Differential binding of calmodulin-related proteins to their targets revealed through high-density *Arabidopsis* protein microarrays

Sorina C. Popescu*, George V. Popescu*, Shawn Bachan*, Zimei Zhang*, Montrell Seay*, Mark Gerstein†, Michael Snyder*†‡, and S. P. Dinesh-Kumar*‡

Departments of *Molecular, Cellular, and Developmental Biology and †Biochemistry and Biophysics, Yale University, 219 Prospect Street, New Haven, CT 06520-8103

Communicated by Joseph R. Ecker, The Salk Institute for Biological Studies, La Jolla, CA, December 29, 2006 (received for review December 16, 2006)

Calmodulins (CaMs) are the most ubiquitous calcium sensors in eukaryotes. A number of CaM-binding proteins have been identified through classical methods, and many proteins have been predicted to bind CaMs based on their structural homology with known targets. However, multicellular organisms typically contain many CaM-like (CML) proteins, and a global identification of their targets and specificity of interaction is lacking. In an effort to develop a platform for large-scale analysis of proteins in plants we have developed a protein microarray and used it to study the global analysis of CaM/CML interactions. An *Arabidopsis thaliana* expression collection containing 1,133 ORFs was generated and used to produce proteins with an optimized medium-throughput plant-based expression system. Protein microarrays were prepared and screened with several CaMs/CMLs. A large number of previously known and novel CaM/CML targets were identified, including transcription factors, receptor and intracellular protein kinases, F-box proteins, RNA-binding proteins, and proteins of unknown function. Multiple CaM/CML proteins bound many binding partners, but the majority of targets were specific to one or a few CaMs/CMLs indicating that different CaM family members function through different targets. Based on our analyses, the emergent CaM/CML interactome is more extensive than previously predicted. Our results suggest that calcium functions through distinct CaM/CML proteins to regulate a wide range of targets and cellular activities.

calmodulin-interacting proteins | protein interaction network

Calcium (Ca^{2+}), a universal secondary messenger among eukaryotes, has a fundamental role in defining cellular responses to various stimuli (1). Calmodulin (CaM) is one of the most well studied calcium sensors, which react by binding free Ca^{2+} and initiating multiple physiological responses. *Arabidopsis* is an ideal system to study the function of CaM-related proteins. Four *Arabidopsis* CaM isoforms are encoded by seven CaM genes, and they share at least 89% identity to the vertebrate CaMs (1). In addition to CaMs, the *Arabidopsis* genome also encodes 50 CaM-like proteins (CMLs), and they contain CaM-like and/or divergent Ca^{2+}-binding domains (1).

An important step in the understanding of CaM-regulated processes is the comprehensive identification of CaM substrates. Because many eukaryotes have multiple CaM-related proteins, it is important to understand whether these different proteins operate through the same or different targets. Traditional approaches, such as yeast two-hybrid assays, expression library screening, and SDS/PAGE overlay with labeled CaM, have identified CaM-binding proteins in plant and animal systems (2). Although ≈40 CaM targets in plants have been identified by using these approaches, it is expected that many more targets are likely to exist (3). A direct analysis of which CaMs/CMLs bind to the different targets is lacking because the methods used for identifying and characterizing CaM/CML-interacting partners are time-consuming and laborious.

In an attempt to identify targets of CaMs/CMLs and determine their specificity of interactions with different partners, we have developed and used protein microarrays. Protein microarrays allow the high-throughput identification and characterization of molecular interactions. Protein microarrays have been used extensively for the investigation of enzymes properties, protein–protein, protein–phospholipid, and protein–nucleic acid interactions in yeast and mammalian systems. Sensitivity, minimal sample consumption, and ease of use are some of the advantages offered by protein microarrays (reviewed in ref. 4).

To investigate CaM/CML proteins, we constructed an *Arabidopsis* protein microarray containing 1,133 proteins. Probing the array with three CaMs and four CMLs revealed >173 novel *in vitro* binding partners. Analysis of these targets revealed remarkable divergence in the binding of many of the CaMs/CMLs, with each protein binding to unique targets. Our results are consistent with a model in which Ca^{2+} functions through distinct CaM/CML proteins to affect a wide range of diverse targets.

Results

Generation of High-Quality *Arabidopsis thaliana* Expression Clones (ATEC). We constructed a plant expression vector, pLIC-C-TAP (Fig. 1*A*), that transiently overexpresses proteins fused to a tandem affinity purification (TAP) tag. This vector allows high-throughput cloning of PCR products by using ligation-independent cloning (LIC) (5). The TAP tag lies downstream of the expressed ORF and is composed of a 9xMYC epitope, His-6, a rhinovirus 3C protease cleavage site, and the 2xIgG binding domain of protein A.

To generate the ATEC collection, ORFs from full-length cDNAs, ORFeome clones (6), or Col-0 ecotype genomic DNA was PCR-amplified with primers designed to eliminate stop codons. Amplified PCR products were cloned into pLIC-C-TAP vector using a modified LIC cloning method as described in *Materials and Methods*. All inserts were sequenced across both ends of the vector-insert junctions.

