Full Length Research Paper

Cloning, localization and phylogenetic analysis of barley putative APETALA 2/ethylene responsive element binding protein (AP2/EREBP) genes

Wei-Tao Li^{1,3}, Ya-Xi Liu¹, Guo-Yue Chen¹, Chun-Ji Liu², Shou-Fen Dai¹, Xiu-Jin Lan¹, Zhi-En Pu¹, You-Liang Zheng^{1,3}, Yu-Ming Wei^{1*}

1Triticeae Research Institute, Sichuan Agricultural University, 211 Huimin Road, Wenjiang, Chengdu, Sichuan, 611130, China.

2CSIRO Plant Industry, 306 Carmody Road, St Lucia, QLD 4067, Australia.

3Key Laboratory of Southwestern Crop Germplasm Utilization, Ministry of Agriculture, Sichuan Agricultural University, 211 Huimin Road, Wenjiang, Chengdu, Sichuan, 611130, China.

Accepted 12 March, 2012

Eight putative barley AP2/EREBP genes were cloned based on the barley HarvEST database. The introns in the translated regions of these genes were all located at the front of the conserved AP2 domain. Among the eight genes, DNA sequences of hv.18885 and hv.17070 were the same between Steptoe and Morex. Of the other six genes, there was very high ratio of 'G' and 'C' bases in the variant DNA sequence regions of three genes between Steptoe and Morex. Thus, only three genes were successfully located on the chromosomes. hv.8737 and hv.15732 were mapped to the short arm of chromosome 7H and hv.2871 to a major quantitative trait loci (QTL) region conferring adult spot blotch tolerance on the short arm of chromosomes 3H. Phylogenetic analysis showed that putative plant AP2/EREBP genes in the AP2 subfamily were completely separated from those in other subfamilies. The putative AP2/EREBP genes in RAV and Soloist subfamilies were clustered together, while those in the ERF and CRT/DRE subfamily were clustered together. Among the main branch-groups of each subfamily in plant, those from wheat, barley and rice had the tendency to cluster together when compared with those from Arabidopsis.

Key words: AP2/EREBP, clone, map, phylogenetic, barley.

INTRODUCTION

AP2/EREBP (APETALA 2/ethylene responsive element binding protein) genes are characterized by the presence of the AP2-DNA binding domain which is about 60 amino acids in length (Okamuro et al., 1997). They are involved in various aspects of plant growth and development such as flower development, hormone signal transduction and cellular differentiation (Wilson et al., 1996; Karim et al., 2009; Krizek, 2009). Meanwhile, they are known to be involved in responses to biotic, pathogenic and environmental stresses (Choi et al., 2002; Yi et al., 2004; Feng et al., 2005; Jung et al., 2007).

Among the more than 2000 transcription factors identified in the Arabidopsis thaliana genome (http:// planttfdb.cbi.edu.cn/), about nine percent belong to the AP2/EREBP family (Dietz et al., 2010). The AP2/EREBP genes are classified into five subfamilies including

^{*}Corresponding author. E-mail: ymwei@sicau.edu.cn. Tel: +86-28-86290909. Fax: +86-28-82650350.

Abbreviations: AP2/EREBP, APETALA 2/ethylene responsive element binding protein; ERF, ethylene response factors; CRT/DRE, C-repeat binding/dehydration-responsive elementbinding factors; RAV, ABI3/VP1; DH, doubled haploid; NJ, neighbour-joining; QTL, quantitative trait loci.

