Molecular Plant Advance Access published October 23, 2007

Molecular Plant • Pages 1–15, 2007

The Development of Protein Microarrays and Their Applications in DNA–Protein and Protein–Protein Interaction Analyses of *Arabidopsis* Transcription Factors

Wei Gong[a,b,*], Kun He[a,b,*], Mike Covington[c], S. P. Dinesh-Kumar[b], Michael Snyder[b], Stacey L. Harmer[c], Yu-Xian Zhu[a] and **Xing Wang Deng[a,b,1]**

a Peking-Yale Joint Center for Plant Molecular Genetics and Agrobiotechnology, College of Life Sciences, and the National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, China
b Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520, USA
c Section of Plant Biology, College of Biological Sciences, University of California, Davis 95616, USA

ABSTRACT We used our collection of *Arabidopsis* transcription factor (TF) ORFeome clones to construct protein microarrays containing as many as 802 TF proteins. These protein microarrays were used for both protein–DNA and protein–protein interaction analyses. For protein–DNA interaction studies, we examined AP2/ERF family TFs and their cognate *cis*-elements. By careful comparison of the DNA-binding specificity of 13 TFs on the protein microarray with previous non-microarray data, we showed that protein microarrays provide an efficient and high throughput tool for genome-wide analysis of TF-DNA interactions. This microarray protein–DNA interaction analysis allowed us to derive a comprehensive view of DNA-binding profiles of AP2/ERF family proteins in *Arabidopsis*. It also revealed four TFs that bound the EE (evening element) and had the expected phased gene expression under clock-regulation, thus providing a basis for further functional analysis of their roles in clock regulation of gene expression. We also developed procedures for detecting protein interactions using this TF protein microarray and discovered four novel partners that interact with HY5, which can be validated by yeast two-hybrid assays. Thus, plant TF protein microarrays offer an attractive high-throughput alternative to traditional techniques for TF functional characterization on a global scale.

INTRODUCTION

Transcription factors (TFs) play important roles in plant cellular and developmental processes by controlling gene expression. Two features stand out when comparing *Arabidopsis* TFs with those of other organisms. First, *Arabidopsis* has a very large group of TFs. In 2000, it was reported that *Arabidopsis* had at least 1533 transcription factors, which was 1.3 times that of *Drosophila* and 1.7 times that of *C. elegans* and yeast (Riechmann, 2000). With the updated *Arabidopsis thaliana* genome annotation, the *Arabidopsis* Transcription Factor database now contains 2304 protein models coded from 1922 loci (Guo et al., 2005; Qu and Zhu, 2006). Secondly, relatively few *Arabidopsis* TFs are well characterized compared with other model organisms, even though it is probably the best studied among plants. Only about 5% of *Arabidopsis* TFs have been characterized genetically in contrast to over 25% of *Drosophila* and *C. elegans* TFs (Ruvkun and Hobert, 1998). This large number of TFs of unknown function calls for a high-throughput means to systematically analyze their roles.

Many studies have shown that plants and animals have evolved independently. Although many developmental processes are logically similar, they are often controlled by non-homologous genes. For example, radial patterning of plant floral organs and *Drosophila* segmental identity are two similar processes that are often used to compare plant and animal development. Both use TFs as master regulators of developmental pattern, but plants mainly use MADS box proteins

[1] To whom correspondence should be addressed. E-mail xingwang.deng@yale.edu.

[*] These authors contributed equally to this work.

doi: 10.1093/mp/ssm009

(Jack, 2001) while *Drosophila* uses homeobox proteins (Veraksa et al., 2000), even though members of both protein families are found in plants and animals. (For a review comparing plant and animal development, see Meyerowitz, 2002.) Therefore, to understand plant regulatory mechanisms, it is necessary to study plant TFs, even if their animal counterparts are well characterized.

Transcription factors consist of, with a few exceptions, a DNA-binding domain and a regulatory domain. *Arabidopsis* TFs are classified into more than 30 families according to the sequence similarities of their DNA-binding domains (Riechmann, 2000). The AP2/ERF (APETALA2 /Ethylene Response Factor), bHLH and MYB families are the three largest.

Functional domains of most *Arabidopsis* TFs have not been characterized. There is thus a pressing need to develop large-scale functional genomics approaches to use the information about their sequences to infer their functions. Many techniques have been used to study TF-DNA interactions, including EMSA (Electrophoretic mobility shift assays, Fried and Crothers, 1981), yeast one-hybrid (Wei et al., 2005), ChIP-chip (Chromatin immunoprecipitation with microarray detection, Liu and Clarke, 2002), and DIP-chip (DNA immunoprecipitation with microarray detection, Liu et al., 2005). However, these methods (except for the yeast one-hybrid) can only be used to determine DNA-binding specificities of individual TFs. In contrast, TF protein microarrays present an approach to examine the DNA-binding activity of the entire TF proteome that is less labor-intensive than the yeast one-hybrid technique (Wei et al., 2005). In addition, protein microarrays have the advantage of a wider variety of applications, since they can be screened for both biochemical activities and protein–protein/DNA/small molecule interaction efficiently, requiring only a few days if a suitable library is available. Thus, while traditional studies of TFs have provided and will continue to provide insightful information, TF protein microarray technology will provide a new and complementary approach to understanding transcriptional regulation.

Protein microarrays have been extensively used to study protein functions in organisms such as bacteria and yeast (Ding et al., 2006; Bertone and Snyder, 2005). For example, a large-scale analysis of the DNA-binding activity of the yeast proteome by probing protein microarrays with labeled yeast genomic DNA was reported in which about 200 positive genes were identified, and eight were tested by the ChIP-chip method (Hall et al., 2004). Recently, a specific protein microarray of 282 yeast TFs was constructed and used for analyzing their interaction with DNA probes (Ho et al., 2006).

The availability of the completed *Arabidopsis* genome sequence (AGI, 2000) allows a proteomic-scale characterization of *Arabidopsis* transcription factors. In a previous report, we generated an ORFeome collection of *Arabidopsis* TF genes and a parallel collection of yeast strains which express the corresponding TFs with several epitope tags fused to their C-terminus (Gong et al., 2004). Here, we report the construction of *Arabidopsis* TF protein microarrays and its use in analyzing TF-DNA binding specificity and TF–protein interactions.

RESULTS AND DISCUSSION

Construction of *Arabidopsis* Transcription Factor Protein Microarrays

The epitope tag fused to the C-terminus of all TF ORFs in our collection includes three copies of Flag, one His6, and two copies of the IgG-binding motif from protein A (Gong et al., 2004). We purified as many as 802 *Arabidopsis* TF proteins from the ordered yeast library using a high-throughput procedure (see Methods). The amount of protein purified from each yeast strain was quantified by Western blot analysis with anti-His antibody. Proteins forming prominent bands were obtained from about 80% of the yeast strains (Figure 1A). The purified proteins were spotted on microarray slides to produce the TF protein microarrays. To routinely check how much fusion

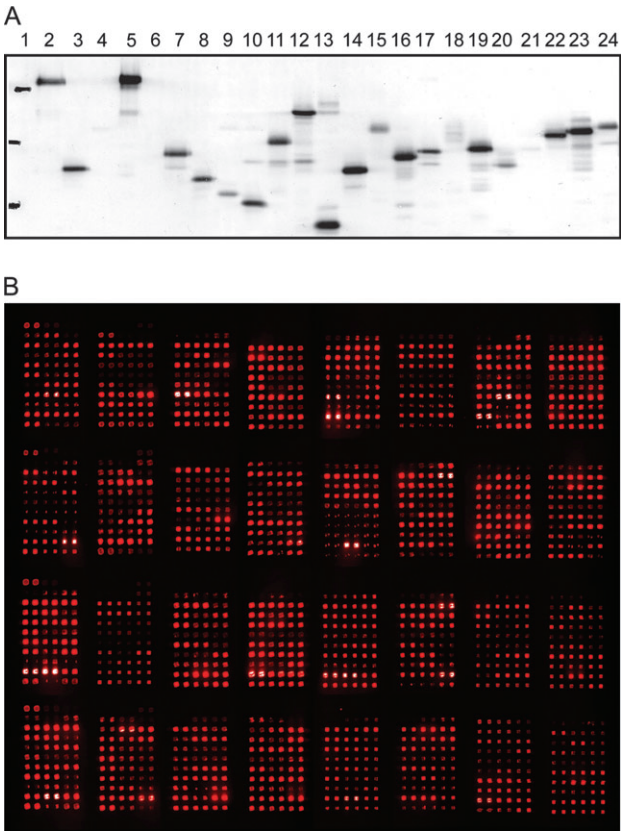

Figure 1. Construction of *Arabidopsis* TF Protein Microarrays.

(A) Detection of TF fusion protein expression by Western blotting. Yeast strains containing TF ORFs were induced for expression of fusion proteins. Each purified fusion protein was subjected to SDS-PAGE and analyzed by Western blotting with anti-His antibody (Methods). Lane 1 is protein standards. Lanes 2–24 are representative samples.

(B) A representative fluorescent image of anti-His antibody probes of the *Arabidopsis*TF protein microarray.

protein was attached to the slides and whether the slides were properly printed, a semi-quantitative assay was done for each batch of printed slides by probing representative microarrays with anti-His antibody (Figure 1B). Normally, more than 85% of protein samples gave signals significantly above background on the microarray slides (the signal to background ratio is $\geqslant 2.0$).

