

## Study of *Arabidopsis thaliana* resistome in response to cucumber mosaic virus infection using whole genome microarray

Rajendra Marathe<sup>1</sup>, Zhong Guan<sup>2</sup>, Radhamani Anandalakshmi<sup>1</sup>, Hongyu Zhao<sup>2</sup> and S.P. Dinesh-Kumar<sup>1,\*</sup>

<sup>1</sup>Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520 (\*author for correspondence; e-mail savithramma.dinesh-kumar@yale.edu); <sup>2</sup>Department of Epidemiology and Public Health, Yale University, New Haven, Connecticut 06520, USA

Received 11 March 2004; accepted in revised form 27 May 2004

**Key words:** CMV, disease resistance, microarray, RCY1, R gene, Signaling

### Abstract

The plant innate immune response is mediated by resistance (*R*) genes and involves hypersensitive response (HR) cell death. During resistance responses, the host undergoes net changes in the transcriptome. To understand these changes, we generated a whole genome transcript profile for *RCY1*-mediated resistance to cucumber mosaic virus strain Y (CMV-Y) in *Arabidopsis*. Using a very stringent selection criterion, we identified 444 putative factors belonging to nine different functional classes that show significant transcript regulation during *Arabidopsis*-CMV-Y interaction. Genes with unknown function formed the largest class. Other functional classes represented in the resistome include kinases and phosphatases, protein degradation machinery/proteases, transcriptional regulators, and others. Interestingly, several of the unknown function genes possess well characterized domains and secondly many genes encode small peptides with less than 100 amino acids. Analysis of 1.1 kb promoter regions of the 444 genes revealed that 9 out of the 12 known cis-binding elements are significantly associated with pathogen responsive cluster. Location and distribution of five prominent binding elements for select group of disease resistance related and unknown function genes is presented. The analysis also revealed 80 defense-responsive genes that might participate in *R* gene-mediated defense against both viral and bacterial pathogens. In addition, chromosome distribution of genes that respond to bacterial and viral pathogens suggests that they are located in small gene clusters and may be transcriptionally co-regulated. Exploring the precise function of the new genes identified in this analysis will offer new insights into plant defense.

### Introduction

The defense responses exhibited by plants against pathogens can be grouped into two major categories: non-host resistance that occurs at the species level and host-specific resistance that occurs at the subspecies level (race/ecotype/variety). Non-host resistance is triggered in response to infection by different pathogen-derived signals, which can be of bacterial, fungal or even plant origin. Host-specific resistance on the other hand, requires the recognition of pathogen-encoded Avr proteins by their cognate plant encoded resistance gene (*R*) products

(Martin *et al.*, 2003). This Avr-R recognition and the subsequent events restrict the pathogen to site of its entry and limit further spread in a resistant plant. Phenotypically, this resistance often manifests itself in the form of the hypersensitive response (HR) at the site of pathogen invasion and enhanced resistance in the other regions of the plant. HR is a form of programmed cell death (PCD) that occurs very early during the infection of a resistant plant by an avirulent pathogen (Lam *et al.*, 2001). HR occurs very rapidly and precedes pathogen movement, thereby limiting systemic spread of the pathogen. A delayed, secondary

response called systemic acquired resistance (SAR) follows PCD and provides a general immunity to entire plant against subsequent attack by the same or other pathogens (Ryals *et al.*, 1996).

*R* genes that confer resistance to viral, fungal, bacterial, nematode and insect pathogens encode structurally similar proteins. *R* genes are classified into eight groups, the largest one being NBS–LRR (Nucleotide Binding Site–Leucine Rich Repeat) proteins with variable N-terminal domains (Hammond-Kosack and Parker, 2003). All the virus specific *R* genes that have been cloned so far namely *N* (against tobacco mosaic virus, TMV), *HRT* (against turnip crinkle virus, TCV), *Rx1* (against potato virus X, PVX) and *RCY1* (against cucumber mosaic virus yellow strain, CMV-Y), encode proteins that contain a centrally located NBS and C-terminal LRR of various lengths (Whitham *et al.*, 1994; Bendahmane *et al.*, 1999; Cooley *et al.*, 2000; Takahashi *et al.*, 2002). These virus specific *R* proteins have different structures at their amino-terminus: *N* contains a toll-interleukin 1 receptor homology region (TIR), while *HRT* and *RCY1* have a coiled-coil (CC) domain, and *Rx1* contains neither. Interestingly, the NBS region of *R* genes share sequence homology with the NBS region of animal cell death genes including *CED4* from *C. elegans*, and Apaf-1, FLASH, CARD4 and NOD1 from humans.

The *R* gene *RCY1* found in *Arabidopsis* C24 ecotype is a single dominant gene that mediates resistance against CMV-Y (Takahashi *et al.*, 1994, 2002). The HR triggered by *RCY1* culminates in death of the infected cells and is seen as necrotic lesions. Takahashi *et al.*, (2002) have shown that this gene is allelic to both *RPP8* from the ecotype Ler and *HRT* from the ecotype Dijon-17, which confer resistance against the fungus *Peronospora parasitica* Emco5 and turnip crinkle virus (TCV), respectively (McDowell *et al.*, 1998; Cooley *et al.*, 2000;). It appears that through intragenic recombination and positive selection *Arabidopsis* ecotypes have evolved at least 6 alleles at this locus (*rpp8*, K15 and *RPH8A* being the other three) (McDowell *et al.*, 1998; J. Miller and S.P.D-K, unpublished).

During plant defense responses, induction of HR and the restriction of pathogen spread are mediated by the activation, repression or steady state maintenance of transcriptional activity of a group of genes or transcriptome that we refer to as

the “resistome”. The transcriptional regulation during *R* gene-mediated resistance and SAR response has been studied using expressed sequence tag (EST)-based cDNA microarrays representing partial *Arabidopsis* genome (Maleck *et al.*, 2000; Schenk *et al.*, 2000). Recently, interactions involving two NBS–LRR type of LRR *R* genes, *RPS2* and *RPM1*, were studied using Affymetrix oligo based GeneChip representing 8000 genes (Tao *et al.*, 2003). They found that 2338 genes were differentially regulated within 9 h of bacterial infection. Similar Affymetrix GeneChips were used to study changes in gene expression in response to five different viruses in *Arabidopsis* (Whitham *et al.*, 2003). This study identified 135 common genes that are differentially regulated following viral infections.

In the present report, we analyzed the *RCY1*-mediated resistance response against the viral pathogen CMV-Y using Affymetrix GeneChips representing the entire genome of *Arabidopsis*. Gene expression profiles at three time points (3, 5 and 18 h) following CMV-Y infection were investigated. We employed 2-way ANOVA and identified a total of 444 genes that are differentially regulated during defense response to CMV-Y infection. The identified genes belong to several functional groups such as kinases, protein degradation machinery, transcription factors, proteins involved in HR cell death and defense response. We also present an analysis of promoters of all 444 genes regulated during *RCY1*-mediated resistance. Frequency and location for the binding sites for most abundant cis factors on the promoters of important genes are presented. Finally we compare the genes identified in our study to those generated by Tao *et al.* (2003) during resistance response to bacterial infections in *Arabidopsis*. Since *RCY1* belongs to NBS-LRR group the whole genome expression analysis presented here may be significant for other resistance pathways governed by NBS–LRR type of *R* genes in *Arabidopsis* and other plant systems.

## Materials and methods

### *Plant growth, virus inoculation, and RNA preparation*

*Arabidopsis thaliana* ecotype C24 plants were grown in a growth chamber maintained at 16 h

light and 8 h dark cycle, 22 °C temperature and 72% relative humidity. The virus (50 µg/ml) was rub-inoculated onto all the leaves of C24 plants at 10–12-leaf stage using carborundum. The tissue was collected at 0, 3, 5, and 18 h pi and quick frozen in liquid nitrogen.

Total RNA was extracted from virus- or mock-inoculated leaves using RNA WIZ (Ambion, TX) and RNA was further purified using RNAeasy columns (Qiagen). GeneChip hybridization and scanning was performed at the Affymetrix GeneChip resource laboratory within the Keck Facility at Yale Medical School, New Haven, CT (<http://keck.med.yale.edu/affymetrix>).

### Microarray data analysis

The GeneChip (ATH1) data of our experiment was processed using DNA-Chip Analyzer (dChip: <http://www.dchip.org/>) which implements Li–Wong’s model (Li and Wong, 2001). We used PM-only model to calculate gene expression values based on GeneChip cell files. Then the gene expression data was normalized using a built-in invariant-set normalization function in dChip and was further analyzed by the following analysis of variance (ANOVA) model

$$y_{ijk}^{(r)} = \mu + A_i + B_j + G_k + (AG)_{ik} + (BG)_{jk} + \beta_{ijk}^{(r)},$$

where  $y_{ijk}^{(r)}$  is the logarithm of the  $r$ th replicate model-based -expression value of gene  $k$  at the  $i$ th time point with the  $j$ th treatment ( $i=1,2,3,4$ ;  $j=1,2$ ;  $k=1,\dots,\sim 22k$ ). The  $R$ -square of this model is  $R^2 = 0.9949531$ . This means that about 99.5% of the variations can be explained by the model. As in many other literatures in which ANOVA model

has been used, we treated all the expression values as independent observations from normal distributions. We only included the important effects and some interactive effects related to genes to keep the model as simple as possible. We have not included the interaction  $(AB)_{ij}$  because in this ANOVA model, the  $R^2$  is calculated based on residual sum of squares which has very large (68,238) degrees of freedom. The interaction  $(AB)_{ij}$  has only three degrees of freedom.

The ANOVA table (Table 1), summarizes some test results about the variable selection in the ANOVA model fitting the microarray data. In this table, DF, SS, MS,  $F$ -ratio and  $p$ -value designate for degree(s) of freedom, sum of squares, mean sum of squares,  $F$ -test statistic and its  $p$ -value, respectively. From the ANOVA table (Table 1), we know that all the effects and interactions are very significant. The fold change for gene  $k$  is estimated as  $fc = \exp\{(bg)_{2k} - \{(bg)_{1k}\}$ , where  $(bg)$  is the least-squares estimate of the interaction (BG) between treatment and gene. This is the change in the gene expression levels which is caused only by the different treatments (virus infected or not). Changes in expression levels caused by other factors such as time and array are filtered and are estimated separately.

The fold change of each gene at each time point is calculated as the average of the ratios of the two replicates. The selected changes at the three time points were transformed to logarithm (base 2) values and analyzed by “Cluster” developed by Michael Eisen (<http://rana.lbl.gov/EisenSoftware.htm>) by performing the complete linkage hierarchical clustering analysis with uncentered correlations. The clustering results are displayed and viewed using ‘TreeView’.

Table 1. Summary of analysis of variance (ANOVA) model\*.