Unigene	Primer	Temperature (°C)	Material	NCBI Number	Length (bp)
hv.17070	F: CGATTGCGAGCTCTAGATACC	55	Morex	HQ647351	3324
	R: CAAAAGAAACACCAAACAGAGAAG		Steptoe	HQ647352	3324
hv.15734	F: CGGCGTGGTATCAGCGGC	60	Morex	HQ647357	2252
	R: ACAGGGTGCTTCCCAATCTCC		Steptoe	HQ647358	2264
hv.7703	F: TCTCGTGGGCTCAAGTCTTCTCT	55	Morex	HQ647355	948
	R: GCATTTGTGATAGTCGTGCTCG		Steptoe	HQ647356	951
hv.17532	F: GCGCACGCCACTTGTACAC	58	Morex	HQ647353	841
	R: CAAAACGAGGAAGGAAGGAGG		Steptoe	HQ647354	841
hv.8737	F: CTCGTGTTCTCWTCTCCATAGC	60	Morex	HQ647361	1165
	R: ACAATACAAACCAACCAATATCAG		Steptoe	HQ647362	1165
hv.18885	F: CCCTTCCTCCCTCCTCTTCC	58	Morex	HQ647359	1662
	R: AGCTTCAGGTTGCAGATAATTTG		Steptoe	HQ647360	1662
hv.2871	F: TCTTCCACACCGACGGCATCT	55	Morex	HQ647363	614
	R: TGCTCGGGAAAACAGACAC		Steptoe	HQ647364	614
hv.15732	F: GGCAACTGCTGCTGGTGAAT	58	Morex	HQ647365	822
	R: ACCCAGCAACCTCAATTCC		Steptoe	HQ647366	822

Table 1. PCR primers designed from conservative regions in the putative barley AP2/EREBP genes.

response factors (ERF), C-repeat binding/dehydrationresponsive element-binding factors (CRT/DRE), ABI3/VP1 (RAV), AP2 and Soloist (Sakuma et al., 2002; Gutterson and Reuber, 2004). The AP2/EREBP genes in species such as poplar, rice, barley and wheat have been studied by comparing them to those of A. thaliana (Nakano et al., 2006; Zhuang et al., 2008, 2010, 2011).

Barley (Hordeum vulgare L.), one of the major food and feed crops in the world, is more widely adaptable to unfavourable growth conditions such as cold, dry and saline environments (Jung et al., 2007). A few AP2/EREBP genes (HvRAF, HvCBF1, HvCBF2 and HvCBF3) have been characterized in barley (Choi et al., 2002; Xue, 2003; Xue and Loveridge, 2004; Jung et al., 2007). Clustering, amino acid sequences analysis, and expression assessment of barley putative AP2/EREBP genes had been done by Zhuang et al. (2011). However, chromosomal gene structure. localization, and phylogenetic analysis of barley AP2/EREBP genes were not reported. In this study, we identified the DNA sequences of eight putative barley AP2/EREBP genes, cloned and mapped several genes to chromosomes, and constructed a phylogenetic tree of putative AP2/EREBP

genes in Arabidopsis, rice, wheat and barley.

MATERIALS AND METHODS

Plant materials and DNA isolation

A doubled haploid (DH) population of barley, produced by a modified bulbosum method from a cross between Steptoe and Morex (Chen and Hayes, 1989), was used in this study. One hundred and fifty lines and their parents were grown in a greenhouse and genomic DNA was extracted from single plants following the CTAB procedure (Sambrook et al., 1989).

Cloning and sequencing the putative barley AP2/EREBP genes

To obtain genomic DNA sequences of the putative barley AP2/EREBP genes, PCR primers (Table 1) were designed using barley EST sequences from the HarvEST database and tested against the two parents of the mapping populations, Steptoe and Morex. The PCR reaction volume was 25 μ l, containing approximately 0.2 μ g template DNA, 2.5 units Taq DNA polymerase with high fidelity (Takara Bio, Inc., Kyoto, Japan), 0.3 μ M each primer (Table 1), 200 μ M each of dNTP (Takara Bio, Inc., Japan), 1.5 mM MgCl2, and 1x PCR buffer. PCR amplifications were carried out in a PTC-240 thermocycler (Genetic Technologies, MJ

Unigene	Primer	Temperature (°C)	Material	Length (bp)
hv.15732	GAATATGACTGTGCCTC <u>T</u> G <u>A</u>	58	Morex	0
	ACCCAGCAACCTCAATTCC		Steptoe	806
hv.2871	TCTTCCACACCGACGGCATCT	55	Morex	0
	CCATGCCTCGGCCAAAG <u>A</u> CA		Steptoe	409
hv.8737	CTCGTGTTCTCWTCTCCATAGC	58	Morex	0
	CCTACACCGCCGACCGTT <u>AA</u>		Steptoe	1070

Table 2. Specific primers designed from SNPs in the putative barley AP2/EREBP genes.

