Southern Illinois University Carbondale
OpenSIUC

Publications | Department of Computer Science

2007

A Computational Approach to Understand *Arabidopsis thaliana* and Soybean Resistance to *Fusarium solani* (Fsg)

Jiazheng Yuan
Southern Illinois University Carbondale

Mengxia Zhu
Southern Illinois University Carbondale, mzhu@cs.siu.edu

M. Javed Iqball
Southern Illinois University Carbondale

Jack Y. Yang
Harvard University

David A. Lightfoot
Southern Illinois University Carbondale, ga4082@siu.edu

Follow this and additional works at: http://opensiuc.lib.siu.edu/cs_pubs

Published in Yuan, J., Zhu, M., Iqbal, M. J., Yang, J. Y., & Lightfoot, D. A. (2007). A computational approach to understand Arabidopsis thaliana and soybean resistance to Fusarium solani (Fsg). Proceedings of the 7th IEEE International Conference on Bioinformatics and Bioengineering, 2007. BIBE 2007, 585 - 592. doi: 10.1109/BIBE.2007.4375620 ©2007 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE. This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder.

Recommended Citation

Yuan, Jiazheng, Zhu, Mengxia, Iqball, M. J., Yang, Jack Y. and Lightfoot, David A. "A Computational Approach to Understand *Arabidopsis thaliana* and Soybean Resistance to *Fusarium solani* (Fsg)." (Jan 2007).

This Article is brought to you for free and open access by the Department of Computer Science at OpenSIUC. It has been accepted for inclusion in Publications by an authorized administrator of OpenSIUC. For more information, please contact opensiuc@lib.siu.edu.

A Computational Approach to Understand *Arabidopsis thaliana* and Soybean Resistance to *Fusarium solani* (Fsg)

Jiazheng Yuan[1,2], Mengxia Zhu[3], M. Javed Iqbal[1], Jack Y. Yang[4], and David A. Lightfoot[1,2]

1. Department Plant, Soil Sciences, and Agriculture System, Southern Illinois University at Carbondale, Carbondale, IL 62901

2. Department of Plant Biology, Southern Illinois University at Carbondale, IL 62901

3. Department of Computer Science, Southern Illinois University at Carbondale, IL 62901

4. Harvard Medical School, Harvard University, Cambridge, MA 02140

jyuan@siu.edu ga4082@siu.edu

Abstract

In this study, we reported the analysis of *Arabidopsis thaliana* microarray gene expression profile of root tissues after the plant was challenged with fungal pathogen *Fusarium solani* f. sp. glycines (Fsg). Our microarray analysis showed that the infection caused 130 transcript abundances (TAs) to increase by more than 2 fold and 32 out of 130 TAs were increased by more than 3 fold in the root tissues. However, only nineteen ESTs were observed with a decrease in TAs by more than 2 fold and 13 of them went down more than 3 fold due to the pathogen infection. In addition, the number of the up-regulated genes was nearly seven times more than that of down-regulated genes. The coordinate regulation of adjacent genes was detected and the distance distribution of the nearest neighbor genes was less likely to be randomly scattered in genome. The results of this study enabled us to decipher the resistance mechanism to Fsg through an integrated computational approach.

1. Introduction

In plants, the same specificity of resistance genes (R genes) has been identified in distantly related species [1, 2]. Seeking the answer may shed light on the reasons that lead to the longevity of R gene specificities and the dynamics of R gene evolution [3]. *Pto* homologs were positioned to syntenic regions in several species such as tomato, potato, and pepper but the Pto locus is mostly conserved between *Lycopersicon esculentum* and *Lycopersicon pimpinellifolium*, suggesting that the locus emerged prior to the divergence of these species [4, 5]. *Lr10* was isolated from hexaploid wheat and encodes a CC-NBS-LRR protein [6]. It is a single-copy gene on wheat chromosome 1AS. Wheat LR10 and Arabidopsis RPM1 proteins possess a significant sequence similarity showing the orthologous conservation of the RPM1-Lr10 type of resistance genes in these species [6].

