipI-3 showed no activation of the artificial promoter (0.49-fold) whereas TaHDZipI-4 and TaHDZipI-5 showed significant activation (18.5-fold ($P<0.001$) and 8.7-fold ($P<0.02$), respectively (Fig. 4).

Wheat γ-clade HD-Zip I TFs exhibit similar DNA-binding specificities

Due to the inability of TaHDZipI-3 to activate transcription from an artificial promoter containing an HD-Zip I *cis*-element, EMSA (electrophoretic mobility shift assays, procedure found in Supporting Information) were performed in the presence of three artificial *cis*-acting elements: HDZ1 3x(5'-CAAT(C/G)ATTG-3'), HDZ2 3x(5'-CAAT(A/T)ATTG-3'), HDZmutant 3x(5'-CAGT(T/A)ACTG -3') and a GCC box 3x(5'-AGCCGCC-3'). Binding preference of the recombinant TaHDZipI-3(107)-GFP protein (Supporting Fig. 3) established that TaHDZipI-3(107) did effectively bind both HDZ1 and HDZ2 *cis*-elements, as indicated by a shift in the positions of biotinylated DNA (Supporting Fig. 3a). Additionally, DPI-ELISA (DNA-Protein-Interaction ELISA, procedure found in Supporting Information) was undertaken and an increase in absorbance, as a result of catalytic activity, confirmed the DNA binding properties of TaHDZipI-3(107)-GFP (Supporting Fig. 4b). Alternatively, TaHDZipI-3(107)-GFP showed weak binding to the HDZ-mutant and the GCC box DNA *cis*-elements (Supporting Fig. 4).

Reporter-based DNA-protein pull-down assays (34, 35) were used to further specify the DNA-protein interaction activity of the three wheat HD-Zip I γ-clade proteins (Fig. 5). The C-termini of full-length wheat HD-Zip I TFs were fused with a 6xHis-tagged CELD domain, which acts as a reporter molecule by hydrolysing 4-nitrophenyl β-D-cellobioside. Biotin-tagged *cis*-elements were immobilised on a Streptavidin matrix and used to pull-down the recombinant CELD-fused proteins (Fig. 5). Strength of DNA interaction was measured by quantifying the hydrolysis of 4-nitrophenyl β-D-cellobioside.

The interaction assay demonstrated in quantitative terms that all three tested HD-Zip I proteins interacted with the HDZ1 and HDZ2 *cis*-elements to a much greater degree than the HDZmutant and GCC box elements (Figs. 5a and 5b). However, TaHDZipI-3 displayed a higher binding affinity for HDZ1 and HDZ2 *cis*-elements compared to TaHDZipI-4 and TaHDZipI-5. The presence of a significantly higher CELD activity against 4-nitrophenyl β-D-cellobioside in the reaction containing either HDZ1 or HDZ2 compared to the control reactions, confirmed that the interactions between wheat γ-clade HD-Zip I proteins and the HDZ1 and HDZ2 elements were highly specific.

Dimerisation specificity of the wheat HD-Zip I γ-clade TFs

It is known that dimerisation can be a distinguishing factor contributing to TF specificity in signalling networks. To demonstrate the dimerisation specificities within the wheat HD-Zip I γ-clade proteins, a reporter-based subunit exchange pull-down assay was used, similar to that reported for the DNA-protein interaction assay (34, 35). These assays allowed an exchange of subunits to be investigated, where HDZipI-CELD could interact with a subunit of HDZipI-biotin after the latter is immobilised on a Streptavidin matrix (Fig. 6). The subunit exchange assay demonstrated that the interactions between HDZipI-CELD and HDZipI-biotin fusions were specific, compared to the control reactions, but that their binding affinities varied as follows. TaHDZipI-3-biotin displayed a 2.5 and 2.1 higher affinity for TaHDZipI-4-CELD and TaHDZipI-5-CELD, respectively, than for TaHDZipI-3-CELD. This suggested

that TaHDZipI-3 prefers hetero-dimerising with either TaHDZipI-4 or TaHDZipI-5 over homo-dimerisation. On the other hand, TaHDZipI-4-biotin displayed 1.5-fold higher affinity for TaHDZipI-5-CELD than for TaHDZipI-4-CELD, but had 2.8-fold higher affinity for TaHDZipI-3-CELD, which again suggested a higher propensity of TaHDZipI-4 for formation of the hetero-dimeric complex over the homo-dimeric one. Finally, TaHDZipI-5-biotin displayed a reduced affinity to TaHDZipI-4-CELD compared to approximately equal propensity to interact with TaHDZipI-3-CELD or TaHDZipI-5-CELD. These data suggested that in contrast to TaHDZipI-4, TaHDZipI-3 formed hetero-dimers, while TaHDZipI-5 had an equal propensity towards forming homo- and hetero-dimers specifically with TaHDZipI-3.

Molecular models of HDs of TaHDZipI-3 in complex with HDZ1, HDZ2 and HDZ1mut *cis*-elements indicate selective binding

DNA binding assays performed with the wheat γ-clade HD-Zip I TFs revealed efficient DNA interaction through both putative HD-Zip I *cis*-elements. To further investigate the molecular mechanism behind the promiscuous interaction, 3D dimeric models of the TaHDZipI-3 HD were constructed in complex with pseudo-palindromic *cis*-elements. The high sequence similarity of the *Drosophila melanogaster* even-skipped HD made the template an ideal candidate to construct these models, as each *cis*-element was expected to have similar dispositions of amino acid residues, when interacting with TFs (Fig. 7a). 3D models were comprised of two HDs of TaHDZipI-3 in complex with three *cis*-elements, HDZ1 (CAAT(C/G)ATTG), HDZ2 (CAAT(A/T)ATTG), and HDZmutant (CAGT(A/T)ACTG) (HDZ1 and HDZmutant in Fig. 7b). Rigorous evaluations of stereo-chemical and energy parameters indicated that the models were reliable (Supporting Table 4). 3D models suggested that the same structural elements as those in the even-skipped HDs from *D. melanogaster* (Hirsch and Aggarwal 1995) were required to form α-helical folds of the TaHDZipI-3 HD (Fig. 7b). In each model, three α-helices were connected by loops with an extended N-terminal loop and one of the α-helices, the recognition α-helix was positioned at the major groove of the DNA helix. These three α-helices formed the core of the HD structure and several inward facing residues in the α-helices carried aromatic side chains, such as Phe30, Phe42, Trp70 and Phe71 (Fig. 7b), similar to even-skipped HDs (Clarke et al. 1994; Hirsch and Aggarwal 1995).