The bold letters with underline at the 3' end means extra mismatched nucleotides.

Research, USA). The PCR reaction conditions were 95°C for 5 min, followed by 35 cycles of 95°C for 1 min, 55 to 60°C for 60 s and 72°C for 1.5 min, with a final extension at 72°C for 7 min. Amplified products were directly visualized by electrophoresis using 1.5% agarose gels. The PCR products were purified, cloned and sequenced in both directions by BGI Life Tech Co., Ltd, Shanghai, China.

The intron/exon structure was determined by blasting cDNA sequence based on the EST sequences from barley HarvEST and NCBI databases and each gene DNA sequence was cloned with DNAman 5.2.2 (http://www.lynnon.com).

Linkage mapping of the putative barley AP2/EREBP genes

DNA sequences between the two barley cultivars Steptoe and Morex were compared by blast with DNAman 5.2.2 (http:// www.lynnon.com) (NCBI No: HQ647351-HQ647366, eight genes from two different cultivars). Based on differences of DNA sequences between Steptoe and Morex, PCR primers (Table 2) were designed to allow the detection of base substitutions between these two cultivars. The primers of five other genes were not successfully designed due to same DNA sequences between Steptoe and Morex or very high ratio of 'G' and 'C' bases in the regions with base substitutions or indels between the two cultivars. The 150 DH lines derived from a cross between these two cultivars were then genotyped with the markers developed, and their map locations were determined by aligning their genotypic profiles with those existing genotypic data of the mapping population (http://wheat.pw.usda.gov/ggpages/map_shortlist.html) (Kleinhofs et al., 1993; Rodrigues et al., 2006; Stein et al., 2007). Linkage maps were constructed using Map Manager Version 3.0 (Lincoln et al., 1993) and drawn with the winQTLCart 2.5 (Wang et al., 2007).

Phylogenetic analysis

Phylogenetic trees of putative AP2/EREBP genes had been constructed in some plant species, such as Arabidopsis, rice, wheat and barley (Nakano et al., 2006; Dietz et al., 2010; Zhuang et al., 2010, 2011). Based on these phylogenetic trees, one random gene from each main cluster in each main branch-group of the five subfamilies was used for multiple protein sequence alignments. Putative AP2/EREBP gene without conserved AP2 domain or absence in some branch-group of each subfamily in these plant species were not used in this study. The sequences analysis was carried out using DNAman 5.2.2 (http://www.lynnon.com) and results obtained were used for phylogenetic analysis. Phylogenetic

trees were constructed using the Neighbour-Joining (NJ) method based on MEGA 4 (Tamura et al., 2007). Secondary structures of predicted proteins were analysed using Lasergene Version 8.0 (Allex, 1999).

RESULTS

Putative AP2/EREBP genes cloned from barley

The unigene code of the HarvEST database of barley was used except for hv.8737 as an identical code for this gene did not exist in the NCBI database. A total of eight barley DNA sequences were obtained. Of these, six (hv.15734, hv.8737, hv.7703, hv.17532, hv.18885 and hv.17070) contained intact exons, and the remaining two (hv.2871 and hv.15732) were partial gene sequences (Figure 1). These eight genes belonged to either the DREB or the ERF subfamily. For the six genes containing intact exons, hv.8737, hv.18885 and hv.7703 possessed one intron in the translated region, while hv.17532 and hv.15734 contained no intron in the translated region. hv. 17070 and hv.15734 possessed two introns in the untranslated region.