Source	DF	SS	MS	f-ratio	$p$ -value
Time	3	47.534	15.845	1056.333	< 1.0e-7
Treatment	1	9.37	9.37	624.667	< 1.0e-7
Gene	22745	201418.6	8.856	590.4	< 1.0e-7
Treatment × Gene	22745	469.401	0.021	1.4	< 1.0e-7
Time × Gene	68235	4607.372	0.068	4.533	< 1.0e-7
Residual	68238	1047.728	0.015		
Total	181967	207600	1.141		

\*Source, DF, SS, MS,  $F$ -statistic,  $p$ -value stand for source of variation, degree(s) of freedom, sum of squares, mean sum of squares,  $F$ -test statistic and its  $p$ -value.

“Treatment × Gene” and “Time × Gene” are sources of variation due to the interaction of gene with treatment and time respectively.

We obtained the promoter sequences for different groups of selected genes from MATDB (<http://mips.gsf.de>). The frequencies and locations of the twelve binding sites were calculated by examining the upstream 1.1 kb region of each gene. The statistically expected frequency was calculated by examining the upstream 1.1 kb regions of all *Arabidopsis* genes.

#### Visual display of the results – A spatial method

To visualize the chromosomal locations of genes with statistically significant changes we counted the number of differentially regulated genes in a series of moving windows along a chromosome. More specifically, for each chromosome, the starting and the end points of the first and the last genes were denoted by  $N_0$  and  $N_1$ , respectively. For a given window size  $w$  and moving step  $k$ , we created  $M = (N_1 - N_0) / k$  moving windows with width  $w$ . For the  $m$ th move, where  $m = 1, 2, \dots, M$ , the window covers  $[(m-1)k+1, (m-1)k+1+w]$ . For each window, we counted the relative frequency  $f(m)$  of the significant genes within the  $m$ th moving, that is,  $f(m)$  is ratio of the number of significant genes to the number of genes in the ATH1 Affymetrix array within the  $m$ th window. In each of the figures, we plotted the relative frequencies  $f(m)$  vs. the starting points  $(m-1)k+1$  of the moving windows for the sets of significant genes described in Tao *et al.* (2003) and those of ours.

## Results

#### Experimental system and gene expression analysis

*Arabidopsis* ecotype C24 contains a single dominant *R* gene *RCY1* that confers resistance to CMV-Y (Takahashi *et al.*, 1994, 2002). The resistance response mediated by *RCY1* is manifested in the form of HR cell death in the inoculated leaves within 48 to 72 h following viral infection (Figure 1). To identify genes that regulate early *RCY1*-CMV resistance signaling, we selected 3, 5 and 18 h post-infection (pi) time points for genome-wide expression profiling using Affymetrix GeneChips with 22,500 probe sets representing approximately 24,000 genes. Total RNA was extracted from the pooled tissue of five independent C24 plants infected with CMV-Y or mock inocu-

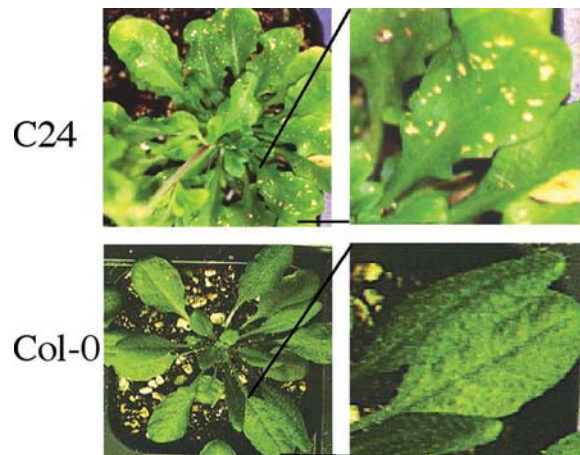


Figure 1. *Arabidopsis* ecotypes C24 and Col-O respectively are resistant and susceptible to CMV-Y. C24 plant infected with CMV-Y exhibiting HR cell death. All the leaves of the plant were hand inoculated by rubbing virus. Close up of HR cell death on inoculated leaf is shown. Col-0 plant infected with CMV-Y showing no HR. Absence of HR is evident in the close up view of the inoculated leaf.

lated at three time points (3, 5 and 18 h pi) and hybridized with Affymetrix GeneChips. Experiments were carried out in two biological replications using a total of ten individually hand inoculated plants per time point.

To evaluate technical or biological variability, we analyzed replication clusters for both mock and virus treated data sets (normalized) and determined that the two replications in each case display same genes clustering in similar orders (supplement Figure 1). This mathematically illustrated that the replications gave highly reproducible data. In addition, the real time RT-PCR analysis on seven genes across all three time points (comparing 21 ratios of virus treated to mock control) using total RNA from a fresh set of similarly treated plants confirmed our microarray data (supplement Table 1).

Gene expression data was analyzed by the ANOVA model (Kerr *et al.*, 2000) to identify differentially expressed genes. False Discovery Rate (FDR) (Storey and Tibshirani, 2003) was calculated based on the  $p$ -value from the ANOVA. We used a stringent cut off at  $FDR = 0.05$  that corresponds to  $p$ -value = 0.001, and selected 444 genes that were differentially regulated in *RCY1*-containing C24 ecotype due to CMV-Y infection compared to mock infected controls (supplement Table 2). The  $FDR = 0.05$  represents approximately 5% of the identified genes to be false positive.

Table 2. Representative genes consisting of defense related, ROI related and *R* gene-like functional groups in the RCY1 resistome against CMV-Y.

ID	AGI	Description	FDR	<i>p</i> -value	Est. Ratio
Defense related					
263382_at	At2g40230	Anthranilate N-hydroxycinnamoyl, putative	0.000279	0.000002	0.661
258805_at	At3g04010	$\beta$ -1,3-glucanase, putative	0.000002	0.000000	0.605
248100_at	At5g55180	$\beta$ -1,3-glucanase-like protein	0.008440	0.000114	0.713
255595_at	At4g01700	Chitinase, putative	0.034100	0.000630	0.741
264514_at	At1g09500	Cinnamyl alcohol dehydrogenase, putative	0.000003	0.000000	0.607
252943_at	At4g39330	Cinnamyl-alcohol dehydrogenase CAD1	0.000000	0.000000	0.562
252984_at	At4g37990	Cinnamyl-alcohol dehydrogenase ELI3-2	0.010400	0.000144	0.717
251895_at	At3g54420	Chitinase, class IV	0.012900	0.000190	0.721
261135_at	At1g19610	Defensin AMP1, putative	0.001240	0.000013	1.466
251879_at	At3g54200	Hin1 protein, putative	0.008710	0.000118	0.714
252317_at	At3g48720	Hsr201 like	0.000563	0.000005	0.671
253104_at	At4g36010	Thaumatococin-like	0.000074	0.000001	0.644
260799_at	At1g78270	UDP-glucose glucosyltransferase, putative	0.001150	0.000012	1.468
256469_at	At1g32540	Zinc-finger protein, putative; Lsd1-like	0.026400	0.000457	1.360
ROI related					
248353_at	At5g52320	Cyt P450	0.000009	0.000000	0.619
256598_at	At3g30180	Cyt P450 homolog	0.001640	0.000018	1.457
253101_at	At4g37430	Cyt P450 monooxygenase	0.000163	0.000001	0.653
253073_at	At4g37410	Cyt P450 monooxygenase-like	0.000001	0.000000	0.593
256874_at	At3g26320	Cyt P450, putative	0.000182	0.000001	1.527
251196_at	At3g62950	Glutaredoxin-like	0.000000	0.000000	1.881
266299_at	At2g29450	Glutathione S-transferase	0.000201	0.000002	0.657
262518_at	At1g17170	Glutathione transferase, putative	0.000280	0.000002	0.661
261037_at	At1g17420	Lipoxygenase	0.000001	0.000000	0.590
264751_at	At1g23020	NADPH oxidase; highly similar to FRO1	0.018800	0.000298	1.373
254098_at	At4g25100	Superoxide dismutase (fragment)	0.000000	0.000000	1.715
<i>R</i> gene-like					
267411_at	At2g34930	LRR; Cf2 like	0.002640	0.000031	0.694
267546_at	At2g32680	LRR; Cf2 like	0.002960	0.000035	0.696
257099s_at	At3g24982	LRR receptor-like	0.000000	0.000000	0.519
259629_at	At1g56510	TIR/NBS/LRR like	0.001750	0.000019	0.688
262383_at	At1g72940	TIR-X	0.010500	0.000148	1.395
262382_at	At1g72920	TIR-X	0.004130	0.000050	1.427
264213_at	At1g65400	TIR-X	0.005570	0.000071	0.706

ID, Affymetrix probe set number; AGI, Arabidopsis gene index; FDR, false discovery rate.

Instead of estimating fold change at each time point, we estimated the overall fold change using the least squares method of ANOVA as the interaction between gene and virus treatment. In this way, we filtered out genes which showed different expression levels across different time points and selected genes for which the changes in the expression levels were induced only by the virus treatment (Details in Experimental Methods section).

#### Analysis of co-regulated genes

To understand the time-dependent expression patterns corresponding to the resistance response to

CMV-Y, we performed a hierarchical clustering of the 444 genes that showed significant differences in expression profile between virus and mock infected plants. The cluster analysis revealed two distinct groups of expression pattern with one showing induction and the other showing overall down regulation (Figure 2). Interestingly, only 106 out of the total 444 differentially expressed genes were induced by virus infections. Whereas, the remaining approximately three fourth were down regulated. Most of the identified genes either remained up- or down- regulated throughout the 18 h period and only a very few (less than 25) altered the direction of expression kinetics within the time schedules.

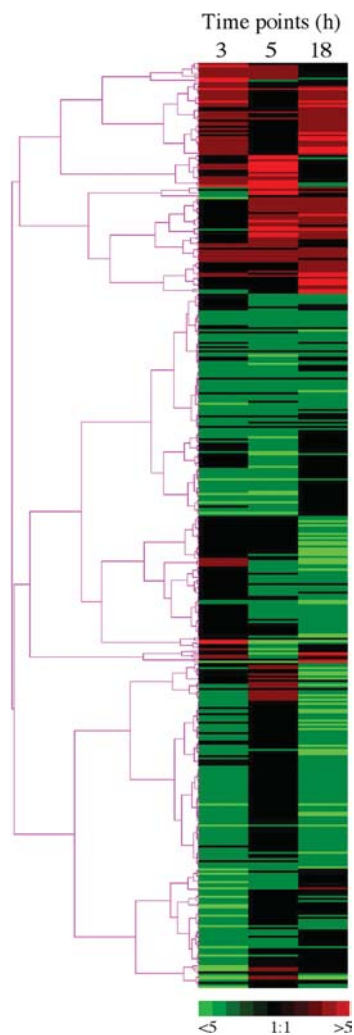


Figure 2. Expression profiles of genes differentially expressed during resistance response to CMV-Y infection in Affymetrix GeneChip of *Arabidopsis* at 3, 5 and 18 h post infection. Red indicates transcriptional activation, green represents repression and black indicates no change in the expression. Scale showing increase or decrease is shown at the bottom.