The AP2/ERF Family as a TF Model Family for DNA Binding Analysis

The function of most TFs involves binding to specific DNA sequences and interacting with other proteins. Interactions between *cis*-acting DNA elements and *trans*-acting transcription factors have been extensively studied. The plant AP2/ERF family has been well characterized, with full-length ORFeome clones available for all but two, and DNA binding properties of its representative subfamily members have been reported using traditional means, so they were selected as a sample group for this study. There are 147 AP2/ERF TFs in the *Arabidopsis* genome (Feng et al., 2005; Nakano et al., 2006), which are usually divided into four subfamilies: AP2, DREB, ERF, and RAV, of which DREB and ERF members account for over 85% of the whole family (Sakuma et al., 2002; Feng et al., 2005). Members of different subfamilies were reported to display distinct DNA-binding activities. For example, several ERF proteins bind to the GCC box: AGCCGCC (Ohme-Takagi and Shinshi, 1995; Buttner and Singh, 1997; Zhou et al., 1997; Hao et al., 1998; Fujimoto et al., 2000; Hao et al., 2002), while some proteins of the DREB subfamily bind the DRE (dehydration response element) or the C-repeat element as defined by A/GCCGAC consensus (Baker et al., 1994; Yamaguchi-Shinozaki and Shinozaki, 1994; Jiang et al., 1996; Stockinger et al., 1997; Thomashow, 1999; Hao et al., 2002; Sakuma et al., 2002). The ANT protein of the AP2 subfamily is reported to bind a consensus sequence: gCAC(A/G)N(A/T)TcCC(a/g)ANG(c/t) (Nole-Wilson and Krizek, 2000; Krizek, 2003). The RAV1 protein of the RAV subfamily binds specifically to bipartite sequences comprising two motifs—CAACA and CACCTG (Kagaya et al., 1999). In total, the DNA-binding activity of 13 AP2/ERF family members has been characterized (Table 1).

Optimization of TF-DNA Binding Analysis Using Protein Microarrays

To evaluate the efficacy of detecting protein–DNA binding activity with protein microarrays, double-stranded DNA probes were designed based on the known binding sites defined for representative members of four AP2/ERF subfamilies. These probes were derived from the previously reported ANT-AP2R1R2, GCC-box, DRE, and RAV1 sequences, respectively. To avoid non-specific binding, for each Cy3-labeled wild-type probe, a Cy5-labeled mutant probe was designed. Mutant probes substituted core motif sequences with alternative As and Ts (see Methods). Thus, four pairs of wild-type (WT) and mutant (Mt) probes were made, designated ABE, MtABE, EBE, MtEBE, DBE, MtDBE, RBE, and MtRBE. Each pair of WT and Mt probes, after proper labeling, was mixed and applied to microarray slides. After proper washing and scanning, the Cy3/Cy5 ratios for each spot were derived for analysis.

After signal normalization of replicated microarray assays, the rank of the ratios for each probe pair was calculated. We found that the rank, rather than the absolute ratio, is a more appropriate indicator, possibly for the following reasons. First, the ratio of each specific DNA probe pair represented the relative binding specificity of the wild-type versus the mutant probe.

Table 1. DNA-Binding Activity of Previously Characterized AP2/ERF Members on Protein Microarrays. Transcription factors were divided into subfamilies according to their amino acid sequence similarities. The fourth column indicates the previously-reported *cis*-element bound by each TF. In the binding probe column, probes which had stronger binding affinities on protein microarrays are indicated in bold.

Locus ID	Gene name	Subfamily	Cognate *cis*-element	Binding probe	Reference
At4g37750	ANT	AP2	ANT-AP2R1R2	**ABE**	Krizek, 2003; Nole-Wilson et al., 2000
At3g23240	ERF1	ERF	GCC-box	**EBE**	Solano et al., 1998
At3g16770	AtEBP	ERF	GCC-box	Not detected	Buttner and Singh, 1997
At4g17500	AtERF-1	ERF	GCC-box	ABE, **EBE**	Hao et al., 1998 Fujimoto et al., 2000
At5g47220	AtERF-2	ERF	GCC-box	**EBE**	Fujimoto et al., 2000
At1g50640	AtERF-3	ERF	GCC-box	Not detected	Fujimoto et al., 2000
At3g15210	AtERF-4	ERF	GCC-box	Not detected	Fujimoto et al., 2000
At5g47230	AtERF-5	ERF	GCC-box	**EBE**, DBE	Fujimoto et al., 2000
At4g25480	CBF3/DREB1A	DREB	DRE	ABE, **DBE**	Sakuma et al., 2002
At4g25490	CBF1/DREB1B	DREB	GCC-box, DRE	**EBE, DBE, RBE**	Stockinger et al., 1997; Hao et al., 2002
At5g11590	TINY2	DREB	GCC-box, DRE	Not detected	Wei et al., 2005
At5g05410	DREB2A	DREB	GCC-box, DRE	Not detected	Sakuma et al., 2002
At1g13260	RAV1	RAV	RAV1BE	DBE, **RBE**	Kagaya et al., 1999

Since different mutant probes were used for each of the four wild-type probes, the ratios are not directly comparable. Secondly, some probes may be specific to one TF subfamily, but others may be bound by multiple TF subfamilies or even families.

We closely examined 13 AP2/ERF TFs whose DNA-binding activities were previously studied with traditional methods. Eight showed DNA-binding activity on protein microarrays with at least one pair of probes (Table 1). Among them, ANT (At4g37750) belongs to the AP2 subfamily; ERF1 (At3g23240), AtERF-1 (At4g17500), AtERF-2 (At5g47220) and AtERF-5 (At5g47230) belong to the ERF subfamily; At4g25480 (CBF3/DREB1A) and At4g25490 (CBF1/DREB1B) belong to the DREB subfamily and RAV1 (At1g13260) belongs to the RAV subfamily. All eight TFs showed similar DNA-binding specificities to previously reported non-microarray data and three bound only one probe. Thus, the proteins in our microarray still have the correct DNA-binding specificity.

In this analysis, five previously studied TFs did not detectably bind any probes. Four of the five had prominent bands of the expected size; thus, it is possible that those four proteins may not fold correctly for specific DNA binding. The TINY2 protein (At5g11590) was not detected in Western blot analysis. It has been shown previously that TINY2 binds the GCC-box weakly and DRE strongly (Wei et al., 2005). Therefore, to test whether the low abundance of TINY2 protein was the reason that it did not detectably bind any probes, we generated a GST-TINY2 recombinant protein in *E. coli*. Both the *E. coli* and yeast recombinant TINY2 proteins were purified, spotted on a microarray slide, and assayed for DNA-binding activity. As shown in Figure 2A, TINY2 purified from *E. coli* bound target elements in a pattern consistent with the previous report (Wei et al., 2005), whereas TINY2 purified from yeast occasionally displayed similar DNA binding but with poor reproducibility (data not shown). The poor reproducibility is likely due to the variability in amounts of TINY2 extracted from yeast.

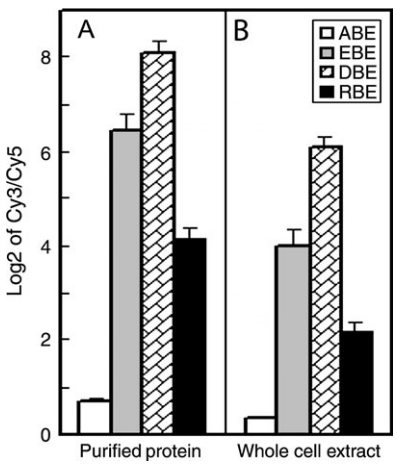

Figure 2. Relative DNA-Binding Affinity of GST-TINY2 Purified from Two Different Hosts.

Immobilized GST-TINY2 purified from *E. coli* **(A)** and *E. coli* whole-cell extract containing GST-TINY2 **(B)** was probed with ABE, EBE, DBE, and RBE. The DNA-binding activities were compared by calculating the log2-based Cy3:Cy5 intensity ratio with different DNA probes.

To examine the effects of protein purity on their behavior on protein microarrays, bacterial whole cell extracts containing TINY2, instead of purified recombinant proteins, were printed on microarray slides and their DNA-binding activity was analyzed. As shown in Figure 2B, the whole cell extracts containing TINY2 have similar DNA-binding specificity as the purified protein preparations. This observation agrees with many EMSA experiments in which TFs bind specific DNA sequences regardless of protein purity (Harmer and Kay, 2005). These findings suggest that the purity of the extracts used to prepare the microarray may not be a limiting factor for protein microarray analysis of DNA–TF interactions.

ANT and RAV1 are the only members of their respective TF subfamilies whose cognate DNA sequences, designated ABE and RBE elements, respectively, have been characterized (Nole-Wilson and Krizek, 2000; Kagaya et al., 1999). Our protein microarray analysis showed that ANT and RAV1 had the highest binding affinity (Figure 3A and 3C) and specificity (Figure 3B and 3D) with ABE and RBE elements, respectively. This result demonstrates that TFs immobilized on the surface of microarray slides can recapitulate the DNA-binding affinities and specificities measured by techniques that use proteins and DNA in solution.