Transcriptional changes play a major role in many plant defense processes [7]. Dissection of alterations in transcript abundance has provided unique opportunities to delve into gene function by the comparison of species, tissue, and time specific transcription of thousands of genes simultaneously [8, 9]. The simultaneous measurement of transcript abundances of thousands of genes in parallel can serve as an important tool in functional genomics. With microarray analysis, the transcript abundances of the annotated genes of Arabidopsis (>27,500 in 2006) can be evaluated in parallel using high-density microarrays of sequenced cDNAs (AGI, 2000) or oligomers [10]. Microarray experiments allow us to detect significant variation in mRNA abundance and improve understanding of the molecular basis of the plant defense responses [11]. Global and simultaneous analysis of expression TA profiles can also be used to find out what variations in mRNA abundances are significant for a certain proportion of the cDNAs and treatments investigated [12, 13].

Hurst et al. [14] showed that coordinated regulation of adjacent genes in chromosomal regions occurred during the regulation of gene expression in several plant species. A microarray approach demonstrated co-expression patterns for adjacent genes in the rice genome [15]. The results suggested that the net number of co-expressed gene models increased proportionally with genome size. In rice, the majority of the coordinately expressed adjacent gene model clusters are within 100 kp of each other and about 10% of the genes in the rice genome demonstrates coordinated expression patterns that correspond with defined chromatin domains [15]. In the co-expressed gene model, gene clusters within the whole chromatin domain may be expressed in a

Authorized licensed use limited to: Southern Illinois University Carbondale. Downloaded on March 18,2010 at 11:13:13 EDT from IEEE Xplore. Restrictions apply.

similar manner because of covalently modification of nucelosomes and thereby, chromatin remodeling around one gene expands along the chromosome to encompass several genes until a boundary element is reached [16].

Sudden death syndrome (SDS) of soybean (*Glycine max*) caused by *Fusarium solani* f. sp. glycines (Fsg). Fsg was recently renamed *F. virguliforme* [17]. Iqbal et al. [11] measured changes in transcript abundances (TAs) of 191 known plant defense and biotic/abiotic stress related genes in soybean roots at five time points over a period of 10 days after Fsg inoculation. These genes were chosen from a soybean root cDNA library [18] and they had at least a two-fold TAs increase from day 3 to day 10 after the pathogen inoculated to a resistance genotype, RIL23 that carried resistance conferred by beneficial alleles at six quantitative trait loci (QTL) derived from inbred lines 'Essex' x 'Forrest' [11].

Here we analyzed the changes that occurred in the abundance of transcripts corresponding to 10,560 *Arabidopsis thaliana* expressed sequence tags (ESTs). Assays used reverse labeled slides from AFGC after inoculation of *A. thaliana* with Fsg. The coordinate regulation of adjacent genes was detected. The null hypothesis was that an apparent random distribution of these genes associated with resistance among the chromosomes should be observed. An ortholog analysis was employed to understand the evolutionary and developmental roles of the regulated genes based on investigation of orthologous relationship between soybean and Arabidopsis. The comparison study on the transcriptional activity between soybean and Arabidopsis after fungal pathogen Fsg pathogenesis allows us to examine the overall impact of evolution on both genomes.

2. Materials and methods

Plant materials

Arabidopsis thaliana cv. Columbia plants were germinated from seeds under conditions at a 16 h photoperiod with 23°C day and 18°C night temperatures and 80 % (v/v) relative humidity under light (500µE/M2/sec) in a growth chamber. Plants were grown on either a soiless Cactus-mix at pH6.5 or MS medium [19]. There were 30 plants per treatment arranged in a randomized complete block with three plants per treatment per block. Twenty-one days after planting, synchronously growing plants were selected and roots were collected from both treatment and control.

Inoculation of roots with Fsg spores

The Fsg isolate 'Mont-1' was obtained from Dr. Shiuxian Li at the National Soybean Research Laboratory. A spore suspension of Fsg isolate 'Mont-1' was prepared as described [11]. The spore suspension, at 5 x 10^4 spore ml^{-1} with sterile distilled water, was made by adding Fsg spores from several Fsg culture plates. Fsg was cultured on potato dextrose agar medium (PDA, Difco, Detroit, MI) supplemented with 80 mg ml^{-1} tetracycline and a few drops of Tween 20 and was continuously stirred on a stir flask to keep a uniform suspension. The spore suspension was poured on the growth medium for the infested plants and the same volume of sterile distilled water was added to non-infested plants.