#### The 'Resistome' is made up of many functional groups

The function for each of the 444 genes was identified based on protein sequence and domain homology and grouped into 10 classes. These include defense related, kinases/phosphatases, protein degradation machinery/proteases, transcriptional regulators, HSPs/chaperones, lipases/hydrolases, transporters, interesting miscellaneous, unknowns and unclassified. Several genes associated with

plant defense were differentially regulated during RCY1-CMV interaction (Table 2). These include defense related genes like  $\beta$ -1-3, glucanase, chitinases, defensin, Hin1-like, Hsr201-like and Lsd1-like; the ROI related genes like Cyt P450, NADPH oxidase, superoxide dismutase, lipoxygenase. We also found differential regulation of genes possessing R protein domains like TIR or NBS or LRR. The seven *R*-like genes (Table 2) were significantly regulated in *RCY1*-mediated resistance response even though *R* genes are generally not known to be regulated at transcriptional level.

Several genes, which encode proteins that are associated with signal transduction pathways, were also differentially regulated (Table 3). These include putative or known kinases/phosphatases such as putative calcium dependent protein kinase (CDPK), receptor kinases, several serine/threonine kinase-like, Pto kinase interactor like and protein phosphatase 2C (PP2C); transcriptional regulators or DNA binding proteins like MYB-like, AP2-like, WRKY, EREBP-like, and ERFs; protein degradation related factors like E2 (UBC17) enzyme, F-box proteins, U-box protein, RING finger proteins, SKP1/ASK1-like; proteases like aspartyl protease, cysteine protease, subtilisin-like protease. Although, only one fourth of the total CMV-Y responsive genes were up regulated, more than 50% of the genes encoding the components of the protein degradation machinery were significantly induced. In comparison, only 4 of the 23 kinases/phosphatases presented in Table 3 showed induction.

Genes like *EDS1* and *PAD4* that encode lipase-like proteins and heat shock protein 90 (Hsp90) are important regulators of *R* gene-mediated plant defense (Falk *et al.*, 1999; Jirage *et al.*, 1999; Hubert *et al.*, 2003; Lu *et al.*, 2003 Liu *et al.*, 2004). In the *RCY1* controlled defense pathway we observed differential regulation of five lipase/hydrolases, 13 HSP/DnaJ domain proteins and eight transporters (Table 4). Except for the gene At5g53550 that encodes an OPT oligopeptide like transporter and At4g13830 that encodes DnaJ-like protein, all others were down regulated. We have also identified an 'Interesting Miscellaneous' class of genes which include putative annexin, cell wall proteins like CER1 and 2, cell death factor Bcl2 interacting protein Nip2, Hsp90 interacting protein/ER chaperone, CCR4-associated factor 1-like and few others (Table 4). These genes do not have precedence as plant defense regulators but based

Table 3. Differentially regulated genes encoding protein kinases and phosphatases, transcriptional regulators, protein degradation machinery/proteases in the RCY1 resistome against CMV-Y.

ID	AGI	Description	FDR	p-value	Est. ratio
Kinases and phosphatases					
262671_at	At1g76040	CDPK, putative	0.000016	0.000000	0.626
261506_at	At1g71697	Choline kinase, putative	0.000282	0.000002	0.661
250556_at	At5g07920	Diacylglycerol kinase	0.015500	0.000238	0.725
254869_at	At4g11890	Kinase-like	0.000000	0.000000	0.539
254255_at	At4g23220	Kinase-like	0.000007	0.000000	0.616
254265s_at	At4g23140	Kinase-like	0.000166	0.000001	0.654
254241_at	At4g23190	Kinase-like	0.000248	0.000002	0.659
254408_at	At4g21390	Kinase-like	0.000442	0.000004	0.667
267550_at	At2g32800	Kinase-like	0.003010	0.000035	0.696
255872_at	At2g30360	Kinase-like	0.003800	0.000046	0.700
255617_at	At4g01330	Kinase-like	0.031700	0.000580	1.352
257978_at	At3g20860	Kinase-like	0.008890	0.000121	0.714
258616_at	At3g02880	Kinase-like	0.049500	0.000987	0.749
256136s_at	At1g48720	Lectin receptor kinase like	0.017900	0.000281	1.375
259426_at	At1g01470	PI-4P-5 kinase, putative	0.000009	0.000000	0.619
260855_at	At1g21920	PI-4P-5 kinase, putative	0.001850	0.000020	1.453
247723_at	At5g59220	PP2C-like ABA induced	0.000000	0.000000	0.540
251017_at	At5g02760	PP2C homolog,	0.000001	0.000000	1.670
258901_at	At3g05640	PP2C, putative	0.000000	0.000000	0.527
251494_at	At3g59350	Pto kinase interactor1 like	0.024700	0.000419	0.734
262360_at	At1g73080	Receptor-like protein kinase	0.038100	0.000710	0.743
249410_at	At5g40380	Receptor-like protein kinase	0.000035	0.000000	0.635
249771_at	At5g24080	Receptor-like protein kinase	0.000943	0.000009	0.678
Transcriptional regulators					
253799_at	At4g28140	AP2-like protein RAP2.4	0.000000	0.000000	0.472
264415_at	At1g43160	AP2-like, putative	0.000000	0.000000	0.558
265418_at	At2g20880	AP2-like, putative	0.000000	0.000000	0.342
259793_at	At1g64380	AP2-like, putative	0.000046	0.000000	0.638
246275_at	At4g36540	bHLH domain	0.048600	0.000968	1.335
267628_at	At2g42280	bHLH protein family	0.000000	0.000000	1.774
266656_at	At2g25900	CCCH-type Zn finger like	0.028600	0.000511	1.356
261088_at	At1g07590	DNA-binding protein	0.010800	0.000153	0.718
248448_at	At5g51190	EREBP, putative	0.000001	0.000000	1.684
247543_at	At5g61600	EREBP-like	0.000019	0.000000	1.593
248794_at	At5g47220	ERF2	0.019800	0.000323	1.370
248799_at	At5g47230	ERF5	0.000009	0.000000	1.615
245250_at	At4g17490	ERF6	0.000033	0.000000	1.577
258133_at	At3g24500	Ethylene-responsive transcriptional coactivator	0.000000	0.000000	0.523
266841_at	At2g26150	Heat shock factor, putative	0.000000	0.000000	0.539
246851_at	At5g26870	MADS-box protein	0.005950	0.000077	0.707
255250_at	At4g05100	MYB-like protein	0.000018	0.000000	0.627
247455_at	At5g62470	MYB96	0.000091	0.000001	0.646
253219_at	At4g34990	MYB-like protein	0.012000	0.000173	0.720
259705_at	At1g77450	NAC domain family	0.002340	0.000027	0.692
261564_at	At1g01720	NAC domain like protein	0.013300	0.000196	0.722
251899_at	At3g54400	Nucleoid DNA-binding-like	0.040500	0.000776	0.745
260037_at	At1g68840	RAV2-like protein	0.044900	0.000885	1.338
261648_at	At1g27730	Salt-tolerance Zn finger protein	0.000323	0.000003	1.509
253405_at	At4g32800	TINY homolog	0.000667	0.000006	0.673
255302_at	At4g04830	Transcriptional regulator-like	0.000012	0.000000	1.608
248611_at	At5g49520	WRKY48	0.000007	0.000000	0.616
265573_at	At2g28200	ZnF_C2H2, putative	0.042400	0.000829	1.340

Table 3. Continued.

ID	AGI	Description	FDR	<i>p</i> -value	Est. ratio
Protein degradation and proteases					
246195_at	At4g36410	E2, UBC17	0.007830	0.000104	1.405
257943_at	At3g21840	SKP1/ASK1-like	0.020100	0.000329	1.370
267117_at	At2g32570	F-box protein	0.000000	0.000000	0.506
256914_at	At3g23880	F-box protein family	0.024100	0.000407	1.363
260287_at	At1g80440	F-box; 2 Kelch repeats	0.016300	0.000252	1.378
262656_at	At1g14200	RING finger protein family	0.002400	0.000028	0.693
257381_at	At2g37950	RING finger	0.000000	0.000000	1.716
259982_at	At1g76410	RING zinc finger like	0.012300	0.000179	1.389
264217_at	At1g60190	RING/U-box-Armadillo	0.000000	0.000000	0.582
253806_at	At4g28270	RING-TM protein, putative	0.000116	0.000001	1.540
255381_at	At4g03510	RMA; RING-TM protein	0.015200	0.000230	1.381
246643s_at	At5g34990	Ulp1 protease family, DUB	0.000171	0.000001	0.654
263302_at	At2g15190	Ulp1 protease family; DUB	0.000000	0.000000	3.222
261253_at	At1g05840	Aspartyl protease	0.025700	0.000438	1.361
261653_at	At1g01900	Subtilisin-like serine protease	0.000265	0.000002	0.660
248918_at	At5g45890	Cysteine protease SAG12	0.000002	0.000000	0.603

ID, Affymetrix probe set number; AGI, *Arabidopsis* gene index; FDR, false discovery rate.

on the function of their homologues in animal or other systems, we suspect they either play direct roles in *RCY1*-mediated defense or if not, simply respond to CMV-Y infection.

We found that almost one fourth of the 444 genes that modulate *Arabidopsis* response to CMV-Y infection have no known functions (Table 5). Curiously 22 of these genes encode very small proteins (less than 100 amino acids) and some as small as 26 or 36 amino acids (Table 5). We used SMART program (<http://smart.embl-heidelberg.de/>) to determine if these predicted proteins contain any known domains that are present in other proteins. Interestingly, some of these predicted proteins contain PDZ domain, coil-coiled (CC) domain, ankyryn repeats, Kelch repeats, MATH and TRAF domain, BTB domain, A20-like zinc finger domain and others. Presence of these domains indicates that they may have a role in plant defense.

#### *Transcription factor binding sites in the RCY1 inducible gene cluster*

The 1.1 kb of sequence upstream from the known or predicted translational start site of all 444 differentially regulated genes in this study were analyzed for the frequency of occurrence of 12 known cis-acting elements (listed in Figure 3; Chen, *et al.*,

2002). We used *p*-value that was calculated value based on the enrichment of binding motifs to determine the statistical significance. In some cases, the enrichment of binding motifs relative to the expected frequency is only two fold, but the *p*-values are extremely low. Even two-fold change could result in a highly statistically significant result because statistical significance is a function of both the magnitude of difference as well as the sample size. When the sample size is very large, a small difference between two samples will result in highly significant results.

