[Plant Signaling & Behavior 2:5, 416-420, September/October 2007]; ©2007 Landes Bioscience

Article Addendum

Arabidopsis Protein Microarrays for the High-Throughput Identification of Protein-Protein Interactions

Sorina C. Popescu[1]
Michael Snyder[1,2]
S.P. Dinesh-Kumar[1,*]

[1]Department of Molecular, Cellular and Developmental Biology; [2]Department of Biochemistry and Biophysics; Yale University; New Haven, Connecticut USA

*Correspondence to: S.P. Dinesh-Kumar; Department of Molecular, Cellular and Developmental Biology; Yale University; 210 Prospect Street; New Haven, Connecticut 06520-8103, USA; Tel.: 203.432.9965; Fax: 203.432.6161; Email: savithramma.dinesh-kumar@yale.edu

Original manuscript submitted: 05/10/07
Manuscript accepted: 05/10/07

Previously published online as a *Plant Signaling & Behavior* E-publication: http://www.landesbioscience.com/journals/psb/article/4416

KEY WORDS

protein microarray, calmodulin, protein-protein interactions, large scale protein production

ACKNOWLEDGEMENTS

This work is supported by Yale University Genomics and Proteomics Center pilot grant and the NSF Arabidopsis 2010 grant (MCB-0313514 and DBI-0519853).

Addendum to:

Differential Binding of Calmodulin Related Proteins to Their Targets Revealed Using High-Density Arabidopsis Protein Microarrays

Popescu SC, Popescu GV, Bachan S, Zhang Z, Seay M, Gerstein M, Snyder M, Dinesh-Kumar SP

Proc Natl Acad Sci 2007; 104:4730–5
PMID: 17360592
DOI: 10.1073/pnas.0611615104.

ABSTRACT

Protein microarray technology has emerged as a powerful new approach for the study of thousands of proteins simultaneously. Protein microarrays have been used for a wide variety of applications for the human and yeast systems. In a recent study, we demonstrated that Arabidopsis functional protein microarrays can be generated and employed to characterize the function of plant proteins. The arrayed proteins were produced using an optimized large-scale plant-based expression system. In a proof-of concept study, 173 known and novel potential substrates of calmodulin (CaM) and calmodulin-like proteins (CML) were identified in an unbiased and high-throughput manner. The information documented here on novel potential CaM targets provides new testable hypotheses in the area of CaM/Ca^{2+}-regulated processes and represents a resource of functional information for the scientific community.

At the foundation of the remarkable complexity of all living organisms, lays an intricate network of interactions between their cellular components—proteins, nucleic acids, sugars and other molecules. Organisms' functions such as growth and development, reproduction, and responses to pathogens and environmental stresses are ultimately the result of constant assembly and disassembly of macromolecular complexes within cells. Therefore, revealing and cataloging these associations will lead to a comprehensive understanding of cellular signaling pathways and regulatory networks of the organism.

Availability of DNA microarray-based high throughput technology has aided in large-scale analyses of organisms' genome as evident by their application in transcript mapping, gene expression profiling, detecting mutations and deletions, and mapping transcription factor binding sites. Although DNA microarray experiments are invaluable, little can be inferred from these studies about the functions of encoded gene products. Recently, protein microarray technology has emerged as a powerful approach for the study of hundreds or thousands of proteins simultaneously in a controlled experimental setup. Yeast and human protein microarrays have been used for a wide variety of applications including the study of protein-protein interactions, protein-nucleic acid interactions, targets of small molecule, antibody specificity, protein kinase substrates, as well as post-translational modifications.[1] An important advantage offered by protein microarrays is the capacity to study, in an unbiased manner, the in vitro behavior of a large number of macromolecules in various types of biochemical reactions. Therefore, novel substrates and subsequently biological functions of proteins can be predicted and then validated through additional methods.

In a recent study, we demonstrated that Arabidopsis functional protein microarrays can be constructed and employed to investigate protein function.[2] Specifically, we used high-density Arabidopsis protein microarrays containing 1,133 proteins to investigate targets of calmodulin (CaM) and calmodulin-like proteins (CML).

