

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

Jian-Xun Feng¹, Di Liu^{1,2}, Yi Pan¹, Wei Gong^{1,3,4}, Li-Geng Ma^{3,4}, Jing-Chu Luo^{1,2}, Xing Wang Deng^{3,4} and Yu-Xian Zhu^{1,3,5,*}

¹The National Laboratory of Protein Engineering and Plant Genetic Engineering Peking University, 100871, Beijing, China (*author for correspondence; e-mail zhuyx@water.pku.edu.cn); ²Center for Bioinformatics, Peking University, 100871, Beijing, China; ³Peking-Yale Joint Center for Plant Molecular Genetics and Agribiotechnology, College of Life Sciences, Peking University, 100871, Beijing, China; ⁴Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut, 05620, USA; ⁵National Plant Gene Research Center (Beijing), 100101, Beijing, China

Received 4 May 2005; accepted in revised form 28 July 2005

Key words: AP2/EREBP, gene family, HMM, microarray, transcription factor

Abstract

AP2/EREBP transcription factors (TFs) play functionally important roles in plant growth and development, especially in hormonal regulation and in response to environmental stress. Here we reported verification and correction of annotation through an exhaustive cDNA cloning and sequence analysis performed on 145 of 147 gene family members. A RACE analysis performed on genes with potential in-frame up-stream ATG codon resulted in identification of At2g28520 as an authentic AP2/EREBP member and corrected ORF annotations for three other members. A further phylogenetic analysis of this updated and likely complete family divided it into three major subfamilies. The expression patterns of the AP2/EREBP family members among the 11 organ or tissue types were examined using an oligo microarray and their hormonal and environmental responsiveness were further characterized using cDNA custom macroarrays. These detailed expression profile results provide strong support for a role for AP2/EREBP family members in development and in response to environmental stimuli, and a foundation for future functional analysis of this gene family.

Introduction

Most cellular and developmental processes are controlled by gene expression resulting from a mobilization of multiple different sets of transcription factors (TFs) (Grandori *et al.*, 2000; Stracke *et al.*, 2001; Dimova *et al.*, 2003; Kohler *et al.*, 2003). It has been reported that a large percentage of *Arabidopsis* open reading frames (ORFs) encode TFs that belong to more than 30 different families,

each possessing a highly conserved and characteristic region recognized as the DNA-binding domain (Riechmann *et al.*, 2000; Ramanathan *et al.*, 2002). Short sequence motifs outside of the DNA-binding domain in different TFs are also conserved within a given subgroup that are involved in both similar and unrelated biological functions (Heim *et al.*, 2003). Initial expression and cDNA cloning of the overall TF genes have been recently reported (Czechowski *et al.*, 2004; Gong *et al.*, 2004).

The *Arabidopsis* genome contains several large TF families with more than 100 members each, such as MADS, bHLH, MYB and AP2/EREBP (Riechmann and Ratcliffe, 2000; Bailey *et al.*, 2003; Heim *et al.*, 2003; Parenicova *et al.*, 2003; Toledo-Ortiz *et al.*, 2003; Jiang *et al.*, 2004). The *APETALA2*/ethylene-responsive element binding protein (AP2/EREBP) family is best characterized by a common AP2 domain of about 60 amino acids that is important for DNA binding (Okamuro *et al.*, 1997; Riechmann and Meyerowitz, 1998; Riechmann *et al.*, 2000; Sakuma *et al.*, 2002; Magnani *et al.*, 2004). AP2/EREBP genes are known to be involved in plant hormone signal transduction as well as in plant's responses to biotic, pathogenic, and environmental stresses (Stockinger *et al.*, 1997; Knight *et al.*, 1999; Brown *et al.*, 2003; Chakravarthy *et al.*, 2003; Gutterson and Reuber, 2004; Magome *et al.*, 2004; Yi *et al.*, 2004). This family was originally divided into two sub-families (AP2 and EREBP), and a third subfamily containing one AP2 and one B3 domain was subsequently added (Riechmann and Meyerowitz, 1998; Kagaya *et al.*, 1999; Riechmann *et al.*, 2000). A total of 144 AP2/EREBP members was recovered from the *Arabidopsis* genome and was divided into five groups: AP2, RAV, DREB, ERF and 'others' (Sakuma *et al.*, 2002).

Recent publication indicated that the family has 145 members and it is no longer plant-specific since homologs have been revealed from the cyanobacterium *Trichodesmium erythraeum*, the ciliate *Tetrahymena thermophila*, and the viruses *Enterobacteria phage Rb49* and *Bacteriophage Felix 01* (Magnani *et al.*, 2004). However, systematic information regarding the expression patterns for the complete AP2/EREBP gene family in various *Arabidopsis* organs or tissue types, or upon hormonal and environmental stresses is not available. Here we report a further in-depth phylogenetic analysis and systematic verification of all annotations by cDNA cloning and a sequencing of most members of this gene family. Both oligo microarray and custom cDNA macroarray were used to examine tissue expression patterns and environmental or hormonal regulations of this important gene family. We found that many members of this TF family were significantly activated upon drought or UV treatment, whereas ethylene and salicylic acid had only a modest effect.

Materials and methods

Plant materials

Arabidopsis thaliana (ecotype Col-0) plants were grown in environmental growth chambers (Conviron, Canada) under 16 h of light and 8 h of darkness photoperiod conditions. Plants were maintained at 23 °C during the light period and 21 °C during the dark period. Methods for collecting different organ or tissue types used in the oligo microarray analysis were described previously (Ma *et al.*, 2005).

All hormonal and environmental treatments were carried out using 6–8 rosette-leaf-stage *Arabidopsis* plants. For ABA treatment, 100 μ M ABA was sprayed on the leaves for 8 h; For NaCl treatment, whole pots were wetted by 300mM NaCl and kept for 8 h; For heat-shock (heat) treatment, plants were pre-warmed at 37 °C for 2 h before transfer to 45 °C for another 2 h; For UV treatment, plants were radiated with ultraviolet light (100 J m⁻²) for 6 h; For drought treatment, the entire plants were uprooted, placed on filter papers and allowed to dry for 6 h; For SA treatment, 4 mM salicylic acid was sprayed on the leaves for 8 h; for ethylene (eth) treatment, plants were placed in a closed chamber containing 100 ppm C₂H₄ for 24 h; cold, plants were placed in a 4 °C cold room for 8 h; For wound treatment, rosette leaves were cut into ~5mm strips and were left in the growth chamber for 8 h before being harvested for RNA isolation. In all cases, parallel and untreated plants of the same stage were used as controls.

Identification of AP2-like domain containing proteins

The *Arabidopsis* genome databases TAIR (Rhee *et al.*, 2003) and MAtdB (Schoof *et al.*, 2002) were searched for unique loci using InterPro IPR001471 accession as the family identifier. Concurrent loci were chosen to yield a multiple sequence alignment using CLUSTAL W 1.8.3 (Thompson *et al.*, 1994). Manually refined alignment was employed to construct a Hidden Markov Model (HMM) profile of the AP2 domains with HMMER 2.3.1 (Eddy, 1998) installed locally. A pattern search with *E* values < 1 was performed against the *Arabidopsis* protein database downloaded from the TAIR ftp

site (as of 2/28/2004, ftp://tairpub:tairpub@ftp.arabidopsis.org/home/tair). The resultant sequences were aligned and checked manually to eliminate false positives.

Phylogenetic analysis

A total of 161 AP2 domains were extracted from 147 hypothetical proteins obtained from a database search. Multiple sequence alignment was performed using CLUSTAL W 1.8.3 (Thompson *et al.*, 1994) and manually adjusted. Data generated from the alignment was used to construct the Bayesian phylogenetic tree of the entire family with MrBayes 3 (Ronquist and Huelsenbeck, 2003) and with the JTT-f model (Jones *et al.*, 1992) of amino acid substitution. The alignments were refined using HMMER 2.3.1 with manual adjustment. A total of 200 001 trees were generated with every 100 trees sampled. The saved trees with a 'burn-in' of 1000 were used to construct the consensus and calculate the posterior probabilities. Visualization tool TreeView (Page, 1996) was used to display the consensus tree.

RNA extraction, reverse transcription and RT-PCR amplifications

Total RNA was isolated either from various *Arabidopsis* tissue or organ types, or from 6 to 8 rosette-leaf-stage *Arabidopsis* seedlings after various hormonal or environmental treatments using the RNeasy plant mini kit (QIAGEN, Germany). Three μg of the RNA was reverse transcribed in a total volume of 20 μl using SuperScriptTM First-Strand Synthesis System for RT-PCR (Invitrogen, USA). Gene-specific primers for different TFs were designed according to annotation available in the *Arabidopsis* genome (STable 1). RT-PCR was performed following a previously described protocol (Ji *et al.*, 2003). PCR products were analyzed by Agarose gels and purified with a gel extraction kit (Clontech, USA), cloned into a pENTR/D/TOPO vector (Invitrogen, USA) and verified by sequencing of the whole insertion from both ends (Gong *et al.*, 2004). Since some of the cloned TFs genes do not possess introns, fractions of all the RNA samples used for RT-PCR were subjected to RNase A digestion to confirm that the PCR products were indeed amplified from cDNA, not from gDNA.