Comparison of the promoters of the 444 genes to those of the entire genome revealed that nine out of these 12 cis-elements analyzed were significantly more abundant in the promoters of genes of the resistome. AP2/EREBP (GCCGCC), RAV1/RAV2, and Myb2 binding elements were represented in the 444 CMV-Y inducible genes to a similar extent as with the whole genome. Genes belonging to lipase/hydrolase class showed only 5 out of 12 elements in their promoters. Three out of these five elements occurred approximately at same frequency as the average of entire genome. We found that except for the promoter regions of genes belonging to *R* gene like and lipase/hydrolases, all others had over two-fold more frequency of WRKY elements compared to the whole genome. WRKY, Myb4 and bZIP (TGA) binding

Table 4. Expression profile of genes encoding lipases and hydrolases, transporters, heat shock proteins (HSPs), and interesting miscellaneous in the RCY1 resistome against CMV-Y.

ID	AGI	Description	FDR	p-value	Est. ratio
Lipases and Hydrolases					
263809_at	At2g04570	GDSL-motif lipase/hydrolase like	0.000000	0.000000	0.548
248921_at	At5g45950	GDSL-motif lipase/hydrolase-like	0.000000	0.000000	0.561
252363_at	At3g48460	Lipase - like	0.000000	0.000000	0.504
259786_at	At1g29660	Lipase/hydrolase, putative	0.027400	0.000482	0.737
249278_at	At5g41900	Hydrolase, alpha/beta fold family	0.000000	0.000000	0.490
Transporters					
251785_at	At3g55130	ABC transporter, similar to BCRP1	0.004550	0.000055	0.702
267008_at	At2g39350	ABC transporter, putative	0.000000	0.000000	0.562
255889_at	At1g17840	ABC transporter, putative	0.004970	0.000061	0.704
262756_at	At1g16370	Transport protein, putative	0.000016	0.000000	0.625
250261_at	At5g13400	Peptide transporter – like	0.000170	0.000001	0.654
252377_at	At3g47960	Peptide transporter, putative	0.000009	0.000000	0.619
263918_at	At2g36590	Proline transporter, putative	0.000000	0.000000	0.433
248276_at	At5g53550	OPT oligopeptide transporter	0.040800	0.000792	1.342
HSPs					
256999_at	At3g14200	DnaJ domain protein	0.000000	0.000000	0.4528
254688_at	At4g13830	DnaJ-like protein	0.031800	0.000579	1.3519
263374_at	At2g20560	Heat shock protein, putative	0.000010	0.000000	0.6196
258979_at	At3g09440	Hsc70-3	0.000000	0.000000	0.492
256245_at	At3g12580	HSP70	0.000000	0.000000	0.5505
248043s_at	At5g56000	HSP81-4	0.000113	0.000001	0.649
248045_at	At5g56030	HSP81-2	0.000560	0.000005	0.6705
248332_at	At5g52640	HSP83	0.000307	0.000003	0.6621
260248_at	At1g74310	HSP101	0.000002	0.000000	0.6012
253614_at	At4g30350	HSP101, putative	0.040100	0.000760	0.7445
262911s_at	At1g59860	HSP17.5, putative	0.005690	0.000072	0.7063
263150_at	At1g54050	Hsp20/alpha crystallin family	0.000078	0.000001	0.6445
254384_at	At4g21870	Hsp20/alpha crystallin family	0.032800	0.000602	0.7404
Interesting miscellaneous					
246028_at	At5g21170	AKIN beta1	0.038800	0.000731	1.344
266418_at	At2g38750	Annexin, putative	0.000217	0.000002	0.658
253173_at	At4g35110	Arabidopsis phospholipase-like	0.017700	0.000277	0.727
267425_at	At2g34810	Berberine bridge enzyme, putative	0.002300	0.000026	0.692
255795_at	At2g33380	Calcium-binding EF-hand like,	0.000000	0.000000	0.560
260881_at	At1g21550	Calcium-binding protein, putative	0.000018	0.000000	0.627
252136_at	At3g50770	Calmodulin-like protein	0.040000	0.000761	1.343
248607_at	At5g49480	Calmodulin-like	0.000011	0.000000	0.621
252679_at	At3g44260	CCR4-associated factor 1-like	0.000000	0.000000	1.777
254189_at	At4g24000	Cellulose synthase, putative	0.000498	0.000004	0.669
264147_at	At1g02205	CER1 protein	0.000016	0.000000	0.625
264146_at	At1g02205	CER1 protein	0.012000	0.000173	0.720
254122_at	At4g24510	CER2	0.000001	0.000000	0.598
259391s_at	At1g06350	Delta 9 desaturase, putative	0.025900	0.000445	0.735
253163_at	At4g35750	E1B 19K/Bcl2-interacting Nip2 like	0.027500	0.000485	1.358
261266_at	At1g26770	Expansin 10	0.001360	0.000014	0.684
255732_at	At1g25450	Fatty acid condensing enzyme CUT1 like	0.029400	0.000528	0.738
252487_at	At3g46660	Glucosyltransferase-like	0.000037	0.000000	0.636
254667_at	At4g18280	Glycine-rich cell wall protein-like	0.019800	0.000321	0.730

Table 4. Continued.

ID	AGI	Description	FDR	p-value	Est. ratio
259037_at	At3g09350	Hsp90 interacting protein; ER chaperone; ARM repeats	0.000376	0.000003	0.665
248392_at	At5g52050	Integral membrane protein-like	0.042200	0.000822	0.746
254805_at	At4g12480	Lipid transfer protein (LTP) family	0.002120	0.000024	1.448
252711_at	At3g43720	Lipid-transfer protein-like	0.003750	0.000045	0.699
256324_at	At1g66760	MATE efflux family protein, putative	0.000000	0.000000	0.588
264289_at	At1g61890	MATE efflux protein family	0.000000	0.000000	0.571
245399_at	At4g17340	Membrane channel like	0.000028	0.000000	0.632
246340s_at	At3g44860	Methyltransferase-related	0.000000	0.000000	0.560
254120_at	At4g24570	Mitochondrial uncoupling protein, like	0.003140	0.000037	1.435
259579_at	At1g28010	P-glycoprotein, putative	0.000026	0.000000	0.631
247776_at	At5g58700	Phosphoinositide phospholipase - like	0.029500	0.000532	0.738
253063_at	At4g37640	Plasma membrane-type calcium ATPase	0.038900	0.000731	0.744
259660_at	At1g55260	Protease inhibitor/LTP family	0.001050	0.000010	0.679
266720s_at	At2g46790	Pseudo-response regulator, APRR9	0.015500	0.000237	0.725
252464_at	At3g47160	RNA-binding protein-like	0.001590	0.000017	1.458
264787_at	At2g17840	Senescence-associated protein 12, like	0.000034	0.000000	0.635
262226_at	At1g53885	Senescence-associated protein -related	0.020000	0.000325	0.730
255543_at	At4g01870	Similar to bacterial tolB unclear if involved in viral transport	0.000202	0.000002	0.657
255649_at	At4g00920	Similarity to mouse microtubule-associated protein 2	0.007600	0.000101	0.711
254839_at	At4g12400	Stress-induced protein st1-like	0.000144	0.000001	0.652
261081_at	At1g07350	Transformer-SR ribonucleoprotein like	0.000029	0.000000	0.633
251221_at	At3g62550	Universal stress protein family	0.011600	0.000166	1.391

ID, Affymetrix probe set number; AGI, Arabidopsis gene index; FDR, false discovery rate.

sites that are in general more represented in the *Arabidopsis* genome are even more abundantly associated with pathogen modulated genes.

#### *Location of Myb4, bZIPs and WRKY binding elements in the RCY1 regulated genes*

Based on the above analysis, binding elements most abundant in pathogenesis related genes are Myb4, three kinds of bZIPs and WRKY. Therefore, we scanned the 1.1 kb region of promoters of all resistance related, *R*-gene like and the unknown classes of genes to locate the exact position and frequency of these five abundant type of binding elements (Figure 4). Not all gene promoters have all the five binding elements in them. Promoters of resistance related genes At1g19610 (putative defensin), At4g39330 (cinnamyl-alcohol dehydrogenase), At5g55180 ( $\beta$ -1,3-glucanase like) and At1g72940 (TIR-X) have all five binding elements. Only two binding sites per gene promoter are present in defense related genes like At1g32540

(LSD1-like), At2g40230 (anthranilate N-hydroxycinnamoyl-like), At4g37990 (cinnamul-alcohol dehydrogenase), At1g56510 (TIR-NBS-LRR) and At2g34930 (Cf2-like) (Figure 4). On the other hand, At1g19610 (defensin) and At5g05250 (unknown) have as many as 11 and 18 binding sites respectively on their promoters (Figure 4). It will be interesting to see how the frequency and types of binding sites on a single promoter determines participation of the protein in different pathways as well as feedback or feed-forward loops that regulate gene expression of a particular gene.

#### *Comparison of resistome of RCY1 with other R gene mediated responses*

Recently, using Affymetrix GeneChip representing 8000 genes, Tao *et al.* (2003) identified 2338 genes that were differentially regulated in *Arabidopsis* in response to the bacterial pathogen *Pseudomonas syringae* in compatible as well as incompatible interactions involving *R* genes

Table 5. Representative genes that encode protein of unknown function in the RCY1 resistome against CMV-Y.

ID	AGI	Description	FDR	p-value	Est. ratio
256442_at	At3g10930		0.00001	0.00000	1.606
260804_at	At1g78410		0.00000	0.00000	1.719
263210_at	At1g10585		0.00412	0.00005	0.701
252597_at	At3g45360		0.00950	0.00013	0.715
256471_at	At1g42580		0.00000	0.00000	0.487
247754_at	At5g59080		0.00780	0.00010	1.405
262049_at	At1g80180		0.00771	0.00010	1.406
253165_at	At4g35320		0.00000	0.00000	0.607
266901_at	At2g34600		0.00004	0.00000	0.637
261247_at	At1g20070		0.00000	0.00000	0.554
256346s_at	At1g54926		0.00000	0.00000	0.525
258830_at	At3g07090		0.00000	0.00000	0.578
250292_at	At5g13220		0.00000	0.00000	0.600
250828_at	At5g05250		0.00001	0.00000	1.604
247474_at	At5g62280		0.00003	0.00000	1.578
260081_at	At1g78170		0.00033	0.00000	0.663
249454_at	At5g39520		0.00048	0.00000	0.668
245724_at	At1g73390		0.00095	0.00001	0.678
256285_at	At3g12510		0.00109	0.00001	0.680
250350_at	At5g12010		0.00524	0.00007	0.705
251725_at	At3g56260		0.00809	0.00011	0.712
265568s_at	At2g05560		0.00936	0.00013	1.399
248302_at	At5g53160		0.01110	0.00016	1.392
256046_at	At1g07135		0.01880	0.00030	1.373
260205_at	At1g70700		0.02120	0.00035	0.731
264238_at	At1g54740		0.02600	0.00045	1.360
265245_at	At2g43060		0.04050	0.00078	1.342
247431_at	At5g62520		0.04050	0.00078	0.745
260420_at	At1g69610		0.04320	0.00085	0.747
266259_at	At2g27830		0.04620	0.00091	1.337
261607_at	At1g49660		0.00000	0.00000	1.640
260357_at	At1g69260		0.01370	0.00020	0.722
247484_at	At5g62110		0.01530	0.00023	1.381
256017_at	At1g19180		0.02720	0.00048	0.736
247646_at	At5g59990		0.03610	0.00067	0.742
253874_at	At4g27450		0.01950	0.00031	1.371
265409_at	At2g16830	26 aa	0.00083	0.00001	1.480
263941_at	At2g35870	36 aa	0.00000	0.00000	0.596
250956_at	At5g03210	41 aa	0.02460	0.00042	0.734
256891_at	At3g19030	60 aa	0.02850	0.00051	1.356
260005_at	At1g67920	67 aa	0.00173	0.00002	0.687
256571_at	At3g30730	67 aa	0.01050	0.00015	1.395
261995_at	At1g33850	70 aa	0.00000	0.00000	0.087
265670s_at	At2g32210	71 aa	0.00551	0.00007	0.706
265674_at	At2g32190	71 aa	0.00084	0.00001	0.676
248904_at	At5g46295	71 aa TM	0.00000	0.00000	2.161
264580_at	At1g05340	72 aa	0.00000	0.00000	0.612
265024_at	At1g24600	72 aa	0.00139	0.00001	0.684
263021_at	At1g23910	73 aa	0.00000	0.00000	1.840
249376_at	At5g40645	73 aa	0.00134	0.00001	0.683
247109_at	At5g65870	77 aa SP	0.02170	0.00036	0.732
245771_at	At1g30250	82 aa	0.01060	0.00015	0.717
267261_at	At2g23120	83 aa	0.00558	0.00007	0.706
267591_at	At2g39705	87 aa	0.00001	0.00000	1.611
264237_at	At1g54700	89 aa; PDZ domain	0.00002	0.00000	1.598
259996_at	At1g67910	91 aa	0.01370	0.00020	1.385
245973_at	At5g32490	94 aa	0.03180	0.00058	0.740