The current ATEC collection includes 1,133 *Arabidopsis* ORFs representing 404 putative and known protein kinases, 291 transcription factors, 113 protein degradation-related proteins, 108 proteins with unknown function, 63 heat-shock proteins, 58 cyto-

Author contributions: S.C.P., M. Snyder, and S.P.D.-K. designed research; S.C.P., G.V.P., S.B., Z.Z., and M. Seay performed research; S.C.P., G.V.P., M.G., M. Snyder, and S.P.D.-K. analyzed data; and S.C.P., G.V.P., M. Snyder, and S.P.D.-K. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: CaM, calmodulin; CML, CaM-like; ATEC, *Arabidopsis thaliana* expression clones; TAP, tandem affinity purification; LIC, ligation-independent cloning; BtCaM, bovine CaM; RLK, receptor-like protein kinase; CDPK, Ca^{2+}-dependent protein kinase; CIPK, CBL-interacting protein kinase.

‡To whom correspondence may be addressed. E-mail: savithramma.dinesh-kumar@yale.edu and michael.snyder@yale.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0611615104/DC1.

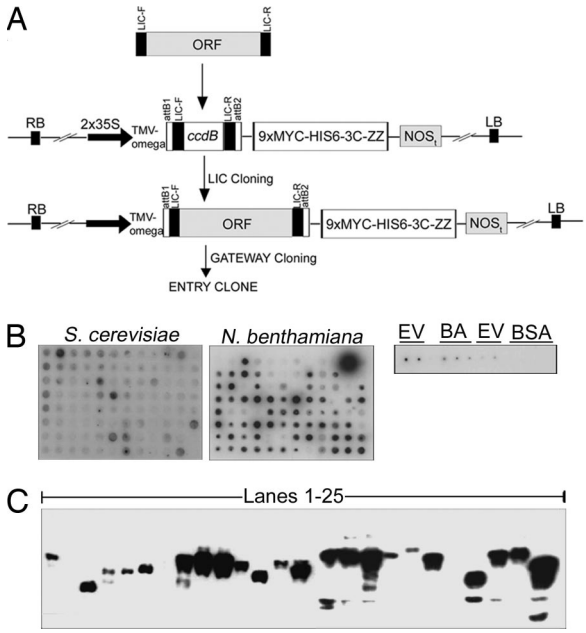

Fig. 1. Expression and purification of recombinant *Arabidopsis* proteins. (*A*) Generation of the ATEC clones. The PCR-amplified ORF without the stop codon flanked by directional LIC forward (LIC-F) and reverse (LIC-R) sequences were cloned in-frame into the TAP tag composed of the 9xMYC epitope–HIS6-3C protease cleavage site–ZZ IgG binding domain. The expression of TAP-tagged ORFs is under the control of duplicated 35S cauliflower mosaic virus promoter (2×35S), a tobacco mosaic virus translational enhancer (TMV-omega), and a NOS terminator (NOS$_t$). RB, T-DNA right border sequence; LB, T-DNA left border sequence. (*B*) (*Left* and *Center*) *In vitro* autophosphorylation activities of 96 protein kinases purified from *S. cerevisiae* (*Left*) and *N. benthamiana* (*Center*). (*Right*) Controls included protein purified from empty-vector-transfected plants (EV), buffer alone (BA), and BSA. (*C*) Western blot analyses of representative recombinant proteins produced in *N. benthamiana* plants using anti-MYC antibodies. Lanes 1–25, different ATEC clones.

chrome P450s, 51 CaMs/CMLs and putative CaM-binding proteins, 35 RNA-binding proteins, and 10 ATP/GTP-binding proteins [see supporting information (SI) Table 3].

Evaluation of Different Expression Systems to Express *Arabidopsis* Proteins. One of the challenges of this work was to employ an expression strategy that resulted in the production of large numbers of high-quality *Arabidopsis* proteins for microarray-based assays. We therefore, initiated a pilot experiment in which a set of 96 protein kinases was produced and purified from a well established yeast expression system (7) and a plant-based expression system.

Immunoblot analyses of purified *Arabidopsis* kinases from yeast revealed that 90% of yeast strains produced detectable fusion proteins (data not shown). However, only 3–5% of the purified *Arabidopsis* kinases from yeast were active in the autophosphorylation assay (Fig. 1*B*). These results suggested that, although we were able to produce proteins using a yeast expression system, they might not be functional for activity.

We therefore examined the expression of the same 96 kinases by using a *Nicotiana benthamiana* transient expression system. ATEC clones were introduced into *Agrobacterium tumefaciens*, and cultures with individual ATEC clones were coinfiltrated with *Agrobacterium* culture containing P19 gene onto the leaves of 4-week-old *N. benthamiana* plants. The P19 protein from tomato bushy stunt virus has been shown to increase transient expression of a transgene (8). Five days after infiltration, tissue was harvested, and soluble recombinant proteins were purified by an affinity purification strategy using IgG-Sepharose beads. The majority (82%) of the protein kinases purified by using this method were found to be active compared with same proteins purified from yeast (Fig. 1*B*). Therefore, we decided to use a plant-based expression system for our studies.

Identification of Putative CaM/CML-Binding Proteins by Using *Arabidopsis* Protein Microarrays. Protein from the 1,133 ATEC clones were expressed in *N. benthamiana* and purified by using the method described above. The quantity and quality of purified proteins was examined by immunoblot analyses using anti-MYC epitope antibodies (Fig. 1*C*, representative blot). The majority of the recombinant proteins (81%) migrated as a single intact band of expected molecular weight, suggesting that full-size proteins were produced. We observed significant differences in abundance among the purified protein. Overall ≈90% of the proteins were detected by immunoblot analyses, with ≈65% of them accumulating at very high levels and the remaining 35% accumulating at moderate levels (data not shown). The differences in protein concentration in the purified samples could be accounted for by such factors as the variation in expression level and/or suboptimal purification conditions. We did not observe any bias toward high-molecular-weight proteins or membrane proteins in terms of expression levels.