Analysis of 5' sequences and rapid amplification of 5'-cDNA ends (RACE)

To find out whether and how many annotated TFs possessed in-frame up-stream ATGs, we downloaded a 500 bp sequence at the 5' upstream of each TF from TAIR and searched for the existence of different translation initiation sites. RACE reactions were performed on all those TFs that showed an in-frame ATG other than the one annotated in TAIR to obtain putative full-length cDNAs according to protocols provided by the manufacturer (Invitrogen, Carlsbad, USA). PCR amplifications of the 5' ends of TFs were performed using GeneRacer 5' primers supplied by the vendor together with gene-specific reverse primers synthesized for each TF (See Stable 1 for all primers used). PCR products were first confirmed using nested reverse gene-specific primers and finally by DNA sequence analysis.

Expression pattern analysis by oligo microarray and RT-PCR

Gene-specific 70mer oligos were designed based on *Arabidopsis* genome annotation data available on Feb. 20, 2002 by Qiagen (http://omad.qiagen.com/download/genelist/arabidopsis_V1_384.prn) and the microarray slides were printed at Yale University as described previously (Ma *et al.*, 2005). The details of this microarray analysis were described in another report (Ma *et al.*, 2005). In brief, the chips were scanned at 532 nm (Cy3) and 635 nm (Cy5) wavelengths with an Axon GenePix 4000B scanner (Axon, Foster City, USA) at 5-nm resolution, and quantified with Axon GenePix Pro 3.0 image analysis. For most genes, each data point represents the average intensities obtained from four hybridizations using independent biological replicates on the array. For those very low level expressers, a minimum of two independent data points (out of the four hybridizations) were required for the gene to be included in the analysis. RT-PCR analyses were performed as a further validation on a number of TFs that showed either tissue- or organ-specificity as revealed by microarray. All gene-specific PCR primers and basic parameters are shown in supplementary Table 1 (STable 1). Geneinvestigator software tools (Zimmermann *et al.*, 2004)

Table 1. Summary of RACE analysis for TF genes with in-frame up-stream translation initiation sites other than the one annotated in TAIR.

Locus ID	Position of in-frame ATG (bp)	RACE results (bp)	New GenBank accession	Comments
At1g28370	-111	-87		Identical to TAIR annotated 5'-UTR
At5g07580	-201	-37		Identical to TAIR annotated 5'-UTR
At3g23220	-33	-52	AY974196	
At5g10510	-45	-508	AY974197	
At2g25820	-177	-270	AY974198	
At3g16280	-165	-189	AY995152	
At1g12630	-12	? ^a	DQ145776 ^b	A different ORF was amplified by RT-PCR
At5g67000	-72	? ^a	DQ145775 ^b	A different ORF was amplified by RT-PCR
At1g50680	-81	? ^a	–	No upstream sequence was amplified by RT-PCR.

^a RACE did not work.

^b PCR product containing the putative new ORF as shown in SFigure 1 was submitted to GenBank after DNA sequence confirmation.

were used to find and to compare expression profiles for those genes that were reported in other related databases.

Preparation of the custom cDNA microarray

A total of 143 AP2/EREBP TFs originally cloned in pENTR/D/TOPO vectors (Gong *et al.*, 2004) were amplified in a 96-well PCR plate in a Perkin-Elmer 9600 thermocycler in 50 μ l reactions using the flanking sequences of the vector as primers (primer 1: CAACTTTGTACAAAAAAGCAGG C; Primer 2: CAACTTTGTA CAAGAAAGCTG GGT). All PCR cycles consisted of one round at 95 °C for 3 min, 35 cycles including 95 °C for 1 min, 55 °C for 20 s and 72 °C for 90 s, with a final extension at 72 °C for 5 min. One tenth of every PCR product was used for restriction endonuclease digestion and insert verification. The remaining DNA was precipitated by two volumes of anhydrous ethanol and redissolved in 10 μ l of 50% DMSO for arraying. Each PCR product was spotted on the slide in triplicates (CapitalBio Corp., Beijing, China; <http://www.capitalbio.com/Product/protocols.htm>). Hybridizations were carried out in triplicate with dye-swap using independent RNA samples. Signal intensities of the housekeeping genes (actin2, tubulin beta-2, 40S ribosomal protein S5A, ribosomal protein L19, the eukaryotic initiation factor 4 α -1, the cytosolic glyceraldehyde-3-phosphate

dehydrogenase gene (GAPDH) and UBQ10) were used to normalize the microarray data. Genes that failed the T-test (at the 0.05 significance level) were considered to be non-expressed. Coefficient of variation (CV) was calculated for all data points that showed higher-than-background expression after the different treatments.

Results

A further annotation update for the AP2/EREBP TF family of the Arabidopsis genome

A total of 147 genes that contained at least one AP2-like domain were retrieved from various *Arabidopsis* genome databases. The initial cDNA cloning of 142 AP2/EREBP TFs was reported previously (Gong *et al.*, 2004), and three more cDNAs genes including At1g79700, At3g23230 and At5g50080 were cloned afterwards. Sequence analysis performed on all of our cDNAs revealed that a 27-bp intron was not recognized in original At5g50080 annotation. We submitted the corrected cDNA sequence to GenBank (accession number AJ580378) after complete sequencing of two independent clones for that gene. We failed to obtain cDNA clones for only two of the 147 genes in the family, At1g16060 and At5g67010, after repeated attempts.

Six TFs might possess in-frame up-stream translation initiation sites other than the ones currently annotated

A total of 135 TFs in the AP2/EREBP family, upon analysis of 500 bp 5' upstream sequence of each member, were found to have in-frame stop-codons that prevented extension of the annotated ORFs reported in TAIR. Of the 12 TFs that contained in-frame up-stream translation initiation codons, three are not suitable for RACE analysis for they possess in-frame ATGs at -3bp (At4g27950) or -6bp (At1g51120 and At5g65130) positions of the annotated translation start sites (data not shown). RACE experiments were carried out on nine TFs and four exhibited ORFs different from the database annotations. A new protein with a more conserved AP2 domain was derived from the re-annotated At2g25820 gene (Table 1). Two of the nine TF genes, At1g28370 and At5g07580, were found to have identical 5'-UTRs as reported in TAIR (Table 1). Since we failed to obtain RACE results for three of the TFs including At5g67000, At1g12630 and At1g50680 after multiple attempts, we designed several pairs of gene-specific primers and performed RT-PCR amplifications on each of these genes. The transcripts of At5g67000 and At1g12630 possessed ORFs different from the ones annotated in the database while no amplification product upstream of the original ORF was observed for At1g50680 (Table 1 and SFigure 1).

A further phylogenetic analysis of the Arabidopsis AP2/EREBP family

As indicated in Figure 1A, AP2/EREBP family was divided into three main clades, the 19-membered AP2, 6-membered RAV and 122-membered EREBP. The characteristic AP2 domains found in both the EREBP and RAV subfamilies were

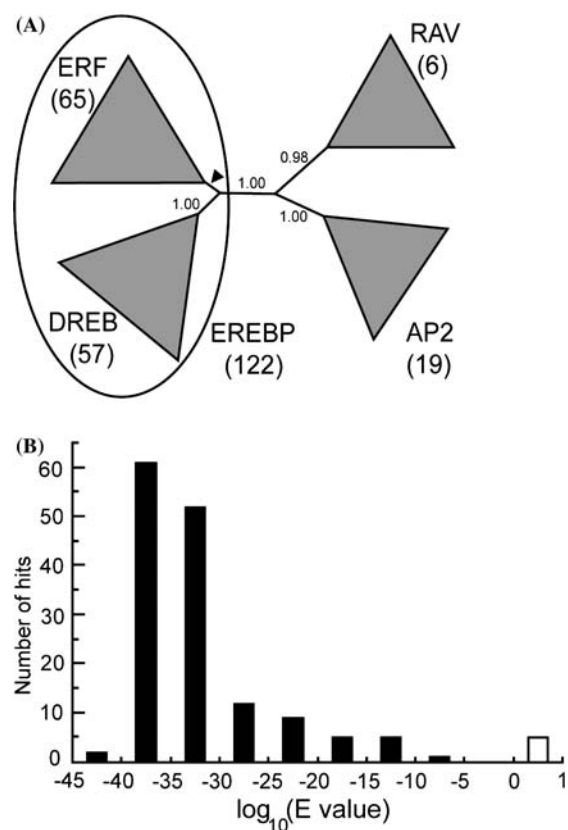


Figure 1. Phylogenetic analyses and chromosomal locations of AP2/EREBP TF genes from the *Arabidopsis* genome. (A) Major clades depict the relationships among different subfamilies within the AP2/EREBP TF family. The topology of the phylogenetic tree was obtained by running MrBayes 3. The number of genes in each subfamily or subgroup is shown in parentheses. The posterior probabilities at the branches indicate the significance. The arrowhead indicates collapsed ERF clade. (B) Results of HMM search for all AP2/EREBP TFs (shown in black bar) with several low homology proteins that are not included in this TF family (shown in empty bar) in the *Arabidopsis* genome.

about 60–70 amino acid residues while that of the AP2 subfamily varied significantly from 41 to 74 (Table 2). Among members of the AP2 subfamily, At4g13040 was the only gene that

Table 2. Summary of Arabidopsis AP2/EREBP transcription factor subfamilies.

Subfamily	No. of proteins	Length of proteins	No. of AP2 domains	Length of AP2 domains
EREBP	122 ^a	122–391	122	62–73
AP2	19	196–581	32 ^b	41–74
RAV	6	333–361	6	63–65

^a Of the total, 57 are in the DREB and 65 are in the ERF subgroups.