Table 5. Continued.

ID	AGI	Description	FDR	p-value	Est. ratio
252793_at	At3g42250	100 aa	0.00255	0.00003	0.693
248218_at	At5g53710	SP	0.00000	0.00000	0.575
248617_at	At5g49590	SP	0.00000	0.00000	0.581
248298_at	At5g53110	SP	0.01900	0.00030	0.729
263647_at	At2g04690	SP	0.01900	0.00030	1.372
263301s_at	At2g15200	SP	0.03970	0.00075	1.344
260167_at	At1g71970	SP	0.02830	0.00050	1.357
260686_at	At1g17620	TM	0.00001	0.00000	0.617
255662_at	At4g00410	TM	0.02640	0.00046	0.736
247933_at	At5g56980	TM	0.00000	0.00000	0.585
266316_at	At2g27080	TM	0.01960	0.00032	0.729
266596_at	At2g46150	TM	0.00004	0.00000	0.638
267230_at	At2g44080	2TM	0.00035	0.00000	1.506
266071_at	At2g18680	5TM	0.00683	0.00009	0.709
264389_at	At1g11960	SP-2TM	0.00001	0.00000	0.623
252866_at	At4g39840	SP-5TM	0.00193	0.00002	0.689
258796_at	At3g04630	CC domain	0.00083	0.00001	0.676
255723_at	At3g29575	CC domain	0.00151	0.00002	0.685
255774_at	At1g18620	CC domain	0.01100	0.00016	1.393
248444_at	At5g51320	DUF321- <i>Arabidopsis</i> specific	0.00083	0.00001	1.480
245629_at	At1g56580	DUF538-plant specific	0.00000	0.00000	0.591
261456_at	At1g21050	DUF617-plant specific	0.00579	0.00007	1.415
265203_at	At2g36630	SP; DUF81, integral membrane family	0.00484	0.00006	0.703
245655_at	At1g56530	Proline-rich family	0.02010	0.00033	0.730
248123_at	At5g54720	2 ANK repeats	0.03110	0.00056	0.739
248713_at	At5g48180	5-Kelch repeats	0.00708	0.00009	0.710
265746_at	At2g06630	A20-like zinc fingers	0.00000	0.00000	0.610
258487_at	At3g02550	lateral organ boundaries LOB domain	0.00004	0.00000	0.636
249872_at	At5g23130	LysM bacterial cell wall degradation	0.01550	0.00024	0.725
263133_at	At1g78450	SP; SOUL heme-binding domain	0.03930	0.00074	1.344
254385s_at	At4g21830	SP-SelR domain to cope with oxidative stress	0.00586	0.00008	0.707
263126_at	At1g78460	SP-SOUL heme-binding domain	0.03890	0.00073	1.345
248466_at	At5g50720	TB2/DP1, HVA22 family	0.00146	0.00002	0.685
245431_at	At4g17080	TM; 5 MORN domains	0.01990	0.00032	1.370
255692_at	At4g00400	TM; PlsC = Phosphate acyltransferases	0.01930	0.00031	0.729
259430_at	At1g01610	3-TM; PlsC = Phosphate acyltransferases	0.00214	0.00002	0.691
257860_at	At3g13062	TM; START domain	0.00038	0.00000	1.503
256021_at	At1g58270	TM; 2MATH domain = meprin and TRAF homology	0.00002	0.00000	0.625
259232_at	At3g11420	TM-DUF604	0.01480	0.00022	0.724
263421_at	At2g17230	TM-Phosphate-induced domain	0.05080	0.00102	0.750
250968_at	At5g02890	transferase family	0.00146	0.00002	0.685
266462_at	At2g47770	TspO/MBR family	0.00568	0.00007	0.706
246097_at	At5g20270	UPF0073-integral membrane	0.00000	0.00000	0.565
260880_at	At1g21380	VHS Domain = in VPS-27; GAT domain	0.03340	0.00062	0.741
262164_at	At1g78070	WD40 repeat	0.00000	0.00000	0.481
264102_at	At1g79270	YTH domain YT521-B-like family splicing	0.00083	0.00001	0.676
252367_at	At3g48360	BTB domain; MEL-26 like	0.00053	0.00000	1.493
257964_at	At3g19850	NPH3 family domain BTB domain	0.00009	0.00000	1.547

ID, Affymetrix probe set number; AGI, *Arabidopsis* gene index; FDR, false discovery rate TM, transmembrane domain; SP, signal peptide; CC, coil-coiled domain.

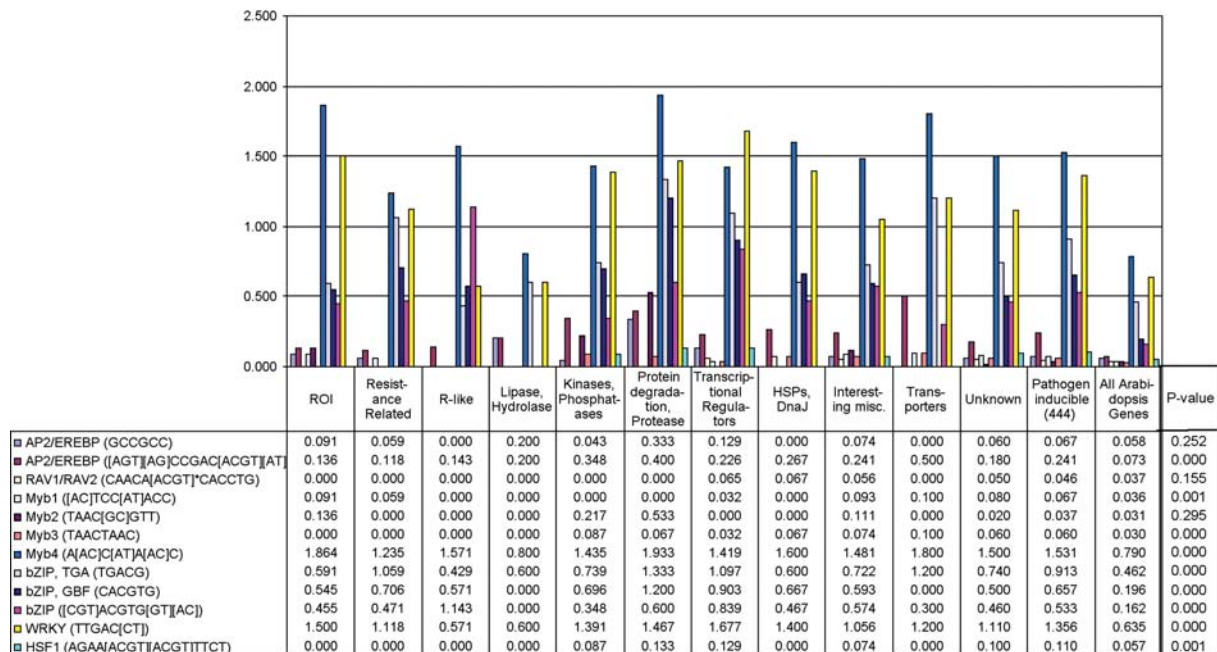


Figure 3. Histograms and data tables showing frequencies of transcription factor binding sites in 1.1 kb promoter sequence of the 444 differentially regulated defense genes. Upstream sequences were obtained from the MIPS database. Consensus binding sites sequence is from (Chen, *et al.*, 2002). Frequencies are shown on the y-axis and genes belonging to different functional classes are shown on the x-axis. Statistical significance of the over-abundance of these binding sites is shown in the last column.

*RPS2* and *RPM1*. Majority of these genes reside on Chromosomes II and IV mainly because of the probe sets used in this early version of GeneChip. In order to contrast the plant response to bacterial versus viral pathogens, we compared the same subset of Chromosomes II and IV specific genes included from the study of Tao *et al.* (2003) with that of ours. They report that a total of 775 + 697 genes (Chromosomes II and IV, respectively) were involved in *RPS2*- and *RPM1*- dependent resistance response. We found that a total of 150 genes from Chromosome II and IV participate in *RCY1* specific defense response and 80 of these also respond to *RPS2*-, and *RPM1*-mediated defense (some of these genes are given in Table 6).

#### Chromosome distribution of the resistome

We were interested in finding out if the genes making up viral and bacterial *R* gene-mediated resistome existed in defined clusters on a chromosome or were randomly distributed. Therefore, we estimated the frequency of occurrence of the

virus-responsive genes in our study and the bacteria-responsive genes from that of Tao *et al.* (2003) that are present on chromosomes II and IV (Figure 5). We used a moving window of 100,000 base pairs that advances every 30,000 base pairs to depict the frequency of occurrence of resistance responsive genes (Figure 5a). Pattern of distribution of genes seen as peaks in moving windows was very similar for bacteria- and virus-responsive genes (Figure 5b). This suggests a clustering of co-regulated genes that respond to bacterial or viral pathogens on a chromosome.