To investigate variation in the binding potential of the pseudo-palindromic *cis*-elements and the ambiguous central nucleotide, TaHDZipI-3 dimeric models were constructed in complex with three *cis*-elements (HDZ1, HDZ2, and HDZmutant). In each of these complexes, nucleotides of the *cis*-elements were not bound through equal interactions with the corresponding recognition α-helices of either chain (Supporting Table 5). This has also been seen in many HD-DNA complexes where binding interactions are not equal even though the sequences of HD and DNA are the same (Hirsch and Aggarwal 1995). Our comparative analyses of the TaHDZipI-3 dimeric HD-HDZ1 and HD-HDZ2 complexes showed that the

binding interactions were comparable and that the ambiguous central nucleotide in the DNA did not affect DNA binding to the HD folds (Fig. 7b). This further supports *in-vitro* evidence presented here as we have demonstrated experimentally that no preference for interaction with either *cis*-element was shown through EMSA and DPI-ELISA (Supporting Fig. 4, Fig. 5). These interactions were formed through the polar amino acid residues of each recognition α-helix of the homo-dimer, yet models suggested differences in interaction of each subunit, specifically through Arg26, Arg29, Arg65 and Asn73 of the α-helix of the first HD, and Arg26, Lys28, Gln72 and Asn73 of the α-helix of the second HD (Supporting Table 5). This observation suggested that *cis*-element binding depended on the orientation relative to HD positions.

As expected, changing the central nucleotide in the HDZ1 and HDZ2 *cis*-elements did not interfere with the binding pattern of the TaHDZipI-3 HD, however, manipulation of key bases (underlined in HDZmutant, CA<u>G</u>T(A/T)A<u>C</u>TG) altered the contact landscape suggesting that HDs could not bind the mutant *cis*-element efficiently, as we have also shown experimentally (Supporting Fig. 4, Fig. 5). The HDZmutant *cis*-element exhibited a different pattern of calculated polar contacts, inferring that these changes modified binding of both HDs to this DNA (Supporting Table 5). The HD-HDZmutant complex also exhibited some anomalous interactions, where specifically Arg26 engaged with the cytosine and guanine nucleotides in chain A and B, respectively. These observations indicated that DNA structural changes are likely to be responsible for the less efficient binding compared with that achieved through HDZ1 and HDZ2.

The presence of water-mediated contacts appeared to be essential for DNA binding in all three dimeric HD-DNA complexes and in the template structure, where Gln66 (and the corresponding Thr44 residues in the *Arabidopsis* template) bound DNA through water-mediated contacts. Additionally, the Gln66 residue formed a contact with the phospho-diester backbone of DNA adding to its suggested significance in stabilising HD-DNA interactions. A comparison of all interactions between the residues of the recognition α-helix and the three *cis*-elements clearly showed the variation in binding between residues and DNA (HDZ1, HDZ2 HDZmutant), as well as in separations between individual residues and DNA (Supporting Table 5).

Examination of electrostatic potentials of modelled HDs revealed positively charged areas around the recognition α-helix (Fig. 7c), likely critical for correct initiation and mediation of shape- and charge-based binding of DNA. Alternatively, neutral surface areas were present at the opposing ends of homo-dimers (Fig. 7c). Residue conservation within HDs examined by ConSurf (Celniker et al. 2013) showed that residues positioned inside the core of each homo-dimer exhibited the highest degree of conservation (Fig. 7d). These residues were essential for maintaining the structural integrity of HDs, such as formation of a salt bridge (Glu38 and Lys74), and hydrophobic and van der Waals forces.

Discussion

While efforts to reveal the molecular basis of the drought response in Arabidopsis have proven fruitful, concerted effort is required to define the molecular response to drought in crop species. Concentrating studies on TFs will help reveal the complex signalling and transcriptional network involved. To begin defining the role of TFs in the molecular drought response several key points of discovery need to be made such as: regulation at the level of transcription, post-transcription and post-translation; *cis*-elements and the promoters of target genes; co-expression of interacting protein partners at the tissue level; and the role in activation or repression of downstream genes. Most of these characteristics remain undefined for the HD-Zip I family.

Identification of wheat γ-clade HD-Zip I TFs

Drought inducible HD-Zip class I TFs have been identified in a variety of species including Arabidopsis, *Medicago truncatula*, poplar, maize, rice, and sunflower from investigations of entire genomic complements or selection for stress inducible genes (Agalou et al. 2008; Ariel et al. 2010; Gago et al. 2002; Hu et al. 2012; Lee and Chun 1998; Zhao et al. 2011). In particular, it is supported by various studies that the γ-clade is strongly regulated by abiotic stresses such as water deficit, salinity, and ABA treatments (52, 59). Yet, few reports have been made on mechanisms of function of these TFs. The aim of the current work has been to isolate drought inducible γ-clade HD-Zip I TF homologues from wheat and to make initial comparisons of function with other γ-clade HD-Zip I TFs. Our investigation suggested that wheat γ-clade HD-Zip I TFs share sequence similarities with other known genes that belong to the γ-clade, however, there are pronounced differences in the response of wheat γ-clade HD-Zip I TFs to abiotic stress related stimuli and their role in transcription.