RNA isolation and microarray procedure

RNA was isolated separately from both inoculated and non-inoculated roots and roots of 30 *Arabidopsis thaliana* plants that were bulked and ground to a fine powder in liquid nitrogen and RNA was extracted with a RNeasy Plant Mini Kit (Qiagen, Valencia, CA) according to the manufacturer's instructions. RNA samples were treated with DNase in order to remove any residual DNA in combination with DNase treatment using the RNase-free DNase kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer's instructions. After DNase treatment, RNA was purified on RNeasy mini spin columns (Qiagen, Valencia, CA). The quantity and quality of the RNA recovered was determined by spectrophotometry at 260-280 nm and electrophoresis on a 1.2 % (w/v) agarose, 20 % (v/v) formaldehyde gel. The microarray hybridization and slide scanning were carried out by the facility at AFGC (http://afgc.stanford.edu). Microarrays (16561.xls and 27314.xls at ftp://smd-ftp.stanford.edu/smd/organisms/AT/) were used in the experiment. The mRNA samples corresponding to treatment (infested) and control (non-infested) was labeled during the cDNA synthesis with Cy3- or Cy5-labeled dUTP and with one technical replicate labeled by reversed dye compared to the first hybridization.

Data analysis

Stringent quality control measures were applied to all stages of data analysis. The Microarray data were normalized by local (local background value was subtracted from the intensity value of each spot) and global metrics. The procedures described by Pevsner [20] followed to adjust for artifact differences in

Authorized licensed use limited to: Southern Illinois University Carbondale. Downloaded on March 18,2010 at 11:13:13 EDT from IEEE Xplore. Restrictions apply.

intensity of the two labels. Coefficients of means and variances on the signal intensities in each channel and ratio of signals from two replicates were calculated by our C++ program, which was also applied to handle the missing and extra data values. The average ratio for a signal microarray from two replicates was computed by the equation of $[\text{Ratio}^{1st} + (1/\text{Ratio}^{2nd})]/2$. The Student's t-test was used to determine the statistical significance for genes considered between and within Microarrays slides ($P < 0.05$). Unless stated otherwise, all analyses were performed using SAS PROC UNIVARIATE (SAS version 9.1 SAS Institute, Inc., Cary, NC) to investigate the normality of distance distribution of the nearest neighbor genes. Classified distance for the nearest neighbor gene was standardized as 0-10kb as A, 10-50kb defines as B, 50-100kb as C, 100-500kb as D, 500-1000kb as E, 1000-2500kb as F, 2500-500kb as G, and >5000kb as H. Data were also subjected to natural logarithm transformation in order to reduce experimental error if any of the data points was not spread symmetrically.

Phylogenetic analyses

Functional sequence analysis was performed on the amino acid sequences that were derived from complete cds using Cluster X [21] with the default settings. Nucleotide sequences were aligned with Clustal W [22]. The results of multiple alignments were subjected to phylogenetic analysis using the algorithm of the MEGA package version 3.1 using the Maximum Parsimony analysis with Kimura two parameter distances [23]. The Maximum-Parsimony was assessed by 500 bootstrap replicates. Only nucleotide sequences were employed in phylogenetic analysis due to higher stringency consideration for the phylogenetic tree construction.

3. Results

Analysis of expression profile

Analysis of the microarray data demonstrated a significant variation within and between the slides after local and global normalization following the instruction of Pevsner [20]. Among the transcripts, 4196 were highly correlated between the two reverse labeled slides from the 10,560 EST assayed. The position and label variations between two replications did not significantly alter the topography between slides. A total 130 transcript abundances (TAs) were increased by more than 2 fold and 32 out of 130 TAs were increased by greater than 3 fold in the root tissues that were infected by the fungal pathogen. In contrast, only 19 transcript abundances went down more than 2 fold and 13 of them were suppressed more than 3 fold due to the treatment (data not shown). The total number of up-regulated genes was nearly seven times greater than that of down-regulated.

Chromosomal distribution of up-regulated genes

To exam whether the chromosomal spatial organization affects gene expression, we calculated distance of neighbored genes in the genome between pairs of the nearest neighbor loci, which was an analogous to a pair of co-expressed pair of genes. Taking an advantage of rich information from the Arabidopsis Information Resource (www.arabidopsis.org), the neighborhood relationships between these 130 up-regulated and 19 down-regulated genes and their chromosomal positions have been obtained (Fig.1 A, B). As showed in Table 1, the mean distance of neighbored genes in each chromosome was ranged as 840~958Kb exception of chromosome II (1,865Kb). Approximately 30% adjacent gene pairs with the genome distance within 100 Kb tended to have similar fold of TA increases, and the range was from 2-5.62. Fifty percent of the gene orientations that were in the same vicinity on the chromosome was more likely to be the same (Table 1) Our results indicated that transcriptional relationship of the nearest neighbor gene showed a tendency to occur in the adjacent position along the chromosome but it is still inconclusive by the position of each gene. Six of eighteen pairs of the nearest neighbor genes had the similar function, another six possessed different functions, and at least one of the pairs in the remaining gene pairs the function was unknown.