AN OPTIMIZED HOMOLOGOUS PROTEIN PRODUCTION SYSTEM FOR CONSTRUCTING FUNCTIONAL ARABIDOPSIS PROTEIN MICROARRAYS

A critical step in the generation of functional protein microarrays is the construction of a high quality expression library from which a large number of distinct protein samples can be produced. For this, we built a high quality expression library of tagged ORFs, ATEC (*Arabidopsis thaliana* expression collection) in a versatile plant expression vector. ATEC collection is represented by multiple gene families including protein kinases, transcription factors, protein with unknown functions, heat-shock proteins, P450 cytochromes, protein degradation-related proteins, CaMs/CMLs and CaM/CML-binding proteins, and others.

Table 1 **Analysis of gene ontology (GO) categories over-represented in the CaM/CML target list**

GO-ID	P-Value	Genes in Test Set	Accession #	Gene Description
48409	5.67E-03	Flower development	AT5G23260	| TT16 | MADS-BOX PROTEIN, TRANSPARENT TESTA16
			AT3G58780	| SHP1 | AGAMOUS-LIKE MADS BOX AGL1 / SHATTERPROOF 1
			AT5G60910	| AGL8 | AGAMOUS-LIKE MADS BOX AGL8 / FRUITFULL
10154	7.29E-04	Fruit development	AT3G61880	| CYP78A9 | CYTOCHROME P450, PUTATIVE
			AT5G60910	| AGL8 | AGAMOUS-LIKE MADS BOX AGL8 / FRUITFULL (AGL8)
48481	4.02E-03	Ovule development	AT5G23260	| TT16 | MADS-BOX PROTEIN, TRANSPARENT TESTA16
			AT3G58780	| SHP1 | AGAMOUS-LIKE MADS BOX AGL1 / SHATTERPROOF 1
48440	7.56E-03	Carpel development	AT5G23260	| TT16 | MADS-BOX PROTEIN, TRANSPARENT TESTA16
			AT3G58780	| SHP1 | AGAMOUS-LIKE MADS BOX AGL1 / SHATTERPROOF 1
48467	8.25E-03	Gynoecium development	AT5G23260	| TT16 | MADS-BOX PROTEIN, TRANSPARENT TESTA16
			AT3G58780	| SHP1 | AGAMOUS-LIKE MADS BOX AGL1 / SHATTERPROOF 1
48413	1.75E-02	Floral whorl development	AT5G23260	| TT16 | MADS-BOX PROTEIN, TRANSPARENT TESTA16
			AT3G58780	| SHP1 | AGAMOUS-LIKE MADS BOX AGL1 / SHATTERPROOF 1
9651	3.77E-03	Response to salt stress	AT3G19290	| ABF4 | ABA-RESPONSIVE ELEMENT-BINDING PROTEIN 2
			AT2G31180	| MYB FAMILY TRANSCRIPTION FACTOR (MYB14)
			AT2G23290	| MYB FAMILY TRANSCRIPTION FACTOR
			AT5G35410	| SOS2 | CBL-INTERACTING PROTEIN KINASE 24 (CIPK24)
6970	6.55E-03	Response to osmotic stress	AT3G19290	| ABF4 | ABA-RESPONSIVE ELEMENT-BINDING PROTEIN 2
			AT2G31180	| MYB FAMILY TRANSCRIPTION FACTOR (MYB14)
			AT2G23290	| MYB FAMILY TRANSCRIPTION FACTOR
			AT5G35410	| SOS2 | CBL-INTERACTING PROTEIN KINASE 24 (CIPK24)
42829	5.84E-03	Physiological defense response	AT4G39950	| CYP79B2 | CYTOCHROME P450 79B2, PUTATIVE (CYP79B2)
			AT4G26070	| MEK1 | MITOGEN-ACTIVATED PROTEIN KINASE KINASE
			AT3G12250	| TGA6 | BZIP FAMILY TRANSCRIPTION FACTOR
			AT1G22070	| TGA3 | BZIP FAMILY TRANSCRIPTION FACTOR
51869	6.55E-03	Physiological response to stimulus	AT4G39950	| CYP79B2 | CYTOCHROME P450 79B2, PUTATIVE (CYP79B2)
			AT4G26070	| MEK1 | MITOGEN-ACTIVATED PROTEIN KINASE KINASE
			AT3G12250	| TGA6 | BZIP FAMILY TRANSCRIPTION FACTOR
			AT1G22070	| TGA3 | BZIP FAMILY TRANSCRIPTION FACTOR
6951	9.60E-04	Response to heat	AT5G37670	| 15.7 KDA CLASS I-RELATED SMALL HEAT SHOCK PROTEIN-LIKE
			AT5G56030	| HSP81 2 | HEAT SHOCK PROTEIN 81-2 (HSP81-2)
			AT5G02500	| HSC70 1 | HEAT SHOCK COGNATE 70 KDA PROTEIN 1
			AT1G11270	| F-BOX FAMILY PROTEIN
9755	1.18E-02	Hormone-mediated signaling	AT3G19290	| ABF4 | ABA-RESPONSIVE ELEMENT-BINDING PROTEIN 2
			AT1G69270	| RPK1 | RECEPTOR LIKE PROTEIN KINASE 1
			AT1G10210	| ATMPK1 | MITOGEN-ACTIVATED PROTEIN KINASE
9814	8.84E-03	Defense response, incompatible interaction	AT4G39950	| CYP79B2 | CYTOCHROME P450 79B2, PUTATIVE (CYP79B2)
			AT3G12250	| TGA6 | BZIP FAMILY TRANSCRIPTION FACTOR
			AT1G22070	| TGA3 | BZIP FAMILY TRANSCRIPTION FACTOR
45087	1.53E-02	Innate immune response	AT4G39950	| CYP79B2 | CYTOCHROME P450 79B2, PUTATIVE (CYP79B2)
			AT3G12250	| TGA6 | BZIP FAMILY TRANSCRIPTION FACTOR
			AT1G22070	| TGA3 | BZIP FAMILY TRANSCRIPTION FACTOR
9627	5.08E-03	Systemic acquired resistance	AT3G12250	| TGA6 | BZIP FAMILY TRANSCRIPTION FACTOR
			AT1G22070	| TGA3 | BZIP FAMILY TRANSCRIPTION FACTOR
9738	1.05E-02	Abscisic acid mediated signaling	AT3G19290	| ABF4 | ABA-RESPONSIVE ELEMENT-BINDING PROTEIN 2
			AT1G69270	| RPK1 | RECEPTOR LIKE PROTEIN KINASE 1
42436	1.19E-02	Indole derivative catabolism	AT4G39950	| CYP79B2 | CYTOCHROME P450 79B2, PUTATIVE (CYP79B2)
6569	1.19E-02	Tryptophan catabolism	AT4G39950	| CYP79B2 | CYTOCHROME P450 79B2, PUTATIVE (CYP79B2)