The identity of wheat γ-clade HD-Zip I TFs was confirmed by phylogenetic analysis through their relatedness to homologous TFs from other species. In addition to amino acid sequence similarity, genes encoding wheat γ-clade HD-Zip I TFs shared conserved exon/intron patterns specific to the clade. The phylogenetic analysis and observation of the conserved motifs of the wheat HD-Zip I γ-clade shows that *TaHDZipI-4* and *TaHDZipI-5* are, however, more diversified in sequence from *TaHDZipI-3*. This diversification is conserved in the maize, rice, sorghum, and *Brachypodium* genomes of the *Poaceae* lineage. While the γ-clade holds true across monocot and dicot genomes, gene duplication has also contributed to the creation of paralogous gene pairs (Agalou et al. 2008; Ariel et al. 2007; Hu et al. 2012; Mukherjee et al. 2009; Zhao et al. 2011). However, no genetic evidence has been presented that indicates the γ1-monocot group members *OsHOX6* or *ZmHDZ9* and *ZmHDZ12* are duplicates of the γ2-monocot subclade gene pairs *OsHOX22*/*OsHOX24* and *ZmHDZ6*/*ZmHDZ24*, respectively. The phylogenetic

analysis performed here also suggested that *TaHDZipI-4* and *TaHDZipI-5* are a paralogous gene pair that has arisen through a duplication event in wheat. *TaHDZipI-3* may have arisen from a duplication event that has occurred much earlier in monocots that has not yet been identified.

Expression of wheat γ-clade TFs under abiotic stresses

It has been previously demonstrated that expression of γ-clade HD-Zip I TFs is induced by ABA, water deficit, and salinity treatments. In response to ABA, *TaHDZipI-3* showed no significant up-regulation in leaves, in contrast to γ-clade HD-Zip I TFs of the dicots, whereas *TaHDZipI-4* and *TaHDZipI-5* showed typical responses (Ariel et al. 2010; Chang et al. 2014; Hu et al. 2012; Lee and Chun 1998; Zhang et al. 2012; Zhao et al. 2011). The cold induction perceived by *TaHDZipI-4* and *TaHDZipI-5* contrasted with the lack of response to cold seen in *HaHB4* and *AtHB7* expression (Gago et al. 2002; Soderman et al. 1996), whereas *TaHDZipI-3* expression decreased during the prolonged cold treatment. In response to a prolonged water deficit, expression of all three wheat γ-clade HD-Zip I TFs increased, although *TaHDZipI-3* expression increased mildly by comparison to the two other wheat HD-Zip I TFs. Similarly in rice, the *TaHDZipI-3* homologue *OsHOX6* was reported to show slight induction in a drought sensitive cultivar but showed no change in expression in a drought tolerant cultivar (Agalou et al. 2008). Conversely, the *TaHDZipI-4* and *TaHDZipI-5* homologues in maize, *ZmHDZ6* and *ZmHDZ12*, showed slight up-regulation of expression in response to water deficit, whereas the *TaHDZipI-3* homologues *ZmHDZ4* and *ZmHDZ9* were strongly up-regulated under the same conditions (Zhao et al. 2011).

Expression of wheat γ-clade TFs in cultivars with contrasting yield responses under cyclic drought

The differences observed in the magnitude of water deficit-induced gene expression showed that, differences were seen not only in the physiological responses and growth habits of the examined wheat cultivars, but they were also observed at the molecular level. However, further analyses showed that the responses of the γ-clade HD-Zip I TFs, particular to each cultivar under water deficit, were not specific to these genes but were also displayed by *TaWZY2* and *TaWcor410* gene induction, indicative of a general stress response. This suggested that at the molecular level, the transcriptional response of the wheat γ-clade HD-Zip I genes is downstream of stress perception and signalling, and could be a part of the general, yet distinct, drought response mechanism displayed by each cultivar. Therefore, the contribution that the wheat HD-Zip I γ-clade is making cannot be conclusively attributed to the marked differences in physiological responses displayed by any of the three cultivars.

There are a number of factors to be considered in the light of cultivar differences in their molecular mechanisms of drought responses. These need to be regarded in discrete phases of the typical crop life cycle: (i) determining pre-stress factors that contribute to stress tolerance and support the development of

the theoretical ideotype, *e.g.* leaf thickness and waxiness, tiller number, and root structure; (ii) *de novo* tissue specific response mechanisms and timing during the season, *i.e.* high ABA levels in the inflorescence are known to reduce spikelet number and fertility, while in the leaf they induce protective measures; (iii) the effect of pre-response and *de novo* mechanisms on the total carbon economy with regards to final yield. It is likely that the cultivar differences observed are due to: (1) stress status during the imposed water deficit, *i.e.* affected by pre-adaptation physiology and then adaptation and, (2) primary signalling involving ABA, *i.e.* strength and timing of the response. Given the emerging story of γ-clade HD-Zip I TFs and their role in transcriptional responses involving ethylene and ABA signalling (Chang et al. 2014; Lü et al. 2014; Manavella et al. 2006; Valdés et al. 2012), they make enticing subjects for extending our understanding of these two key phytohormones in crop physiological and molecular responses to drought. To this end an investigation into the dynamics of TF function of the three wheat HD-Zip I γ-clade TFs was undertaken.

The wheat HD-Zip I γ-clade interacts with the putative HD-Zip I *cis*-element

There is now accumulating evidence indicating that HD-Zip I TFs bind both variants of the pseudo palindromic *cis*-element often with similar efficacy as presented here. However this evidence is still largely *in-vitro* and *in-vivo* studies have both supported the putative *cis*-element and presented new targets. Our attempts to investigate *cis*-element protein interactions by *in vitro* methods have shown that each of the three wheat γ-clade HD-Zip I TFs bind both *cis*-element types and further *in-silico* analyses presented a model that suggested the central nucleotide plays little involvement in this interaction. Ultimately, discerning the downstream targets of these TFs will reveal *cis*-element targets and structural data will reveal the nature of the interaction.