#### Discussion

In this study, we have generated the expression profile of *Arabidopsis* ecotype C24 that shows *RCY1*-mediated resistance to CMV-Y infection using the Affymetrix GeneChip representing the entire *Arabidopsis* genome. To ensure that plant-to-plant differences were well represented in our tissue sampling, we used two biological replications for each time point with each replication

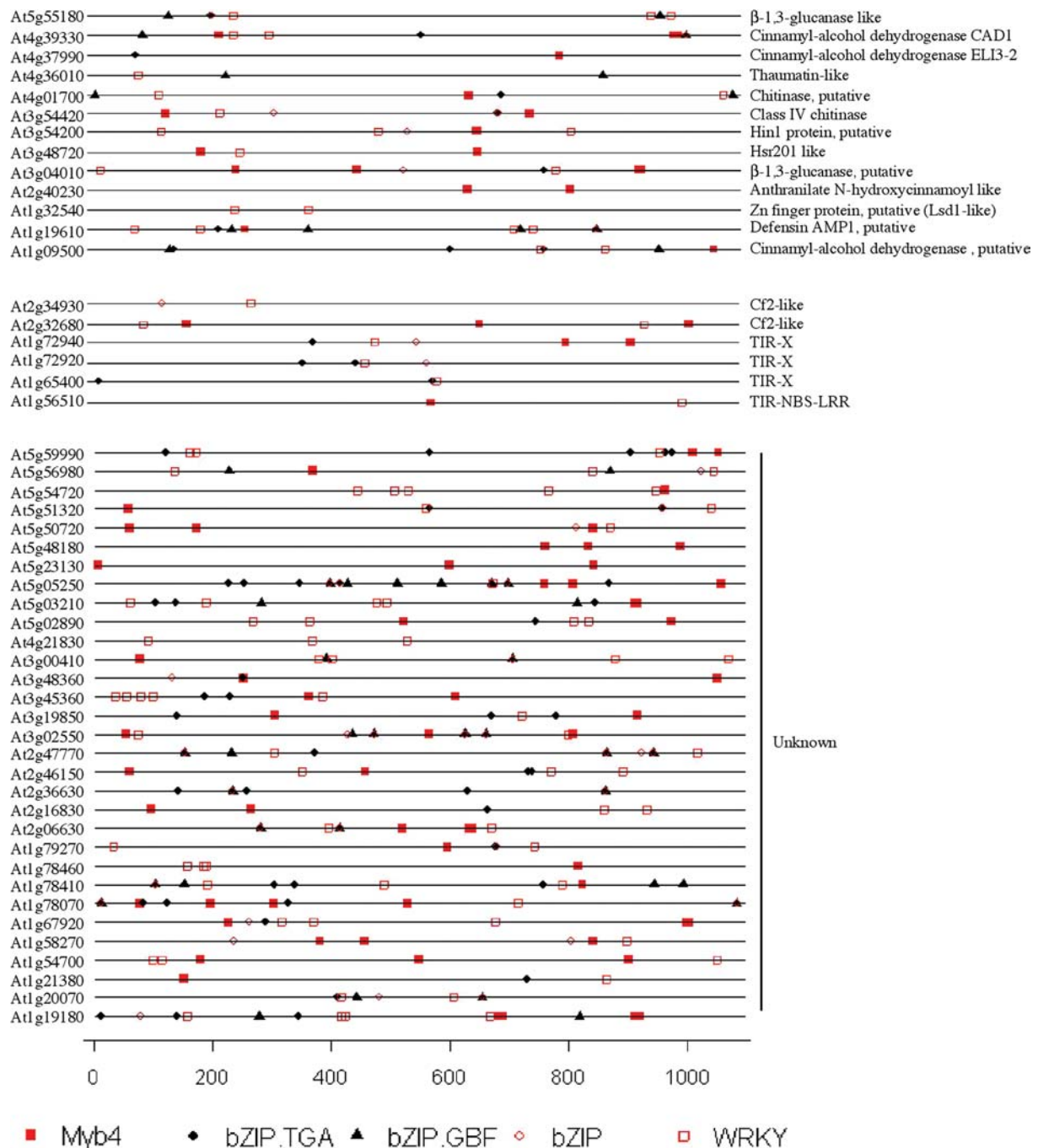


Figure 4. Cartoon diagram showing the location of binding sites Myb4, bZIP (TGA), bZIP (GBF), bZIP and WRKY within the 1.1 kb promoters of defense related, R gene like, and unknown genes. Consensus binding site sequence is as shown in Figure 3.

consisting of tissue pooled from five independently infected plants. Replication clusters for both mock and virus treated data set (after normalization) indicated high reproducibility and consistency in the data. The authenticity of the data was further

confirmed by analyzing real time RT-PCR for representative genes for all the time points.

RCY1 belongs to NBS-LRR class of R proteins which is the largest group of R-proteins conferring resistance to viral, bacterial, and fungal

Table 6. Representative of the common set of 80 genes responding to bacterial and viral pathogens.

ID	AGI	Description
<b>Defense Related</b>		
253277_at	At4g34230	Cinnamyl alcohol dehydrogenase, putative
252943_at	At4g39330	Cinnamyl-alcohol dehydrogenase CAD1
252984_at	At4g37990	Cinnamyl-alcohol dehydrogenase ELI3-2
266299_at	At2g29450	Glutathione S-transferase
253104_at	At4g36010	Thaumatococin-like
267546_at	At2g32680	Cf2 like
267411_at	At2g34930	Cf2 like
<b>Kinases</b>		
254869_at	At4g11890	Kinase-like
254408_at	At4g21390	Kinase-like
254255_at	At4g23220	Kinase-like
255872_at	At2g30360	Kinase-like
267550_at	At2g32800	Kinase-like
255617_at	At4g01330	Kinase-like
<b>Protein degradation</b>		
267117_at	At2g32570	F-box protein
253806_at	At4g28270	RING-TM protein, putative
255381_at	At4g03510	RMA; RING-TM protein
<b>Transcriptional regulators</b>		
265418_at	At2g20880	AP2-like, putative
245078_at	At2g23340	AP2-like, putative
267628_at	At2g42280	bHLH protein family
246275_at	At4g36540	bHLH protein family
266656_at	At2g25900	CCCH-type Zn finger protein, putative
245250_at	At4g17490	ERF6
253219_at	At4g34990	MYB-like protein
255302_at	At4g04830	Transcriptional regulator, putative
265573_at	At2g28200	ZnF_C2H2, putative
<b>Unknown</b>		
253165_at	At4g35320	Unknown
265670s_at	At2g32210	Unknown 71 aa
266316_at	At2g27080	Unknown TM
267230_at	At2g44080	Unknown 2TM
266071_at	At2g18680	Unknown 5TM
263647_at	At2g04690	Unknown SP
252866_at	At4g39840	Unknown SP-5TM
265203_at	At2g36630	Unknown SP; DUF81, integral membrane family
255692_at	At4g00400	Unknown TM; PlsC = Phosphate acyltransferases
263421_at	At2g17230	Unknown TM-Phosphate-induced protein 1 domain
266462_at	At2g47770	Unknown TspO/MBR family
254385s_at	At4g21830	Unknown SP-SelR domain to cope with oxidative stress
<b>Others</b>		
267008_at	At2g39350	ABC transporter, putative
266418_at	At2g38750	Annexin, putative
267425_at	At2g34810	Berberine bridge enzyme, putative
255795_at	At2g33380	Calcium-binding EF-hand protein, putative
254122_at	At4g24510	CER2
253163_at	At4g35750	E1B 19K/Bcl-2-interacting protein Nip2, putative
263809_at	At2g04570	GDSL-motif lipase/hydrolase
253778_at	At4g28480	Heat-shock protein
254688_at	At4g13830	DnaJ-like protein
266415_at	At2g38530	Nonspecific lipid-transfer protein, putative
254805_at	At4g12480	Lipid transfer protein (LTP) family

Table 6. Continued.

ID	AGI	Description
266977_at	At2g39420	Phospholipase, putative
264787_at	At2g17840	Senescence-associated protein 12, putative
254839_at	At4g12400	Stress-induced protein sti1 -like

ID, Affymetrix probe set number; AGI, Arabidopsis gene index; Viral pathogen induced gene list derived from this work; Bacterial pathogen induced gene list derived from Tao et al. (2003).

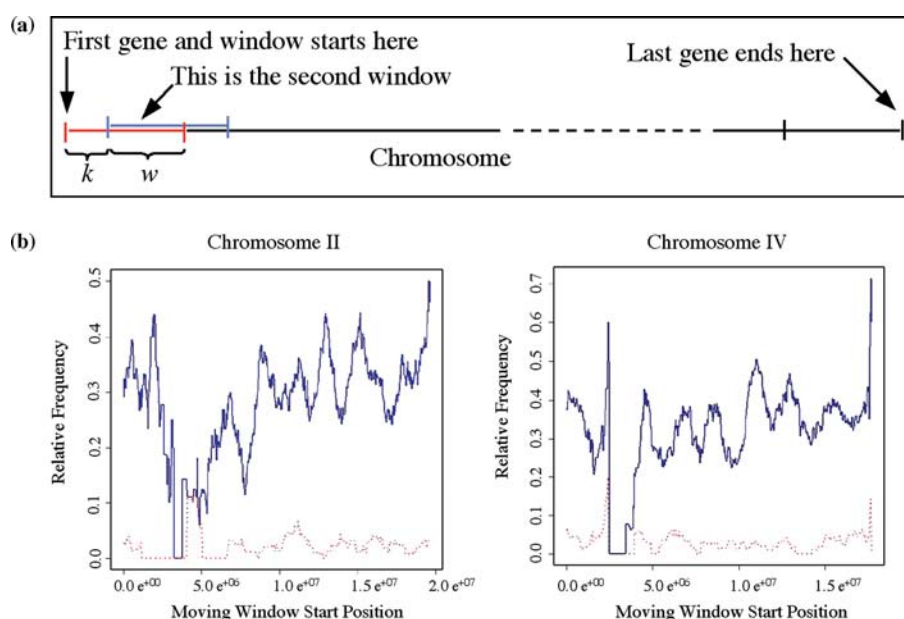


Figure 5. Distribution of differentially regulated genes in *Arabidopsis* interactions with CMV-Y and those with *Ps. syringae*. (a) Cartoon diagram explaining the concept of “moving windows” along the *Arabidopsis* chromosome used to calculate frequency of resistosome forming genes. (b) For each of the *Arabidopsis* chromosomes, the frequency of CMV-Y-responsive (red line) or bacteria-responsive (blue line) genes is shown in a 100,000 bp window that moves continuously through 30,000 bp along the length of chromosome. The  $y$ -axis represents frequency and  $x$ -axis the position of window along the chromosome length (see Material and Methods for more details).

pathogens. Only four viral *R* genes have been cloned from plants so far, and they all encode NBS-LRR proteins. Thus, many of the 444 genes recovered in this study may have relevance to general HR resistance response in multiple plant-pathogen systems. In fact, within the subset of genes on chromosomes II and IV, we identified 80 genes that participate in the defense signaling governed by viral (RCY1) and bacterial (RPM1 and RPS2) *R* proteins. Extrapolating this observation, if the entire genome is considered, the number of genes shared by both systems would most likely be higher. Interestingly, the number of

genes responding to viral or bacterial invasion is similarly distributed in a given window moving over the chromosome. This suggests that there are co-regulated gene clusters on a chromosome that responds to pathogen attack in general.