To prepare *Arabidopsis* protein microarrays, the 1,133 recombinant protein preparations were printed in duplicate onto nitrocellulose-coated glass slides. Buffer lacking protein was also printed at multiple positions on the slide to serve as a negative control. For detection of the proteins on the microarray, the slides were probed with antibodies to a MYC epitope tag, which crudely indicates the level of each fusion protein (Fig. 2*A*).

To identify CaM/CML-interacting proteins, the protein microarrays were probed in duplicate in the presence of calcium with CaM1, CaM6, CaM7, CML8, CML9, CML10, and CML12, which were amino-conjugated with an Alexa Fluor 647. The *Arabidopsis* protein microarrays were also probed with commercially available Alexa Flour 594-conjugated bovine CaM (BtCaM). Using the scoring system described in *Materials and Methods*, we found a total of 173 different proteins that bound the three CaMs and four CMLs (see SI Table 4). Representative CaM- and CML-interacting proteins are shown in Table 1 and Fig. 2*B*. Of the 173 total targets, 122 proteins interacted with CaM1, 99 with CaM6, 117 with CaM7, 77 with CML8, 113 with CML9, 102 with CML10, 86 with CML12, and 84 with bovine CaM (Fig. 2*C*; see also SI Table 4). Approximately 25% (44 of 173) of the proteins interacted with all CaMs/CMLs, whereas the same percentage of proteins interacted with only one CaM/CML (Fig. 2*D*; see also SI Table 4). The remaining 50% of the proteins bound to two or more CaMs/CMLs. To determine the specificity of binding of CaM/CMLs to their targets, we tested eight proteins that bind all CaM/CMLs in the presence and absence of EGTA. At least in seven cases, binding was reduced significantly in the presence of EGTA (data not shown).

Our analyses identified six of nine previously known *Arabidopsis* homologs of CaM targets, including SAUR protein (At5g20810), a phosphofructokinase (At5g47810), a diacylglycerol kinase (At5g63770), Hsp70–1 (At5g02500), TGA3 (At1g22070), and WRKY21 (At2g30590) (2). However, three targets (At4g30360, At5g60390, and At3g09440) that are predicted to bind CaM were not detected in our analyses. In addition to known targets, we also identified a significant number of CaM/CML targets that were previously uncharacterized (SI Tables 5–7). These targets include 70 putative intracellular and receptor protein kinases, 60 transcription factors, and 43 other classes of proteins. The transcription factors included MADS box, bZIP, MYB, WRKY, Scarecrow, NAM, AUX/IAA, and SAUR_B family proteins. Other classes of proteins included cell-cycle-specific proteins, F-box proteins, RNA-binding proteins, cytochrome P450s, Ca^{2+}-binding EF-hand proteins, and proteins of unknown function.

Arabidopsis CaMs (CaM1–CaM7) show remarkable sequence similarity (1). CaM1 differs from CaM7 by four amino acids, and CaM6 differs from CaM7 by only a single amino acid. Interestingly,

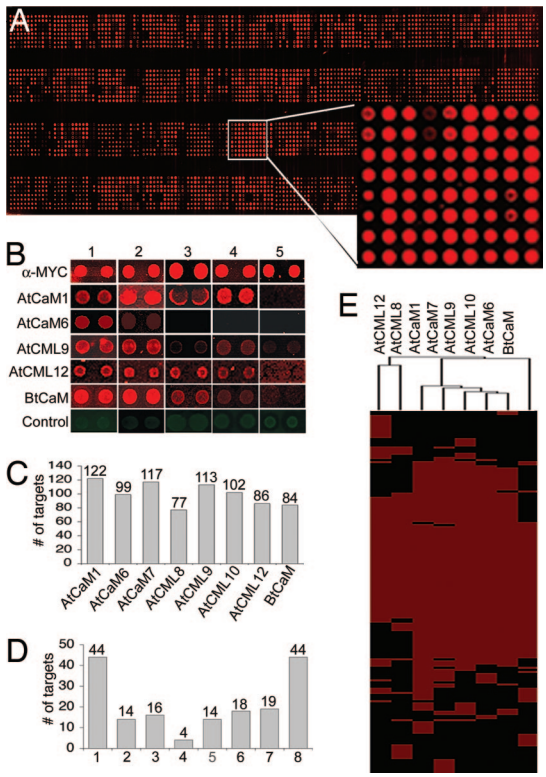

Fig. 2. Characterization of the putative CaM/CML substrates on the *Arabidopsis* protein microarrays. (*A*) Protein microarray containing 1,133 purified protein preparations arrayed in duplicate spots on FAST slides was probed with anti-MYC antibodies. An enlarged image of one block is shown to the right of the protein microarray. (*B*) Representative targets of CaM/CML proteins. The top row shows the amount of each protein preparation on the microarray as detected with anti-MYC antibodies. The bottom row (negative control) shows the signal detected on the microarray for the same protein preparations after probing with the AtCaM1 in the presence of the Ca^{2+} chelator EGTA. Lanes: 1, At2g23080; 2, At1g54610; 3, At5g20810; 4, At4g23650; 5, At1g74740. (*C*) Representation of the total number of CaM/CML targets identified on the protein microarrays. Estimated numbers of substrates binding to each CaM/CML protein are shown above each column. (*D*) Representation of CaM/CML target specificity. Estimated numbers of substrates predicted to bind a number of CaM/CML are shown above each column. (*E*) Hierarchical clustering analysis of predicted CaM/CML-interacting proteins. Columns represent various CaMs/CMLs, and rows represent CaM-binding proteins (see SI Table 8 for gene names in rows). The resultant cladogram, generated by single linkage hierarchical clustering, is shown at the top. Red depicts interaction, and black indicates no interaction. An enlarged high-resolution image is shown as SI Fig. 5.