TaHDZipI-3 acts as a negative regulator of transcription

The transient expression assays presented here, in the *T. monococcum* cell culture, showed that TaHDZipI-4 and TaHDZipI-5 could activate transcription from an artificial promoter bearing an HD-Zip I *cis*-element, whereas TaHDZipI-3 did not activate transcription. This presented the possibility that TaHDZipI-3 could act as a negative regulator of transcription. The γ-clade *MtHB1* was shown to negatively regulate lateral root emergence through suppressing the activity of the *MtLBD1* promoter, while the β_2-clade *AtHB5* and *AtHB6* were shown to suppress *BDL* expression in protoplasts (Ariel et al. 2010; De Smet et al. 2013). The HD-Zip II family is known to include members that act as transcriptional repressors by virtue of their EAR-like (ERF-associated Amphiphilic Repression) motifs in the N-terminal domain (Ciarbelli et al. 2008). The γ-clade TFs are smaller in molecular size than other HD-Zip class I TFs and lack an acidic domain common to many HD-Zip I TFs in the N-terminal domain (Gago et al.

2002). However, it has been determined that the γ-clade AtHB12 has an acidic C-terminal region that is necessary for transactivation (Lee et al. 2001). The C-termini of the wheat γ-clade TFs have acidic regions that may constitute acidic activation domains and tryptophan/phenylalanine abundant regions reminiscent of AHA motifs (Capella et al. 2014), which consist of Trp and Phe residues surrounded by hydrophobic residues embedded within acidic regions (Doring et al. 2000; Kotak et al. 2004). Also, TaHDZipI-3 does not possess an apparent repression motif, such as domain I of the Aux/IAA proteins LXLXLX, that is associated with suppression of transcription found in many of the HD-Zip II TFs (Ciarbelli et al. 2008; Tiwari et al. 2004). This analysis presented the possibility that repression could occur by either of two mechanisms: (i) TaHDZipI-3 is actively inhibiting transcription through interaction with the transcription complex after dimerisation and DNA binding or, (ii) TaHDZipI-3, may be unable to bind DNA as a homo-dimer, sequester other HD-Zip I TFs through hetero-dimerisation, consequently inhibiting their binding to DNA. To investigate DNA binding by TaHDZipI-3, EMSA was initially performed which confirmed binding to either of two HDZ *cis* elements and this was further supported by DPI-ELISA. This suggested that the mechanism of suppression was active inhibition of transcriptional machinery.

Dimerisation within the HD-Zip I γ-clade

In this work we have shown that the γ-clade HD-Zip I TFs are able to interact in CELD assays suggesting a role for hetero-dimerisation as a regulatory mechanism of their *trans*-activation function. TaHDZipI-3 interaction was stronger with TaHDZipI-4 and TaHDZipI-5 than with itself as a homo-dimer and, likewise, TaHDZipI-4 interacted more strongly with TaHDZipI-3 than with itself or TaHDZipI-5. However, TaHDZipI-5 displayed an equally strong response with TaHDZipI-3 and itself as a homo-dimer and much more weakly with TaHDZipI-4. The impact that dimerisation has on expression levels of target genes of the γ-clade has yet to be investigated but, as with the bZip family (Marco Llorca et al. 2014), it is likely to play a crucial role in the specificity of target gene interaction and expression as discussed below.

Regulation of target genes by the HD-Zip I γ-clade

As monomers HD-Zip proteins have a very weak affinity for DNA interaction implying that they need high concentrations to participate in transcription as dimers (Palena et al. 1999). Further to this, the analysis presented here of transcript expression levels, *cis*-element interaction, transcriptional activation, and dimerisation has led us to propose a model whereby levels of TaHDZipI-3 protein act as a regulator of γ-clade function (Fig. 8). Under non-stressed conditions γ-clade expression is relatively low in leaves with TaHDZipI-3 and TaHDZip-5 being the most abundant transcripts (Figs. 3 and 4). The low expression levels and strong affinity of TaHDZipI-3 for dimerisation with TaHDZipI-4 and TaHDZip-5

over homo-dimerisation (Fig. 6) would result in minimal activation of target gene promoters due to the proposed role of TaHDZipI-3 in transcription (Fig. 5). Under conditions of abiotic stress where TaHDZipI-4 and TaHDZip-5 transcript expression is elevated, the balance would be tipped towards dimers with the potential to increase activation of downstream target genes. However, highlighting the complexity of TF networks, the impact that the abundance of each monomer will contribute is further complicated by the stronger affinity of TaHDZipI-3 homo-dimers for DNA over TaHDZipI-4 and TaHDZipI-5 homo-dimers and the weaker promoter activation displayed by TaHDZipI-5 when compared with TaHDZipI-4. Ultimately, identifying the downstream target genes of the γ-clade will enable validation of this model of gene regulation.

There is a plethora of reports describing that the manipulation of TF genes has potential for improving drought tolerance in plants. However, the knowledge gained in this area needs to be further applied to field crop species to provide evidence for the efficacy of this approach. The success of this approach in wheat relies on the extensive characterisation of TF regulation at the translational and post-translational level, identification of downstream targets of TFs, and the role of TFs and their interacting partners in regulation of transcription of target genes. A better understanding of the molecular drought response and pre-drought mechanisms will enable us to engineer more efficient drought response strategies. This would increase the potential of allele selection and gene modulation in crop plants to enable the tailoring of beneficial physiological responses to improve yield potential under water limited conditions.