The clustering of the 444 differentially regulated genes in this study resulted in two prominent groups of genes, those that were consistently either induced or down regulated across all the three time points constituting the 18 h time window. This in large part may be attributed to the high stringency we used for selection. The stringent criterion results in tightly controlled false positives giving

fewer but more reliable genes in the final analysis. However, when we consider the two-fold or more change in expression levels between virus and mock infection, many more genes (~1,600, see supplement Table 3) were found regulated during CMV infection. The real-time RT-PCR analysis on 25 of these genes that are excluded from the 444-gene resistome obtained from the high stringency analysis show that these genes in fact change their expression kinetics across the three time points (see supplement Table 1). The drawback of the high stringent selection therefore is higher rate of false negatives that may result in exclusion of some genes that may truly influence RCY1 mediated response.

We divided the 444-gene resistome into functional classes based on the domain homology in order to gain insight into the type of factors that may be regulating different mechanisms that eventually culminate into HR cell death and ensuing defense response. RCY1 invokes HR cell death to contain CMV-Y to the inoculated leaf of *Arabidopsis*. Reactive oxygen species (ROI) are important signaling molecules that regulate the onset of HR cell death (Levine *et al.*, 1994). The differential regulation of NADPH oxidase, peroxidases, superoxide dismutase, Cytochrome P450-like, and other similar proteins show that ROI signaling constitutes an important part of the RCY1-mediated resistance response and shares many common factors with other resistance pathways.

Several lines of direct and indirect evidence indicate that HR cell death shares many similarities with PCD or apoptosis (Lam *et al.*, 2001). HR appears as distinct patches of necrotic tissue in the infected tissue, which indicates tight regulation of genes involved both in the induction as well as inhibition of PCD. We find differential regulation of several genes involved in cell death as well as its inhibition in leaves undergoing HR. We see differential regulation of LSD1-like, Hsr201-like, A20- (an inhibitor of cell death)-like zinc finger protein, annexin, E1B19K/BCL2 interacting protein Nip2, Hsp90 interacting ARM repeat containing protein/ER chaperone, all of which are associated with apoptosis or programmed cell death. LSD1 is a negative regulator of cell death in plants (Dietrich *et al.*, 1997). Nip2 is a pro-apoptotic factor in animals that interacts with anti apoptosis factor BCL2 as well as the adenovirus E1B 19K protein (Boyd *et al.*, 1994). Nip2 has

been shown to induce apoptosis via caspase dependent mitochondrial activation. Although there are no caspases present in the sequenced plant genomes of *Arabidopsis* and rice, related metacaspases are present (Uren *et al.*, 2000). It will be interesting to see what role Nip2 like gene plays in plant hypersensitive response. Bax inhibitor proteins have been shown to interfere with HR mediated cell death (Huckelhoven *et al.*, 2003). Hsr201 are hypersensitive related proteins that are induced during bacterial infections of plant tissues (Czernic *et al.*, 1996). Hsp90 a highly conserved ATP dependent molecular chaperone has recently been shown to play a critical role for R gene mediated resistance (Hubert *et al.*, 2003; Lu *et al.*, 2003; Liu *et al.*, 2004). Thus differential regulation of the above genes in an RCY1-dependent manner indicates their importance in R gene mediated defense.

HR-induced cell death resulting from the R-Avr recognition needs to be controlled by transient signaling to prevent the death of the host plant. To achieve this, the ubiquitin-proteasome-dependent degradation machinery of the cell may selectively target the destruction of the receptor-ligand complex and other down-stream regulatory proteins. Ubiquitination of a protein requires the sequential activity of the three classes of enzymes: E1, E2 and E3 (Deshaies, 1999). In the SCF (Skp1-Cullin-F box) type E3 ubiquitin ligase complex, the F-box protein functions as a receptor for the substrate protein and hence determines the specificity of ubiquitination. Recently, SGT1 has been shown to associate with Skp1 of the SCF-type E3 ubiquitin ligase and with Rar1, a key protein in the *Mlo*-mediated resistance to barley powdery mildew and *N* gene-mediated resistance to TMV (Azevedo *et al.*, 2002; Liu *et al.*, 2002; Peart *et al.*, 2002). SGT1 may be regulating the targeting of different key proteins involved in disease resistance to the SCF complex, for degradation. We observed variation in the expression levels of E2 enzymes, U-box proteins and several other RING finger and F-box containing proteins (Table 3) in our study. These F-box and other proteins may mediate the degradation of components of the CMV-Y resistance pathway.

Proteases are important effectors of stress and pathogen response. Golldack *et al.* (2003) reported increased expression of subtilisin-like serine proteases in plants treated with jasmonic acid and

cadmium. A papain-like cysteine endoprotease encoded by *Rcr3*, is essential for the function of tomato *Cf-2* gene that confers resistance to *Cladosporium fulvum* (Kruger *et al.*, 2002). Several genes that encode putative proteases like subtilisin, cysteine proteinase, aspartyl protease, etc have been identified in our study. The precise role of these proteases in the *RCY1*-CMV pathway needs to be investigated. It is likely that the target proteins of these proteases play essential roles in defense.

Protein kinases and phosphatases are vital in disease signaling (Zhang and Klessig, 2001) as they control the phosphorylation and dephosphorylation of proteins involved in the perception of a signal and its subsequent transduction. We identified a number of gene encoding receptor protein kinases, CDPKs, kinase-like, Pto-kinase interactor-like proteins and phosphatases differentially regulated in the *RCY1*-mediated resistance to CMV-Y. It is interesting that Pto kinase interactor like proteins also participate in *RCY1*-mediated resistance which show convergence in signaling pathways down stream of pathogen recognition in viral and bacterial resistance signaling.

Kinases and phosphatases control signaling mechanisms by activating or inactivating the protein functions, and the proteasome machinery by regulating their abundance. Signal transduction pathways are also regulated at transcriptional level. Several classes of transcription factors that showed variation in their expression levels in *RCY1*-dependent manner include proteins containing WRKY, AP2-like domain, EREBP-like domains, ERFs, MADS box proteins and others. In addition to these transcription factor involvements, we also performed a systematic analysis of 1.1 kb upstream sequence of all the 444 gene-coding sequences. This analysis revealed 9 out of 12 known DNA binding domains are significantly represented in the defense related cluster of 444 genes and the promoters are in fact super-enriched for six of them. We also made a limited attempt to localize these elements within the promoter regions of selected disease resistance related or R-protein like genes. At this point we cannot make any specific conclusions but we feel that a genome wide location search combined with more robust pattern searching algorithms will yield better results providing insights into transcriptional regulation of defense response genes.

Our study analyzes early response of C24 following infection by CMV-Y. The time points for tissue sampling were carefully selected to match gene expression changes associated with the HR response that is mediated by *RCY1* of C24. We realize that not all genes differentially regulated in our study are exclusively *RCY1*-dependent and many in fact may be responding to the presence of virus. These *RCY1*-independent CMV responding genes that show significant regulation are also potentially important. For instance, the “interesting miscellaneous” category consists of genes that have not been shown to have a role in *R* gene-mediated defense previously (for example RNA-binding proteins) or their role in HR cell death does not have precedence in plant systems (Annexin). RNA binding proteins may be involved in RNAi or post-transcriptional gene silencing (PTGS) which in plants is recognized to be an adaptive antiviral defense system. CMV encoded 2b protein is a known suppressor of gene silencing and is essential for the virus to overcome the effects of SA-mediated defense (Ji and Ding, 2001). Thus, RNA silencing based antiviral defense and *R* gene-dependent immunity are likely to have common elements. Annexin has a role in engulfment of apoptotic cells (Arur *et al.*, 2003). There is no precedence in plants for active elimination of dead cells through their engulfment. It will be interesting to identify precise role of Annexin-like protein in *RCY1*-mediated resistance to CMV-Y.

In addition to known or putative protein factors, we have uncovered many genes in this study with unknown function. A closer examination of unknown class gene sequences revealed that many of them possess known domains. These include at least 14 proteins with transmembrane domains or nine proteins with signal peptide. The unknown gene At2g06630 has an A20-like zinc finger domain which may have a role in the restriction of HR cell death. In mammalian systems, A20 possess functionally two distinct domains, at N-terminus it can bind TRAF domains of TRAF1-TRAF2 and the C-terminal domain can mediate inhibition of NF-kappaB activation and cell death (Song *et al.*, 1996). At1g58270 gene classified as unknown, possess TRAF homology domain. The protein encoded by At1g21380 contains VHS and GAT domains. VHS domain marks the proteins that are involved in endocytosis and vesicular traf-

ficking and some of which may be involved in stage-specific assembly of the endocytic machinery (Lohi *et al.*, 2002). GAT domain is found on GGA proteins (that also have VHS domain) and may have a role in mediating protein trafficking (Dell'Angelica *et al.*, 2000). Some of the genes in the unknown class encode proteins with less than 100 amino acids. Some of these small proteins possess transmembrane or signal peptide domains. A protein consisting of only 89 amino acids encoded by At1g54700 contains PDZ domain. PDZ is a well conserved domain occurring in bacteria, vertebrate, and plant proteins involved in signaling (Ponting *et al.*, 1997). These proteins are often localized to highly specialized submembranous sites and may function in cellular junction formation, receptor or channel clustering and intracellular signaling events. Possibly there is a host of small proteins not yet studied which perform distinctive functions in plant defense. We view this unknown class as un-chartered territory that needs to be explored and mined.

In conclusion, we used whole genome *Arabidopsis* Affymetrix GeneChip and identified a number of genes that are differentially regulated during CMV-Y resistance. In addition to the known or predicted genes involved in defense, a number of cell death and anti-cell death genes, that take part in PCD in animals, have been discovered in our study. These will be very useful in comparing PCD and HR-mediated cell death. Dissection of *RCY1*-mediated signaling is vital because CMV has the largest host range and infects over 1200 plant species. Many of the CMV host plants include vegetable and other agricultural crops making CMV a very economically important pathogen. The resistome analysis presented here will significantly contribute in the quick elucidation of viral resistance mechanisms.

### Acknowledgements

We thank the members of S.P.D-K lab for critical reading of the manuscript, Ning Sun for providing the promoter sequence data set and the Perl Script, Shinya Tsuda for CMV. We thank Janet Hager, Shrikant Mane, Sheila Westman and other members of the Yale NIH/NIDDK Microarray Biotechnology Center. The NIH Grant RO1

GM62625 to S.P.D-K and NSF-DMS 0241160 to H.Z supported this work.

