The technique of selected and amplified binding-sequence (SAAB) is an important approach to study the DNA-binding activity of transcription factors (Blackwell and Weintraub 1990). Here we describe a novel SAAB method based on the electrophoretic mobility shift assay (EMSA) and PCR-assisted oligonucleotide selection to characterize the DNA-binding specificity of a plant transcription factor KANADI (KAN). This strategy provides a clean, quick and reliable way to identify consensus sequence for DNA-protein interaction.

The DNA-binding domain (DBD) is a common structural feature of transcription factors. (Ref.). Studying the DNA binding specificity and capacity of the DBD is essential for elucidating the function of a transcription factor. As a prevailing approach for this purpose, SAAB has been extensively utilized in both animal (Ref.) and plant systems (Ref.). Currently, there are two major types of SAAB. One is EMSA-based random oligonucleotide selection(REF.). The other is affinity purification-based selection(REF.). Both of them utilize oligonucleotides containing random central sequences. These random oligos are incubated with the protein of interest. Oligos bound to the protein are isolated and amplified by PCR. The PCR products are reiteratively re-bound to the protein and re-amplified until the nonspecific-binding oligos no longer contribute a significant fraction. In the final step, PCR products (putative protein-binding oligos) are sequenced to identify the binding site of protein (Ref.). The difference between these two methods is the strategy for random oligonucleotide selection. In the former method, DNA-protein complex is recognized by a retarded DNA band in EMSA. Oligos interacting with the protein are isolated from this band. In the latter, DNA-protein complex is pulled down by a ligand immobilized to a solid support (usually sepharose

beads). The complex is released from the ligand and protein-binding oligos can be extracted from the elute. Compared to EMSA-based selection, affinity purification-based strategy saves time and labor by avoiding electrophoresis and membrane-blotting in the selection process. It is also much safer without using radioactive reactions. However, as a trade for speed and safety, it is potentially less specific due to non-specific binding of the beads to target DNA. Moreover, this method lacks visible detection of DNA-protein interaction by EMSA, which makes it more vulnerable to random errors during selection. Based on these factors, researchers are seeking alternative techniques to circumvent the disadvantages of both methods. For this purpose, we applied the following strategy in our study to identify the DNA-binding property of KAN *in vitro*, trying to compromise the limiting factors of both methods and provide a more efficient way for SAAB.

In this strategy, we firstly expressed and purified KAN DBD protein(a truncated KAN full length protein containing the putative DBD) in *Escherichia. coli* cells. A plasmid incorporating DNA of KAN DBD was utilized for protein expression. In order to construct this plasmid, PCR was performed to amplify the DNA of KAN DBD from *KAN* cDNA. Two primers, 5'ATTCGGATCCAAGATGCCGACAAAGCGAAGC-3' and 5'-AAGCGAATTCCTTGTTAGTGGTCTTAACAGTTCG-3', were used. PCR conditions were 94°C for 20s, 54°C for 20s, 72°C for 15s for 34 cycles. PCR product was digested by BamH1 and EcoR1 and cloned into *E. coli* vector pGEX-2TK (Amersham Biosciences, Piscataway, NJ) where the coding region of KAN DBD was fused with a glutathione S-transferase (GST) tag. Plasmid construct pGEX-DBD was transformed into *E.coli* BL21 (DE3) strain. Expression of KAN DBD was induced by 0.1mM IPTG in the transformed cell culture. After IPTG induction, *E.col*i cells were collected and lysed with

sonication. After centrifuge, the soluble fraction of cell lysate was utilized for purification using MicroSpin GST Purification Module (Amersham Biosciences, Piscataway, NJ). KAN DBD purification was performed according to the manufacture's instruction. This purified protein was analyzed with SDS-PAGE eletrophoresis. A single protein band with proper molecular weight was detected in the gel after Coomassie Blue (Bio-rad, Hercules, CA) staining.**(FIG.)**. Before SAAB, proteins were dialyzed in the dialysis buffer containing 20mM Tris-HCl, pH 8.0 and 80mM KCl to remove reduced glutathione and treated with 0.07 U DNase I (Fermentas, Hanover, MD) on ice for 1 hour to remove any contaminating *E. coli* DNA followed by EDTA addition to 2 mM to inactivate DNaseI. Pretreated proteins were applied to the *in vitro* DNA binding site selection described in Protocol 1.This protocol was adapted from a random oligo-selection strategy by Hosoda, et.al. (2002).