Continued on next page

Table 1 **Analysis of gene ontology (GO) categories over-represented in the CaM/CML target list (continued)**

46218	1.19E-02	Indolalkylamine catabolism	AT4G39950	| CYP79B2 | CYTOCHROME P450 79B2, PUTATIVE (CYP79B2)
42402	1.19E-02	Biogenic amine catabolism	AT4G39950	| CYP79B2 | CYTOCHROME P450 79B2, PUTATIVE (CYP79B2)
42219	1.78E-02	Amino acid derivative catabolism	AT4G39950	| CYP79B2 | CYTOCHROME P450 79B2, PUTATIVE (CYP79B2)
9074	1.78E-02	Aromatic amino acid family catabolism	AT4G39950	| CYP79B2 | CYTOCHROME P450 79B2, PUTATIVE (CYP79B2)
10120	1.78E-02	Camalexin biosynthesis	AT4G39950	| CYP79B2 | CYTOCHROME P450 79B2, PUTATIVE (CYP79B2)
6006	9.11E-04	Glucose metabolism	AT5G47810	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT1G20950	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
			AT4G26270	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT4G37870	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
6092	1.62E-03	Main pathways of carbohydrate metabolism	AT5G47810	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT1G20950	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
			AT4G26270	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT4G37870	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
19318	2.74E-03	Hexose metabolism	AT5G47810	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT1G20950	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
			AT4G26270	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT4G37870	| PHOSPHOENOLPYRUVATE CARBOXYKINASE (ATP)
15980	3.52E-03	Energy derivation by oxidation of organic compounds	AT5G47810	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT1G20950	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
			AT4G26270	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT4G37870	| PHOSPHOENOLPYRUVATE CARBOXYKINASE (ATP)
6096	2.01E-03	Glycolysis	AT5G47810	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT1G20950	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
			AT4G26270	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
6007	8.18E-03	Glucose catabolism	AT5G47810	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT1G20950	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
			AT4G26270	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
46365	8.51E-03	Monosaccharide catabolism	AT5G47810	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT1G20950	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
			AT4G26270	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
19320	8.51E-03	Hexose catabolism	AT5G47810	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT1G20950	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
			AT4G26270	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
46164	9.19E-03	Alcohol catabolism	AT5G47810	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT1G20950	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
			AT4G26270	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
16052	1.35E-02	Carbohydrate catabolism	AT5G47810	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT1G20950	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
			AT4G26270	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
44275	1.35E-02	Cellular carbohydrate catabolism	AT5G47810	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
			AT1G20950	| PYROPHOSPHATE—FRUCTOSE-6-PHOSPHOTRANSFERASE
			AT4G26270	| PHOSPHOFRUCTOKINASE FAMILY PROTEIN
45788	5.97E-03	Regulation of cell shape	AT5G23260	| TT16 | MADS-BOX PROTEIN, TRANSPARENT TESTA16