The distance distribution of the nearest neighbor genes was less likely to be randomly scatted in the genome. From Shapiro-Wilk test for normality, the distance distribution showed not normal on each chromosome ($W=0.63\text{-}0.83$, $P<0.001$). However, the distribution appeared to correspond better to the normal distribution than that of the distance of neighbor gene if the genome distances were classified based on the standardized length of the fragment

Authorized licensed use limited to: Southern Illinois University Carbondale. Downloaded on March 18,2010 at 11:13:13 EDT from IEEE Xplore. Restrictions apply.

Table 1. Relationship of the nearest neighbor genes. The paired neighbor genes were shown in alternated colors. W values (Shapiro-Wilk test) were consisted of two procedures: the nearest neighbor distance and arbitrarily cataloged distance (classified).

Locus (Nearst neighbor <100kb)	Distance (kb)	S-W test; Pr<W (Distance)	S-W test; Pr<W (Classified)	Orietation	Description	Fold change
AT1G07240				forward	UDP-glucoronosyl/UDP-glucosyl transferase	3.06
AT1G07240	0			forward	ATPK7_serine/threonine-protein kinase 7	3.03
AT1G27640				reverse	unkown	2.47
AT1G27760	53.7			forward	interferon-related protein / IFRD protein family	2.54
AT1G62390				reverse	octicosapeptide/Phox/Bem1p (PB1)	2.08
AT1G62570	82			forward	flavin-containing monooxygenase	2.07
AT1G63650				forward	bHLH protein (Atmyc-2)	2.63
AT1G63780	61.9			reverse	IMP4	2.04
AT1G67820				forward	protein phosphatase 2C (PP2C)	4.25
AT1G68060	79.4			reverse	ATMAP70-1	2.09
Mean distance (Chr1)	928	0.63; 0.0001	0.92; 0.04			
AT2G02080				reverse	C2H2-type zinc finger protein	2.32
AT2G02130	18.3			forward	putative protease inhibitor II	2.12
AT2G24570				reverse	WRKY17	3.16
AT2G24580	5.5			reverse	sarcosine oxidase family protein	2.45
AT2G33810				forward	SPL3; transcription factor	2.43
AT2G33830	3.4			reverse	dormancy/auxin protein	2.34
Mean distance (Chr2)	1865	0.73; 0.0018	0.87; 0.09			
AT3G09770				reverse	zinc finger (C3HC4-type RING finger)	2.06
AT3G09780	2.6			reverse	protein kinase family protein	4.8
AT3G45640				forward	mitogen-activated protein kinase 3	3.06
AT3G45710	23.9			forward	proton-dependent oligopeptide transport	2.9
AT3G47860				reverse	apolipoprotein D-related	2.16
AT3G47965	49.6			reverse	unkown	2.03
AT3G48690				reverse	unkown	2.2
AT3G48720	8.1			forward	transferase family protein	4.22
Mean distance (Chr3)	840	0.69; 0.0004	0.92; 0.05			
AT4G27260				forward	indole-3-acetic acid amido synthetase	2
AT4G27450	71.3			reverse	unkown	2.01
AT4G27450				reverse	unkown	2.01
AT4G27700	97.5			reverse	rhodanese-like domain-containing protein	2.03
Mean distance (Chr4)	958	0.83; 0.0036	0.86; 0.02			
AT5G06120				forward	ran-binding protein	2.05
AT5G06330	82.1			reverse	hairpin-responsive protein	5.62
AT5G15530				forward	biotin carboxyl carrier protein 2	2.26
AT5G15530	0			forward	BCCP2 (biotin carboxyl carrier protein 2)	2
AT5G39950				reverse	ATTRX2	2.29
AT5G39970	6.1			forward	HIPL2	2.46
AT5G62280				forward	unkown	2.31
AT5G62530	80.7			reverse	ALDH12A1 (Aldehyde dehydrogenase 12A1)	2.36
Mean distance (Chr5)	905	0.68; 0.0001	0.89; 0.01			

Authorized licensed use limited to: Southern Illinois University Carbondale. Downloaded on March 18,2010 at 11:13:13 EDT from IEEE Xplore. Restrictions apply.