^b Thirteen of these proteins contain two AP2 domains, and 6 of them contain only one AP2 domain.

possessed a single R2 domain (Figure 1A, Table 2). The E value for this protein was 8.6×10^{-8} with a score of 38.2 (Figure 1B and data not shown). Five genes in the Arabidopsis genome, including the high-affinity nitrate transporter ACH1 (At1g08090), glycosyl hydrolase family 1 proteins (At3g18070, At3g18080), extra-large G-protein-related (At4g01090) and the copia-like retrotransposon family member (At2g23330) displayed certain sequence similarities with AP2/EREBP family members. Multiple sequence alignment revealed substantial conservation among the AP2 domains obtained from the entire family, and especially within each subfamily (SFigure 2).

Expressional analysis of the Arabidopsis AP2/EREBP family in different tissue or organ types

Long oligo (70mer) microarray was used to profile the expression patterns of 139 AP2/EREBP family members among 11 different tissue or organ types (Figure 2, Right). A neighbor-joining phylogenetic tree was generated for the entire gene family according to the sequence conservation patterns observed from the common AP2 domain (Figure 2, Left). Detailed search of the literature revealed that the expression patterns of 31 genes (as indicated by a red star following the individual locus ID in Figure 2) out of the 139 characterized in our tissue-specific microarray studies did not

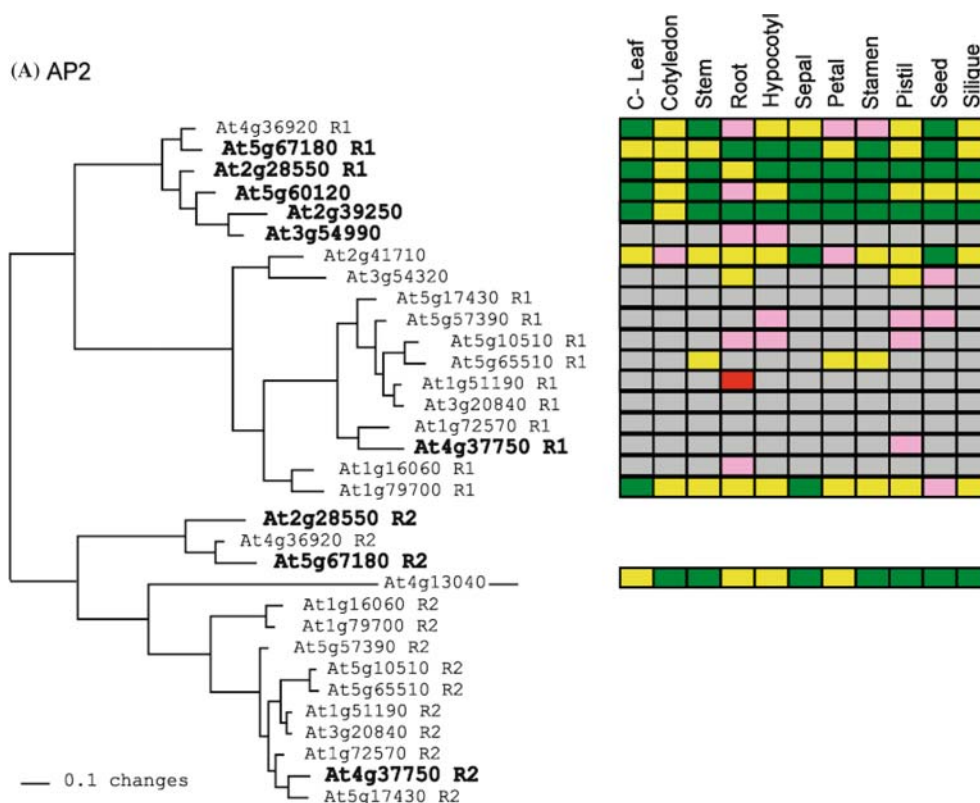


Figure 2. Construction of protein maximum likelihood trees for the AP2/EREBP family and expressional profiling obtained from oligo microarray. Unrooted trees for each sub-family are shown on the left and results obtained from oligo microarray analysis of 139 TFs in various *Arabidopsis* tissue or organ types are shown on the right. Ratios obtained by dividing the hybridizing intensity from a particular tissue or organ type with that of rosette leaves for a particular TF was subjected to logarithmic transformation and used as a graphic unit. Red bars, genes increased more than 8-fold; pink bars, genes increased 2- to 8-fold; yellow bars, genes changed between ± 2 -fold; green bars, genes decreased more than 2-fold. Genes that either failed the T-test (at the 0.05 significance level) or produced total signal intensities below threshold (equal to or less than 200) are shown in gray and were considered as non-expressers. The median CV for all expressed data points were 0.288. nd, genes not included in the oligo microarray. Boldface locus IDs denote genes previously studied in the literature.

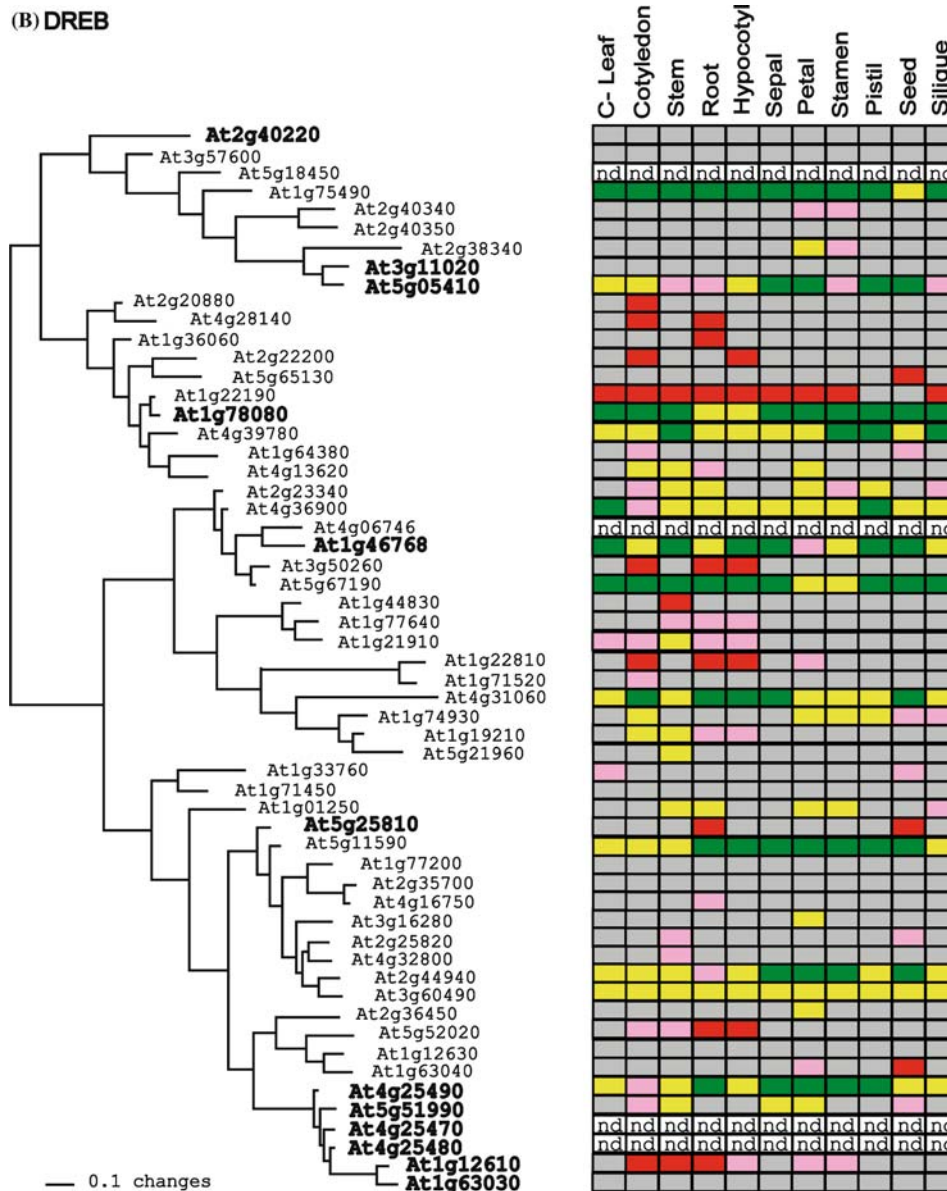
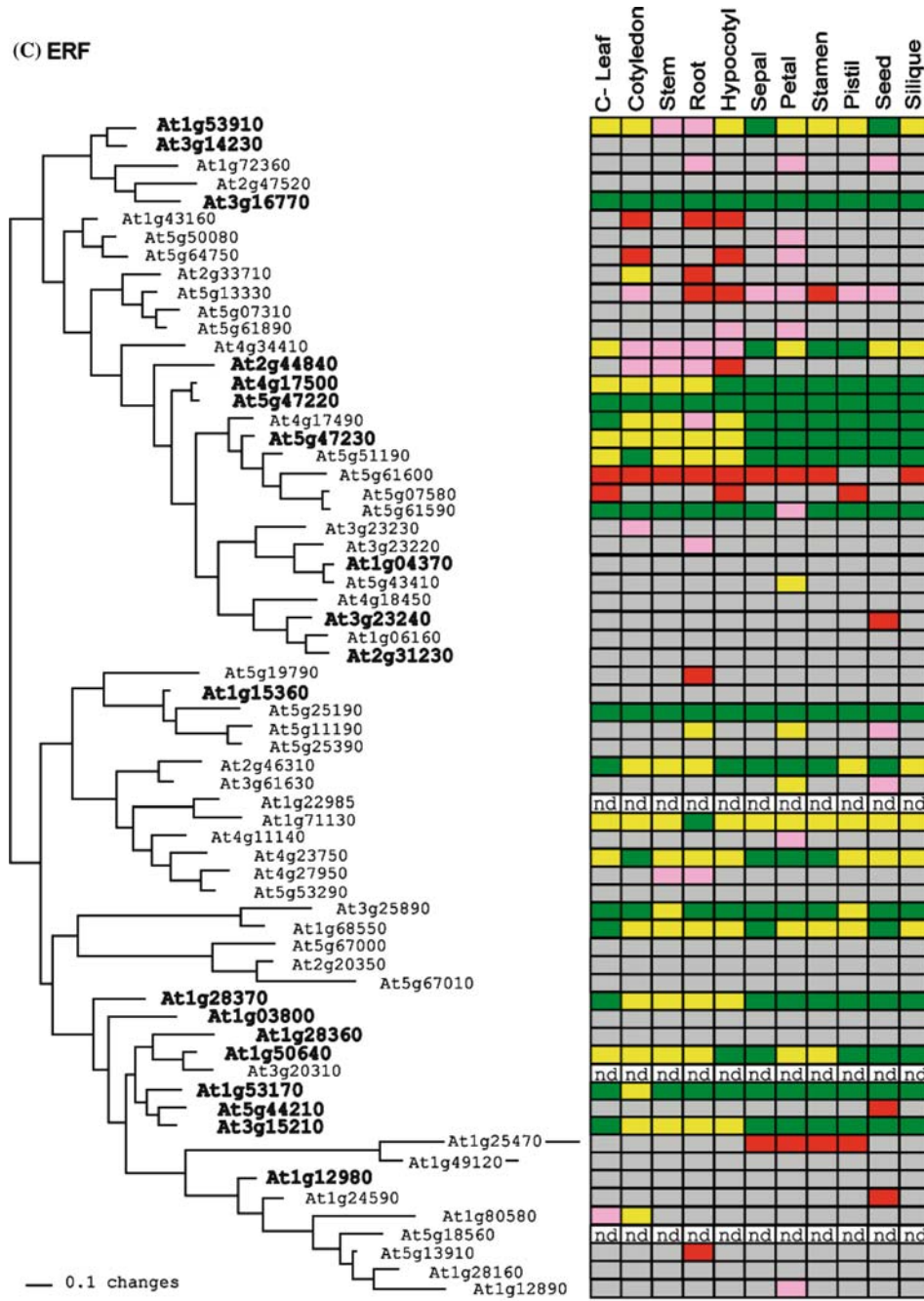