Another important requirement to construct functional protein microarrays is that the purified proteins have to be functional, that is, they should retain their cellular activity and act similarly to their cellular counter-parts, when on the protein chip. Therefore, proteins must be produced and purified in a biological system that allows native folding and post-translational modifications to occur, and provides necessary cofactors for protein activity. In this respect, we tested the performance of two eukaryotic systems for protein expression, *Saccharomyces cerevisiae* and *Nicotiana benthamiana*. Our experiments clearly demonstrated that although proteins were

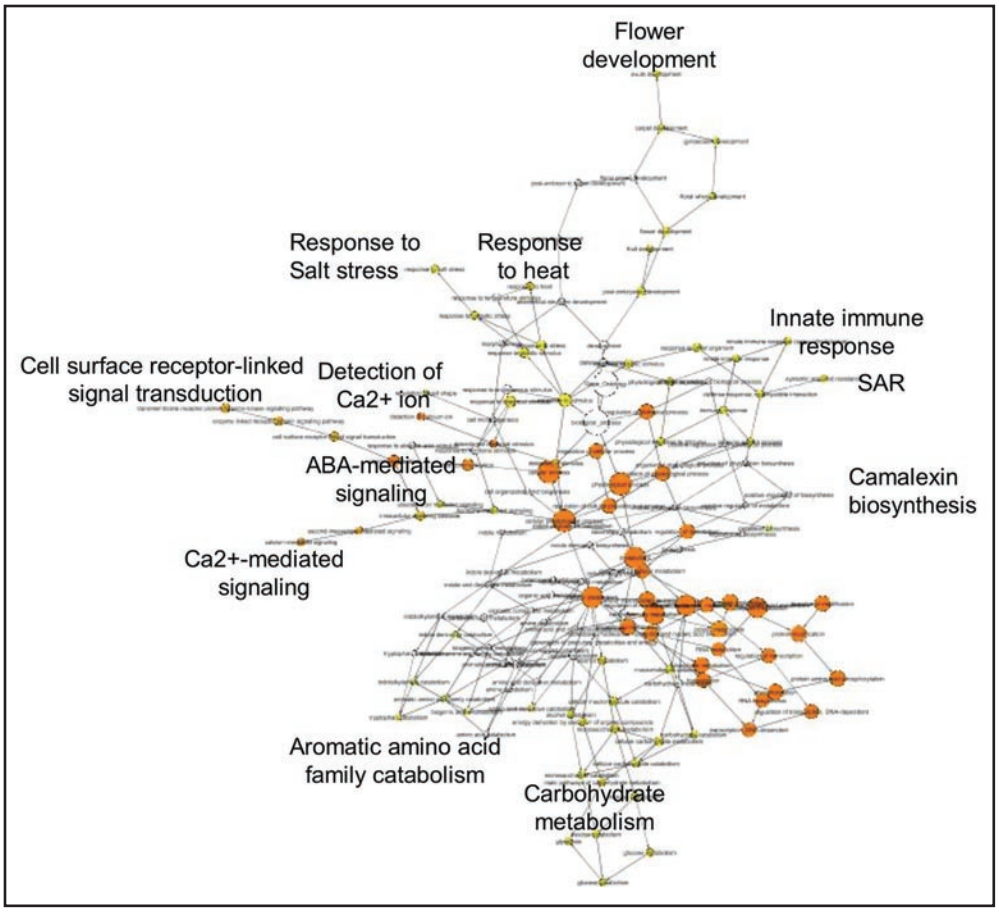

Figure 1. Targets of CaMs/CMLs identified from the protein microarrays belongs to different biological processes based on the Gene Ontology (GO) analyses.