(W=0.86-0.92, P=0.09-0.01) and found no prominent characteristic changes among the chromosomes (data not shown). The distribution of classified distance demonstrated higher level of the randomness throughout each of the five Arabidopsis chromosomes, suggesting that this long-range correlation may prevail within the genome. There was no significant difference among the distance distribution before and after the distance value transformed by natural logarithm (data not shown). Consequently, the null hypothesis was rejected that there was an apparent non-random distribution of these genes associated with the nearest neighbor genes among the chromosomes. Furthermore, assumptions were not altered after the preprocessing step of data natural logarithm transformation.

Phylogenetic analyses homologous genes involved in the resistance

The transcript abundances in the Arabidopsis microarray data shared homology to the genes involved in resistance, signal transduction, plant defense and transport of metabolites et al. Those TAs were functionally clustered into three major groups (clades, Fig. 2) analyzed using CLUSTAL X [21]. Maximum Parsimony (MP) analysis compared the up-regulated Arabidopsis TAs with soybean ESTs from RIL23 that carried resistance conferred by six quantitative trait loci (QTL) [11]. Twenty-eight pairs of the genes from both species were grouped together and 13 pairs possessed at least one bootstrapping value with various Kimura two parameter distances (Table 2). Seven soybean ESTs showed higher function identities with those of the Arabidopsis counterparties, implying that those genes in soybean and Arabidopsis might derived from common ancestors and shared similar function. These genes shared 1 to 1 ratio of the functional relationship in those seven paired: NM_11869, AF386961, NM_113674, NM_111048, NM_126269, P49078, NM_119619 of Arabidopsis and AI495119, AI44393, AI440894, AI494910,

Table 2. Thirteen of twenty-eight pairs taxoned together genes from both Arabidopsis and soybean after MP and pairwise distance analyses using MEGA 3.1 program [23]. The pair grouped genes were shown in alternated colors.

Accession number	Description	Bootstrapping value	Kimura-2 distance
NM_113160	Protease inhibitor	12	1.32
gm_AI438014	Epoxide hydrolase		
NM_125647	Aldehyde dehydrogenase 12A1	1	1.03
gm_AI442455	Endo-1,4-beta-glucanse		
AA712319	Unknown	33	N/C
gm_AI437885	G-box binding factor		
NM_105450	Protein phosphatase 2C (PP2C)	7	N/C
gm_AI442632	Root specific RB7-5A		
NM_104935	Flavin-containing monooxygenase	90	1.87
gm_AI443454	Putative water channel protein		
NM_118763	ATP binding / arginine-tRNA ligase	2	0.88
gm_AI443444	Proline-rich 14Kda protein		
T46355	Transferase family protein	1	1.23
gm_AI440615	Extacellular dermal glycoprotein precursor		
N65599	Unknown	6	N/C
gm_AI443248	Calnexin		
T42268	Unknown	10	N/C
gm_AI437503	Indole-3-acetate		
AY120750	TSA1	1	1.65
gm_AI443953	MAP kinase		
NM_117628	transferase family protein	11	N/C
gm_AI437977	Protein disulfide isomerase		
NM_120694	Ran-binding protein	23	1.44
gm_AI416624	13-Lipoxygenase		
NM_111869.3	Phenylalanine ammonia-lyase	2	0.87
gm_AI494672	Phenylalanine ammonia-lyase class II		

Authorized licensed use limited to: Southern Illinois University Carbondale. Downloaded on March 18,2010 at 11:13:13 EDT from IEEE Xplore. Restrictions apply.

BI345412, AI443448, AI496621 of soybean. However, their bootstrapping values and Kimura two

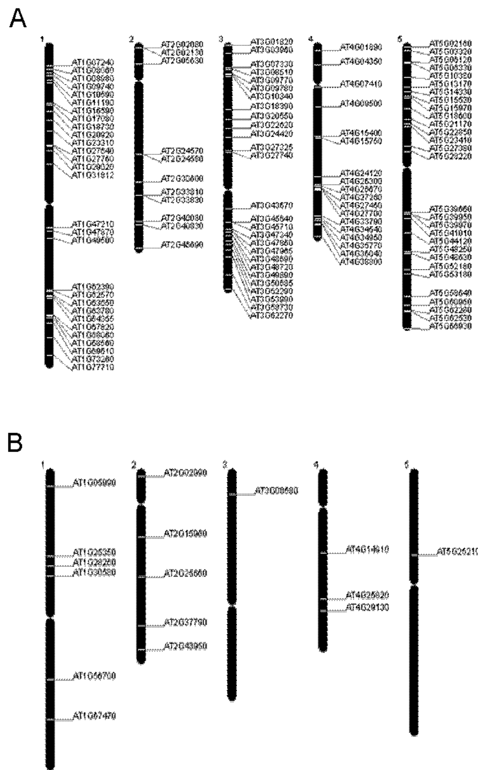

Figure 1. The neighborhood relationships among 130 up-regulated (A) and 19 down-regulated genes (B) and their chromosomal positions in Arabidopsis five chromosomes.

