FUNCTIONAL GENOMIC ANALYSIS OF WALNUT-NEMATODE INTERACTIONS

Monica T. Britton, Charles A. Leslie, Gale H. McGranahan, and Abhaya M. Dandekar

ABSTRACT

Plant parasitic nematodes can devastate walnut orchards and are able to remain dormant deep in the soil profile for several years, beyond where traditional pesticides can penetrate. The only effective pesticides are often very toxic to beneficial organisms and are detrimental to the environment; several of the most effective are in the process of being phased-out. To develop novel methods of diagnosis and control, it is necessary to characterize pathogens and their interactions with host and non-host plants by performing genomic and functional genetic analyses. This project focuses on the interaction between walnut roots and *Pratylenchus vulnus*, an important nematode parasite. The primary goal is to develop a knowledge base of the genes involved in the interaction between the nematode and the walnut root on which it feeds. Genomics tools, including DNA sequencing and microarrays, are being utilized to create a database available to the walnut research community. The information is then analyzed to identify walnut genes whose expression is changed by nematode infection. This study will establish relationships between nematode infection of walnut roots and the expression of specific genes in the nematode and the walnut host and will allow diagnostic tests to be developed to identify the presence and population of specific parasites within the plant's rhizosphere. In addition, functional genomic techniques are being used to suppress the expression of specific nematode genes to identify those that may be targets for novel pesticides.

OBJECTIVES

1. To survey the pattern of genes expressed in walnut roots and in the feeding of nematodes that determine the outcome of an infection.
   a) Generation of expressed sequence tags (ESTs) by sequencing cDNA libraries prepared from mRNA isolated from *P. vulnus*, and from infected and uninfected walnut plants.
   b) Construction of a database of nematode and walnut genomic information compiled from the ESTs and made available to all researchers.
   c) Genomic analysis of the nematode and walnut databases to assign putative functions to particular genes and identify candidates for functional analysis.

2. To identify and validate nematode and walnut genes associated with the infection process in healthy and sick plants.
   a) Construction and analysis of walnut gene chips to compare expression levels between uninfected and infected plants.
   b) Confirmation of differentially expressed walnut genes by Taqman® real-time quantitative RT-PCR.
   c) Functional analysis of nematode genes using RNA interference in vitro.
   d) Functional analysis of nematode genes using carrot disc assays.

2007 was the second year of this four-year project. Objective 1a was completed in 2006, and was described in the January 2007 research report. During 2007, objectives 1b, 1c, 2a, 2b, and 2c were primarily addressed.

PROCEDURES

Objective 1b. Construction of a database of nematode and walnut genomic information compiled from the ESTs and made available to all researchers.

In 2004, one *Pratylenchus vulnus* cDNA library and two walnut cDNA libraries were built. In 2006, one additional *P. vulnus* cDNA library and four additional walnut cDNA libraries were constructed. All walnut plants used in this research are a clone Px1 of the interspecific hybrid Paradox (*Juglans regia* x *J. hindsii*). Expressed sequence tags (ESTs) were sequenced from each of the 2006 libraries at the UC Davis CAES Genome Facility. Table 1 details the libraries which have been constructed and the ESTs generated from each library.

It should be noted that any attempt to physically remove the nematodes would have affected gene expression in the infected walnut roots. Therefore, it was expected that the infected walnut root cDNA libraries and associated ESTs would contain both walnut and *P. vulnus* sequences. These have been separated primarily by BLAST analysis (comparing the ESTs against both nematode and plant sequences). From this process, 112 non-plant sequences were removed from consideration, while 1650 sequences that could not be definitively assigned to plant or nematode were withheld from database submission (and are not included in Table 1). These sequences will be compared to walnut sequences generated by other research (including the current Walnut Genome Analysis project) and will be submitted after confirmation can be made.

These ESTs were processed to validate the length of good sequence reads and to generate contigs which determine the number of unique genes represented. Data processing and submission to NCBI was performed at the UC Davis CAES Genome Facility. The resulting 12,797 new walnut ESTs were combined with the 5042 existing *Juglans* sequences in the NCBI database (primarily *J. regia* ESTs) and processed using CAP3 to generate a set of 8622 consensus sequences representing unique genes. The *P. vulnus* ESTs have also been clustered into 1176 consensus sequences which are available through NEMBASE at http://nematodes.org/NeglectedGenomes/NEMATODA/Pratylenchus_vulnus/.

Objective 2a. Construction and analysis of walnut gene chips to compare expression levels between uninfected and infected plants.

Microarray analysis utilizes the NimbleGen NimbleChip 4-plex gene expression array (Roche NimbleGen Inc., Madison, WI). These chips can be fabricated quickly, contain probes for sequences specified by the user, and are relatively inexpensive, making them ideal for studying gene expression in non-model organisms like walnut that do not have extensive sequence data available.

After the walnut ESTs were sequenced and contigs processed, a probeset for a NimbleChip 4-plex walnut array was developed from 8457 of the walnut consensus sequences. The arrays

include sequences obtained in this project as well as those from *J. regia* and *J. hindsii* (parental species of Paradox) already available in the public database. Twelve samples were initially processed to validate the arrays. Leaf and root RNA (which had been isolated in 2006 as part of Objective 1a) were separately converted to cDNA using an Invitrogen Superscript II cDNA Synthesis kit. This cDNA was shipped to NimbleGen for labeling and array hybridization. Three biological replicates were processed from each of the two tissues (leaves and roots) of infected and uninfected Px1 plants, with one plant representing a single biological replicate.