Figure 2. (Continued).

match with those reported in the Genevestigator database (<https://www.genevestigator.ethz.ch/>). The transcript levels of quite a few newly identified TFs were found to increase more than 64-fold in one or more *Arabidopsis* tissue or organ types. A similar number of genes were suppressed more than 16-fold when compared to that of rosette leaves (Figure 3A). At5g52020 (DREB subgroup) was detected only in vegetative tissue with the highest amount found in roots while At1g25470 (ERF subgroup) was mainly expressed in the

reproductive parts with residual amounts found in seeds and siliques (Figure 3B, C). We inserted RT-PCR results above each graph for validation of the microarray data. The tissue specific expression profiles of one previously characterized AP2 subfamily member (At3g54990) and eight new members, including four from DREB (At2g38340, At2g40340, At1g33760 and At4g16750), two from the ERF subgroup (At5g13910 and At1g25490) and one each from the AP2 (At1g51190) and RAV subfamilies (At3g25730) were confirmed by

(C) ERF



(D) RAV



Figure 2. (Continued).

RT-PCR analysis (Figure 3D). To complete the expression profile for the whole family, we carried out RT-PCR analysis on 8 of the TFs that were missing on the microarray probed with various *Arabidopsis* tissue and also on 4 of the TFs missing on the microarray probed with hormone-treated or environmental stressed RNA samples (Figure 4).

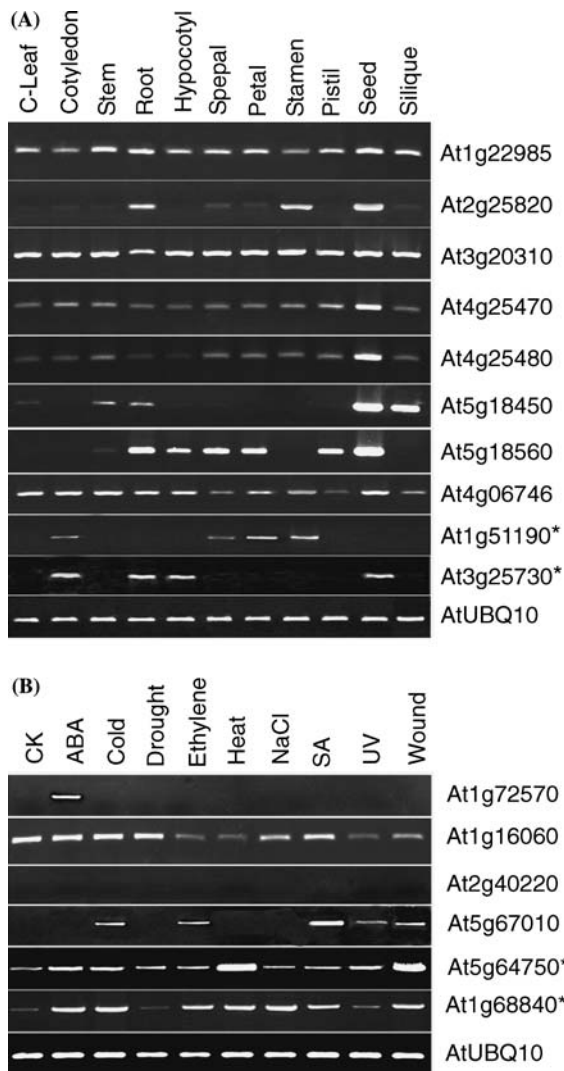


Figure 4. RT-PCR analysis of TFs that were not included on the microarrays. (A) Expression profiling of 8 TFs in various *Arabidopsis* tissue or organ types. (B) Expression profiling of 4 TFs upon different hormonal or environmental treatments. * denotes genes that were represented on the respective microarrays. We produced RT-PCR results of these genes to assist a direct comparison of the two different data sets.

Identification of hormonal or environmental responsive TFs

The ORF fragments of 143 AP2/EREBP TFs were PCR amplified and custom-printed on macroarray slides, and were probed with RNA samples prepared from *Arabidopsis* seedlings after various hormonal or environmental treatments. A large percentage of the TFs were found expressed at lower than 1% of the values obtained for the actin2 internal control gene in untreated *Arabidopsis* seedlings. Drought, UV, cold and NaCl treatments resulted in the activation of many more TF genes than other treatment (Tables 3 and 4). Thirteen TFs responded to drought with more than 8-fold increase in signal intensity, with nine TFs responding to UV, five TFs to NaCl and one TF to heat (Table 3, shown in bold). At3g23240 (ERF1) was mainly induced by UV treatment with modest responses to ethylene, NaCl and drought treatments. At1g04370 (AtERF14) responded significantly to UV and drought while At4g25480 (DREB1A/CBF3) was activated only under drought conditions with modest responses observed after NaCl and UV. Five of the 7 internal control genes were found to be expressed at consistent levels after different hormonal or environmental treatments, with UBQ10 and GAPDH appearing to be saturated under all or most of the experimental conditions (STable 2). A systematic search using all the Locus IDs available currently was carried out to summarize the previously published literature regarding this gene family (STable 3).

Discussion

Based on the presence of the conserved AP2-like domains, 147 proteins were identified as belonging to the AP2/EREBP family (Figures 1 and 2) from the *Arabidopsis* genome in this analysis, compared with the previously reported 141 (Alonso *et al.*, 2003), 144 (Sakuma *et al.*, 2002) or 145 (Magnani *et al.*, 2004). At1g22190 and At1g63040 were not included in any previous studies and were recognized in our analysis as members of the DREB subfamily (Figure 2 and SFigure 2). The putative polypeptide encoded by At2g25820 (a member of the DREB subgroup) contains only 29 amino acid residues in the conserved AP2 domain as

Table 3. Profiling of 143 AP2/EREBP TFs with reference to the absolute signal intensities of the actin2 gene.