Protein Microarray Analysis Detects Specific DNA-Binding Activity of Many Previously Uncharacterized AP2/ERF Family Members

In addition to the eight previously reported TFs, our microarray analysis also detected specific DNA-binding by 48 previously uncharacterized members of the AP2/ERF family of transcription factors (Table S1). This new group of TFs with demonstrated DNA-binding activity includes well described genes such as ABI4 (ABSCISIC ACID-INSENSITIVE PROTEIN 4, At2g40220, Finkelstein et al., 1998), RAV2 (related to ABI3/VP1 2, At1g68840, Kagaya et al., 1999), ESR1 (DORNROSCHEN/ENHANCER OF SHOOT REGENERATION1, At1g12980, Banno et al., 2001), DRNL (DORNROSCHEN-LIKE, ESR2, At1g24590, Ikeda et al., 2006), PLT2 (PLETHORA 2, At1g51190, Aida et al., 2004) and DDF1 (DWARF AND DELAYED-FLOWERING 1, At1g12610, Magome et al., 2004). Figures 4A and 4B show sample images of TF-DNA microarray binding assays for representative previously uncharacterized ERF and DREB members illustrating specific DNA-binding activity.

It has been reported previously that some TFs such as DREB2A and CBF1 bind both the GCC-box and DRE (Liu et al., 1998; Sakuma et al., 2002; Hao et al., 2002). Here, our microarray analysis also revealed that many of the previously uncharacterized members of the DREB and ERF subfamilies bound both the GCC-box and DRE probes. This may occur because the GCC-box and DRE motif are similar sequence elements, sharing a common core CCGNC. Indeed, in our microarray analysis, CBF1 (At4g25490) binds both EBE and

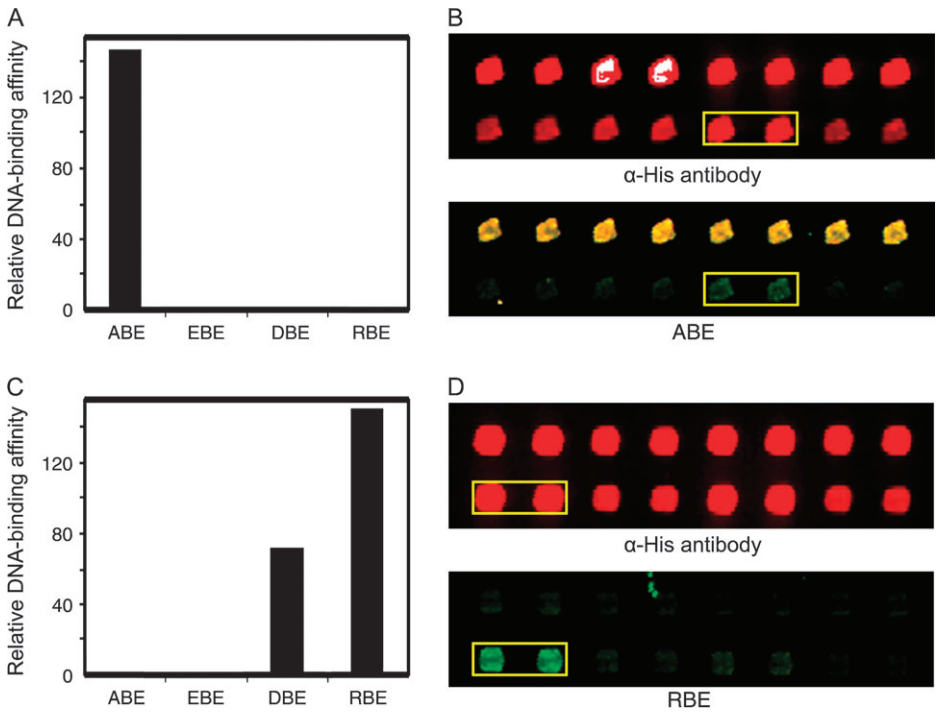

Figure 3. DNA-Binding Affinity and Specificity of ANT and RAV1.

(A, C) Relative DNA-binding affinity of ANT and RAV1 with four DNA probes.

(B, D) Images of ABE-binding and RBE-binding assays. Immobilized proteins printed in duplicate were probed with ABE/MtABE and RBE/MtRBE, respectively (bottom row of each panel). The same protein preparations were probed with anti-His antibody as control (top row of each panel). Green spots show positive signals and yellow spots show nonspecific binding signals. ANT and RAV1 proteins are boxed in yellow.

DBE probes with a stronger affinity for DBE (Figure 4C), consistent with a previous study using a conventional approach (Hao et al., 2002).

The Correlation between DNA-Binding Specificity and DNA-Binding Domain

To obtain a better view of the DNA-binding profile of the entire AP2/ERF family, we carried out a clustering analysis of those TFs with detectable specific DNA-binding activity in the microarray analysis. To this end, we assigned binding levels 0–5 to the AP2/EREBP family TFs (see Methods), where 0 means that a TF does not bind a specific probe, while level 5 means the binding activity is very strong. Using the linkage hierarchical clustering algorithm, we illustrated the binding pattern of the AP2/EREBP family for the four representative DNA elements (Figure 5A). Of 137 AP2/ERF family members examined in our microarray, 56 (41%) TFs detectably bound at least one probe when compared to the mutated version of the probe. These TFs were easily clustered into five groups, based on their DNA-binding specificities, since each group displays very specific binding patterns (Figure 5A). Groups I through IV preferentially bound probes ABE, EBE, DBE, and RAV, respectively, whereas members of group V bound probes EBE and DBE.

Examination of the subfamilies that the members of each group belonged to allowed us to evaluate the relationship between DNA-binding specificity and DNA-binding domain sequence conservation. As shown in Figure 5B, 14 out of 17 TFs in group II are from the ERF subfamily; 12 out of 13 TFs in group III are from the DRE subfamily; four out of five TFs in group IV are members of the RAV subfamily; and all TFs in group V belong to either the ERF or the DREB subfamily. These results indicate that in these members of the AP2/ERF family, DNA-binding specificity is highly correlated with the similarity of DNA-binding domain sequence conservation, which is quite consistent with the prevalent view (see Table 1). Surprisingly, in the case of group I that preferentially bound the ABE probe, more than half were from the ERF and DREB subfamilies rather than from the AP2 subfamily. This indicated that the ABE sequence may contain certain sequence features that can be recognized by some proteins of the ERF and DRE subfamilies. This may be due to the fact that a core ABE sequence element—CCG—is also present in known binding sites of ERF and DREB subfamily members.

We further examined 15 AP2 subfamily members that were on the microarray. Five showed DNA-binding activity, and all five very specifically bound the ABE probe with high affinity (Figure 6A). AP2 subfamily members all possess two AP2 domains called AP2-R1 (AP2 repeat 1) and AP2-R2 (AP2 repeat 2) (Nole-Wilson and Krizek, 2000; Krizek, 2003). It was suggested that the two AP2 domains bind specifically and

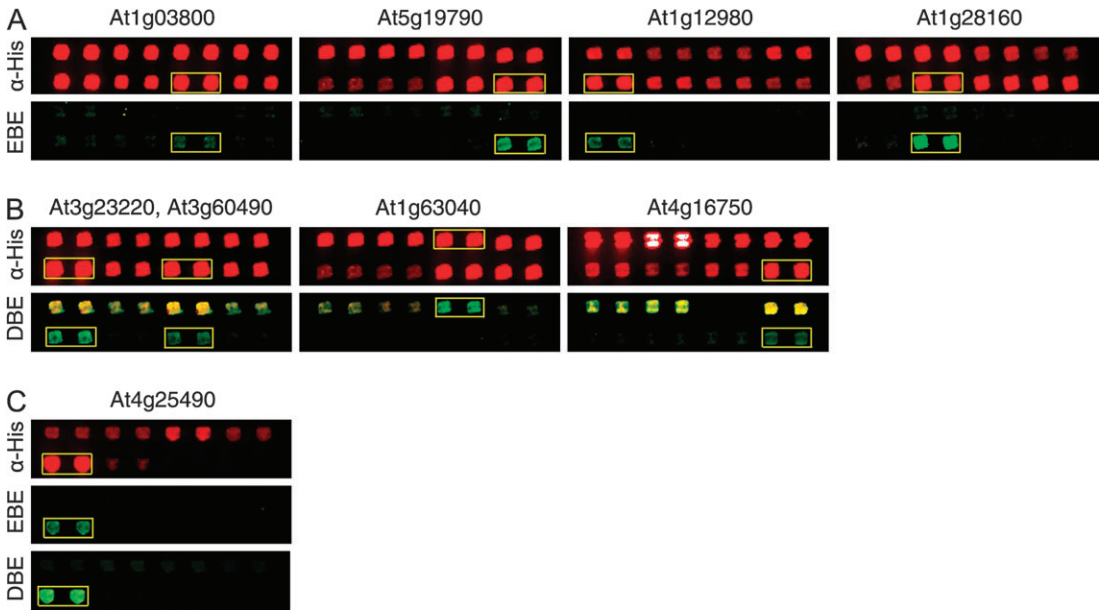

Figure 4. Examples of Protein–DNA Binding Assays on Protein Microarrays.

Immobilized proteins printed in duplicate were probed with EBE and DBE, respectively (bottom row of each panel). The same protein preparations were probed with anti-His antibody as control (top row of each panel). Positive signals (green) and corresponding anti-His signals are boxed in yellow.

(A) Representative EBE-binding images of ERF members.

(B) Representative DBE-binding images of DREB members.

(C) CBF1 (At4g25490) is a DREB member which binds to both EBE and DBE.

cooperatively to bipartite recognition sequences composed of two motifs (Krizek, 2003). We constructed phylogenetic trees of 15 AP2 subfamily members with each of the two AP2 domains, R1 or R2, respectively (Figure 6B and 6C). In both phylogenetic trees, the five positive TFs were grouped into one branch. This result indicated again that domain structural similarity played an important role in determining DNA-binding affinity and specificity.