Acknowledgements: We thank to Dr Gang Ping Xue (University of Queensland, Australia) for providing plasmids and for suggestions of protein-protein interaction studies. This work was supported by the grants from the Australian Research Council (to MH and SL), the Grains Research & Development Corporation and the Government of South Australia. JH was supported by the Australian Postgraduate Award scholarship from the University of Adelaide.

Author contributions statement: JCH and PS have designed and performed most of the experiments. MT and MH performed molecular modelling. NB, UB, BL and PL assisted with experiments. MH and SL guided research. JCH, SL and MH wrote the manuscript.

Compliance with ethical standards: Conflict of interest Authors declare that they have no conflict of interest.

Figure legends

Fig. 1 Sequence conservation of HD-Zip I TFs.

(**a**) An unrooted radial phylogenetic tree displaying evolutionary relationships of HD-Zip I TFs from *Arabidopsis* and selected monocots. Two subclades within the monocots are proposed and named γ_1-monocot and γ_2-monocot, while the remainder of sequences are named γ-dicot. Full-length protein sequences were aligned using MAFFT algorithm, and phylogenetic reconstruction was performed with the Neighbour Joining algorithm and 1000 boot strap replications in MEGA5. Bootstrap support values are indicated at the nodes and an estimated five residue substitutions per 1000 are represented by a scale bar. Two letter gene prefixes identify species of origin: At, *Arabidopsis thaliana*; Bd, *Brachypodium distachyon*; Os, *Oryza sativum*; Sb, *Sorghum bicolor*; Ta, *Triticum aestivum*; Td, *Triticum durum*; Zm, *Zea mays*. GenBank accession numbers of each entry are included. (**b**) Diagram of the genomic structure of γ-clade HD-Zip I TF coding sequences showing the position of the conserved intron. Intron and exon regions are represented by a line and boxes respectively. Encoding sequences are represented by white, dark grey and pale grey boxes which encode the N- and C-termini, HD, and interrupted LZ, respectively.

Fig. 2 Expression levels of the γ-clade *HD-Zip I* genes in the wheat cultivar RAC875 subjected to ABA treatment and environmental stresses.

(**a**) Response to 200 μM ABA dissolved in dimethyl sulfoxide (DMSO) or equal amount of DMSO (negative control). (**b**) Response to cold treatment; temperature is plotted on secondary vertical axis. (**c**), Response to increasing water deficit; water potential is plotted on secondary vertical axis. White, pale grey, and dark grey bars represent *TaHDZipI-3*, *TaHDZipI-4*, and *TaHDZipI-5* expression, respectively. Error bars represent standard deviation of biological triplicates

Fig. 3 Expression of wheat γ-clade HD-Zip I TFs, LEA, and dehydrin genes in the flagleaf of three wheat cultivars under cyclic drought.

QPCR analysis in RAC875, Kukri, and Excalibur subjected to cyclic drought conditions showing expression of (A) wheat γ-clade HD-Zip I genes, and (B) wheat LEA gene *TaCOR410*, and dehydrin *TaWZY2* gene. Rewatering occurred when plants showed wilting as an indication of drought stress (Supporting Fig. 2). Error bars represent standard deviations of biological triplicates.

Fig. 4 Differential activation of an artificial promoter bearing an HD-Zip I *cis*-element by the wheat HD-Zip I γ-clade TFs.

An HDZ2 synthetic promoter-GUS construct was co-bombarded with each of three constructs for constitutive expression of an HD-Zip I TF. Activation levels of the HDZ2 artificial promoter by wheat γ-clade HD-Zip I TFs is shown as a number of GUS foci. The GUS construct bombarded with GFP was used as a negative control

Fig. 5 DNA-binding specificities of TaHDZipI-3, TaHDZipI-4 and TaHDZipI-5 proteins.

(**a**) Interaction analysis based on catalytic tagging of TaHDZipI-3, TaHDZipI-4 and TaHDZipI-5 with biotinylated HDZ1 3x(5'-CAATCATTG-3'/3'-CAATCATTG)-5'), HDZ2 3x(5'-CAATAATTG-3'/3'-CAATTATTG-5'), HDZmutant 3x(5'-CAGTTACTC-3'/3'-GAGTAACTG -5') and the GCC box (5'-AGCCGCC-3'/3'-GGCGGCT-5') *cis*-elements. Equimolar concentrations of proteins and *cis*-elements were used, and proteins without biotinylated *cis*-elements served as negative controls. Negative control values were subtracted from experimental values. (**b**) Absorbance values are means of triplicate assays with standard deviations

Fig. 6 Confirmation and quantification of homo- and hetero-dimerisation of TaHDZipI-3, TaHDZipI-4 and TaHDZipI-5 proteins using the CELD method.

(**a**) Oligomerisation patterns of TaHDZipI-3, TaHDZipI-4 and TaHDZipI-5 TFs were analysed by interaction analyses, based on catalytic tagging. Equimolar concentrations of proteins were used and proteins without biotinylated fusions served as negative controls. Negative control values were subtracted from experimental values. (**b**) Absorbance values are means of triplicate assays with standard deviations

Fig. 7 Molecular features of HDs of TaHDZipI-3 in homo-dimeric forms and with DNA *cis*-elements.