in our cluster analyses, CaM1 and CaM7 targets cluster together relative to CaM6 targets (Fig. 2*E*; see also SI Fig. 5). Among the targets of CMLs, CML8 and CML12 targets cluster together, and CML9 and CML10 targets cluster together (Fig. 2*E*). The targets of CML8 and CML12 showed the highest divergence compared with the other CaMs/CMLs (Fig. 2*E*). Thus, the amino acid differences of the different CaMs/CMLs confer different target specificities.

The Interaction Network of CaMs/CMLs Targets Revealed Distinct Subnetworks. We used Cytoscape software (9) to generate a bipartite interaction network map between different CaMs/CMLs and their targets (Fig. 3; see also SI Fig. 6). Our analyses identified four hubs containing different CaMs/CMLs based on the similarity in their target preference. The largest hub contains CaM1, CaM6, CaM7, CML9, and CML10. These CaMs/CMLs share more common targets and form a highly interconnected cluster. Hubs containing CML8 and CML12 interacted with a larger number of unique targets or had a more restricted target set. A separate hub in the network represents the BtCaM that showed substrate similarity with the other *Arabidopsis* CaMs/CMLs.

We generated subnetworks corresponding to CaM targets belonging to receptor-like protein kinases, WRKY and TGA transcription factors, and calcium-dependent protein kinases (Fig. 3 *B–D*; see also SI Fig. 7). Some of the predicted targets in the subnetwork analysis were shown previously to bind CaMs. Other predicted targets contain known CaM-binding motifs (SI Fig. 8). These results indicate that protein microarray-based approaches can be used to identify CaM/CML target interaction networks.

***In Vivo* Validation of CaM/CML Targets.** To validate the interactions identified on the protein microarray, *in vivo* coimmunoprecipitation assays were performed for many of the CaMs and their targets. HA-tagged CaM1, CaM6, CaM7, and CML9 were coexpressed in *N. benthamiana* plants with 20 targets tagged with MYC epitope (Table 2). Immunoprecipitation of total protein extracts derived from the cotransfected leaves was performed by using anti-HA antibody, and the complexes were analyzed by immunoblot analysis using anti-MYC antibodies (Fig. 4). Our experiments confirmed 53% (52 of 80) of the interactions that were detected on the protein microarray (Fig. 4 and Table 2). We were able to coimmunoprecipitate 17 targets with at least one CaM or CML. In addition to known targets like SAUR_B and TGA3, several interactions that we confirmed by coimmunoprecipitation analyses were previously unknown. These targets include CIPK6, CIPK24, CDC2-like protein kinases, and an unknown F-box protein, WRKY43, WRKY53 and IAA31 transcription factors, and putative peptidyl-prolyl-*cis*-transisomerases.

Discussion

Functional characterization of proteins depends on the identification of molecules that associate with them. The proteome interaction networks or interactomes will provide not only detailed information on how protein complexes form in the cell but will also help to understand how various cellular pathways communicate and influence each other. Extensive protein interaction maps have been generated for *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, and the human proteome by using automated yeast two-hybrid analyses and TAP followed by mass spectrometry analyses (10–12). Recently, protein microarrays containing the entire proteome of *S. cerevisiae* have been used to study protein–protein, protein–phospholipid, and protein–nucleic acid interactions and glycosylation (reviewed in ref. 4).

In this report, we developed and used a medium-throughput method to express and purify high-quality *Arabidopsis* proteins from a plant expression system. An important characteristic of our protein production method is the use of a homologous expression system for protein expression and purification. The large-scale production of purified recombinant proteins is an important step in many genome-wide approaches for analyzing protein function. In general, prokaryotic organisms are not well suited for eukaryotic protein expression. In addition to the protein solubility and stability challenges, the lack of eukaryotic posttranslational modifications and chaperones restrict the use of the *Escherichia coli* expression system in high-throughput proteomic studies. Only 48% of the 10,167 *C. elegans* ORFs were expressed in the *E. coli* expression system, and of these only 15% were soluble (13). Similar results were obtained for a small group of human proteins expressed in *E. coli* (14). Likewise we found that for protein kinases, the yeast expression system, although suitable for expression of proteins, did not produce them in an active state. Our results demonstrate that a homologous expression system is better for retaining the activity of purified proteins.