Locus ID	Cold	NaCl	ABA	Wound	Drought	Ethylene	SA	Heat	UV	CK
Atlg01250	49±7	138±42	109±14	74±8	125±33	51±12	104±20	50±8	46±9	84±23
Atlg03800	4±3	5±4	4±4	3±3	10±2	0±0	8±5	4±5	10±1	6±5
Atlg04370	7±4	24±16	4±3	2±2	122±29	7±7	25±45	8±12	448±28	10±8
Atlg06160	12±2	12±7	4±2	5±2	7±4	45±20	41±13	5±2	70±6	21±10
Atlg12610	2±2	21±12	3±5	2±2	57±20	1±1	2±3	1±3	3±3	5±3
Atlg12630	5±3	4±3	3±3	4±2	4±2	1±1	5±7	2±3	3±3	5±3
Atlg12890	328±34	288±57	236±20	263±4	193±67	310±50	291±37	158±34	381±37	272±50
Atlg12980	2±3	5±3	3±1	1±1	3±2	0±0	3±4	6±14	3±3	3±4
Atlg13260	548±105	380±67	80±3	191±25	150±49	390±87	331±30	119±20	206±10	321±77
Atlg15360	2±2	2±4	3±3	2±1	5±1	1±2	3±4	3±1	2±4	3±3
Atlg19210	4±2	79±22	5±4	10±3	186±52	3±5	10±2	3±3	21±4	9±6
Atlg21910	65±6	156±57	126±9	16±5	114±32	186±43	45±5	122±18	19±6	110±29
Atlg22190	128±5	351±60	209±4	130±21	620±61	232±44	193±29	477±80	407±26	229±48
Atlg22810	4±3	202±48	13±2	8±2	222±67	11±9	14±5	2±4	66±7	5±5
Atlg22985	3±2	1±2	4±1	1±2	3±1	1±2	2±2	3±2	1±3	2±3
Atlg24590	6±5	6±6	3±4	3±3	9±3	2±2	8±11	6±3	5±4	6±5
Atlg25470	33±3	20±10	16±5	20±3	17±8	20±8	24±5	25±8	31±5	30±9
Atlg25560	815±91	632±108	434±35	284±24	299±34	626±137	561±107	221±21	524±29	555±128
Atlg28160	8±3	26±11	6±2	6±1	38±19	8±7	12±7	9±8	31±6	12±10
Atlg28360	6±4	27±13	10±3	9±2	26±8	8±6	18±7	4±2	13±4	15±8
Atlg28370	21±5	178±41	47±6	28±4	202±70	142±37	73±17	10±4	207±18	39±10
Atlg33760	3±3	4±2	3±3	3±2	5±4	2±3	4±5	2±4	3±2	3±5
Atlg36060	28±8	40±12	26±9	23±2	62±8	85±14	24±6	93±19	89±5	70±18
Atlg43160	17±7	135±34	28±11	96±10	626±233	113±14	105±22	34±16	354±19	17±6
Atlg44830	6±3	28±15	4±3	3±1	24±13	38±10	11±5	17±5	6±5	12±8
Atlg46768	130±19	18±8	37±5	13±3	45±13	15±9	20±9	32±7	41±7	22±7
Atlg49120	10±1	22±6	13±3	9±3	13±3	9±9	18±4	4±6	28±6	16±8
Atlg50640	121±11	128±37	103±5	76±5	162±50	76±25	107±14	153±6	218±11	120±30
Atlg50680	4±3	5±6	4±4	1±2	7±4	1±1	4±4	3±5	6±2	4±4
Atlg51120	3±3	6±5	2±3	1±3	12±15	1±1	7±4	3±3	4±4	5±6
Atlg51190	7±4	7±3	5±3	7±2	10±2	4±4	8±4	8±5	10±7	8±4
Atlg53170	56±4	97±19	43±4	34±4	267±108	64±15	65±7	12±11	127±12	65±14
Atlg53910	707±73	454±36	677±60	385±42	581±184	531±97	475±47	456±30	477±25	484±89
Atlg63030	4±2	9±5	7±6	2±3	12±2	1±3	3±3	4±3	2±3	3±5
Atlg63040	8±4	22±16	4±3	15±7	23±7	3±3	8±6	4±5	5±2	10±15
Atlg64380	33±5	195±61	203±13	109±11	531±51	91±20	107±30	16±7	53±8	42±11
Atlg68550	113±17	120±36	74±7	73±5	119±27	119±27	105±13	153±35	149±8	91±23
Atlg68840	536±55	288±17	150±9	76±10	237±72	383±94	182±22	127±17	365±4	523±100
Atlg71130	135±12	127±35	131±6	83±5	153±34	57±15	102±32	64±16	224±13	129±26
Atlg71450	3±3	4±3	3±4	1±3	7±2	1±1	1±2	1±2	3±3	3±4
Atlg71520	1±2	5±2	5±2	0±1	4±4	0±2	3±5	1±2	11±3	1±3
Atlg72360	461±48	58±21	65±8	49±4	87±20	155±23	53±12	52±10	189±22	85±17
Atlg72570	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Atlg74930	9±4	181±53	9±2	41±4	134±22	17±7	31±11	6±6	79±5	48±14
Atlg75490	35±3	16±7	5±4	5±2	30±3	21±6	19±6	9±5	65±9	61±18
Atlg77200	7±4	5±3	5±4	1±3	9±4	2±2	6±4	1±4	5±2	6±4
Atlg77640	27±7	62±31	46±5	12±5	50±9	38±17	25±7	15±14	9±4	29±11
Atlg78080	168±13	397±60	224±13	267±34	570±115	290±45	332±39	330±50	444±75	263±55
Atlg79700	339±9	28±7	65±6	24±2	140±32	50±10	31±5	113±11	149±9	406±59
Atlg80580	6±3	9±3	5±2	6±1	14±5	3±4	7±6	5±3	11±4	8±4
At2g20350	174±28	89±12	106±4	73±10	142±20	109±22	76±19	76±17	128±11	92±16
At2g20880	16±6	75±24	21±8	48±6	1310±260	58±12	61±9	25±7	201±18	56±14
At2g22200	7±4	20±9	13±5	27±2	64±6	13±4	36±6	5±8	21±4	15±7
At2g23340	4±2	2±3	4±2	2±2	2±1	1±1	1±2	2±4	1±3	2±3
At2g25820	15±3	62±12	16±3	27±4	104±8	17±7	54±10	12±3	32±5	12±5

Table 3. (Continued).