(**a**) A sequence alignment of HDs of TaHDZipI-3 (target), TaHDZipI-4 and TaHDZipI-5, and 1JGG:A (template). Identical residues in sequences are in grey boxes. Conservation is also specified above the alignment, where number 9 (brown) indicates absolutely conserved residues. (**b**) Ribbon representations show the disposition of secondary structural elements, where the residues (cpk colours) of recognition α-helices, facilitating water-mediated contacts with *cis*-elements, are shown in dotted lines. Ribbons of chains A and B are coloured in blue (TaHDZipI-3) and cyan (1JGG). Black arrows point to N-termini. The coding strands of HDZ1 (5'-CAATCATTGC/3'-GCAATGATTG) and HDZmutant (5'-CAGTTACTGC/3'-GCAGTAACTG; mutated bases in a forward strand are in red) are shown in cartoon/stick representations in atomic colours. Bases of HDZ1 (bold black) and HDZmutant (bold/red black) coding strands are labelled. (**c**) Surface morphology, coded by electrostatic potential, of

TaHDZipI-3 in complex with HDZ1, Blue and red colours indicate electropositive and electronegative areas contoured at ±3 kTe^{-1}. (**d**) HD of TaHDZipI-3 coded by degrees of conservation, analysed by ConSurf. Highly conserved residues are coloured in deep magenta, and the less conserved residues are in cyan to white

Fig. 8 A proposed model for γ-clade function in gene regulation.
TaHDZipI-3, TaHDZipI-4, and TaHDZipI-5 monomers are represented by dark grey circles, pale grey ovals and white ovals, respectively. The formation of hetero- or homo-dimers is dependent upon the levels of protein expression and the strength of dimerisation interaction. Under non-stressed conditions the prevailing transcript is TaHDZipI-3 which binds more strongly with the less expressed TaHDZipI-4 and TaHDZipI-5 than itself. Due to the promoter suppression character of TaHDZipI-3, expression of the genes regulated by the γ-clade network is decreased overall. Under conditions of abiotic stress elevated levels of TaHDZipI-4 and TaHDZipI-5 leads to an overall increase in expression of γ-clade regulated gene transcripts

References

Agalou A, Purwantomo S, Overnaes E, Johannesson H, Zhu X, Estiati A, de Kam R, Engstroem P, Slamet-Loedin I, Zhu Z, Wang M, Xiong L, Meijer A, Ouwerkerk, P (2008) A genome-wide survey of HD-Zip genes in rice and analysis of drought-responsive family members. Plant Mol Biol 66:87-103

Ariel F, Diet A, Verdenaud M, Gruber V, Frugier F, Chan R, Crespi M (2010) Environmental regulation of lateral root emergence in *Medicago truncatula r*equires the HD-Zip I transcription factor *HB1*. Plant Cell 22:2171-2183

Ariel F, Manavella P, Dezar C, Chan R (2007) The true story of the HD-Zip family. Trends Plant Sci 12:419-426

Cabello J, Dezar C, Manavella P, Chan R (2007) The intron of the *Arabidopsis thaliana COX5c* gene is able to improve the drought tolerance conferred by the sunflower *Hahb-4* transcription factor. Planta 226:1143-1154

Capella M, Ré DA, Arce AL, Chan RL (2014) Plant homeodomain-leucine zipper I transcription factors exhibit different functional AHA motifs that selectively interact with TBP or/and TFIIB. Plant Cell Rep 33:1-13

Cattivelli L, Rizza F, Badeck FW, Mazzucotelli E, Mastrangelo AM, Francia E, Mare C, Tondelli A, Stanca AM (2008) Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. Field Crop Res 105:1-14 doi:10.1016/j.fcr.2007.07.004

Celniker G, Nimrod G, Ashkenazy H, Glaser F, Martz E, Mayrose I, Pupko T, Ben-Tal N (2013) ConSurf: using evolutionary data to raise testable hypotheses about protein function. Isr J Chem 53:199-206 doi:10.1002/ijch.201200096

Chang X, Donnelly L, Sun D, Rao J, Reid MS, Jiang C-Z (2014) A petunia homeodomain-leucine zipper protein, PhHD-Zip, plays an important role in flower senescence. Plos One 9 doi:10.1371/journal.pone.0088320

Ciarbelli AR, Ciolfi A, Salvucci S, Ruzza V, Possenti M, Carabelli M, Fruscalzo A, Sessa G, Morelli G, Ruberti I (2008) The Arabidopsis Homeodomain-leucine Zipper II gene family: diversity and redundancy Plant Mol. Biol. 68:465-478 doi:10.1007/s11103-008-9383-8

Clarke ND, Kissinger CR, Desjarlais J, Gilliland GL, Pabo CO (1994) Structural studies of the ENGRAILED homeodomain. Prot Sci 3:1779-1787

Comelli RN, Welchen E, Kim HJ, Hong JC, Gonzalez DH (2012) Delta subclass HD-Zip proteins and a B-3 AP2/ERF transcription factor interact with promoter elements required for expression of the Arabidopsis *cytochrome c oxidase 5b-1* gene. Plant Mol Biol 80:157-167 doi:10.1007/s11103-012-9935-9

Curtis MD, Grossniklaus U (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol 133:462-469 doi:10.1104/pp.103.027979

De Smet I, Lau S, Ehrismann JS, Axiotis I, Kolb M, Kientz M, Weijers D, Juergens G (2013) Transcriptional repression of BODENLOS by HD-Zip transcription factor *HB5* in *Arabidopsis thaliana*. J Exp Bot 64:3009-3019 doi:10.1093/jxb/ert137

Dezar C, Gago G, Gonzalez D, Chan R (2005) *Hahb-4*, a sunflower homeobox-leucine zipper gene, is a developmental regulator and confers drought tolerance to *Arabidopsis thaliana* plants. Transgenic Res 14:429-440

Doring P, Treuter E, Kistner C, Lyck R, Chen A, Nover L (2000) The role of AHA motifs in the activator function of tomato heat stress transcription factors *HsfA1* and *HsfA2*. Plant Cell 12:265-278

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-1797 doi:10.1093/nar/gkh340

Eini O, Yang N, Pyvovarenko T, Pillman K, Bazanova N, Tikhomirov N, Eliby S, Shirley N, Sivasankar S, Tingey S, Langridge P, Hrmova M, Lopato S (2013) Complex regulation by Apetala2 domain-containing transcription factors revealed through analysis of the stress-responsive *TdCor410b* promoter from durum wheat. Plos One 8 doi:10.1371/journal.pone.0058713