The intron sizes of the putative barley AP2/EREBP genes ranged from 76 to 1078 bp. All the introns in the translated region were located in the front of the conserved AP2 domain. Of the eight putative barley AP2/EREBP gene sequences, differences between Morex and Steptoe were found for neither hv.17070 nor hv.18885. Differences in base substitutions or indels for the other six were identified in the translated or untranslated regions. For instance, there were four base substitutions and two indels in intron I of hv.15734 between Morex and Steptoe (HQ647357 and HQ647358). Among the six putative AP2/EREBP genes with base substitutions or indels between Morex and Steptoe, only two (hv.8737 and hv.15732) had different amino acid sequences. For hv.8737, the residue was 'G' at position293 for Morex but 'V' for Steptoe. For hv.15732, the residue was 'G' at position 97 for Morex but 'E' for Steptoe. However, the difference in amino acid

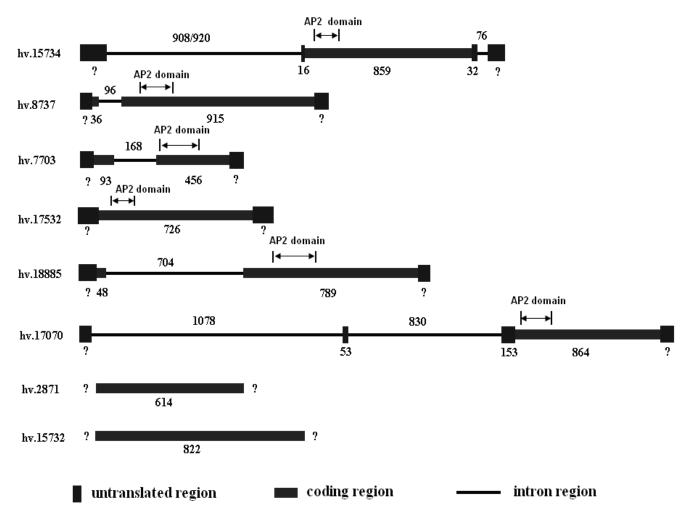


Figure 1. Exon/intron structures of eight putative barley *AP2/EREBP* genes. The lengths of the introns each gene was shown above the each gene-figure, while the lengths of untranslated and coding regions were shown under the each gene-figure. Fragment length of Steptoe is the front in two numbers parted by "/", while that of Morex is the latter. Rectangle: exon; line: intron; '?': unknown number of DNA base.

sequences between Steptoe and Morex did not result in the change of their secondary structures by analysis using Lasergene Version 8.0.

Mapping barley putative AP2/EREBP genes

Locus-specific primers were successfully designed for three of the putative barley AP2/EREBP genes based on the differences of DNA sequences between Steptoe and Morex (Table 2). Sequence differences for two (hv.18885 and hv.17070) of the other five genes were not present between these two varieties and those for the remainder (hv.15734, hv.7703 and hv.17532) contained very high ratio of 'G' and 'C' bases in the regions with base substitutions or indels. One of these genes for which primers were obtained, hv.2871, was mapped at the distal end of the short arm of chromosome 3H with a distance of 13.7 cm from the marker MWG571C. The other two genes, hv.8737 and hv.15732, were both mapped on the distal end of the short arm of chromosome 7H. The distance between these two genes was 11.1 cm, and, hv.8737 was more distally located (Figure 2).

Phylogenetic tree of putative AP2/EREBP genes in plant

Putative AP2/EREBP genes in plant could be clearly divided into five subfamilies, which were consistent with the denoted group related to APETALA2 (Figure 3). The putative AP2/EREBP genes in the AP2 subfamily were completely separated from others, whereas those in the