Locus ID	Cold	NaCl	ABA	Wound	Drought	Ethylene	SA	Heat	UV	CK
At2g28550	562 ± 48	260 ± 30	254 ± 16	220 ± 22	265 ± 56	145 ± 24	222 ± 37	214 ± 49	101 ± 10	220 ± 39
At2g31230	100 ± 7	112 ± 41	61 ± 8	102 ± 12	96 ± 24	178 ± 29	154 ± 28	30 ± 11	611 ± 30	77 ± 19
At2g33710	31 ± 9	55 ± 19	27 ± 5	19 ± 2	56 ± 15	35 ± 7	38 ± 8	14 ± 7	150 ± 15	28 ± 7
At2g35700	9 ± 3	17 ± 7	7 ± 3	11 ± 3	25 ± 11	7 ± 9	14 ± 7	5 ± 5	4 ± 2	9 ± 5
At2g36450	74 ± 7	82 ± 29	58 ± 4	60 ± 3	116 ± 23	59 ± 12	88 ± 20	41 ± 12	199 ± 15	78 ± 17
At2g38340	158 ± 19	713 ± 146	389 ± 28	762 ± 85	616 ± 129	422 ± 80	667 ± 79	63 ± 9	537 ± 18	358 ± 84
At2g39250	11 ± 4	35 ± 12	31 ± 8	20 ± 6	23 ± 6	14 ± 6	26 ± 6	5 ± 3	8 ± 4	16 ± 6
At2g40220	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
At2g40340	103 ± 10	98 ± 25	95 ± 12	77 ± 8	273 ± 74	170 ± 32	102 ± 19	240 ± 53	236 ± 17	49 ± 14
At2g40350	26 ± 5	17 ± 3	15 ± 8	11 ± 2	27 ± 8	25 ± 11	22 ± 7	21 ± 13	21 ± 2	13 ± 7
At2g41710	251 ± 20	148 ± 30	143 ± 8	94 ± 5	141 ± 31	199 ± 33	144 ± 19	200 ± 20	252 ± 18	218 ± 43
At2g44840	8 ± 4	85 ± 9	22 ± 4	24 ± 6	43 ± 7	17 ± 6	27 ± 4	3 ± 3	52 ± 5	19 ± 8
At2g44940	55 ± 10	64 ± 15	49 ± 9	76 ± 5	44 ± 6	26 ± 6	65 ± 15	5 ± 7	29 ± 4	31 ± 9
At2g46310	35 ± 4	28 ± 12	30 ± 2	22 ± 2	29 ± 5	33 ± 5	24 ± 8	25 ± 11	30 ± 6	36 ± 10
At2g47520	11 ± 15	11 ± 3	4 ± 2	3 ± 1	13 ± 2	6 ± 5	13 ± 9	55 ± 32	129 ± 22	5 ± 4
At3g11020	61 ± 4	71 ± 21	27 ± 3	26 ± 4	164 ± 43	43 ± 10	38 ± 7	142 ± 11	142 ± 15	38 ± 11
At3g14230	537 ± 82	193 ± 29	245 ± 14	146 ± 11	215 ± 32	387 ± 88	158 ± 26	429 ± 79	370 ± 30	207 ± 33
At3g15210	111 ± 10	188 ± 54	67 ± 7	78 ± 8	283 ± 57	90 ± 13	130 ± 22	40 ± 10	255 ± 15	84 ± 19
At3g16280	9 ± 2	17 ± 11	7 ± 1	10 ± 1	27 ± 7	6 ± 6	24 ± 16	3 ± 4	75 ± 8	11 ± 5
At3g16770	994 ± 43	425 ± 33	313 ± 22	205 ± 20	797 ± 197	633 ± 70	275 ± 50	444 ± 56	703 ± 54	705 ± 165
At3g20310	81 ± 11	70 ± 18	58 ± 6	41 ± 3	105 ± 32	50 ± 9	62 ± 6	49 ± 16	76 ± 8	77 ± 18
At3g20840	7 ± 2	6 ± 4	3 ± 4	4 ± 2	14 ± 6	4 ± 2	7 ± 3	7 ± 5	10 ± 4	7 ± 5
At3g23220	6 ± 3	83 ± 16	2 ± 2	4 ± 2	13 ± 4	4 ± 3	5 ± 3	4 ± 2	81 ± 19	8 ± 5
At3g23230	32 ± 8	50 ± 8	16 ± 5	13 ± 5	40 ± 7	14 ± 6	16 ± 5	16 ± 11	158 ± 14	21 ± 7
At3g23240	22 ± 4	116 ± 32	30 ± 2	17 ± 5	69 ± 16	135 ± 30	41 ± 10	4 ± 3	473 ± 70	32 ± 8
At3g25730	157 ± 29	136 ± 37	27 ± 5	121 ± 9	198 ± 68	166 ± 36	191 ± 25	52 ± 12	325 ± 20	71 ± 20
At3g25890	232 ± 34	175 ± 48	85 ± 8	131 ± 10	103 ± 28	265 ± 67	134 ± 12	245 ± 58	231 ± 17	177 ± 36
At3g50260	248 ± 24	134 ± 28	63 ± 11	76 ± 10	368 ± 85	137 ± 27	148 ± 10	119 ± 12	263 ± 23	62 ± 14
At3g54320	5 ± 2	6 ± 2	3 ± 2	5 ± 2	5 ± 3	2 ± 2	3 ± 4	7 ± 2	8 ± 5	5 ± 3
At3g54990	63 ± 9	49 ± 14	38 ± 9	57 ± 2	45 ± 4	23 ± 12	59 ± 14	8 ± 5	44 ± 7	24 ± 9
At3g57600	29 ± 4	42 ± 23	16 ± 7	26 ± 3	67 ± 8	19 ± 10	35 ± 7	33 ± 11	31 ± 6	37 ± 12
At3g60490	51 ± 9	77 ± 24	17 ± 6	72 ± 5	49 ± 10	56 ± 9	103 ± 25	46 ± 11	150 ± 21	52 ± 14
At3g61630	100 ± 9	157 ± 49	67 ± 6	47 ± 3	200 ± 64	109 ± 13	83 ± 16	119 ± 20	229 ± 15	84 ± 17
At4g06746	6 ± 4	4 ± 4	5 ± 2	2 ± 2	16 ± 18	1 ± 1	8 ± 12	4 ± 2	7 ± 5	8 ± 8
At4g11140	40 ± 9	52 ± 10	22 ± 4	29 ± 3	72 ± 14	30 ± 12	35 ± 6	24 ± 4	49 ± 9	43 ± 13
At4g13040	81 ± 6	69 ± 19	86 ± 11	54 ± 5	99 ± 25	113 ± 28	70 ± 12	49 ± 1	125 ± 14	129 ± 25
At4g13620	7 ± 4	7 ± 2	6 ± 5	4 ± 3	8 ± 4	3 ± 2	5 ± 4	41 ± 27	9 ± 1	6 ± 5
At4g16750	18 ± 4	36 ± 11	42 ± 4	27 ± 8	37 ± 7	11 ± 4	20 ± 9	4 ± 2	9 ± 5	15 ± 8
At4g17490	157 ± 18	80 ± 15	77 ± 11	62 ± 8	84 ± 24	76 ± 15	75 ± 19	48 ± 13	92 ± 7	110 ± 24
At4g17500	87 ± 14	174 ± 38	155 ± 9	97 ± 18	99 ± 14	182 ± 39	167 ± 20	23 ± 4	417 ± 66	152 ± 37
At4g18450	9 ± 4	4 ± 3	5 ± 1	4 ± 2	12 ± 4	3 ± 3	5 ± 2	1 ± 4	5 ± 2	12 ± 5
At4g23750	50 ± 2	64 ± 10	38 ± 5	47 ± 6	72 ± 10	43 ± 12	50 ± 9	49 ± 10	43 ± 4	73 ± 12
At4g25470	4 ± 4	6 ± 8	3 ± 3	2 ± 2	4 ± 3	1 ± 1	3 ± 3	3 ± 3	5 ± 2	4 ± 4
At4g25480	37 ± 5	72 ± 8	13 ± 2	30 ± 5	132 ± 29	12 ± 3	15 ± 4	5 ± 2	42 ± 2	12 ± 5
At4g25490	15 ± 5	26 ± 11	11 ± 6	5 ± 2	47 ± 15	8 ± 3	6 ± 3	4 ± 3	13 ± 3	7 ± 6
At4g27950	21 ± 4	34 ± 15	8 ± 3	15 ± 5	80 ± 10	17 ± 12	29 ± 7	22 ± 7	44 ± 8	21 ± 8
At4g28140	15 ± 3	112 ± 31	26 ± 3	105 ± 6	894 ± 142	94 ± 19	118 ± 25	24 ± 8	201 ± 10	36 ± 11
At4g31060	18 ± 5	14 ± 9	15 ± 4	16 ± 2	22 ± 22	14 ± 8	23 ± 12	21 ± 9	16 ± 3	31 ± 13
At4g32800	31 ± 2	98 ± 38	52 ± 5	40 ± 4	77 ± 19	54 ± 13	81 ± 21	35 ± 14	35 ± 4	44 ± 12
At4g34410	8 ± 3	286 ± 58	17 ± 5	18 ± 5	519 ± 107	77 ± 22	48 ± 12	6 ± 6	127 ± 15	27 ± 12
At4g36900	178 ± 18	144 ± 41	157 ± 9	73 ± 9	266 ± 53	113 ± 15	110 ± 16	204 ± 36	113 ± 14	135 ± 29
At4g36920	205 ± 29	92 ± 17	84 ± 6	86 ± 9	72 ± 16	76 ± 17	95 ± 18	33 ± 8	56 ± 5	75 ± 15
At4g37750	74 ± 7	64 ± 13	56 ± 5	86 ± 11	54 ± 5	49 ± 12	63 ± 9	69 ± 18	45 ± 10	94 ± 15
At4g39780	28 ± 5	57 ± 19	16 ± 3	34 ± 4	77 ± 22	41 ± 12	69 ± 18	12 ± 6	113 ± 15	54 ± 11
At5g05410	29 ± 3	161 ± 38	85 ± 10	48 ± 6	487 ± 125	94 ± 19	86 ± 25	325 ± 72	116 ± 5	110 ± 22
At5g07310	5 ± 3	15 ± 10	5 ± 2	6 ± 2	25 ± 8	4 ± 4	6 ± 2	2 ± 2	12 ± 4	6 ± 4

Table 3. (Continued).

Locus ID	Cold	NaCl	ABA	Wound	Drought	Ethylene	SA	Heat	UV	CK
At5g07580	790±45	506±94	429±25	312±15	162±36	775±98	504±40	297±29	230±19	440±112
At5g10510	60±6	66±14	49±3	67±6	92±24	38±11	60±17	57±14	110±10	48±10
At5g11190	16±3	9±6	7±3	6±2	16±8	3±2	6±3	9±4	7±4	9±5
At5g11590	22±4	21±13	8±4	10±3	9±3	5±4	15±7	3±2	11±4	14±7
At5g13330	137±10	200±30	265±22	169±23	453±103	209±22	265±36	52±10	409±27	100±25
At5g13910	11±5	13±6	9±3	13±4	16±6	9±6	23±12	16±11	20±5	12±5
At5g17430	66±5	126±34	116±11	107±15	57±18	87±16	108±24	77±21	62±6	82±20
At5g18450	13±3	21±15	11±4	11±5	36±13	12±9	11±5	11±3	18±10	14±8
At5g18560	30±5	51±18	30±5	21±5	55±11	30±14	34±6	29±5	37±5	33±11
At5g19790	19±6	26±13	14±4	22±3	44±5	24±11	24±6	80±17	26±9	27±9
At5g21960	1±2	21±5	2±3	6±1	12±2	1±2	3±4	3±1	11±3	2±5
At5g25190	31±6	38±12	17±2	17±4	41±8	90±16	30±10	16±5	39±7	155±39
At5g25390	5±2	6±2	8±3	5±2	5±3	5±4	4±3	2±1	5±3	4±4
At5g25810	52±7	35±18	33±8	24±1	53±12	22±17	31±7	37±11	28±4	31±10
At5g43410	4±3	6±4	4±2	2±2	6±2	3±4	4±2	5±5	91±8	4±3
At5g44210	58±4	37±11	31±5	24±4	252±67	61±12	30±7	26±5	54±7	42±10
At5g47220	316±6	387±74	268±16	111±8	228±75	484±56	355±28	48±13	927±46	396±91
At5g47230	108±6	186±16	157±9	88±7	253±75	141±23	143±30	24±7	302±41	189±41
At5g50080	8±2	9±6	7±3	5±2	16±4	9±6	7±2	6±3	25±11	9±5
At5g51190	74±6	121±17	127±11	31±5	103±23	116±19	66±18	9±5	170±16	153±30
At5g51990	14±7	15±8	8±2	6±3	20±10	6±6	18±15	4±2	14±1	10±7
At5g52020	10±3	16±3	8±2	9±6	66±23	8±7	13±3	14±8	15±5	10±8
At5g53290	28±9	61±11	20±7	41±4	115±24	25±8	60±9	22±7	59±11	23±8
At5g57390	7±6	8±4	6±5	5±3	13±3	4±4	8±2	7±3	7±3	7±4
At5g60120	81±6	77±17	76±9	60±5	69±17	47±11	64±14	91±14	48±6	72±15
At5g61590	914±40	245±43	242±15	140±23	162±56	527±47	289±19	326±25	204±18	562±135
At5g61600	135±7	179±34	210±22	81±9	95±10	152±22	109±23	44±9	117±8	271±55
At5g61890	10±2	98±24	59±7	96±8	269±57	61±16	66±33	8±4	35±6	15±6
At5g64750	188±11	279±50	63±3	156±12	1418±441	182±32	229±38	121±17	525±41	52±11
At5g65130	6±4	8±4	4±3	2±2	8±3	2±2	7±4	4±2	9±4	6±4
At5g65510	40±6	29±10	30±4	21±5	44±7	28±9	23±5	20±4	64±4	36±11
At5g67000	5±2	6±3	4±3	3±2	13±8	2±3	4±6	5±3	5±3	7±4
At5g67180	23±8	27±13	32±3	20±3	49±8	15±7	27±15	9±6	13±5	20±8
At5g67190	184±20	53±20	39±3	34±3	121±23	51±8	57±11	83±16	82±6	87±19