We noted that some TFs other than AP2/ERF family members also bound our DNA probe designed for this family of TFs. Although we have not ruled out false positives in this situation, this observation is in line with an emerging hypothesis that transcription factors may bind to other sequences in addition to their well defined target motifs. For example, in addition to their expected target sites, E2F (Weinmann et al., 2002) and NF-κB-factor (Martone et al., 2003) were reported to bind many non-consensus sequences.

Using Protein Microarrays to Identify Transcription Factors that Bind the 'Evening Element'

The 'evening element' (EE) is a *cis*-element which was identified as overrepresented in the promoters of evening-phased genes, and multimerized EE is sufficient to confer evening-phased rhythms on a reporter gene (Harmer et al., 2000; Harmer and Kay, 2005). CCA1 (CIRCADIAN CLOCK ASSOCIATED 1) and LHY (LATE ELONGATED HYPOCOTYL) are well characterized dawn-phased genes that both regulate EE-mediated gene expression by binding to EE in the promoter region of their target genes (Harmer and Kay., 2005). They both belong to the CCA1 subfamily of Myb-like transcription factors, which has 10 members: CCA1, LHY, and RVE1 to RVE8 (Yanhui et al., 2006). In addition to CCA1 and LHY, RVE1 (At5g17300) was also shown to bind to EE. EMSA Analysis using crude *Arabidopsis* extracts found that extracts made from plant samples harvested in the afternoon (CT6) and evening (CT12) had much stronger binding activity than samples harvested at dawn and in the morning (Harmer and Kay, 2005). It was thus reasoned that there must be transcriptional regulators other than CCA1 and LHY that are able to bind EE.

To search for novel EE-binding TFs, we used the EE element as a probe to screen a microarray containing 440 TFs coming mainly from the Myb, Myb-like, and AP2/ERF families, together with several small TF families. Forty-one TFs that reproducibly bound to EE were identified. Of the 440 TFs spotted on this microarray, 119 (27%) were Myb proteins and 47 (11%) were Myb-related. By contrast, 15 (37%) and 11 (27%) of the positive TFs were Myb or Myb-related proteins, respectively, showing that these two types of TF were enriched in the pool of EE-binding candidates. To identify TFs within this pool that were expressed in the evening phase, we searched a DNA microarray database for clock-regulated genes in *Arabidopsis* which has expression data for 207 of the 440 proteins used in the microarray (Covington and Harmer, 2007) PLoS Biol 5(8): e222. From the 41 EE-binding candidates, we identified

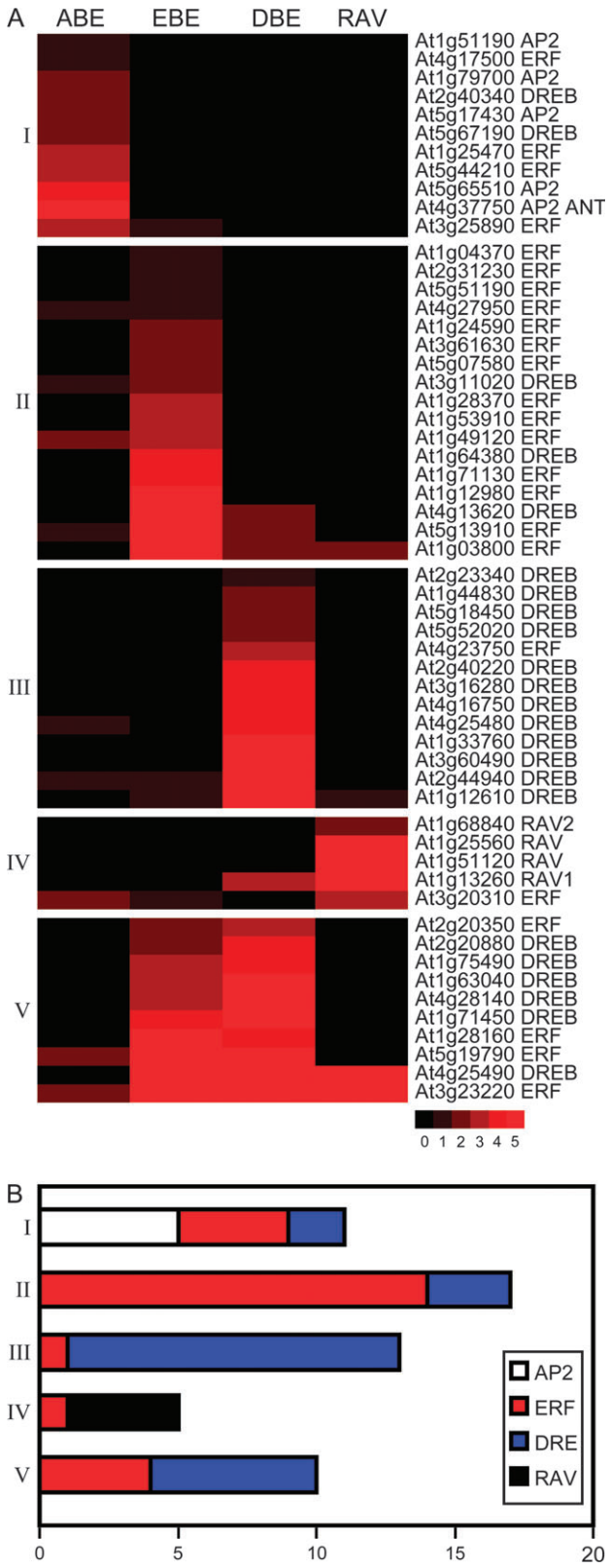

Figure 5. Correlation between DNA-Binding Specificity and DNA-Binding Domain.

(A) Cluster diagram of 56 AP2/ERF TFs which showed positive binding with at least one probe. Panels from left to right are the microarray results with different DNA probes. Panels from top to bottom are the five groups.

(B) Subfamily distribution of the five groups indicated on the left. AP2, ERF, DRE, and RAV subfamilies are indicated as white, red, blue, and black, respectively.

11 TFs displaying clock-regulated expression, including 1 Myb protein and 8 Myb-related proteins (Table 2 and Figure 7). Of the 11 TFs that showed both EE-binding and clock-regulated expression, expression of six peaked at dawn, including CCA1 and LHY. All six dawn-phase genes belong to the CCA1 subfamily

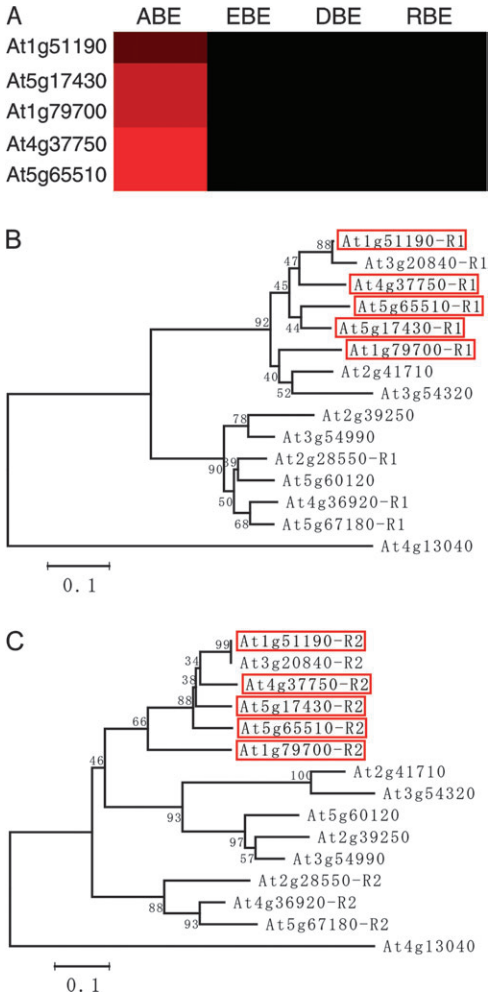

Figure 6. DNA-Binding and Phylogenetic Analysis of AP2 Subfamily Members.

(A) DNA-binding profiles of five positive AP2 subfamily members. Of 15 AP2 subfamily members immobilized on slides, five gave positive signals in TF–DNA binding assays. Panels from left to right indicate the microarray results with different DNA probes. Locus IDs of five TFs are indicated on the left.

(B, C) Unrooted phylogenetic trees of 15 AP2 subfamilies using separated AP2 domains. Protein maximum likelihood trees were constructed according to sequence conservation of AP2-R1 domain (B) and AP2-R2 domain (C). Scale bar corresponds to 0.1 amino acid substitutions per residue. Five positive DNA-binding TFs noted in (A) are boxed in red.

Table 2. Eleven Clock-Regulated TFs that Bind the EE (Evening Element) on Protein Microarrays.