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126-2132 doi:10.1107/s0907444904019158

Fleury D, Jefferies S, Kuchel H, Langridge P (2010) Genetic and genomic tools to improve drought tolerance in wheat. J Exp Bot 61:3211-3222 doi:10.1093/jxb/erq152

Fujita Y, Fujita M, Shinozaki K, Yamaguchi-Shinozaki K (2011) ABA-mediated transcriptional regulation in response to osmotic stress in plants. J Plant Res 124:509-525 doi:10.1007/s10265-011-0412-3

Gago G, Almoguera C, Jordano J, Gonzalez D, Chan R (2002) *Hahb-4*, a homeobox-leucine zipper gene potentially involved in abscisic acid-dependent responses to water stress in sunflower. Plant Cell Environ 25:633-640

Harris JC, Hrmova M, Lopato S, Langridge P (2011) Modulation of plant growth by HD-Zip class I and II transcription factors in response to environmental stimuli. New Phytol 190:823-837 doi:10.1111/j.1469-8137.2011.03733.x

Henriksson E, Olsson ASB, Johannesson H, Johansson H, Hanson J, Engstrom P, Soderman E (2005) Homeodomain leucine zipper class I genes in Arabidopsis. Expression patterns and phylogenetic relationships. Plant Physiol 139:509-518 doi:10.1104/pp.105.063461.

Hirsch JA, Aggarwal AK (1995) Structure of the Even-skipped homeodomain complexed to AT-rich DNA: new perspectives on homeodomain specificity. EMBO J 14:6280-6291

Hjellstrom M, Olsson A, Engstrom O, Soderman E (2003) Constitutive expression of the water deficit-inducible homeobox gene *ATHB7* in transgenic Arabidopsis causes a suppression of stem elongation growth. Plant Cell Environ 26:1127-1136

Hrmova M, Lopato S (2014) Enhancing abiotic stress tolerance in plants by modulating properties of stress responsive transcription factors. In: Tuberosa R, Graner A, Frison E (eds) Genomics of Plant Genetic Resources. Springer Netherlands, pp 291-316. doi:10.1007/978-94-007-7575-6_12

Hu R, Chi X, Chai G, Kong Y, He G, Wang X, Shi D, Zhang D, Zhou G (2012) Genome-wide identification, evolutionary expansion, and expression profile of homeodomain-leucine zipper gene family in poplar (*Populus trichocarpa*). Plos One 7:e31149

Izanloo A, Condon AG, Langridge P, Tester M, Schnurbusch T (2008) Different mechanisms of adaptation to cyclic water stress in two South Australian bread wheat cultivars. J Exp Bot 59:3327-3346 doi:10.1093/jxb/ern199

Johannesson H, Wang Y, Engstrom P (2001) DNA-binding and dimerization preferences of Arabidopsis homeodomain-leucine zipper transcription factors *in vitro*. Plant Mol Biol 45:63-73

Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059-3066 doi:10.1093/nar/gkf436

Kelley LA, Sternberg MJE (2009) Protein structure prediction on the web: a case study using the Phyre server. Nature Prot 4:363-371

Kotak S, Port M, Ganguli A, Bicker F, von Koskull-Doring P (2004) Characterization of C-terminal domains of Arabidopsis heat stress transcription factors (Hsfs) and identification of a new

signature combination of plant class A Hsfs with AHA and NES motifs essential for activator function and intracellular localization. Plant J 39:98-112 doi:10.1111/j.1365-313X.2004.02111.x

Kovalchuk N, Smith J, Pallotta M, Singh R, Ismagul A, Eliby S, Bazanova N, Milligan AS, Hrmova M, Langridge P, Lopato S. (2009) Characterization of the wheat endosperm transfer cell-specific protein TaPR60. Plant Mol Biol 71:81-98

Kovalchuk N, Jia W, Eini O, Morran S, Pyvovarenko T, Fletcher S, Bazanova N, Harris J, Beck-Oldach K, Shavrukov Y, Langridge P, Lopato S (2013) Optimization of *TaDREB3* gene expression in transgenic barley using cold-inducible promoters. Plant Biotech J 11:659-670 doi:10.1111/pbi.12056

Krieger E, Koraimann G, Vriend G (2002) Increasing the precision of comparative models with YASARA NOVA - a self-parameterizing force field. Proteins 47:393-402 doi:10.1002/prot.10104

Larkin MA et al. (2007) Clustal W and clustal X version 2.0. Bioinformatics 23:2947-2948 doi:10.1093/bioinformatics/btm404

Laskowski RA, Macarthur MW, Moss DS, Thornton JM (1993) PROCHECK - a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283-291 doi:10.1107/s0021889892009944

Lee Y, Chun J (1998) A new homeodomain-leucine zipper gene from *Arabidopsis thaliana* induced by water stress and abscisic acid treatment. Plant Mol Biol 37:377-384

Lee Y, Oh H, Cheon C, Hwang I, Kim Y, Chun J (2001) Structure and expression of the *Arabidopsis thaliana* homeobox gene *Athb-12*. Biochem Biophys Res Commun 284:133-141

Lin Z, Hong Y, Yin M, Li C, Zhang K, Grierson D (2008) A tomato HD-Zip homeobox protein, LeHB-1, plays an important role in floral organogenesis and ripening. Plant J 55:301-310 doi:10.1111/j.1365-313X.2008.03505.x

Liu W, Fu R, Li Q, Li J, Wang L, Ren Z (2013) Genome-wide identification and expression profile of homeodomain-leucine zipper class I gene family in *Cucumis sativus.* Gene 531:279-287