The signal intensities of all TFs in the nine different treatments as well as in untreated seedlings (CK) are expressed as the permillage of actin2^a. The numbers for each treatment are the means of six hybridizations (including dye-swap)±SD and the numbers for CK represented 54 data points obtained from hybridizations of all nine chips and six repeats. Boldface numbers indicate that the signal intensities were increased more than 8-fold under particular conditions. The median CV for all expressed TFs in these arrays was 0.17.

^a The average signal intensities calculated from nine different treatments plus the CK for actin2 were 30,307±5052.

Table 4. Statistics of the relative expression levels of all expressed TFs with reference to the absolute signal intensities of actin2.

	Cold	NaCl	ABA	Wound	Drought	Eth	SA	Heat	UV	CK
< 1%	47	29	50	47	20	47	35	60	32	39
1–10%	60	71	66	75	69	60	72	58	58	74
10–40%	26	38	24	20	41	29	32	21	40	23
> 40%	10	5	3	1	13	7	4	4	13	7
Total signal intensities*	14.0	13.7	9.1	7.8	20.1	12.2	10.9	8.3	17.5	11.5
Average intensities (%)	9.8	9.5	6.3	5.5	14.1	8.6	7.6	5.8	12.2	8.0
Maximum intensities (%)	99.4	71.3	67.7	76.2	141.8	77.5	66.7	47.7	92.7	70.5

* The total signal intensities for all 143 TFs divided by that of the internal actin2 control.

annotated in TAIR. It was always a questionable member of this gene family since the missing β -sheet-forming sequence has been proposed to have a key role in DNA binding activity (Allen *et al.*, 1998). Here we found that the original database-annotated initiation codon ATG (See also AAT68354, AAK44156 or AAX23824) was preceded by an in-frame up-stream ATG that resulted in a perfect AP2 domain (Table 4 and data not shown). Our RACE results matched perfectly with GenBank protein entry B84653 (Lin *et al.*, 1999), clearly suggesting that the upstream ATG is the one used.

Using established criteria (Ma *et al.*, 2005), approximately 50% (104/209 total data points for AP2 and 33/66 for RAV subfamily) to 61% in the EREBP subfamily (761/1254) were considered as not expressed (Figure 2). Kim and coworkers amplified most of the predicted genes on *Arabidopsis* chromosome 2 and discovered that about 84% of the amplicons were expressed in certain growth stages or under certain environmental conditions (Kim *et al.*, 2003). Of the 1864 DNA amplicons representing experimentally confirmed or predicted *Arabidopsis* transcription factor genes printed on the microarray, approximately 1400 or 75% were expressed (Jiao *et al.*, 2003). Discrepancies occurred almost exclusively on low-level expressers with only 6 genes (At3g57600, At1g77200, At3g14230, At1g06160, At1g15360 and At1g12980) that could be considered as intermediate level expressors. Ethylene treatment did not result in over-all activation of expression for the AP2/EREBP family of TF genes, but affected specific individual gene expression (Table 4). By contrast, the cumulative ratio value for all the cloned TF genes on the custom macroarray increased from 11.5 in untreated seedlings to 20.1 and 17.5 upon drought and UV treatment, respectively, when compared to that of the internal actin2 control (Table 4). The responsiveness of a large number of AP2/EREBP TF genes to drought and UV treatment coincided with the involvement of this gene family in the mediation of environmental stress responses (Okamoto *et al.*, 1997; Chakravarthy *et al.*, 2003). We failed to detect significant hybridizing signals for the two root-specific (At3g20840 and At1g28160) and three shoot-specific (At3g57600, At5g47220 and At5g25390) genes reported by Czechowski *et al.* (2004) using high-sensitivity real-time RT-PCR

analysis. The seed-specific At1g75490 and seed-, root- and *pet al*-specific At5g11190 gene reported in our microarray (Figure 2) did not match with results obtained by the same group (Czechowski *et al.*, 2004).

When the hormonal or environmental responsiveness of the AP2/EREBP TFs were compared to the Genevestigator database (<https://www.genevestigator.ethz.ch/>) using Genevestigator software tools (Zimmermann *et al.*, 2004), we found that the expression patterns of most genes matched well between the two data sets. Detailed analysis revealed that a total of 20 TFs showed different expression for one or two treatments with only one gene (At5g07580) differed in three treatments. Similarly, a vast majority of the TFs showed similar expression profiles between the two data sets with 15 genes displayed unmatched profiles in 3 or 4 tissue or organ types. At3g14230 and At1g53710 exhibited different expression profiles in all tissue and organ types (data not shown). We attribute these minor inconsistencies to different growth conditions or treatments, and the distinct sensitivity of the assays used.

AP2/EREBP genes have been implicated in controlling *Arabidopsis* flowering and seed development, especially in the specification of organ and meristem identity (Riechmann and Meyerowitz, 1998). Several AP2 subfamily members (At2g28550, At3g54990 and At5g60120) that were known to be the targets of *miRNA172* and were part of the *Arabidopsis* flowering time control mechanism (Aukerman and Sakai, 2003; Schmid *et al.*, 2003), were detected mainly in vegetative tissue (Figure 2). At1g12610, a DREB subgroup member that conferred delayed-flowering phenotype when overexpressed in *Arabidopsis* plants (Magome *et al.*, 2004), displayed high levels in stems, roots and cotyledons (Figure 2). Hormonal or environmental activation patterns for several known TFs, such as At4g25490, At3g23240, At3g11020, At4g25480, and At5g05410 (Table 3) are also supported by previous findings in the literature for CBF1/DREB1B, ERF1, DREB2B, DERB1A/CBF3 and DREB2A, respectively (Stockinger *et al.*, 1997; Solano *et al.*, 1998; Sakuma *et al.*, 2002). The discovery of organ-specific or environmental responsive AP2/EREBP TFs may lead to new tools for understanding the molecular mechanism(s) of developmental regulation.

In conclusion, we cloned 145 of 147 putative AP2/EREBP TF genes and performed systematic annotation and phylogenetic updates on this important gene family. RACE analysis on genes with potential in-frame upstream ATGs indicated that At2g28520, with its newly corrected ORF, could be an authentic AP2/EREBP gene. We believe that the current comprehensive analysis provides a likely complete membership list, a family tree and detailed expression profiling from representative organ and tissue types as well as that subsequent to a whole range of hormonal and environmental treatments.

Acknowledgements

We thank Hong-Ya Gu and Kun He for help with the phylogenetic analysis. This work was supported by a grant from China National Natural Science Foundation (Grant 30221120261) and the China National Basic Research Program (Grant 2004CB117302).