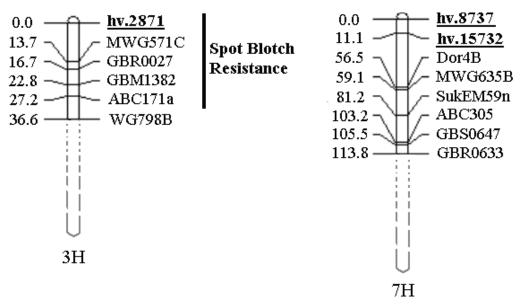


Figure 2. Map locations of the putative barley *AP2/EREBP* genes with MAPMAKER (LOD=3.0). The genotypes of each putative *AP2/EREBP* gene in this mapping population were compared with mapping data (SNP, SSR, RFLP, genes and SSAP). The map location of putative *AP2/EREBP* genes is underlined. The dark bar on chromosomes (3H) indicate quantitative trait locus (QTL) from barley cv Steptoe × barley cv Morex mapping population (Bilgic et al., 2005).

RAV and Soloist subfamilies were clustered together. Those in the ERF and CRT/DRE subfamilies were also clustered together. Among the main branch-groups of each subfamily in plant, those from wheat, barley and rice had the tendency to cluster together as compared to those from Arabidopsis. However, it was found that two putative AP2/EREBP genes (ta.50437 and ta.56957) in CRT/DRE subfamily from wheat were not grouped together with others in CRT/DRE subfamily.

DISCUSSION

Previous scientists analyzed and clustered the 147 AP2/EREBP genes in Arabidopsis, 164 in rice and 117 in wheat (Sakuma et al., 2002; Nakano et al., 2006; Riano et al., 2007; Zhuang et al., 2011), and only 53 in barley (Zhuang et al., 2011). Of these four species, genes belonging to the Soloist subfamily have not been detected in rice. Though the numbers of the AP2/EREBP genes differ among the four species, the percentage of genes in each of the subfamilies (excluding those for the Soloist subfamily) was similar (Zhuang et al., 2010). As the barley genome is approximately 40 times larger than that of Arabidopsis and 10 times larger than that of rice (Arumuganathan and Earle, 1991), it is thus expected that additional AP2/EREBP genes could exist in barley.

Though sizes and cleavage sites of introns of putative barley AP2/EREBP genes cloned were varietal, the locations of introns in regulated region were all on the front of conversed AP2 domain, not inside. The characteristic of intron was also found in those universal stress proteins in plant (Li et al., 2010). The reason may be to avoid disappearance of conserved AP2 domain when false cleavage of introns occurred.

The three putative barley AP2/EREBP genes mapped in this study were located on two different chromosome regions. On the basis of previous quantitative trait loci (QTL) mapping of the 'Morex × Steptoe' population, linkage analysis showed that only hv.2871 falls into a major QTL region conferring adult spot blotch tolerance on chromosome 3H (Bilgic et al., 2005). The mapping results could also supply a molecular basis for developing candidate genes related with biotic or abiotic tolerance and development of plant lines by combining QTL results obtained in the future. Additionally, the three genes could be used for Marker-Assisted Selection after identifying their biological function. On the basis of the phylogenetic tree obtained for Arabidopsis, wheat, rice and barley, the putative AP2/EREBP genes in monocots were clustered together in main branch-groups of each subfamily by comparing with those in dicots. It is inferred that there were species specificities in each subfamily of the putative AP2/EREBP genes, which also is proved by similar clustering tendency of each subfamilies of Arabidopsis, wheat, barley and rice.

On the basis of clustering results, the molecular evolution of the AP2/EREBP genes in AP2 subfamily might be independent, whereas those between RAV and Soloist subfamilies might be related with each other as