### References

- Arur, S., Uche, U.E., Rezaul, K., Fong, M., Scranton, V., Cowan, A.E., Mohler, W. and Han, D.K. 2003. Annexin I is an endogenous ligand that mediates apoptotic cell engulfment. *Dev. Cell* 4: 587–598.
- Azevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A., Shirasu, K. and Schulze-Lefert, P. 2002. The RAR1 interactor SGT1, an essential component of *R* gene-triggered disease resistance. *Science* 295: 2073–2076.
- Bendahmane, A., Kanyuka, K. and Baulcombe, D.C. 1999. The *Rx* gene from potato controls separate virus resistance and cell death responses. *Plant Cell* 11: 781–791.
- Boyd, J.M., Malstrom, S., Subramanian, T., Venkatesh, L.K., Schaeper, U., Elangovan, B., D'Sa-Eipper, C. and Chinnadurai, G. 1994. Adenovirus E1B 19 kDa and Bcl-2 proteins interact with a common set of cellular proteins. *Cell* 79: 341–351.
- Chen, W., Provart, N.J., Glazebrook, J., Katagiri, F., Chang, H.S., Eulgem, T., Mauch, F., Luan, S., Zou, G., Whitham, S.A., Budworth, P.R., Tao, Y., Xie, Z., Chen, X., Lam, S., Kreps, J.A., Harper, J.F., Si-Ammour, A., Mauch-Mani, B., Heinlein, M., Kobayashi, K., Hohn, T., Dangl, J.L., Wang, X., Zhu, T. 2002. Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative functions in response to environmental stresses. *Plant Cell* 14: 559–574.
- Cooley, M.B., Pathirana, S., Wu, H.-J., Kachroo, P. and Klessig, D.F. 2000. Members of the *Arabidopsis* HRT/RPP8 family of resistance genes confer resistance to both viral and oomycete pathogens. *Plant Cell* 12: 663–676.
- Czerniec, P., Huang, H.C. and Marco, Y. 1996. Characterization of *hrs201* and *hrs515*, two tobacco genes preferentially expressed during the hypersensitive reaction provoked by phytopathogenic bacteria. *Plant Mol Biol* 31: 255–265.
- Dell'Angelica, E.C., Puertollano, R., Mullins, C., Aguilar, R.C., Vargas, J.D., Hartnell, L.M. and Bonifacino, J.S. 2000. GGAs: a family of ADP ribosylation factor-binding proteins related to adaptors and associated with the Golgi complex. *J. Cell Biol.* 149: 81–94.
- Deshaies, R.J. 1999. SCF and Cullin/Ring H2-based ubiquitin ligases. *Annu. Rev. Cell Dev. Biol.* 15: 435–467.
- Dietrich, R.A., Richberg, M.H., Schmidt, R., Dean, C. and Dangl, J.L. 1997. A novel zinc finger protein is encoded by the *Arabidopsis* *LSD1* gene and functions as a negative regulator of plant cell death. *Cell* 88: 685–694.
- Falk, A., Feys, B.J., Frost, L.N., Jones, J.D.G., Daniels, M.J. and Parker, J.E. 1999. EDS1, an essential component of *R* gene-mediated disease resistance in *Arabidopsis* has homology to eukaryotic lipases. *Proc. Natl. Acad. Sci.* 96: 3292–3297.
- Gollack, D., Vera, P. and Dietz, K.J. 2003. Expression of subtilisin-like serine proteases in *Arabidopsis thaliana* is cell-specific and responds to jasmonic acid and heavy metals with developmental differences. *Plant Physiol.* 118: 64–73.

- Hammond-Kosack, K.E. and Parker, J.E. 2003. Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. *Curr. Op. Biotechnol.* 14: 177–193.
- Hubert, D.A., Tornero, P., Belkhadir, Y., Krishna, P., Takahashi, A., Shirasu, K. and Dangl, J.L. 2003. Cytosolic HSP90 associates with and modulates the *Arabidopsis* RPM1 disease resistance protein. *EMBO J.* 22: 5679–5689.
- Huckelhoven, R., Dechert, C. and Kogel, K.H. 2003. Overexpression of barley BAX inhibitor 1 induces breakdown of mlo-mediated penetration resistance to *Blumeria graminis*. *Proc. Natl. Acad. Sci.* 100: 5555–5560.
- Ji, L.H. and Ding, S.W. 2001. The suppressor of transgene RNA silencing encoded by Cucumber mosaic virus interferes with salicylic acid-mediated virus resistance. *Mol. Plant Microbe Interact.* 14: 715–724.
- Jirage, D., Tootle, T.L., Reuber, L., Frost, L., Feys, B.J., Parker, J.E., Ausubel, F.M. and Galzebrook, J. 1999. *Arabidopsis thaliana* *PAD4* encodes a lipase like gene that is important for salicylic acid signaling. *Proc. Natl. Acad. Sci.* 96: 13583–13588.
- Kerr, M.K., Martin, M. and Churchill, G.A. 2000. Analysis of variance for gene expression microarray data. *J. Comput. Biol.* 7: 819–837.
- Kruger, J., Thomas, C.M., Golstein, C., Dixon, M.S., Smoker, M., Tang, S., Mulder, L. and Jones, J.D. 2002. A tomato cysteine protease required for Cf-2-dependent disease resistance and suppression of autonecrosis. *Science* 296: 744–747.
- Lam, E., Kato, N. and Lawton, M. 2001. Programmed cell death, mitochondria and the plant hypersensitive response. *Nature* 411: 848–853.
- Levine, A., Tenhaken, R., Dixon, R. and Lamb, C. 1994. H<sub>2</sub>O<sub>2</sub> from the oxidative burst orchestrates the plant hypersensitive disease resistance response. *Cell* 79: 583–593.
- Li, C., and Wong, W.H. 2001. Model-based analysis of oligonucleotide arrays: Expression index computation and outlier detection. *Proc. Natl. Acad. Sci.* 98: 31–36.
- Liu, Y., Burch-Smith, T., Schiff, M., Feng, S. and Dinesh-Kumar, S.P. 2004. Molecular chaperone Hsp90 associates with resistance protein N and its signaling proteins SGT1 and Rar1 to modulate an innate immune response in plants. *J. Biol. Chem.* 279: 2101–2108.
- Liu, Y., Schiff, M., Serino, G., Deng, X.-W. and Dinesh-Kumar, S.P. 2002. Role of SCF ubiquitin-ligase and the COP9 signalosome in the *N* gene-mediated resistance response to tobacco mosaic virus. *Plant Cell* 14: 1483–1496.
- Lohi, O., Poussu, A., Mao, Y., Quiocho, F. and Lehto, V.P. 2002. VHS domain—a longshoreman of vesicle lines. *FEBS Lett.* 513: 19–23.
- Lu, R., Malcuit, I., Moffett, P., Ruiz, M.T., Peart, J., Wu, A.J., Rathjen, J.P., Bendahmane, A., Day, L. and Baulcombe, D.C. 2003. High throughput virus-induced gene silencing implicates heat shock protein 90 in plant disease resistance. *EMBO J.* 22: 5690–5699.
- Maleck, K., Levine, A., Eulgem, T., Morgan, A., Schmid, J., Lawton, K.A., Dangl, J.L. and Dietrich, R.A. 2000. The transcriptome of *Arabidopsis thaliana* during systemic acquired resistance. *Nat. Genet.* 26: 403–410.
- Martin, G.B., Bogdanove, A.J. and Sessa, G. 2003. Understanding the functions of plant disease resistance proteins. *Annu. Rev. Plant Biol.* 54: 23–61.
- McDowell, J.M., Dhandaydham, M., Long, T.A., Aarts, M.G., Goff, S., Holub, E.B. and Dangl, J.L. 1998. Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the *RPP8* locus of *Arabidopsis*. *Plant Cell* 10: 1861–1874.
- Peart, J.R., Lu, R., Sadanandom, A., Malcuit, I., Moffett, P., Brice, D.C., Schausser, L., D.A., J., Xiao, S., M.J., C., Dow, M., Jones, J.D., Shirasu, K. and Baulcombe, D.C. 2002. Ubiquitin ligase-associated protein SGT1 is required for host and non-host disease resistance in plants. *Proc. Natl. Acad. Sci.* 99: 10865–10869.
- Ponting, C.P., Phillips, C., Davies, K.E. and Blake, D.J. 1997. PDZ domains: targeting signaling molecules to sub-membranous sites. *Bioessays* 19: 469–479.
- Ryals, J.A., Neuenschwander, U.H., Willits, M.G., Molina, A., Steiner, H.-Y. and Hunt, M.D. 1996. Systemic acquired resistance. *Plant Cell* 8: 1809–1819.
- Schenk, P.M., Kazan, K., Wilson, I., Anderson, J.P., Richmond, T., Somerville, S.C. and Manners, J.M. 2000. Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis. *Proc. Natl. Acad. Sci.* 97: 11655–11660.
- Song, H.Y., Rothe, M. and Goettel, D.V. 1996. The tumor necrosis factor-inducible zinc finger protein A20 interacts with TRAF1/TRAF2 and inhibits NF-kappaB activation. *Proc. Natl. Acad. Sci.* 93: 6721–6725.
- Storey, J.D. and Tibshirani, R. 2003. Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci.* 100: 9440–9445.
- Takahashi, H., Goto, N. and Ehara, Y. 1994. Hypersensitive response in cucumber mosaic-virus infected *Arabidopsis thaliana*. *Plant J.* 6: 369–378.
- Takahashi, H., Miller, J., Nozaki, Y., Sukamto., Takeda, M., Shah, J., Hase, S., Ikegami, M., Ehara, Y. and Dinesh-Kumar, S.P. 2002. *Arabidopsis thaliana* *RPP8/HRT* family resistance gene, conferring resistance to cucumber mosaic virus requires salicylic acid, ethylene and a novel signal transduction mechanism. *Plant J.* 32: 655–667.
- Tao, Y., Xie, Z., Chen, W., Glazebrook, J., Chang, H.S., Han, B., Zhu, T., Zou, G. and Katagiri, F. 2003. Quantitative nature of *Arabidopsis* responses during compatible and incompatible interactions with the bacterial pathogen *Pseudomonas syringae*. *Plant Cell* 15: 317–330.
- Uren, A.G., O'Rourke, K., Aravind, L.A., Pisabarro, M.T., Seshagiri, S., Koonin, E.V. and Dixit, V.M. 2000. Identification of paracaspases and metacaspases: two ancient families of caspase-like proteins, one of which plays a key role in MALT lymphoma. *Mol. Cell* 6: 961–967.
- Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C. and Baker, B. 1994. The product of the tobacco mosaic virus resistance gene *N*: similarity to toll and the interleukin-1 receptor. *Cell* 78: 1101–1115.
- Whitham, S.A., Quan, S., Chang, H.S., Cooper, B., Estes, B., Zhu, T., Wang, X. and Hou, Y.M. 2003. Diverse RNA viruses elicit the expression of common sets of genes in susceptible *Arabidopsis thaliana* plants. *Plant J.* 33: 271–283.
- Zhang, S. and Klessig, D.F. 2001. MAPK cascade in plant defense signaling. *Trends Plant Sci.* 6: 520–527.