We used an *Arabidopsis* protein microarray containing 1,133 protein preparations to study protein–protein interactions involving

Table 1. Representative CaM- and CML-binding proteins identified on protein microarrays

At no.	Description	CaM1	CaM6	CaM7	CML8	CML9	CML10	CML12	BtCaM	No. of hits
At3g12250	TGA6	X								1
At3g44200	Protein kinase	X								1
At3g06030	MAPKKK12		X							1
At1g48150	MADS-box family			X						1
At2g30590	WRKY21			X						1
At1g29340	E3 ligase							X		1
At1g10210	MAPK1	X							X	2
At2g26730	RLK	X		X						2
At4g26270	Phosphofructokinase	X	X			X				3
At1g07530	SCL14	X						X	X	3
At4g18650	TGA2.1		X	X		X				3
At3g05050	CDC2-like Kinase			X	X	X				3
At2g24540	F-box family protein	X		X		X	X			4
At3g01970	WRKY45	X	X	X		X	X			5
At5g66770	SCL4	X	X	X		X	X			5
At1g12680	CDPK-like	X	X	X		X	X		X	6
At4g26070	MAPKK1	X	X	X		X	X		X	6
At3g50310	MAPKKK20	X	X	X	X	X	X		X	7
At1g22070	TGA3	X	X	X		X	X	X	X	7
At2g23080	Casein kinase II	X	X	X	X	X	X	X	X	8
At3g17600	IAA31	X	X	X	X	X	X	X	X	8
At5g20810	SAUR_B	X	X	X	X	X	X	X	X	8

X represents an interaction found in the protein microarray analyses.

seven CaMs/CMLs. In contrast with only a single isoform present in vertebrates, plants contain several CaM/CML proteins classified according to their similarity to the mammalian counterpart (1). Considering the functionally diverse range of plant CaM-binding proteins, it is expected that the variability in plant CaM sequences translates to variability in their target preferences. Indeed, a large fraction of the CaM targets described in the literature are unique to plant CaMs (2, 3).

CaM Interaction with Transcription Factors. A large number (60 of 173) of CaM/CML-interacting proteins were transcription factors. Consistent with these findings, binding of CaM to nuclear factors has been shown to affect their DNA-binding characteristics (15). Our analyses indicated that CaM binds to several TGA and WRKY transcription factors that play a role in the activation of stress or defense pathways. Interestingly, TGA transcription factors were found to interact with the AtCaMs/CMLs but not with the BtCaM. TGA6 associated with only CaM1, whereas TGA3 interacted with CaM1, CaM6, CaM7, CML9, and CML10. It is possible that the observed differences in the binding specificities of various TGAs stem from the diversity in their posttranscriptional modifications and their interaction abilities. Six WRKY proteins, WRKY21, WRKY43, WRKY45, WRKY50, and WRKY53, interacted with one or several CaMs/CMLs.

Our analyses detected associations between CaMs/CMLs and GRAS and AUX/IAA transcription factor family proteins that were not previously known to bind CaM. GRAS family members are known to play a role in root development and gibberellin signaling (16). The AUX/IAA transcription factors regulate many developmental processes governed by auxin (17). Ca^{2+} and CaM roles in auxin signal transduction have been suggested previously based on the observation that CaM interacts with the SAUR protein in several plant species (2). The AtSAUR_B protein on our protein array interacted with all CaMs/CMLs tested. Furthermore, the CaM-binding domain previously mapped in AtSAUR_A, Zm-SAUR1, and SAUR 10A5 is conserved in AtSAUR_B (SI Fig. 8*A*).

CaM Interaction with Receptor-Like Protein Kinases (RLKs). Twenty RLKs and RLK-like proteins were identified as substrates of various CaMs/CMLs. RLKs are involved in mediating developmental and defense-related cellular responses by binding extracellular ligands and activating downstream signaling pathways (18). Moreover, interactions with CaM, phosphorylation, or dephosphorylation are processes shown to activate some animal RLKs (19). *Arabidopsis* Clavata 1, RLK4, SFR1, BRI1, and CRCK1 were shown to be CaM targets (20). Sequence alignment of the twenty RLKs that bind CaMs/CMLs in our analyses indicated that 14 of them contain known CaM-binding domains present in BoSRK (SI Fig. 8*B*).

CaM Interaction with Ca^{2+}-Dependent Protein Kinases (CDPKs) and CBL-Interacting Protein Kinases (CIPKs). CDPKs are regulators of metabolic enzymes and abiotic and biotic stress signal transduction pathways (21). *Arabidopsis* CRK1; tobacco CaMK3, CBK1, and CBK2; maize MCK1, rice CBK; and the apple homolog of the mammalian CaM kinase II were shown to interact with one or more CaMs (2). An *Arabidopsis* homolog, At3g19100 of the NtCaMK3 present on the protein array, interacted with five CaM isoforms. Interestingly, two other CDPKs (CDPK6 and CDPK30) interacted with the CaMs. In addition to CDPKs, three CIPKs (CIPK6, CIPK10, and CIPK24) interacted with various CaM isoforms. Sequence alignment analysis of the newfound targets of CaMs from our analyses with the known CaM-binding domain present in NtCBK1 and AtCRK1 resulted in the identification of a similar domain in the *Arabidopsis* homologs (SI Fig. 8*C*).

CaM Interactions with Proteins Involved in the Cell Cycle. Changes in the intracellular Ca^{2+} concentration were found to have a crucial role in the regulation and progression of cell division. CaM has emerged as the most important Ca^{2+} sensor protein during the cell cycle (22). Studies with mammalian cells indicate that CaM carries out its control either by interacting with cell-cycle-specific proteins or by modulating protein phosphorylation. In mammalian cells, Ca^{2+}/CaM was shown to modulate the activity of cdc2-type CDKs and cyclin B proteins (23). Similarly, plant CDKs perform crucial roles during cell division but are also activated after abiotic stress, wounding, and hormone treatment in a cell-cycle-independent manner (24). We found that a plant-specific B-class CDK, CDKB2, interacts strongly with all CaMs. Additionally, two other CDK-like proteins interacted with various CaMs on the array. These results advocate for involvement of CaM in the regulation of plant cell