expressed at comparable levels in both systems, the plant-produced proteins retained their enzymatic function after purification compared to yeast expression system.[2]

Based on these results, a large-scale plant-based protein expression and production system was optimized and employed to generate 1133 Arabidopsis proteins. These purified recombinant proteins were used to generate first high density Arabidopsis protein microarrays. To demonstrate the utility, these protein microarrays were probed with several CaM and CML proteins to identify their binding partners.

FUNCTIONAL ANALYSIS OF SEVERAL ARABIDOPSIS CAMS/CMLS TARGETS USING HIGH-DENSITY PROTEIN MICROARRAYS

In contrast to the yeast and vertebrates, plants contain a large family of CaM and CaM-like proteins.[3] The diversity of plant CaM/CMLs structures translates in variability in their target preference. Plant CaM/CML bind a wide variety of substrates.[4] Some of these targets are homologous to animal CaM targets, but some of the targets are unique to plants. To identify other targets of plant CaMs/CMLs, protein microarray containing 1133 proteins were probed with 3 CaMs (CaM1, CaM6 and CaM7) and 4 CMLs (CML8, CML9, CML10 and CML12). Our analysis identified 173 putative CaM/CML substrates, belonging to various protein families. Interestingly, we found that the majority of the targets were specific to one or a small subset of CaMs/CMLs.[2]

Analysis of Gene Ontology (GO) categories in our set of predicted CaM/CML targets identified several over-represented biological processes (Fig. 1 and Table 1). There is a large repertoire of previously identified CaM-binding proteins that are linked to plant response to various environmental stresses. Also, genes encoding specific plant CaM and CaM-related proteins are induced by abiotic factors such as cold, and biotic stressors like fungal proteins, viruses, and pathogenic bacteria.[5]

Protein array screening predicted CaM/CML to interact to a group of proteins with recognized roles in processes related to plant response to biotic stresses such as innate immune response, systemic acquired resistance (SAR) and camalexin biosynthesis. Two of the bZIP transcription factors, TGA3 and TGA6 that bind CaMs are critical components of salicylic acid (SA)-mediated signal transduction pathway.[6,7] CaM is known to enhance the ability of TGA3 to bind a C/G-box promoter element.[8] Another interesting putative target is represented by CYP79B2, a Cytochrome P450 enzyme that functions in the conversion of tryptophan to indole alkaloids (phytoalexins). Phytoalexins such as camalexins are secondary metabolites that accumulate in response to pathogen attack. It has been hypothesized that CaM/Ca^{2+} regulate the synthesis of phytoalexins through at least one component, the CaM-binding IQD1 transcription factor.[9] Another putative CaM-binding protein, MAPKK1, is known to be activated by MEKK1 in response to wounding,[10] and participates in defense responses to bacterial elicitor flagellin.[11]

Heat-shock proteins (Hsp), mediators of cellular responses to elevated temperatures and other environmental factors, are well-known CaM targets. Mammalian CaM binds Hsp proteins.[12] CaM also interacts with the cytosolic maize Hsp70 and inhibits its intrinsic ATPase activity.[13] In our assays, several Arabidopsis heat-shock proteins interacted with CaM/CMLs. Two members of the Hsp90 family, the Heat shock 70 kDa protein 1 (Hsp70-1) and Hsp81-2 interacted with all CaM isoforms.

A group of seven receptor and receptor-like kinases (RLK) putative targets identified on protein arrays are known to function in cell-surface receptor-linked signal transduction (Table 1). LRR-type RLKs are involved in mediating developmental and defense-related cellular responses via binding extra cellular ligands and activating downstream signaling pathways. Moreover, interactions with CaM, phosphorylation or dephosphorylation are processes shown to activate animal and plant RLKs.[14-16]

CONCLUSIONS

Previous research has discovered a large and varied repertoire of CaM/CMLs substrates in plants, thus has led to understanding CaM/CML regulation of many biological processes. Here we demonstrated that use of protein arrays for dissecting CaM/CML-regulated signaling in Arabidopsis. Known and novel potential CaM/CML substrates

were identified in an unbiased and high-throughput manner. The information documented here on novel potential CaM targets provides new testable hypotheses in the area of CaM/Ca^{2+}-regulated processes and represents a resource of functional information for the scientific community.