Locus ID	Gene name	Family	Subfamily	Peak phase
At5g17300	RVE1	MYB like	CCA1	0.476
At3g09600	RVE8	MYB like	CCA1	0.48
At1g01060	LHY	MYB like	CCA1	0.702
At2g46830	CCA1	MYB like	CCA1	0.705
At5g02840	RVE4	MYB like	CCA1	1.458
At1g18330	RVE7	MYB like	CCA1	6.82
At5g18680	TLP11	TLP	/	9.272
At4g18390	Unknown	TCP	/	9.842
At5g08520	Unknown	MYB	2R	12.85
At5g37260	RVE2	MYB like	CCA1	20.311
At1g01520	RVE3	MYB like	CCA1	23.766

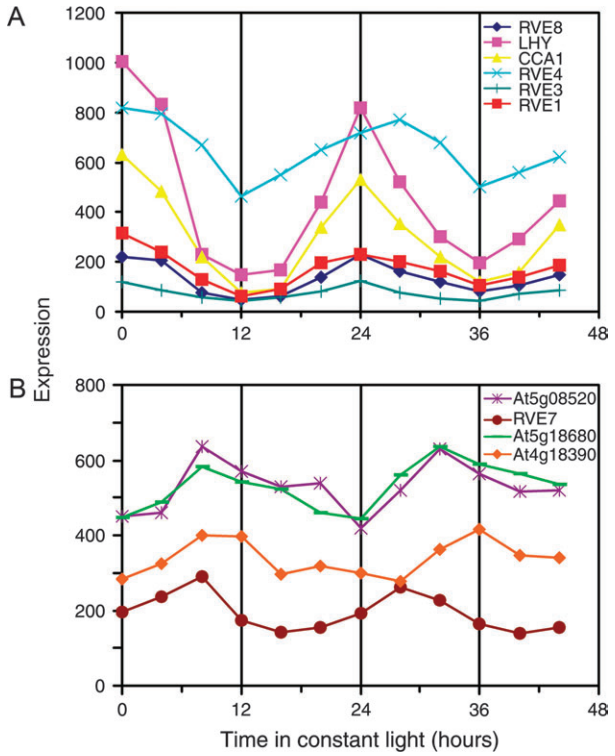

Figure 7. Transcription Patterns of 10 EE-Binding Myb TFs under Constant Light.

(A) Six Myb TFs cycled with dawn phase.
(B) Four Myb TFs cycled with evening phase.
More information about these TFs is given in Table 2.

of the Myb-related family (Figure 7A). Expression of four TFs peaked in the afternoon to evening (Figure 7B), and may represent good candidates for contributing to the high evening phase EE-binding activity observed in *Arabidopsis* extracts. Expression of the last one—RVE2—peaked in between.

Using the Protein Microarrays to Detect TF Interactions with Other Proteins

To develop a protein microarray-based assay for detecting interactions of TFs with other proteins, we selected representative factors from light regulatory pathways. Light is one of the most influential environmental factors that regulate plant development, and the products of many genes including TFs function in light signal transduction pathways (Deng and Quail, 1999; Jiao et al., 2007). One of the proteins we chose is COP1 (CONSTITUTIVE PHOTOMOPHOGENIC 1)—a key repressor of photomorphogenesis. COP1 is known to interact with several TFs, including HY5 (LONG HYPOCOTYL 5, Ang et al., 1998), HYH (Holm et al., 2002), HFR1 (Duek et al., 2004), and others. Our second choice was HY5—a bZIP family transcription factor which is a positive regulator of photomorphogenesis and binds directly to the promoters of many light-responsive genes to affect their transcription (Chattopadhyay et al., 1998; Lee et al., 2007). Besides interacting with COP1, HY5 can also act as a homodimer, heterodimer with HYH (Holm et al., 2002) and interact with SPA1 (Saijo et al., 2003). Thus, both proteins are known to interact with multiple partners in light-signaling pathways.

We first evaluated the effectiveness of detecting TF–protein interactions on protein microarrays. Six genes encoding proteins known to interact with COP1 or HY5 based on prior studies were cloned into *E. coli* expression vectors adding a peptide tag. The recombinant fusion proteins were purified and used to construct a mini-microarray. Corresponding peptide tags were spotted next to target proteins as negative controls. We then used this mini-microarray to optimize the assay procedure based on the known protein–protein interactions. As shown in Figure 8, of the nine anticipated interactions among the six proteins, four (COP1-HFR1, COP1-COP1, HY5-HYH, and HY5-HY5) showed strong binding, three (COP1-HYH, COP1-SPA1, and HY5-COP1) showed weak binding, and the last two (COP1-PHYA and HY5-SPA1) displayed no significant binding using an optimized protein microarray assay procedure (see Methods).

We then used this optimized procedure to probe protein microarrays containing 802 TFs with biotinylated GST-HY5 and found 20 candidates that reproducibly bound HY5 (Table 3). To validate their interaction with HY5, we randomly picked 10 candidates and performed yeast two-hybrid assays with a series of truncated HY5 proteins (for unknown reasons, we were unable to express full-length LEXA-HY5 fusions in yeasts). As shown in Figure 9, interaction with HY5 can be validated by the yeast assay for at least four of the 10 candidates. Interestingly, this analysis also revealed that the HY5 bZIP domain, which was presumed to perform dimerization and DNA-binding (Ang et al., 1998), seemed to play an important role in HY5 binding, with the four transcription factors showing positive interactions in our yeast assay. In a previous study, the N-terminal 77 amino acids of HY5 were shown to be necessary and sufficient to mediate direct interaction with COP1 (Ang et al.,

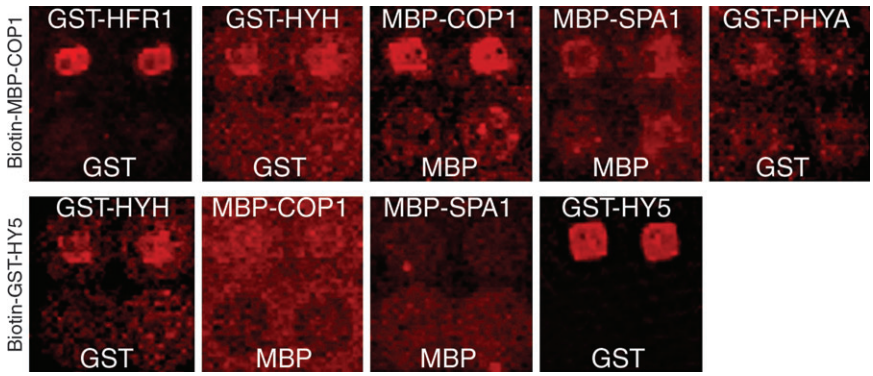

Figure 8. Detecting Interactions between Known Proteins on Protein Microarrays.

Top panels: Images of protein–COP1 binding. Tagged HFR1, HYH, COP1, SPA1, and PHYA were immobilized on slides next to their corresponding tags as indicated.

Bottom panels: Images of protein–HY5 binding. Tagged HYH, COP1, SPA1, and HY5 were immobilized on slides next to their corresponding tags as indicated.

Table 3. TFs that Bind HY5 on Protein Microarrays.

Locus ID	Family
At5g19790	AP2/ERF
At2g41130	bHLH
At5g21960	AP2/ERF
At1g25330	bHLH
At4g13480	MYB
At3g56980	bHLH
At5g18450	AP2/ERF
At5g57150	bHLH
At5g04150	bHLH
At1g73870	C2C2-co-like
At3g47640	bHLH
At3g23230	AP2/ERF
At5g48560	bHLH
At4g01120	bZIP
At4g25560	MYB
At1g09530	bHLH
At3g56770	bHLH
At5g46760	bHLH
At4g13040	AP2/ERF
At4g11140	AP2/ERF

1998). Our result suggested that HY5 can interact with different target proteins using different motifs, which is consistent with the fact that HY5 is a key regulator which integrates input from several different photoreceptors in the light signaling network (Lee et al., 2007).

CONCLUSION

In this study we demonstrated the utility of TF protein microarrays for detecting both TF–DNA and TF–protein interactions. This is the first time that a plant protein microarray has been constructed and used to study TF-DNA and TF–protein interactions on a global scale. We have presented evidence that protein microarrays can be powerful tools for studying protein DNA-binding activities as well as protein–protein interactions. We have also provided data confirming that the amino acid sequences of the DNA-binding domains of AP2/ERF family TFs largely determine their DNA-binding specificities. For the AP2/ERF TF family, our microarray analysis has characterized DNA binding properties of 49 new TFs for the first time. Further, we have used this TF protein microarray to identify four candidate TFs which bound the EE and have the proper phased gene expression pattern. We have also developed procedures for detecting protein interactions using TF protein microarrays and identified four novel proteins that bind HY5, which can be validated by yeast two-hybrid assays. Plant TF protein microarrays thus offer an attractive high-throughput alternative to traditional techniques for functional characterization of TFs on a global scale.

METHODS

High-Throughput TF Protein Expression and Purification from Yeast

Yeast fusion proteins were expressed in 96-well format, as described by Zhu et al. (2000), with minor modifications. Please see supplemental materials and methods for the purification protocol.

Western-Blotting of Purified Fusion TF Proteins

Each purified fusion protein was quantified by Western blotting according to Gong et al. (2004).