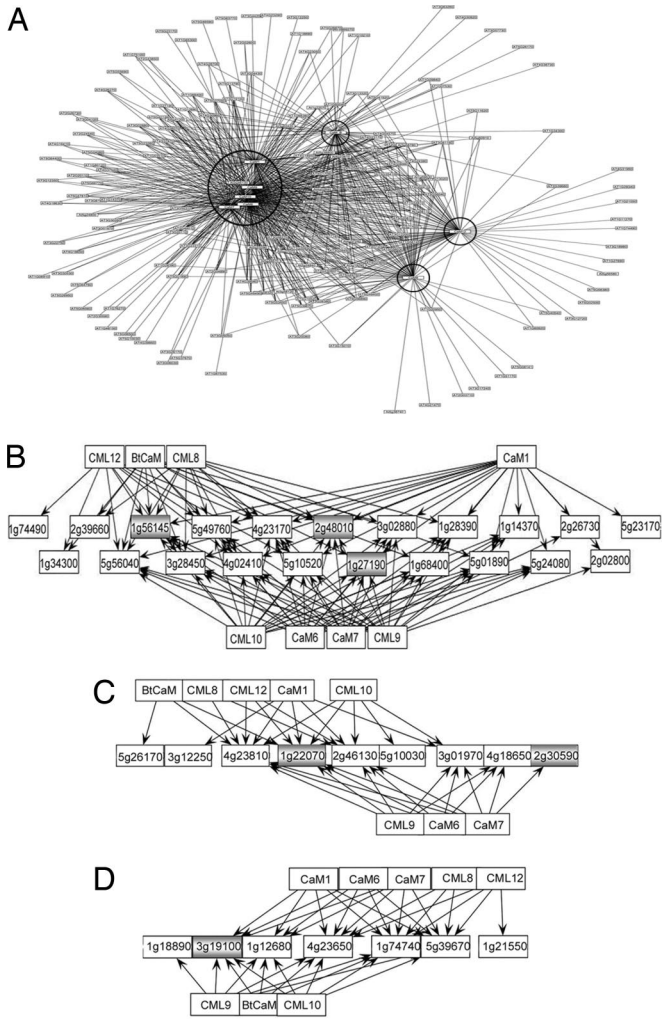

Fig. 3. The interaction network of CaM/CMLs and their predicted substrates. (*A*) A bipartite interaction network constructed with Cytoscape. The network contains 181 nodes representing the eight CaM/CMLs and 173 predicted substrates. The nodes are represented by rectangles and contain The *Arabidopsis* Information Resource database accession numbers of the predicted targets or the protein symbols of CaM/CMLs. The network contains 799 edges, which are depicted by arrows that connect the nodes. The inside network circles pinpoint the hubs of the CaM/CML interaction network. The CaM/CMLs that constitute a hub are most similar in terms of their substrate preference. An enlarged high-resolution image is shown as SI Fig. 6. (*B–D*) Bipartite subnetworks of CaM/CML-binding RLKs (*B*), transcription factors (*C*), and calcium-dependent protein kinases (*D*). Shadowed rectangles represent previously shown interaction with CaMs. An enlarged high-resolution image is shown as SI Fig. 7.

cycle events specifically at G_2 to mitosis transition, possibly through a direct interaction with CDKs and/or CDK/cyclin complexes.

Conclusions

The results presented here not only reaffirm the known functions of plant CaMs/CMLs but also greatly expand our current knowledge of their role in growth, development, and during stress and defense responses. The information on novel potential CaM/CML targets generated in this project provides testable hypotheses in the area of CaM/Ca^{2+}-regulated processes and represents a great resource of functional information. Identification of several known plant and homologs of mammalian CaM-binding proteins in our screens demonstrated the suitability of our protein expression system and the successful application of algorithms for protein–protein interaction prediction. However, further analyses are necessary to validate these predicted interactions, and a major challenge will be to experimentally characterize them *in planta*. Finally, our work demonstrates that plant protein microarrays are valid tools for the large-scale characterization of protein–protein interactions and that our system could be used further in developing the interactome of the entire *Arabidopsis* proteome.

Table 2. Confirmed CaM/CML targets by coimmunoprecipitation assays

At no.	Description	CaM1	CaM6	CaM7	CML9
At2g23080	Casein kinase II alpha	X			
At4g30960	CBL-interacting protein kinase	X			
At5g35410	CBL-interacting protein kinase	X		X	X
At1g54610	CDC2-like kinase family	X		X	
At1g74740	Calcium-dependent protein kinase	X		X	X
At1g20930	CDC2-like kinase family	X	X	X	X
At4g23650	CDPK		X	X	X
At4g17080	PI-4-phosphate-5-kinase-related				
At5g20810	SAUR_B, putative	X	X	X	X
At5g56030	Heat-shock protein 81 and 82	X	X	X	X
At3g56070	Peptidyl-prolyl-*cis-trans* isomerase	X	X	X	X
At3g50310	MAPKKK20; similar to NPK1	X	X		X
At3g59150	Cyclin-like F-box family protein	X			
At4g08980	F-box family protein				
At4g24390	F-box family protein				
At5g39670	Calcium-binding EF-hand family	X			
At2g46130	WRKY43	X			
At4g23810	WRKY53	X	X		X
At3g17600	IAA31		X		
At1g22070	TGA3	X		X	X

X represents confirmed interaction.