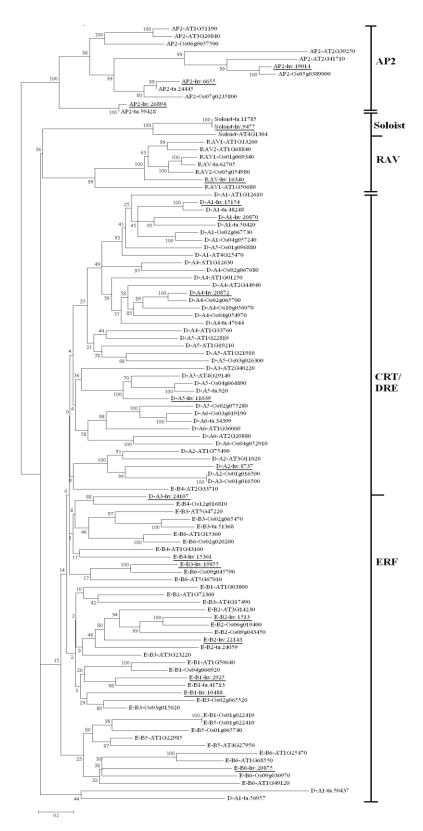


Figure 3. Phylogenetic tree constructed with the Neighbour-Joining method and Poisson correction of the *AP2/EREBP* in *Arabidopsis*, rice, wheat and barley. D, E, R, AP2 and Soloist represent CRT/DRE-, ERF-, RAV-, AP2- and Soloist-types, respectively. The barley *AP2/EREBP* genes are underlined.

well as those between ERF and CRT/DRE subfamilies. Additionally, ta.50437 and ta.56957 were separated from other genes of CRT/DRE subfamily in Arabidopsis, wheat, rice and barley, which may be caused by different genotype. It is inferred that there were some specific AP2/EREBP genes in CRT/DRE subfamily of hexaploid wheat.

REFERENCES

- Allex CF (1999). Computational Methods for Fast and Accurate DNA Fragment Assembly. Ph.D. thesis, Department of Computer Sci. University of Wisconsin-Madison.
- Bilgic H, Steffenson BJ, Hayes PM (2005). Comprehensive genetic analyses reveal differential expression of spot blotch resistance in four populations of barley. Theor. Appl. Genet. 111: 1238-1250.
- Chen FQ, Hayes PM (1989). A comparison of Hordeum bulbosum mediated haploid producion efficiency in barley using in vitro floret and tiller culture. Theor. Appl. Genet. 77: 701-704.
- Choi DW, Rodriguez EM, Close TJ (2002). Barley cbf3 gene identification, expression pattern, and map location. Plant Physiol. 129: 1781-1787.
- Dietz KJ, Vogel MO, Viehhauser A (2010). AP2/EREBP transcription factors are part of gene regulatory networks and integrate metabolic, hormonal and environmental signals in stress acclimation and retrograde signalling. Protoplasma, 245: 3-14.
- Feng JX, Liu D, Pan Y, Gong W, Ma LG, Luo JC, Deng XW, Zhu YX (2005). An annotation update via cDNA sequence analysis and comprehensive profiling of developmental, hormonal or environmental responsiveness of the Arabidopsis AP2/EREBP transcription factor gene family. Plant Mol. Biol. 59: 835-868.
- Gutterson N, Reuber TL (2004). Regulation of disease resistance pathways by AP2/ERF transcription factors. Curr. Opin. Plant, Biol. 7: 465-471.
- Jung J, Won SY, Suh SC, Kim HR, Wing R, Jeong Y, Hwang I, Kim M (2007). The barley ERF-type transcription factor HvRAF confers enhanced pathogen resistance and salt tolerance in Arabidopsis. Planta, 225: 575-588.
- Karim MR, Hirota A, Kwiatkowska D, Tasaka M, Aida M (2009). A role for Arabidopsis PUCHI in floral meristem identity and bract suppression. Plant Cell, 21: 1360-1372.
- Kleinhofs A, Kilian A, Saghai-Maroof MA, Biyashev RM, Hayes P, Chen FQ, Lapitan N, Fenwick A, Blake TK, Kanazin V, Ananiew E, Dahleen L, Kudrna D, Bollinger J, Knapp SJ, Lin B, Sorrells M, Henn M, Franckowiak JD, Hoffman D, Skadsen R, Steffenson BJ (1993). A molecular, isozyme and morphological map of the barley (Hordeum vulgare) genome. Theor. Appl. Genet. 86: 705-712.
- Krizek BA (2009). Aintegumenta and Aintegumentalike6 act redundantly to regulate Arabidopsis floral growth and patterning. Plant Physiol. 150: 1916-1929.
- Li WT, Wei YM, Wang JR, Liu CJ, Lan XJ, Jiang QT, Pu ZE, Zheng YL (2010). Identification, localization and characterization of putative USP genes in barley. Theor. Appl. Genet. 121: 907-917.
- Lincoln SE, Daly MJ, Lander ES (1993). A Whitehead Institute for Biomedical Res. Technical Report, Third Edition. Nakano T, Suzuki K, Fujimura T, Shinshi H (2006). Genome-wide
- Nakano T, Suzuki K, Fujimura T, Shinshi H (2006). Genome-wide analysis of the ERF gene family in Arabidopsis and rice. Plant Physiol. 140: 411-432.