Construction of Protein Microarrays

Purified fusion proteins in 96-well plates were printed on FAST slides (Schleicher and Schuel, Florham Park, NJ, USA) in

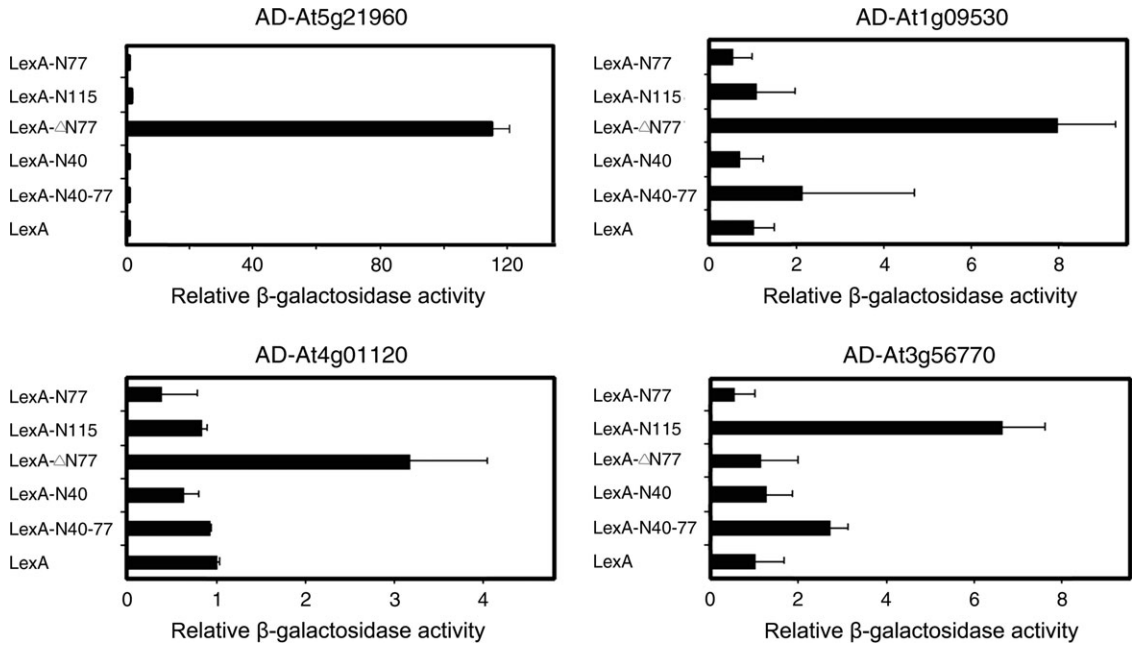

Figure 9. Yeast Two-Hybrid Interactions between Truncated HY5 Proteins and Candidate Partners Detected by Protein Microarrays.
β-galactosidase activities of colonies expressing LexA fused to truncated HY5 proteins and candidate TFs fused to the Activating Domain. Data are averages of four to six individual primary colonies. Error bars represent standard deviations.

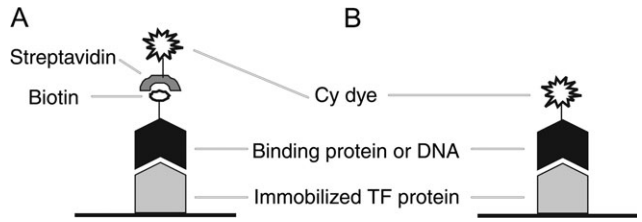

Figure 10. Strategies for Detecting Protein–DNA and Protein–Protein Interactions.
(A) Sandwich assay for protein–protein interactions.
(B) Direct assay for protein–DNA interactions.

duplicate with a VersArray ChipWriter system (BioRad, Hercules, CA, USA).

Recombinant Plasmids

DNA fragments were subcloned into pEG202 and pJG4.5 (Ausubel et al., 1994) to generate LexA-HY5 and AD-TF constructs. The LexA-Hy5 deletions, as well as GST-HY5 constructs, were generated as described previously (Ang et al., 1998). Full-length COP1 coding sequence was cloned into pMAL-c2X (NEB, Ipswich, MA, USA) to express the MBP–COP1 fusion protein.

Preparation of Probes

DNA probes: Biotinylated oligonucleotides were obtained from the W. M. Keck Facility (http://keck.med.yale.edu/). All oligonucleotides were 50 bases long and each contained two or four copies of a conserved binding motif. The sense strand sequences of binding motifs and DNA probes are shown in Table S2. Double-stranded DNA probes were prepared by denaturing and annealing equimolar amounts of biotinylated oligos (100 μM) into double-stranded DNA (50 μM, about 2 μg/μl).

Protein probes: GST-HY5 and MBP-COP1 were expressed in *E. coli* and purified with Glutathione Sepharose 4 Fast Flow (Amersham, Piscataway, NJ, USA) and amylose resin (NEB, Ipswich, MA, USA), respectively, according to the manufacturer's protocol. Just before eluting proteins from beads, BSA and EZ-Link Sulfo-NHS-LC-LC-Biotin (Pierce, Rockford, IL, USA) were added to the slurry to a final concentration of 2 mg/ml and 4 μg/ml, respectively. After incubating at 4°C for 2 h, resin was washed and protein was eluted. Purified protein samples were used for probing slides immediately or stored at –80°C.

Protein Attachment and Detection Optimization in Protein Microarrays

Slide solid supports can influence the efficiency of protein attachment as well as the degree of nonspecific binding. Since various slide substrata have been developed in the past few years, we compared nitrocellulose-coated slides and aldehyde-treated slides using the semi-quantitative assay. Our comparison demonstrated that nitrocellulose-coated slides had much stronger signal intensity (due to its high protein-binding capacity) but lower signal-to-noise ratio (nitrocellulose may cause light scatter with laser scanner detection methods, Kukar et al., 2002). After extensive experimentation,

we optimized binding buffer components and binding conditions so that we reduced the background to a reasonable level and kept the foreground signal level intact. Therefore, we used nitrocellulose-coated slides in the subsequent experiments.

Generally, there are two detection strategies for functional protein microarrays. Sandwich assays use biotinylated probes and subsequently detect them with fluorescent dye-labeled streptavidin (Figure 10A), while direct assays use probes labeled directly with fluorescent dyes (Figure 10B). We chose different strategies for the TF–DNA binding and TF–protein binding assays because of the different properties of the probes. DNA probes were relatively pure and highly uniform, so we used the direct labeling method to detect binding signals, taking advantage of the fact that direct detection has higher specificity and fewer incubation steps. In contrast, the efficiency of protein labeling is low, which is likely due to protein degradation and cross-labeling of nonspecific proteins. Moreover, in a complex solution, the sensitivity of protein–protein interactions is lower than protein–DNA interactions because both proteins must be folded properly to interact. We therefore used the indirect sandwich strategy with protein probes to increase the binding sensitivity by multiple signal amplification steps and to lower unspecific binding by additional stringent wash steps.

Protein Array Data Analysis

Positive signals in the semi-quantitative assays all have a signal-to-background ratio of $\geqslant$2.0. Local background intensity was calculated by the Genepix software; then the limma package of the *R* statistical project was used for the data analysis (Smyth, 2004). After within-slide normalization, the spots with intensities lower than the negative control spots were removed. The retained spots were fitted to a single linear model by the function lmFit to obtain each probe's log-transformed Cy3/Cy5 ratio. To better demonstrate the relative binding affinities of TF and DNA probes, the rank of the Cy3/Cy5 ratio was calculated for each TF; higher rank value indicates higher relative DNA-binding affinity (Figure3). Then, all the AP2/ERF TFs that showed positive binding signals were evenly divided into five groups according to the low-to-high order of ratio ranks and assigned level values from 1 to 5, respectively, and the other AP2/ERF TFs were assigned level 0. To investigate the DNA probe-binding specificities within the AP2/ERF family, the levels of AP2/ERF TFs that showed positive binding signal with at least one probe were clustered using the CLUSTER program with an average-linkage hierarchical clustering algorithm (Eisen et al., 1998).

In the TINY2–DNA binding assays, we compared the relative binding affinities for four probes using the Cy3/Cy5, but not the rank ratio, because the purified target protein abundance from *E. coli* is much higher than that from yeast, so the ratios of proteins purified from different organisms are not comparable.

Yeast Two-Hybrid Analysis

Truncated LexA-HY5 and AD-TF constructs were transformed into EGY48 and L40 strains, respectively. Relative β-galactosidase activities were calculated according to Ausubel et al. (1994). For unknown reasons, we could not stably express LexA-HY5 protein after many attempts; therefore, we could not test the full-length HY5-TF interactions in yeast.

Supplementary Material

The following supplementary material is available for this article online:

Yeast Protein Purification in 96-Well Format

Yeast Protein Purification in 96-Well Format

This protocol was developed from the protocol described by Zhu et al. (2000).

1. The frozen culture in a 96-well box was transferred from –80°C to ice and 100ul of 0.5mm acid-washed glass beads (Sigma, St Louis, MO, USA) was added to each well. While the culture was still frozen, 250 ul of lysis buffer containing fresh protease inhibitors was added. A 96-Well square format silicone capmat (Whatman, Florham Park, NJ, USA) was used to seal each well. After thawing for 15 min on ice, the cells in the 96-well box were broken by shaking for one minute in a paint-shaker at 4°C four times at 2 min intervals.

2. After centrifuging at 3000 rpm for 5 min, the supernatants were collected and transferred to an 800 ul 96-well unifilter plate (Whatman, Florham Park, NJ, USA) placed on top of a 96-well deep well microplate (VWR, West Chester, PA, USA).

3. To obtain more proteins, 250 ul of lysis buffer containing fresh protease inhibitors was added to the cell debris, and the cells were lysed a second time and the lysates were transferred to the same 96-well unifilter plate.

4. The combined cell lysates were centrifuged through the filter plate into the 96-well deep well microplate for 2 min at 3000 rpm. The volume of filtered lysate in each well was roughly 500 ul.

5. Meanwhile, Ni-NTA Agarose (30 ul of beads per sample) (Qiagen, Valencia, CA, USA) was washed four times with cold lysis buffer without protease inhibitors, and finally resuspended in 5X of its original volume with lysis buffer containing fresh protease inhibitors.

6. 6 100 ul of washed Ni-NTA Agarose was added to each well and sealed tightly with a cap mat (VWR, West Chester, PA, USA). The beads and the lysates were rotated end-over-end at 4°C for two hours.