Protocol1 in vitro DNA binding site selection of KAN DBD

1. **Random oligonucleotides synthesis.** A mixture of 54-base oligonucleotides in which the middle 16 bases are composed of random sequences with each of the 4 nucleotides at equal molar concentration was synthesized (IDT, Coralville, IA). These oligos were converted into dsDNA by extending from a primer 5'-CGACGCTCTGACTCGAGG-3' binding to their 3'- ends using Klenow fragment (Fermentas, Hanover, MD).
2. **DNA and protein interaction.** The dsDNA (0.5nmol) was incubated with purified KAN DBD protein (50 pmol) in dialysis buffer for 1 hour at 4°C.
3. **EMSA.** The mixture of DNA and protein was loaded to 9% native polyacrylamide gel in 0.5X Tris-borate-EDTA (TBE) buffer at 4°C.Two separate gels were performed

simultaneously. One gel was stained with 1XSYBR Safe™ (Invitrogen, Eugene, OR) for 30 minutes and viewed under 306nm UV transilluminator equipped in The KODAK Gel Logic 200 Imaging System (Eastman KODAK Company, Rochester, NY) with 535nm WB50 filter to visualize DNA in the shifted DNA-protein complex. The other gel was stained with E-Zinc™ reversible stain kit (Pierce, Rockford, IL) following the procedure provided by the manufacturer and viewed under white light epi-illuminator in the same device described above to visualize protein in the complex. Gel images were captured using KODAK molecular imaging software (version 4.0) (Eastman KODAK Company, Rochester, NY).

4. **Isolation of DNA interacting with KAN DBD.** The shifted DNA-protein complex in EMSA was excised from the gel and DNA in the complex was purified with phenol/chloroform extraction followed by NaOAc/ethanol precipitation.
5. **PCR.** Purified DNA was subjected to PCR amplification using a pair of primers binding to the ends of 54-oligos (5'-GCTGAGTCTGAACAAGCTTG-3' and 5'-CGACGCTCTGACTCGAGG-3'). PCR condition was …PCR products were precipitated and used in the subsequent selections as described above.
6. **Repeat of selection for 5 additional cycles.** In two parallel samples, DNA from SYBR Safe™ staining was always used in selection with SYBR Safe™ and the same with E-Zinc™ stain.
7. **Cloning of KAN DBD-binding DNA and data analysis** After 6 cycles of selection, the resulting DNA fragments were cloned into pGEM-T Easy (Promega, Madison, WI) and sequenced. Sequence comparison and motif identification utilized the Gibbs Motif Sampler Homepage (http://bayesweb.wadsworth.org/gibbs/gibbs.html).

Analysis of sequences by Gibbs Sampler identified a 6 bp motif GAATAW in 54 instances. To reduce the random variation in the Gibbs sampling, manual alignment of the sequences was conducted based on this motif **(FIG?).** A slightly different motif GNATDW (where N is any nucleotide, D is A, G, or T, and W is A or T) was identified. This motif was found at least once in each of 50 oligonucleotides and 28 oligonucleotides contained a second binding site. The probability matrix for manual alignment is shown in **(TABLE1)**. This consensus sequence was subsequently confirmed by gel-shift assay in which KAN DBD exhibited higher affinity to this core-binding site than to mutant variants (data not shown). Furthermore, statistical analysis of promoter sequence in putative KAN target genes (revealed by microarray and quantitative PCR) shows significant enrichment of this motif element compared to other unrelated genes (data not shown). These results indicate the biological significance of this *in vitro* binding site for the interaction between KAN DBD with its targets.

Table1

A	C	G	T
0	0	78	0
48	11	11	8
78	0	0	0
0	0	0	78
62	0	7	9
32	0	0	46

In this strategy, we applied two independent methods to detect the DNA-protein complex in EMSA. In both SYBR Safe^{TM} and E-Zinc^{TM} staining, a shifted band relative to free

DNA or free protein was observed with the same migration distance to the bottom of loading well **(FIG)**. Moreover, the KAN DBD-binding oligos identified in both methods are strictly consistent. These results indicate that these two staining strategies can serve as a good control for each other. Although double stain of DNA and protein with non-radioactive dyes is not a novel approach in EMSA, such as the EMSA kit E33075(Invitrogen, Eugene, OR), applying this strategy to DNA-binding site selection hasn't been reported so far. The SYBR Safe™ stain for DNA has low mutagenicity, which makes it much safer than autoradiography and ethidium bromide (EtBr) stain (Ref.) Furthermore, it provides greater sensitivity than EtBr and the abundance of DNA bound to protein can be quantitated. The E-Zinc™ stain is a reversible stain technique for protein. The formation of zinc hydroxide on the gel surface produces a semiopaque background which makes protein bands transparent in the gel (Ref.). The addition of imidazole significantly increases its sensitivity to the level equivalent to silver stain (0.25ng)(Ref.). E-Zinc™ stain is rapid (only 15minutes) compared to Coomassie stain and silver stain and there is no requirement for fixation before staining. So all the molecules (DNA and protein) remain active after elution from the gel, which makes it feasible to isolate DNA from the retarded band and use it for subsequent binding to KAN DBD protein in our approach.

SYBR Safe™ and E-Zinc™ utilize two distinct mechanisms to detect DNA and protein interaction. Combination of these two methods provides a reliable pool of protein-targeting oligos with high specificity. This strategy is more rapid than traditional EMSA-based selection and provides a clean experimental condition by avoiding hazardous reagents. All these features indicate the potential of this modified SAAB to be broadly

applicable to identify target sequence of any DNA-binding protein that can be expressed and purified *in vitro*.