distances did not always agree with the functional similarity. Interestingly, fair numbers of EST sequences were not found in the other organism, even in closely related gene families that were associated with the fungal resistance. Exception of several taxa had high bootstrap values, and the majority families or subfamilies in the higher nodes tended to give low bootstrap values (data not shown).

4. Discussion

Several studies showed an association between functionally related genes with neighboring positions along the chromosomal regions defines as chromatin domains [14]. The chromosomes of eukaryotes display non-homogenous structure with condensed and packed manner while transcription performs spread unevenly over the chromosomes based on each individual organization of the chromosome [24]. Most of the coordinately expressed neighbor gene model clusters are about 100Kb of genomic sequence in rice [15]. These neighbor genes can be organized into a co-expression group. Approximate 10% of the rice genome shows a coordinated expression pattern [15]. However, the mechanism for this co-expression pattern in the genome is still debatable. To decipher the relationships between gene co-expression and spatial position along the chromosomes after pathogen treatment, we examined the distance distribution of the nearest neighbor gene pairs throughout the genome of these 130 up-regulated Arabidopsis genes. Nearly 30% adjacent gene pairs whose genome distance was within 100 Kb tended to have similar fold of TA increases and some of those genes contained a similar function and gene architecture. The distance distribution of the nearest neighbor resistance genes did not appeared randomly scattered in genome on each chromosome. Almost equal number of the nearest neighbor genes had either the same or different orientations. Our results suggested that functional relationships displayed a tendency to occur in neighbor positions along the chromosomes but it was still inconclusive with regard to the position of each gene.

Seven gene groups were identified according to the outcome of functional analysis in this study, showing both functional and orthologous relationships in these two species. The 1 to 1 ratio of the orthologous function relationship was found among those seven paired groups, indicating that these gene groups were descended from a common ancestor and corresponded to well-conserved functions. These 1 to 1 orthologue classes were presumed to represent conserved functions in Arabidopsis and soybean, but they shared diverse bootstrapping values and distances. Whether or not it might have begun to diversify nucleotide sequence in one species as a result of gene duplication is still unknown. Further, only a few of them were the nearest neighbors. The n to n orthologue classes were also found from the phylogenetic tree (data not shown). This information may represent functions that have begun to diversify in both species. Thirteen of twenty-eight pairs of the genes from both species were grouped together and each pair possessed unique bootstrapping value and Kimura's phylogenetic distance (Table 2). Based on the phylogenetic analysis, the organization of soybean phenylalanine ammonia-lyase (PAL) protein was very similar to that of the Arabidopsis PAL protein, suggesting the orthologous relationship of the two PAL proteins. Exception of several taxa appeared with high bootstrap values, majority families or subfamilies in the higher nodes tended to give low bootstrap values.

Authorized licensed use limited to: Southern Illinois University Carbondale. Downloaded on March 18,2010 at 11:13:13 EDT from IEEE Xplore. Restrictions apply.

Gene transcriptional changes play a major role in many plant defense processes [7]. The intensively characterized species such as *Arabidopsis thaliana*