7. The beads were collected by centrifuging at 2000 rpm for 2 minutes and the supernatant was discarded. Beads were washed twice with 200ul wash buffer, then the beads were resuspended in 200 ul wash buffer and the slurry was

transferred to a cold filter plate and wash buffer was removed by centrifuging for 1 minutes at 2000 rpm.

8. 8 After complete removal of the buffer, 35 ul of elution buffer was added to each well and incubated with beads for half an hour at 4° C.

9. The filter plate was placed onto a 96-well PCR plate then the eluate was collected by centrifuging for 1 min at 3000 rpm (how many gs?).

10. Each purified protein was aliquoted into three 96-well PCR plates and immediately stored in a –80°C freezer.

Probing protein microarrays

For the semi-quantitative assays slides were blocked with PBS containing 1% BSA and probed with Penta His Antibody (Qiagen, Valencia, CA, USA) at 1 μg/ml for 1 h. After washing three times, Fluorolink Cy5 labeled goat anti-mouse IgG (H+L) (Amersham, Piscataway, NJ, USA) was added to PBS (plus 0.5% Tween 20) at 250 ng/ml and incubated with slides for 1 h. Slides were then washed three times, dried and imaged with a GenePix 4000B scanner ((Axon™ Instruments, Union City, CA, USA).

For TF-DNA binding assays, slides were blocked with 1% BSA and 0.1 μg/μl poly $(dA\text{-}dT)_n$ for 1 h. Pairwise combinations of wild-type probe and an equal amount of respective mutant probe were added into DNA-binding buffer (25 mM HEPES pH 7.9, 40 mM KCl, 0.5 mM EDTA, 1% BSA, 0.5 mM DTT, 0.1% Tween 20, 0.1 μg/μl poly (dA-dT)) at a concentration of 5 μg/ml and incubated with slides for 2 h. Slides were then washed five times, dried and imaged with a GenePix 4000B scanner.

For TF-protein binding assays slides were blocked with 1% BSA and probed with biotinylated protein probes for 2 h at room temperature. After washing four times, fluorolink Cy 5 labeled streptavidin (Amersham, Piscataway, NJ, USA) was added to binding buffer at a concentration of 10 ng/ml and incubated with slides for 1 h. Slides were then washed, dried and imaged with a GenePix 4000B scanner. A negative control assay was done in parallel with TF-HY5 binding assay by omitting biotinylated GST-HY5 from the binding buffer. Proteins showing strong binding signals on negative control slides were not taken into account.

HY5 binding buffer: 50 mM Tris (pH 7.5), 150 mM NaCl, 1% BSA, 0.5 mM DTT and 10% Glycerol.

HY5 wash buffer: 50 mM Tris (pH 7.5), 150 mM NaCl, 0.5 mM DTT and 0.1% Tween 20.

COP1 binding buffer: 50 mM Tris (pH 7.5), 200 mM NaCl, 5 mM $MgCl_2$, 120 μM $ZnCl_2$, 1% BSA, 0.5 mM DTT and 10% Glycerol.

COP1 wash buffer: 50 mM Tris (pH 7.5), 200 mM NaCl, 5 mM $MgCl_2$, 120 μM $ZnCl_2$, 0.5 mM DTT and 0.1% Tween 20.

Before probing slides, biotinylated MBP-COP1 protein was pre-incubated in binding buffer for 45 min at 22°C. All procedures were carried out at room temperature and protected from light. Slides were washed for 10 min at each wash step. Before blocking, slides were dipped in Ni-NTA elution buffer for 1–2 min until the nitrocellulose coating was totally wet, then were dipped in binding buffer several times to remove excess elution buffer. All buffers used in binding assays were filtered. We placed the chamber containing the slides on an orbital shaker with gentle agitation (80 rpm) during blocking and hybridization procedures.

ACKNOWLEDGMENTS

We thank William Terzaghi for reading and commenting on this manuscript. This work was supported by the Chinese National Natural Science Foundation and in part by The National Science Foundation 2010 Program (US) and by the National Institutes of Health grant number GM069418 to SLH.

REFERENCES

AGI. The *Arabidopsis* Genome Initiative (2000). Analysis of genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408,** 796–815.

Aida, M., Beis, D., Heidstra, R., Willemsen, V., Blilou, I., Galinha, C., Nussaume, L., Noh, Y.S., Amasino, R., and Scheres, B. (2004). The PLETHORA genes mediate patterning of the *Arabidopsis* root stem cell niche. Cell **1,** 109–120.

Ang, L.H., Chattopadhyay, S., Wei, N., Oyama, T., Okada, K., Batschauer, A., and Deng, X.W. (1998). Molecular interaction between COP1 and HY5 defines a regulatory switch for light control of *Arabidopsis* development. Mol. Cell **1,** 213–222.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K., eds (1994). *Saccharomyces cerevisiae*: Current Protocols in Molecular Biology (New York: John Wiley and Sons), Suppl, pp. 13.6.2–13.6.4.

Baker, S.S., Wilhelm, K.S., and Thomashow, M.F. (1994). The 59-region of *Arabidopsis thaliana cor15α* has *cis*-acting elements that confer cold-, drought- and ABA-regulated gene expression. Plant Mol. Biol. **24,** 701–713.

Banno, H., Ikeda, Y., Niu, Q.W., and Chua, N.H. (2001). Overexpression of *Arabidopsis* ESR1 induces initiation of shoot regeneration. Plant Cell **13,** 2609–2618.

Bertone, P., and Snyder, M. (2005). Advances in functional protein microarray technology. FEBS J **272,** 5400–5411.

Buttner, M., and Singh, K.B. (1997). *Arabidopsis thaliana* ethylene-responsive element binding protein (AtEBP), an ethylene-inducible, GCC box DNA-binding protein interacts with an ocs element binding protein. Proc. Natl Acad. Sci. USA **94,** 5961–5966.

Chattopadhyay, S., Ang, L.H., Puente, P., Deng, X.W., and Wei, N. (1998). *Arabidopsis* bZIP protein HY5 directly interacts with light-responsive promoters in mediating light control of gene expression. Plant Cell **10,** 673–683.

Deng, X.W., and Quail, P.H. (1999). Signaling in light-controlled development. Semin Cell Dev. Biol. **10,** 121–129.

Ding, X.Z., Paulsen, I.T., Bhattacharjee, A.K., Nikolich, M.P., Myers, G., and Hoover, D.L. (2006). A high efficiency cloning and expression system for proteomic analysis. Proteomics **6,** 4038–4046.

Duek, P.D., Elmer, M.V., van Oosten, V.R., and Fankhauser, C. (2004). The degradation of HFR1, a putative bHLH class transcription

factor involved in light signaling, is regulated by phosphorylation and requires COP1. Curr. Biol. **14,** 2296–2301.

Eisen, M.B.S., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc. Natl Acad. Sci. **95,** 14863–14868.

Feng, J.X., Liu, D., Pan, Y., Gong, W., Ma, L.G., Luo, J.C., Deng, X.W., and Zhu, Y.X. (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,** 853–868.

Finkelstein, R.R., Wang, M.L., Lynch, T.J., Rao, S., and Goodman, H.M. (1998). The *Arabidopsis* abscisic acid response locus ABI4 encodes an APETALA 2 domain protein. Plant Cell **10,** 1043–1054.

Fried, M., and Crothers, D.M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucl. Acids Res. **9,** 6505–6525.

Fujimoto, S.Y., Ohta, M., Usui, A., Shinshi, H., and Ohme-Takagi, M. (2000). *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. Plant Cell **12,** 393–404.

Gong, W., et al. (2004). Genome-wide ORFeome cloning and analysis of *Arabidopsis* transcription factor genes. Plant Physiol **135,** 773–782.

Guo, A.Y., He, K., Liu, D., Bai, S.N., Gu, X.C., Wei, L.P., and Luo, J.C. (2005). DATF: a database of *Arabidopsis* transcription factors. Bioinformatics **21,** 2568–2569.

Hall, D.A., Zhu, H., Zhu, X.W., Royce, T., Gerstein, M., and Snyder, M. (2004). Regulation of gene expression by a metabolic enzyme. Science **306,** 482–484.

Hao, D., Ohme-Takagi, M., and Sarai, A. (1998). Unique mode of GCC box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF domain) in plant. J. Biol. Chem. **273,** 26857–26861.

Hao, D., Yamasaki, K., Sarai, A., and Ohme-Takagi, M. (2002). Determinants in the sequence specific binding of two plant transcription factors, CBF1 and NtERF2, to the DRE and GCC motifs. Biochemistry **41,** 4202–4208.

Harmer, S.L., and Kay, S.A. (2005). Positive and negative factors confer phase-specific circadian regulation of transcription in *Arabidopsis*. The Plant Cell **17,** 1926–1940.

Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.S., Han, B., Zhu, T., Wang, X., Kreps, J.A., and Kay, S.A. (2000). Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. Science **290,** 2110–2113.

Ho, S.W., Jona, G., Chen, C.T.L., Johnston, M., and Snyder, M. (2006). Linking DNA-binding proteins to their recognition sequences by using protein microarrays. Proc. Natl Acad. Sci. USA **103,** 9940–9945.

Holm, M., Ma, L.G., Qu, L.J., and Deng, X.W. (2002). Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in *Arabidopsis*. Genes Dev. **16,** 1247–1259.

Ikeda, Y., Banno, H., Niu, Q.W., Howell, S.H., and Chua, N.H. (2006). The ENHANCER OF SHOOT REGENERATION 2 gene in *Arabidopsis* regulates CUP-SHAPED COTYLEDON 1 at the transcriptional level and controls cotyledon development. Plant Cell Physiol **47,** 1443–1156.