References

- Allen, M.D., Yamasaki, K., Ohme-Takagi, M., Tateno, M. and Suzuki, M. 1998. A novel mode of DNA recognition by a B-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *EMBO J.* 17: 5484–5496.
- Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C. and Ecker, J.R. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653–657.
- Arabidopsis Genome Initiative 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815.
- Aukerman, M.J. and Sakai, H. 2003. Regulation of flowering time and floral organ identity by a microRNA and its *APET ALA2*-like target genes. *Plant Cell* 15: 2730–2741.
- Bailey, P.C., Martin, C., Toledo-Ortiz, G., Quail, P.H., Huq, E., Heim, M.A., Jakoby, M., Werber, M. and Weisshaar, B. 2003. Update on the basic helix-loop-helix transcription factor gene family in *Arabidopsis thaliana*. *Plant Cell* 15: 2497–2501.
- Brown, R.L., Kazan, K., McGrath, K.C., Maclean, D.J. and Manners, J.M. 2003. A role for the GCC-box in jasmonate-mediated activation of the PDF1.2 gene of *Arabidopsis*. *Plant Physiol.* 132: 1020–1032.
- Chakravarthy, S., Tuori, R.P., D'Ascenzo, M.D., Fobert, P.R., Despres, C. and Martin, G.B. 2003. The tomato transcription factor Pti4 regulates defense-related gene expression via GCC box and non-GCC box cis elements. *Plant Cell* 15: 3033–3050.
- Czechowski, T., Bari, R.P., Stitt, M., Scheible, W. and Udvardi, M.K. 2004. Real-time RT-PCR profiling of over 1400 *Arabidopsis* transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. *Plant J.* 38: 366–379.
- Dimova, D.K., Stevaux, O., Frolov, M.V. and Dyson, N.J. 2003. Cell cycle-dependent and cell cycle-independent control of transcription by the *Drosophila* E2F/RB pathway. *Genes Dev.* 17: 2308–2320.
- Eddy, S.R. 1998. Profile hidden Markov models. *Bioinformatics* 14: 755–763.
- Gong, W., Sun, Y.P., Ma, L.G., Pan, Y., Du, Y.L., Wang, D.H., Yang, J.Y., Hu, L.D., Liu, X.F., Dong, C.X., Ma, L., Chen, Y.H., Yang, X.Y., Gao, Y., Zhu, D., Tan, X., Mu, J.Y., Zhang, D.-B., Liu, Y.L., Dinesh-Kumar, S.P., Li, Y., Wang, X.-P., Gu, H.-Y., Qu, L.-J., Bai, S.-N., Lu, Y.-T., Li, J.-Y., Zhao, J.-D., Zuo, J., Huang, H., Deng, X.W. and Zhu, Y.-X. 2004. Genome-wide ORFome cloning and analysis of *Arabidopsis* transcription factor genes. *Plant Physiol.* 135: 773–782.
- Grandori, C., Cowley, S.M., James, L.P. and Eisenman, R.N. 2000. The Myc/Max/Mad network and the transcriptional control of cell behavior. *Annu. Rev. Cell Dev. Biol.* 16: 653–699.
- Gutterson, N. and Reuber, T.L. 2004. Regulation of disease resistance pathways by AP2/ERF transcription factors. *Curr. Opin. Plant Biol.* 7: 465–471.
- Heim, M.A., Jakoby, M., Werber, M., Martin, C., Weisshaar, B. and Bailey, P.C. 2003. The basic helix-loop-helix transcription factor family in plants: A genome-wide study of protein structure and functional diversity. *Mol. Biol. Evol.* 20: 735–747.
- Ji, S.J., Lu, Y.C., Feng, J.X., Wei, G., Li, J., Shi, Y.H., Fu, Q., Liu, D., Luo, J.C. and Zhu, Y.-X. 2003. Isolation and Analyses of Genes Preferentially Expressed During Early Cotton Fiber Development by Subtractive PCR and cDNA Array. *Nucleic Acids Res.* 31: 2534–2543.
- Jiang, C., Gu, X. and Peterson, T. 2004. Identification of conserved gene structures and carboxy-terminal motifs in the Myb gene family of *Arabidopsis* and *Oryza sativa* L. ssp. *indica*. *Genome Biol.* 5: R46.
- Jiao, Y., Yang, H., Ma, L., Sun, N., Yu, H., Liu, T., Gao, Y., Gu, H., Chen, Z., Wada, M., Gerstein, M., Zhao, H., Qu, L.-J. and Deng, X.W. 2003. A genome-wide analysis of blue-light regulation of *Arabidopsis* transcription factor gene expression during seedling development. *Plant Physiol.* 133: 1480–1493.
- Jones, D.T., Taylor, W.R. and Thornton, J.M. 1992. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 8: 275–282.
- Kagaya, Y., Ohmiya, K. and Hattori, T. 1999. RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res.* 27: 470–478.
- Kim, H., Snesrud, E.C., Haas, B., Cheung, F., Town, C.D. and Quackenbush, J. 2003. Gene expression analyses of *Arabidopsis* chromosome 2 using a genomic DNA amplicon microarray. *Genome Res.* 13: 327–340.

- Knight, H., Veale, E.L., Warren, G.J. and Knight, M.R. 1999. The *sfr6* mutation in *Arabidopsis* suppresses low-temperature induction of genes dependent on the CRT/DRE sequence motif. *Plant Cell* 11: 875–886.
- Kohler, C., Hennig, L., Spillane, C., Pien, S., Grissem, W. and Grossniklaus, U. 2003. The polycomb-group protein ME-DEA regulates seed development by controlling expression of the MADS-box gene PHERES1. *Genes Dev.* 17: 1540–1553.
- Lin, X., Kaul, S., Rounsley, S., Shea, T.P., Benito, M.-I., Town, C.D., Fuji, C.Y., Mason, T., Bowman, C.L., Barnstead, M., Feldblyum, T.V., Buell, C.R., Ketchum, K.A., Lee, J., Ronning, C.M., Koo, H.L., Moffat, K.S., Cronin, L.A., Shen, M., Pai, G., Aken, S.V., Umayam, L., Tallon, L.J., Gill, J.E., Adams, M.D., Carrera, A.J., Creasy, T.H., Goodman, H.M., Somerville, C.R., Copenhaver, G.P., Preuss, D., Niernan, W.C., White, O., Eisen, J.A., Salzberg, S.L., Fraser, C.M. and Venter, J.C. 1999. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* 402: 761–768.
- Ma, L., Sun, N., Liu, X., Jiao, Y., Zhao, H.Y. and Deng, X.W. 2005. Organ-specific expression of *Arabidopsis* genome during development. *Plant Physiol.* 138: 1–12.
- Magnani, E., Sjolander, K. and Hake, S. 2004. From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. *Plant Cell* 16: 2265–2277.
- Magome, H., Yamaguchi, S., Hanada, A., Kamiya, Y. and Oda, K. 2004. dwarf and delayed-flowering 1, a novel *Arabidopsis* mutant deficient in gibberellin biosynthesis because of overexpression of a putative AP2 transcription factor. *Plant J.* 37: 720–729.
- Okamoto, J.K., Caster, B., Villarreal, R., van Montagu, M. and Jofuku, K.D. 1997. The AP2 domain of *APET ALA2* defines a large new family of DNA binding proteins in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* 94: 7076–7081.
- Page, R.D.M. 1996. TreeView: An application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12: 357–358.
- Parenicova, L., de Folter, S., Kieffer, M., Horner, D.S., Favalli, C., Busscher, J., Cook, H.E., Ingram, R.M., Kater, M.M., Davies, B., Angenent, G.C. and Colombo, L. 2003. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS world. *Plant Cell* 15: 1538–1551.
- Ramanathan, Y., Zhang, H., Aris, V., Soteropoulos, P., Aaronson, S.A. and Tolia, P.P. 2002. Functional cloning, sorting, and expression profiling of nucleic acid-binding proteins. *Genome Res.* 12: 1175–1184.
- Rhee, S.Y., Beavis, W., Berardini, T.Z., Chen, G., Dixon, D.A., Doyle, A., Garcia-Hernandez, M., Huala, E., Lander, G., Montoya, M., Miller, N., Mueller, L.A., Mundodi, S., Reiser, L., Tacklind, J., Weems, D.C., Wu, Y., Xu, I., Yoo, D., Yoon, J. and Zhang, P. 2003. The *Arabidopsis* information resource (TAIR): a model organism database providing a centralized, curated gateway to *Arabidopsis* biology, research materials and community. *Nucleic Acids Res.* 31: 224–228.
- Riechmann, J.L. and Meyerowitz, E.M. 1998. The AP2/EREBP family of plant transcription factors. *Biol. Chem.* 379: 633–646.
- Riechmann, J.L. and Ratcliffe, O.J. 2000. A genomic perspective on plant transcription factors. *Curr. Opin. Plant Biol.* 290: 2105–2110.
- Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C.-Z., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J.Z., Ghandehari, D., Sherman, B.K. and Yu, G.-L. 2000. *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290: 2105–2110.
- Ronquist, F. and Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
- Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K. and Yamaguchi-Shinozaki, K. 2002. DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochem. Biophys. Res. Comm.* 290: 998–1009.
- Schmid, M., Uhlenhaut, N.H., Godard, F., Demar, M., Bressan, R., Weigel, D. and Lohmann, U. 2003. Dissection of floral induction pathways using global expression analysis. *Development* 130: 6001–6012.
- Schoof, H., Zaccaria, P., Gundlach, H., Lemcke, K., Rudd, S., Kolesov, G., Arnold, R., Mewes, H.W. and Mayer, K.F.X. 2002. MIPS *Arabidopsis thaliana* database (MA-TDB): an integrated biological knowledge resource based on the first complete plant genome. *Nucleic Acids Res.* 30: 91–93.
- Solano, R., Stepanova, A., Chao, Q. and Ecker, J.R. 1998. Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. *Genes & Dev.* 12: 3703–3714.
- Stockinger, E.J., Gilmour, S.J. and Thomashow, M.F. 1997. *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc. Natl. Acad. Sci. USA* 94: 1035–1040.
- Stracke, R., Werber, M. and Weisshaar, B. 2001. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* 4: 447–456.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673–4680.
- Toledo-Ortiz, G., Huq, E. and Quail, P.H. 2003. The *Arabidopsis* basic/helix-loop-helix transcription factor family. *Plant Cell* 15: 1749–1770.
- Yi, S.Y., Kim, J.-H., Joung, Y.-H., Lee, S., Kim, W.-T., Yu, S.H. and Choi, D. 2004. The pepper transcription factor C_oPF1 confers pathogen and freezing tolerance in *Arabidopsis*. *Plant Physiol.* 136: 2862–2874.
- Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. and Grissem, W. 2004. GENEVESTIGATOR. *Arabidopsis* microarray database and analysis toolbox. *Plant Physiol.* 136: 2621–2632.