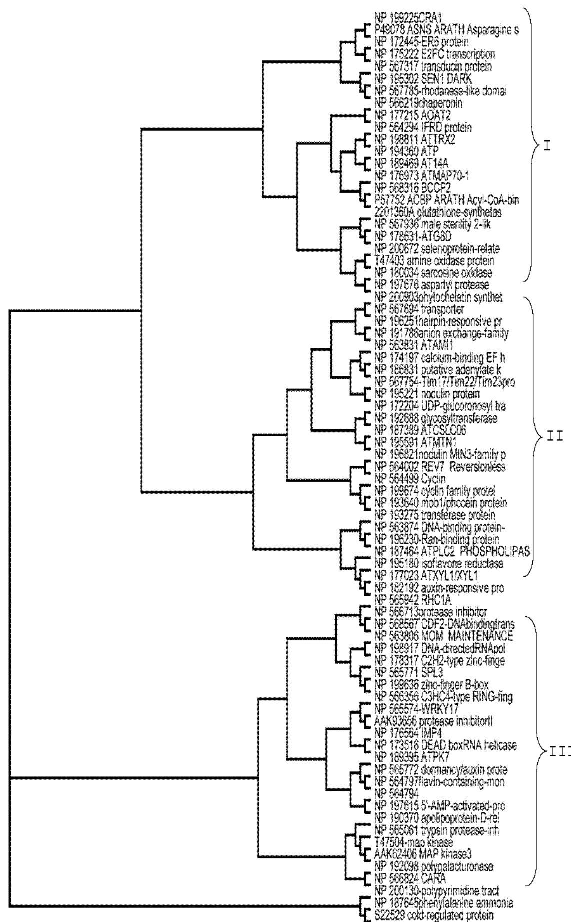

Figure 2. Amino acid sequence alignment of up-regulated Arabidopsis gene products using Clustal X [21]. Clades were marked by single braces.

provides an excellent platform for studying gene/QTL interactions with resistance mechanisms. Several studies suggested that disease resistance genes shared the same specificity identified in distantly related plant species [1, 2, 26] and the reason why the specificity was maintained was perhaps due to balancing selection in lineages leading to multiple plant species [3]. Soybean SDS consists of root infection and leaf scorch. Studies root tissues for TA analysis in both model species will enable us to explore the processes of resistance to SDS. Investigation on differentially expressed genes of Arabidopsis and soybean in response to Fsg can lead to a better understanding on the mechanisms of resistance in crops for certain diseases. Microarray experiment allows interrogation of tens of thousands of genes simultaneously. Orthologous genes with related evolutionary descent should play similar developmental or physiological roles. To investigate the molecular interactions involved in Fsg resistance, we integrated results from a soybean Fsg resistance study with our Arabidopsis DNA microarray analysis through a computational approach. The results of this study can be used as a model system to improve our understanding of plant resistance to Fsg.

5. References

[1] J.L. Dangl, J.L., Ritter, C., Gibbon, M.J., Mur, L.A., Wood, J.R., Goss, S., Mansfield, J., Taylor, and A. Vivian., “Functional homologs of the Arabidopsis RPM1 disease resistance gene in bean and pea”, *Plant Cell*, ASPB, USA, 1992, 4, pp. 1359–1369.

[2] M.T. Simonich and R.W. Innes, “A disease resistance gene in Arabidopsis with specificity for the avrPph3 gene of *Pseudomonas syrmgae* pv phaseolrcola” *Mol. Plant-Microbe lnteract*, APS, USA, 1995, 8, pp. 637-640.

[3] T. Ashfield, L.E. Ong, K. Nobuta, C.M. Schneider, and R.W. Innes, “Convergent evolution of disease resistance gene specificity in two flowering plant families”, *Plant Cell*, ASPB, USA, 2004, 16, pp. 309–318.

[4] D. Leister, A. Ballvora, F. Salamini, and C. Gebhardt, “A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants”, *Nature Genet.*, NPG, UK, 1996, 14, pp. 421-429.

[5] R. C. Grube, E. R. Radwanski, and M. Jahn, “Comparative Genetics of Disease Resistance within the Solanaceae” *Genetics*, GSA, USA, 2000, 155, pp. 873 - 887.

[6] S. Feuillet, N. Travella, L. Stein, L. Albar, L. Nublat, and B. Keller, “Map-based isolation of the leaf rust disease resistance gene Lr10 from the hexaploid wheat (*Triticum aestivum* L.) genome”, *Proc Natl Acad Sci USA*, PNAS, USA, 2003, 100, pp. 15253–15258.

[7] N. Journot-Catalino, I.E. Somssich, D. Roby, and T. Kroj, “The transcription factors WRKY11 and WRKY17 Act as negative regulators of basal resistance in Arabidopsis thaliana”, *Plant Cell*, ASPB, USA, 2006, 18, pp. 3289-3302.

[8] P.M. Schenk, K. Kazan, I. Wilson, J.P. Anderson, T. Richmond, S.C. Somerville & J.M. Manners, “Coordinated plant defense responses in Arabidopsis revealed by microarray analysis”, *Proc Natl Acad Sci USA*, PNAS, USA, 2000, 97, pp. 11655–11660.