Jack, T. (2001). Plant development going MADS. Plant Mol. Biol. **46,** 515–520.

Jiang, C., Iu, B., and Singh, J. (1996). Requirement of a CCGAC *cis*-acting element for cold induction of the BN115 gene from winter Brassica napus. Plant Mol. Biol. **30,** 679–684.

Jiao, Y., Lau, O.S., and Deng, X.W. (2007). Light-regulated transcriptional networks in higher plants. Nat. Rev. Genet. **8,** 217–230.

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.

Krizek, B.A. (2003). AINTEGUMENTA utilizes a mode of DNA recognition distinct from that used by proteins containing a single AP2 domain. Nucleic Acids Res. **31,** 1859–1868.

Kukar, T., Eckenrode, S., Gu, Y., Lian, W., Megginson, M., She, J.X., and Wu, D. (2002). Protein microarrays to detect protein-protein interactions using red and green fluorescent proteins. Anal Biochem. **306,** 50–54.

Lee, J., He, K., Stolc, V., Lee, H., Figueroa, P., Gao, Y., Tongprasit, W., Zhao, H., Lee, I., and Deng, X.W. (2007). Analysis of transcription factor HY5 genomic binding sites revealed its hierarchical role in light regulation of development. The Plant Cell **19,** 731–749.

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1998). Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low- temperature-responsive gene expression, respectively, in *Arabidopsis*. Plant Cell **10,** 1391–1406.

Liu, X., Noll, D.N., Lieb, J.D., and Clarke, N.D. (2005). DIP-chip: rapid and accurate determination of DNA-binding specificity. Genome Res. **15,** 421–427.

Liu, X., and Clarke, N.D. (2002). Rationalization of gene regulation by a eukaryotic transcription factor: Calculation of regulatory region occupancy from predicted binding affinities. J. Mol. Biol. **323,** 1–8.

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.

Martone, R., et al. (2003). Distribution of NF-kappaB-binding sites across human chromosome 22. Proc. Natl Acad. Sci. USA **100,** 12247–12252.

Meyerowitz, E.M. (2002). Plants compared to animals: the broadest comparative study of development. Science **295,** 1482–1485.

Nakano, T., Suzuki, K., Fujimura, T., and Shinshi, H. (2006). Genome-wide analysis of the *ERF* gene family in *Arabidopsis* and Rice. Plant Physiol **140,** 411–432.

Nole-Wilson, S., and Krizek, B.A. (2000). DNA binding properties of the *Arabidopsis* floral development protein AINTEGUMENTA. Nucleic Acids Res. **28,** 4076–4082.

Ohme-Takagi, M., and Shinshi, H. (1995). Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. Plant Cell **7,** 173–182.

Qu, L.-J., and Zhu, Y.-X. (2006). Transcription factor families in *Arabidopsis*: major progress and outstanding issues for future research. Curr. Opin. Plant Biol. **9,** 544–549.

Riechmann, J.L., et al. (2000). Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science **290,** 2105–2110.

Ruvkun, G., and Hobert, O. (1998). The taxonomy of developmental control in *Caenorhabditis elegans*. Science **282,** 2033–2041.

Saijo, Y., Sullivan, J.A., Wang, H.Y., Yang, J., Shen, Y.P., Rubio, V., Ma, L.G., Hoecker, U., and Deng, X.W. (2003). The COP1-SPA1 interaction defines a critical step in phytochrome A-mediated regulation of HY5 activity. Genes Dev. **17,** 2642–2647.

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. Commun **290,** 998–1009.

Smyth, G.K. (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Statistical Applications in Genetics and Molecular Biology **3,** Article 3.

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.

Thomashow, M.F. (1999). Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. Annu. Rev. Plant Physiol. Plant Mol. Biol. **50,** 571–599.

Veraksa, A., Del Campo, M., and McGinnis, W. (2000). Developmental patterning genes and their conserved functions: from model organisms to humans. Mol. Genet. Metab **69,** 85–100.

Weinmann, A.S., Yan, P.S., Oberley, M.J., Huang, T.H., and Farnham, P.J. (2002). Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. Genes Dev. **16,** 235–244.

Wei, G., Pan, Y., Lei, J., and Zhu, Y.X. (2005). Molecular cloning, phylogenetic analysis, expressional profiling and in vitro studies of TINY2 from *Arabidopsis thaliana*. J. Biochem. Mol. Biol. **38,** 440–446.

Yamaguchi-Shinozaki, K., and Shinozaki, K. (1994). A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. Plant Cell **6,** 251–264.

Yanhui, Y.H., et al. (2006). The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol. Biol. **60,** 107–123. http://www.springerlink.com/content/k4405850q9150p64/.

Zhou, J., Tang, X., and Martin, G.B. (1997). The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that bind a *cis*-element of pathogenesis-related genes. EMBO J **16,** 3207–3218.

Zhu, H., et al. (2000). Global analysis of protein activities using proteome chips. Science **293,** 2101–2105.

Table S1. DNA probe design.

(A) Conserved *cis*-element motifs used as templates in DNA probe design.

Probe	Sequence	Template for probe
ANT-AP2R1R2	g**CAC(A**/G)N(**A**/T)**TcCC(**a/g)ANG(c/t)	ABE
GCC-box	TAAGA**GCCGCC**	EBE
DRE	T**ACCGAC**AT	DBE
RAV1 binding element	g**CaACA(**g/t) (a/t) $(N)_n$**caCCTG(**a/g)	RBE
CCR2_EE	AAAATATCT	EE

All motifs were obtained from previously-reported experiments. Uppercase letters indicate conserved positions; lowercase letters indicate relatively less-conserved positions. N indicates positions for which no particular base was preferred. Core motifs, which play the most important role in DNA recognition, are indicated in bold.

(B) Sequences of DNA probes

Probe name	Probe sequence
ABE	TTGGTGCACATATCCCGATGCTTACATTGGTGCACATATCCCGATGCTTA
MtABE	TTGGTGATATTTATATGATGCTTACATTGGTGATATTTATATGATGCTTA
EBE	TAAGAGCCGCCTAAGAGCCGCCTAAGAGCCGCCGGCGGCTCTTA
MtEBE	TAAGAATATATTAAGAATATATTAAGAATATATATATATTCTTA
DBE	GATATACTACCGACATGAGTTCCAAGATATACTACCGACATGAGTTCCAA
MtDBE	GATATACTATATATATGAGTTCCAAGATATACTATATATATGAGTTCCAA
RBE	TGGCAACAGTAAACACCTGACTCAGTGGCAACAGTAAACACCTGACTCAG
MtRBE	TGGTTTATGTAAAATATATTCTCAGTGGTTTATGTAAAATATATTCTCAG
EE	AAACCTAGAAAATATCTAAACCTTGAAACCTAGAAAATATCTAAACCTTG
MtEE	AAACCTAGAAAATCGAGAAACCTTGAAACCTAGAAAATCGAGAAACCTTG

ABE, EBE, DBE, and RBE were derived from ANT-AP2R1R2, GCC-box, DRE, and RAV1 binding elements, respectively. Mutant probes were designed by substituting core motif sequences of wild type probes with alternative As and Ts.

Table S2. Relative DNA-binding values of 56 positive AP2/ERF TFs.

Locus ID	DNA-binding values with DNA probes			
	ABE	EBE	DBE	RBE
At1g04370	0	13	0	0
At1g12980	0	146	0	0
At1g24590	0	40	0	0
At1g28370	0	61	0	0
At1g49120	58	80	0	0
At1g53910	0	60	0	0
At1g64380	0	101	0	0
At1g71130	0	104	0	0
At2g31230	0	24	0	0
At3g11020	16	34	0	0
At3g61630	0	55	0	0
At5g07580	0	45	0	0
At5g13910	17	141	30	0
At5g51190	0	15	0	0
At1g12610	0	18	135	10
At1g33760	0	0	137	0
At1g44830	0	0	51	0
At2g20880	0	38	110	0
At2g23340	0	0	19	0
At2g40220	0	0	111	0
At2g44940	2	1	131	0
At3g16280	0	0	92	0
At3g60490	0	0	134	0
At4g16750	0	0	97	0
At4g23750	0	0	62	0
At4g25480	15	0	115	0
At5g18450	0	0	31	0
At5g52020	0	0	47	0
At1g13260	0	0	71	149
At1g25560	0	0	0	131
At1g51120	0	0	0	141
At1g68840	0	0	0	42
At3g20310	33	29	0	62
At1g03800	0	132	54	34
At1g28160	0	148	112	0
At1g63040	0	84	146	0
At1g71450	0	93	138	0
At1g75490	0	72	104	0
At2g20350	0	57	80	0
At3g23220	37	142	140	137
At4g13620	0	123	59	0
At4g25490	0	139	148	122
At4g28140	0	86	144	0
At5g19790	30	149	121	0
At1g25470	86	0	0	0
At1g51190	20	0	0	0
At1g79700	41	0	0	0
At2g40340	54	0	0	0
At3g25890	65	12	0	0
At4g17500	23	0	0	0
At4g27950	26	14	0	0
At4g37750	146	0	0	0
At5g17430	55	0	0	0
At5g44210	66	0	0	0
At5g65510	107	0	0	0
At5g67190	49	0	0	0

Table S2. Relative DNA-binding values of 56 positive AP2/ERF TFs.

High values indicate strong DNA-binding affinities. Zero indicates no positive binding signals were detected on the protein microarrays.