Materials and Methods

Construction of ATEC Clones. Each *Arabidopsis* ORF was amplified by using gene-specific primers containing LIC-compatible extensions, either from cDNA or ORFeome (6) collections or from Col-0 genomic DNA. The primers included the first 25 or 26 nucleotides of the ORF, including the ATG start codon, and the last 25 or 26 nucleotides, excluding the stop codon. The LIC forward (5′-GCACAAGAAGGTCC-3′) and reverse (5′-GAAGAAGAGAGGTGC-3′) sequences were added at the 5′ ends of the each primer set. The PCR conditions used to clone ORFs are available upon request.

Purified PCR product (50 ng) and pLIC-C-TAP StuI-cut vector were treated with T4 DNA polymerase in the presence of 25 mM dATP and dTTP, respectively, for 30 min at 22°C followed by heat inactivation at 75°C for 20 min. A 10-μl mixture of treated PCR fragments and vector was incubated at 65°C for 2 min and 22°C for 10 min and transformed into DH10B *E. coli* cells. Plasmid DNA prepared from the transformants was sequenced from both ends of the insert and verified against The *Arabidopsis* Information Resource database.

Protein Expression in Plants and Purification. Sequence-verified ATEC clones were transformed into GV2260 *A. tumefaciens*. A single *Agrobacterium* containing ATEC construct was cultured for 14–16 h at 26°C. The bacteria were pelleted at 3,000 × *g* for 20 min, resuspended in infiltration media (10 mM MES/10 mM $MgSO_4$/200 μM acetosyringone) and adjusted to an OD_{600} of 0.5 and incubated

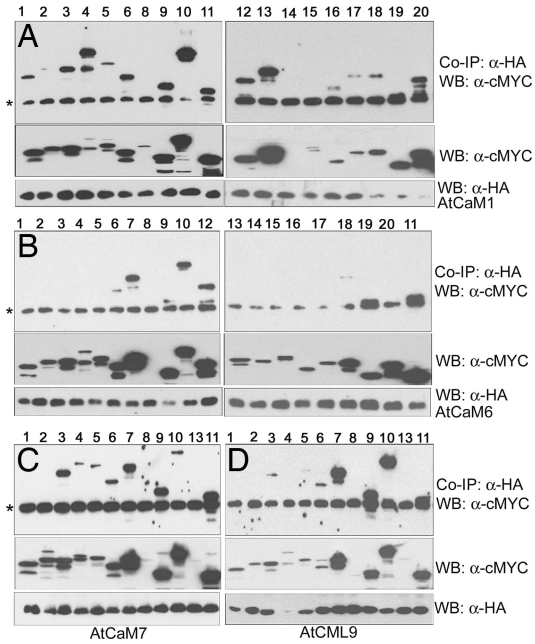

Fig. 4. Coimmunoprecipitation analysis of selected CaMs/CMLs and their targets. Dual combinations of CaM1 (*A*), CaM6 (*B*), CaM7 (*C*), or CML9 (*D*) tagged with HA epitope and various targets tagged with MYC epitope were coexpressed in *N. benthamiana* plants. Coimmunoprecipitation (Co-IP) assays were performed on total soluble protein by using anti-HA antibody (α-HA), and Western blot (WB) analyses were performed by using anti-MYC antibodies (α-MYC). Asterisks mark the nonspecific reaction of the antibody with the large subunit of IgG. Numbers on the top of each blot correspond to CaM/CML targets listed in Table 2.

at room temperature (RT) for at least 2 h. *Agrobacterium* cultures transformed with P19 were resuspended in infiltration media to an OD_{600} of 0.8–0.9. Equal volumes of ATEC and P19 cultures were mixed and infiltrated into *N. benthamiana* plants by using a 1-ml needleless syringe. Four days after infiltration, the leaves were collected and stored at −80°C.

Tissue was ground in liquid nitrogen and mixed 1:1 (vol/wt) with an extraction buffer [100 mm Tris·HCl (pH 7.5)/150 mM NaCl/5 mM EDTA/10 mM 2-mercaptoethanol/10% glycerol/0.1% Triton X-100/1× Complete protease inhibitors (Roche, St. Louis, MO)/ 1 mM PMSF] and 50 μl of 0.5-mm glass beads (Biospec Products, Bartlesville, OK). The mixture was processed by beating in a paint shaker (5G-HD; Harbil, Wheeling, IL) three times for 1 min each, and lysates were centrifuged for 10 min at 14,000 × *g*. The supernatants were loaded onto a 96-well, 2-ml glass-filled polypropylene box (GF/C UniFilter, Whatman, Brentford, United Kingdom) and centrifuged for 15 min at 1,000 × *g*. The supernatants were collected in a 96-well PCR plate (ABgene, Epsom, United Kingdom) and IgG Sepharose 6 Fast Flow beads (Amersham Biosciences, Uppsala, Sweden) was added and incubated at 4°C on shaker for 2 h. After centrifugation at 3,000 × *g* for 1 min, IgG beads were washed three times in washing buffer (extraction buffer plus 350 mM NaCl) and incubated with 1 unit of PreScission protease (Amersham Biosciences) at 4°C for 10 h. Mixtures were centrifuged at 3,000 × *g* for 1 min, and the supernatants containing recombinant proteins were transferred to new 96-well PCR plates, mixed with glycerol to a final concentration of 30%, and stored at −80°C.