- Okamuro JK, Caster B, Villarroel R, Van Montagu M, Jofuku KD (1997). The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in Arabidopsis. Proc. Natl. Acad. Sci. USA, 94: 7076-7081.
- Riano PDM, Ruzicic S, Dreyer I, Mueller RB (2007). PInTFDB: an integrative plant transcription factor database. BMC Bioinform. 8: p. 42.
- Rodriguez MO, Sullivan D, Donini P, Papa R, Chipparino E, Leigh F, Attene G (2006). Integration of retrotransposons-based markers in a linkage map of barley. Mol. Breed. 17: 173-184.
- Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, Yamaguchi-Shinozaki K (2002). DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydrationand cold-inducible gene expression. Biochem. Biophys. Res. Commun. 290: 998-1009.
- Sambrook J, Fritsch EF, Maniatis T (1989). Molecular cloning: A laboratory manual. New York: Cold Spring Harbor Laboratory Press.
- Stein N, Prasad M, Scholz U, Thiel T, Zhang H, Wolf M, et al. (2007). A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. Theor. Appl. Genet. 114: 823-839.
- Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24(8): 1596-1599.
- Wang S, Basten CJ, Zeng ZB (2007). Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. http://statgen.ncsu.edu/qtlcart/WQTLCart.htm.
- Wilson K, Long D, Swinburne J, Coupland G (1996). A dissociation insertion causes a semidominant mutation that increases expression of TINY, an Arabidopsis gene related to APETALA2. Plant Cell, 8: 659-671.
- Xue GP (2003). The DNA-binding activity of an AP2 transcriptional activator HvCBF2 involved in regulation of low-temperature responsive genes in barley is modulated by temperature. The Plant J, 33: 373-383.
- Xue GP, Loveridge CW (2004). HvDRF1 is involved in abscisic acidmediated gene regulation in barley and produces two forms of AP2 transcriptional activators, interacting preferably with a CT-rich element. Plant J. 37: 326-339.
- Yi SY, Kim JH, Joung YH, Lee S, Kim WT, Yu SH, Choi D (2004). The pepper transcription factor CaPF1 confers pathogen and freezing tolerance in Arabidopsis. Plant Physiol. 136: 2862-2874.
- Zhuang J, Cai B, Peng RH, Zhu B, Jin XF, Xue Y, Gao F, Fu XY, Tian YS, Zhao W, Qiao YS, Zhang Z, Xiong AS, Yao QH (2008). Genomewide analysis of the AP2/ERF gene family in *Populus trichocarpa*. Biochem. Biophys. Res. Commun. 371: 468-474.
- Zhuang J, Chen JM, Yao QH, Fei X, Sun CC, Zhou XR, Zhang J, Xiong AS (2010). Discovery and expression profile analysis of AP2/ERF family genes from Triticum aestivum. Mol. Biol. Rep. 38: 745-753.
- Zhuang J, Anyia A, Vidmar J, Xiong AS, Zhang J (2011). Discovery and expression assessment of the AP2-like genes in Hordeum vulgare. Acta. Physiol. Plant, 33: 1639-1649.