[9] J. Wan, F.M. Dunning & A.F. Bent, “Probing plant–pathogen interactions and downstream defense signaling

Authorized licensed use limited to: Southern Illinois University Carbondale. Downloaded on March 18,2010 at 11:13:13 EDT from IEEE Xplore. Restrictions apply.

using DNA microarrays", *Funct Integr Genomics*, Springer, Germany, 2002, 2, pp. 259–273.

[10] D.J. Lockhart and E.A. Winzeler, "Genomics, gene expression and DNA arrays", *Nature* NPG, UK, 2000, 405, pp. 827−836.

[11] M.J. Iqbal, S. Yaegashi, R. Ahsan, K. Shopinski, D.A. Lightfoot, "Root response to *F. solani* f. sp. glycines: Temporal accumulation of transcripts in partially resistant and susceptible soybean". *Theoretical and Applied Genetics*, Springer, Germany, 2005, 110, pp. 1429-1438.

[12] R.A. Gutierrez, R.M. Ewing, J.M. Cherry, P.J. Green, "Identification of unstable transcripts in Arabidopsis by cDNA microarray analysis: rapid decay is associated with a group of touch- and specific clock-controlled genes", *Proc Natl Acad Sci USA*, PNAS, USA, 2002, 99, pp. 11513–11518.

[13] J.C. Lloyd and O.V. Zakhleniuk, "Responses of primary and secondary metabolism to sugar accumulation revealed by microarray expression analysis of the Arabidopsis mutant, *pho3*", *J Exp Bot,,* Oxford University Press, UK, 2004, 55, pp. 1221–1230.

[14] L.D. Hurst, C. Pal, and M.J. Lercher, "The dynamics of eukaryotic gene order", *Nat. Rev. Genet.*, NPG, UK, 2004, 5 pp. 299–310.

[15] L.G. Ma et al., "A microarray analysis of the rice transcriptome and its comparison to Arabidopsis", *Genome Res*, Cold Spring Harbor, USA, 2005, 15, pp. 1274–1283.

[16] M. Labrador, and V.G. Corces, "Setting the boundaries of chromatin domains and nuclear organization", *Cell*, Cell Press, USA, 2002, 111, pp. 151–154.

[17] T. Aoki, K. O'Donnell, Y. Homma, A.R. Lattanzi, "Sudden death syndrome of soybean is caused by two morphologically and phylogenetically distinct species within the *Fusarium solani* species complex—*F. virguliforme* in North America and *F. tucumaniae* in South America", *Mycologia*, MSA, USA, 2003, 95, pp. 660–684.

[18] R. Shoemaker et al., "A compilation of soybean ESTs: generation and analysis", *Genome*, NRC Research Press, Canada, 2002, 45, pp. 329–338.

[19] T. Murashige and F. Skoog, "A revised medium for rapid growth and bioassays with tobacco tissue cultures", Physiol. Plant. ASPB, USA, 1962, 15, pp. 473-497.

[20] J. Pevsner, *Bioinformatics and Functional Genomics*, John Wiley & Sons Inc., USA, 2003.

[21] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins, "The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools", *Nucleic Acids Research,* Oxford University Press, UK, 1997, 25, pp. 4876–4882.

[22] J.D. Thompson, D.G. Higgins, and T.J. Gibson, "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gp penalties and weight matrix choice", *Nucleic Acids Research*, Oxford University Press, UK, 1994, 22, pp. 4673 -4680.

[23] S. Kumar, K. Tamura and M. Nei, "MEGA 3: integrated software for Molecular Evolutionary Genetics Analysis sequence alignment", *Brief Bioinform,* Oxford University Press, UK, 2004, 5, pp. 150–163.

[24] M. Szentirmay, and M. Sawadogo, "Spatial organization of RNA polymerase II transcription in the nucleus", *Nucleic Acids Research*, Oxford University Press, UK, 2000, 28, pp. 2019-2025.

[25] R. Hückelhoven, "Powdery mildew susceptibility and biotrophic infection strategies", *FEMS Microbiology Letters,* Elsevier Inc, Netherlands, 2 005, 245, pp. 9-17.

[26] S. Gabriëls, et al., "An NB-LRR protein required for HR signalling mediated by both extra- and intracellular resistance proteins", The Plant Journal, 2007, 50, pp. 14-28.

Authorized licensed use limited to: Southern Illinois University Carbondale. Downloaded on March 18,2010 at 11:13:13 EDT from IEEE Xplore. Restrictions apply.