Protein Microarray Printing and Probing. Each protein was arrayed in duplicate spots onto FAST slides (Schleicher & Schuell, Keene, NH) with 48-pin contact printer (Chip Writer Pro; Bio-Rad, Hercules, CA). The arrayed proteins were air-dried at 4°C and transferred to −80°C for storage. Slides were blocked in SuperBlock buffer (Pierce, Rockford, IL) for 1–2 h at RT and incubated with a 1:2,500 dilution of MYC epitope antibodies (Santa Cruz Biotechnology, Santa Cruz, CA) for 1 h. Slides were washed three times in TBS-T (20 mM Tris/137 mM NaCl/0.1% Tween 20) and incubated in the Cy5-conjugated anti-mouse IgG (Jackson ImmunoResearch Laboratories, West Grove, PA) for 1 h at RT. Slides were washed three times in TBS-T, spun-dried, and scanned in a Genepix 4200A scanner (Axon Instruments, Sunnyvale, CA).

Protein microarrays were probed with Alexa Fluor 647-labeled AtCaMs/CMLs or Alexa Fluor 594-conjugated BtCaM (Molecular Probes, Carlsbad, CA). Probing buffer (200 μl; 50 mM Tris·HCl, pH 7.5/150 mM NaCl/ 5 mM $CaCl_2$) containing 50 μl of purified labeled CaM/CML was applied onto each slide and covered with HybriSlip. The slides were incubated for 1 h in a humid chamber at RT, washed three times in a probing buffer containing 0.1% Tween 20 and 1% glycerol. Slides were spun-dried and scanned in a Genepix 4200A scanner.

Bioinformatic Tools. The scanned microarray images were subsequently preprocessed by using the Genepix software to obtain report files containing the mean, median, and standard deviation of array spots and background regions. The data from Genepix report files were further analyzed with Matlab to evaluate the efficiency of the protein binding on slides. Microarray data were then processed with Perl and Awk scripts to obtain protein–protein interaction data. Protein interaction prediction, hypothesis testing, and network analysis were performed by using Matlab scripts. CaM protein interaction profiles were grouped with Cluster and visualized by using TreeView. The protein interaction network was visualized with Cytoscape (11). CaM/CML target predictions were performed by using the methods described in *SI Materials and Methods*.

We thank the *Arabidopsis* Biological Resource Center for providing Drs. J. Ecker and A. Theologis's cDNA and ORFeome clones. This work was supported by a Yale University Genomics and Proteomics Center Pilot Grant and by National Science Foundation Grants MCB-0313514 and DBI-0519853.

1. McCormack E, Tsai, YC, Braam J (2005) *Trends Plants Sci* 10:383–389.
2. Bouche N, Yellin A, Snedden WA, Fromm H (2005) *Annu Rev Plant Biol* 56:435–466.
3. Reddy VS, Ali GS, Reddy ASN (2002) *J Biol Chem* 277:9840–9852.
4. Kung LA, Snyder M (2006) *Nat Rev Mol Cell Biol* 7:617–622.
5. Dieckman L, Gu M, Stols L, Donnelly MI, Collart FR (2002) *Protein Expr Purif* 25:1–7.
6. Yamada K, Lim J, Dale JM, Chen H, Shinn P, Palm CJ, Southwick AM, Wu HC, Kim C, Nguyen M, *et al.* (2003) *Science* 302:842–846.
7. Zhu H, Klemic JF, Chang S, Bertone P, Casamayor A, Klemic KG, Smith D, Gerstein M, Reed MA, Snyder M (2000) *Nat Genet* 26:283–289.
8. Voinnet O, Rivas S, Mestre P, Baulcombe D (2003) *Plant J* 33:949–956.
9. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) *Genome Res* 13:2498–2504.
10. Parrish JR, Gulyas KD, Finley RL, Jr (2006) *Curr Opin Biotechnol* 17:387–393.
11. Gagneur J, David L, Steinmetz LM (2006) *Trends Microbiol* 14:336–339.
12. Andrews BJ, Bader GD, Boone C (2003) *Nat Biotechnol* 21:1297–1299.
13. Luan C-H, Qiu S, Finley JB, Carson M, Gray RJ, Huang W, Johnson D, Tsao J, Reboul J, Vaglio P, *et al.* (2004) *Genome Res* 14:2102–2110.
14. Dyson M, Shadbolt S, Vincent K, Perera R, McCafferty J (2004) *BMC Biotechnol* 4:32.
15. Onions J, Hermann S, Grundstrom T (2000) *Biochemistry* 39:4366–4374.
16. Pysh LD, Wysocka-Diller JW, Camilleri C, Bouchez D, Benfey PN (1999) *Plant J* 18:111–119.
17. Reed JW (2001) *Trends Plants Sci* 6:420–425.
18. Boller T (2005) *Curr Opin Cell Biol* 17:116–122.
19. Martin-Nieto J, Villalobo A (1998) *Biochemistry* 37:227–236.
20. Yang T, Chaudhuri S, Yang L, Chen Y, Poovaiah BW (2004) *J Biol Chem* 279:42552–42559.
21. Hrabak EM, Chan CM, Gribskov M, Harper JF, Choi JH, Halford N, Kudla J, Luan S, Nimmo HG, Sussman MR, *et al.* (2003) *Plant Physiol* 132:666–680.
22. Santella L (1998) *Biochem Biophys Res Commun* 244:317–324.
23. Colomer J, Lopez-Girona A, Agell N, Bachs O (1994) *Biochem Biophys Res Commun* 200:306–312.
24. Zhiponova MK, Pettko-Szandtner A, Stelkovics E, Neer Z, Bottka S, Krenacs T, Dudits D, Feher A, Szilak L (2006) *Plant Physiol* 140:693–703.