Lü P, Zhang C, Liu J, Liu X, Jiang G, Jiang X, Khan MA, Wang L, Hong B, Gao J (2014) *RhHB1* mediates the antagonism of gibberellins to ABA and ethylene during rose (*Rosa hybrida*) petal senescence. Plant J 78:578-590 doi:10.1111/tpj.12494

Manavella P, Arce A, Dezar C, Bitton F, Renou J, Crespi M, Chan R (2006) Cross-talk between ethylene and drought signalling pathways is mediated by the sunflower *Hahb-4* transcription factor. Plant J 48:125-137

Marco Llorca C, Potschin M, Zentgraf U (2014) bZIPs and WRKYs: two large transcription factor families executing two different functional strategies. Front Plant Sci 5 doi:10.3389/fpls.2014.00169

Mattsson J, Soderman E, Svenson M, Borkird C, Engstrom P (1992) A new homeobox-leucine zipper gene from *Arabidopsis thaliana.* Plant Mol Biol 18:1019-1022

Meijer AH, de Kam RJ, d'Erfurth I, Shen W, Hoge JHC (2000) HD-Zip proteins of families I and II from rice: interactions and functional properties. Mol Gen Genet 263:12-21

Mukherjee K, Brocchieri L, Burglin TR (2009) A comprehensive classification and evolutionary analysis of plant homeobox genes. Mol Biol and Evol 26:2775-2794 doi:10.1093/molbev/msp201

Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol 302:205-217 doi:10.1006/jmbi.2000.4042

Olsson A, Engstrom P, Soderman E (2004) The homeobox genes *ATHB12* and *ATHB7* encode potential regulators of growth in response to water deficit in Arabidopsis. Plant Mol Biol 55:663-677

Palena C, Gonzalez D, Chan R (1999) A monomer-dimer equilibrium modulates the interaction of the sunflower homeodomain leucine-zipper protein *Hahb-4* with DNA. Biochem J 341:81-87

Palena C, Tron A, Bertoncini C, Gonzalez D, Chan R (2001) Positively charged residues at the N-terminal arm of the homeodomain are required for efficient DNA binding by homeodomain-leucine zipper proteins. J Mol Biol 308:39-47

Pei J, Kim B-H, Grishin NV (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res 36:2295-2300 doi:10.1093/nar/gkn072

Rice P, Longden I, Bleasby A (2000) EMBOSS: The European molecular biology open software suite. Trends in Genetics 16:276-277 doi:10.1016/s0168-9525(00)02024-2

Ruberti I, Sessa G, Lucchetti S, Morelli G (1991) A novel class of plant-proteins containing a homeodomain with a closely linked leucine zipper motif. EMBO J 10:1787-1791

Sali A, Blundell TL (1993) Comparative protein modeling by satisfaction of spatial restraints. J Mol Biol 234:779-815 doi:10.1006/jmbi.1993.1626

Schena M, Davis RW (1992) HD-Zip proteins: members of an Arabidopsis homeodomain protein superfamily. Proc Natl Acad Sci U.S.A. 89:3894-3898

Sessa G, Morelli G, Ruberti I (1993) The *ATHB-1* and *-2* HD-Zip domains homodimerize forming complexes of different DNA-binding specificities. EMBO J 12:3507-3517

Sessa G, Morelli G, Ruberti I (1997) DNA-binding specificity of the homeodomain leucine zipper domain. J Mol Biol 274:303-309

Shen M-Y, Sali A (2006) Statistical potential for assessment and prediction of protein structures. Prot Sci 15:2507-2524 doi:10.1110/ps.062416606

Sippl MJ (1993) Recognition of errors in 3-dimensional structures of proteins. Proteins 17:355-362 doi:10.1002/prot.340170404

Soderman E, Mattsson J, Engstrom P (1996) The Arabidopsis homeobox gene *ATHB-7* is induced by water deficit and by abscisic acid. Plant J 10:375-381

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol and Evol 24:1596-1599 doi:10.1093/molbev/msm092

Tiwari SB, Hagen G, Guilfoyle TJ (2004) Aux/IAA proteins contain a potent transcriptional repression domain. Plant Cell 16:533-543 doi:10.1105/tpc.017384

Tron A, Comelli R, Gonzalez D (2005) Structure of homeodomain-leucine zipper/DNA complexes studied using hydroxyl radical cleavage of DNA and methylation interference. Biochemistry 44:16796-16803

Valdés AE, Övernäs E, Johansson H, Rada-Iglesias A, Engström P (2012) Homeodomain-leucine zipper (HD-Zip) class I transcription factors *ATHB7* and *ATHB12* modulate abscisic acid signalling by regulating protein phosphatase 2C and abscisic acid receptor gene activities. Plant Mol Biol doi:DOI 10.1007/s11103-012-9956-4

Xue GP (2005) A CELD-fusion method for rapid determination of the DNA-binding sequence specificity of novel plant DNA-binding proteins. Plant J 41:638-649 doi:10.1111/j.1365-313X.2004.02323.x

Xue GP, Bower NI, McIntyre CL, Riding GA, Kazan K, Shorter R (2006) *TaNAC69* from the NAC superfamily of transcription factors is up-regulated by abiotic stresses in wheat and recognises two consensus DNA-binding sequences. Funct Plant Biol 33:43-57 doi:10.1071/fp05161

Zhang S, Haider I, Kohlen W, Jiang L, Bouwmeester H, Meijer AH, Schluepmann H, Liu CM, Ouwerkerk PBF (2012) Function of the HD-Zip I gene *Oshox22* in ABA-mediated drought and salt tolerances in rice. Plant Mol Biol doi:10.1007/s11103-012-9967-1

Zhao Y, Zhou Y, Jiang H, Li X, Gan D, Peng X, Zhu S, Cheng B (2011) Systematic analysis of sequences and expression patterns of drought-responsive members of the HD-Zip gene family in maize. PLoS One 6:e28488 doi: 10.1371/journal.pone.0028488