References

1. Kung LA, Snyder M. Proteome chips for whole-organism assays. Nat Rev Mol Cell Biol 2006; 7:617-22.
2. Popescu SC, Popescu GV, Bachan S, Zhang Z, Seay M, Gerstein M, Snyder M, Dinesh-Kumar SP. Differential binding of calmodulin-related proteins to their targets revealed through high-density Arabidopsis protein microarrays. Proc Natl Acad Sci 2007; 104:4730-5.
3. McCormack E, Tsai YC, Braam J. Handling calcium signaling: Arabidopsis CaMs and CMLs. Trends Plant Sci 2005; 10:383-9.
4. Reddy VS, Ali GS, Reddy ASN. Genes encoding calmodulin-binding proteins in the *Arabidopsis* genome. J Biol Chem 2002; 277:9840-52.
5. Bouche N, Yellin A, Snedden WA, Fromm H. Plant-specific calmodulin-binding proteins. Annu Rev Plant Biol 2005; 56:435-66.
6. Zhang Y, Tessaro MJ, Lassner M, Li X. Knockout analysis of Arabidopsis transcription factors TGA2, TGA5, and TGA6 reveals their redundant and essential roles in systemic acquired resistance. Plant Cell 2003; 15:2647-53.
7. Zhou JM, Trifa Y, Silva H, Pontier D, Lam E, Shah J, Klessig DF. NPR1 differentially interacts with members of the *TGA/OBF* family of transcription factors that bind an element of the *PR-1* gene required for induction by salicylic acid. Mol Plant Microbe Interact 2000; 13:191-202.
8. Szymanski DB, Liao B, Zielinski RE. Calmodulin isoforms differentially enhance the binding of cauliflower nuclear proteins and recombinant TGA3 to a region derived from the Arabidopsis Cam-3 promoter. Plant Cell 1996; 8:1069-77.
9. Nafisi M, Sønderby IE, Hansen BG, Geu-Flores F, Nour-Eldin HH, Nørholm MHH, Jensen NB, Halkier BA. Cytochromes P450 in the biosynthesis of glucosinolates and indole alkaloids. Phytochem Rev 2006; 5:331-46.
10. Hadiarto T, Nanmori T, Matsuoka D, Iwasaki T, Sato K, Fukami Y, Azuma T, Yasuda T. Activation of Arabidopsis MAPK kinase kinase (AtMEKK1) and induction of AtMEKK1-AtMEK1 pathway by wounding. Planta 2006; 223:708-13.
11. Meszaros T, Helfer A, Hatzimasoura E, Magyar Z, Serazetdinova L, Rios G, Bardóczy V, Teige M, Koncz C, Peck S, Bögre L. The Arabidopsis MAP kinase kinase MKK1 participates in defence responses to the bacterial elicitor flagellin. Plant J 2006; 48:485-98.
12. Moriya M, Ochiai M, Yuasa HJ, Suzuki N, Yazawa M. Identification of Ca^{2+}-dependent calmodulin-binding proteins in rat spermatogenic cells as complexes of the heat-shock proteins. Mol Reprod Dev 2004; 69:316-24.
13. Sun XT, Li B, Zhou GM, Tang WQ, Bai J, Sun DY, Zhou RG. Binding of the maize cytosolic Hsp70 to calmodulin, and identification of calmodulin-binding site in Hsp70. Plant Cell Physiol 2000; 41:804-10.
14. Charpenteau M, Jaworski K, Ramirez BC, Tretyn A, Ranjeva R, Ranty B. A receptor-like kinase from *Arabidopsis thaliana* is a calmodulin-binding protein. Biochem J 2004; 379:841-8.
15. Yang T, Chaudhuri S, Yang L, Chen Y, Poovaiah BW. Calcium/calmodulin up-regulates a cytoplasmic receptor-like kinase in plants. J Biol Chem 2004; 279:42552-9.
16. Vanoosthuyse V, Tichtinsky G, Dumas C, Gaude T, Cock JM. Interaction of calmodulin, a sorting nexin and kinase-associated protein phosphatase with the *Brassica oleracea* S locus receptor kinase. Plant Physiol 2003; 133:919-29.