Probe-level data were analyzed and normalized to confirm that the quality of the array data was acceptable. Gene-level data were then processed using the Robust Multichip Average (RMA) algorithm. Using the R package LMGene, a one-way ANOVA model was fit for each gene and the standard errors were adjusted by an empirical Bayes method. In addition, all p-values were adjusted for multiple hypotheses testing. Genes with adjusted p-values less than 0.05 were considered differentially expressed between infected and control samples. The lists of differentially expressed genes in leaves and roots were then annotated by performing BLAST searches against both the TAIR7 (Arabidopsis) annotated database and the more comprehensive (but less well-annotated) NCBI nr database. These are shown in tables 2 and 3.

In 2007, a longitudinal study was conducted to determine the time course of a nematode infection on walnut. In March 2007, 80 dormant Px1 clonal plants were removed from cold-treatment, transplanted into tree-pots, and moved into the greenhouse. However, only 37 successfully emerged from dormancy. Of these, four were randomly chosen as pre-inoculation controls. The leaves and roots of each plant were harvested and frozen on the day of inoculation. 17 of the plants were then inoculated with approximately 6000 nematodes per plant, and the remaining 16 plants were mock-inoculated to serve as uninfected controls. At each timepoint—one week post inoculation (wpi), 2 wpi, 4 wpi, and 8 wpi—four infected and four uninfected plants were randomly chosen and their leaves and roots harvested and frozen. (The seventeenth infected plant was also harvested at the 8 wpi timepoint.)

Total RNA was isolated separately from the roots and leaves of each of the 37 plants using the hot borate method (Wilkins and Smart, 1996) and purified with a Qiagen RNeasy mini kit. Three plants from each treatment/timepoint were chosen for microarray analysis (for a total of 54 samples). The leaf and root RNA was separately converted to cDNA using an Invitrogen Superscript II cDNA Synthesis kit. This cDNA was shipped to NimbleGen for labeling and array hybridization. The data resulting from these arrays is currently being statistically analyzed to determine which genes are differentially expressed between infected and uninfected plants over the eight week experimental time course.

Also in 2007, a similar longitudinal study was performed using a model plant, *Medicago truncatula*, which is also a host for *P. vulnus*. This experiment was conducted in the greenhouse, using 66 pots, each containing two *M. truncatula* A17 seedlings and representing one biological replicate. Six weeks after germination, 41 of the pots were randomly chosen and inoculated with approximately 4200 mixed-stage nematodes. The remaining pots were mock-inoculated (no nematodes) to serve as controls. Roots and leaves were harvested and frozen from five pots pre-inoculation to serve as time zero controls. At each of four timepoints—7 days post inoculation (dpi), 16 dpi, 8 weeks post inoculation (wpi) and 12 wpi—plants from five inoculated and five

uninoculated pots were harvested with roots and leaves frozen. Periodically throughout the experimental time period, roots of infected plants were stained and observed microscopically to monitor the course of infection.

Total RNA was isolated from the roots of the ten biological replicates (infected and controls) from the 16 dpi time point using the hot borate method (Wilkins and Smart, 1996) and purified with a Qiagen RNeasy mini kit. This RNA was processed into labeled cRNA using the Affymetrix one-cycle labeling kit, and was subsequently hybridized to Affymetrix Medicago genome arrays. Statistical analysis of the data derived from this experiment will provide a list of differentially expressed Medicago genes. By comparing expression levels of Medicago genes to those on the walnut array, it may be possible to better annotate the homologous walnut sequences.

Objective 2c. Functional analysis of nematode genes using RNA interference in vitro.

RNA interference is being used as a functional genomics tool to suppress the expression of specific *P. vulnus* genes. A detailed explanation of the procedures can be found in the January 2007 report for this project. In the lab, populations of *P. vulnus* are maintained in vitro on roots of germinated Paradox (line Px1) somatic embryos. Nematodes are then subsequently transferred to carrot discs in vitro for population multiplication, after which *P. vulnus* are first isolated from carrot discs by washing with sterile water. 2000 individuals are used per replication for each gene to be silenced. After 24 hours of soaking in an optimized solution containing in vitro transcribed dsRNA, the nematodes are frozen for transcript analysis.

Transcript analysis utilizes Taqman® real-time quantitative RT-PCR to determine if transcription of the targeted gene has been suppressed in comparison with that of nematodes which did not undergo dsRNA soaking and nematodes that were subjected to soaking in dsRNA of negative control genes.

RESULTS AND DISCUSSION

Objective 1a. Generation of ESTs by sequencing cDNA libraries prepared from mRNA isolated from *P. vulnus*, and from infected and uninfected walnut plants.
Status: completed (see previous year's report)

Objective 1b. Construction of a database of nematode and walnut genomic information compiled from the ESTs and made available to all researchers.

As described above, all walnut and *P. vulnus* EST sequences have been deposited in the NCBI database (dbEST). Additionally, clustering of both sets of ESTs has been completed, with the P. vulnus clusters available through NEMBASE at
http://nematodes.org/NeglectedGenomes/NEMATODA/Pratylenchus_vulnus/.

Objective 1c. Genomic analysis of the nematode and walnut databases to assign putative functions to particular genes and identify candidates for functional analysis.

A first pass analysis of the walnut consensus sequences was performed by a BLAST comparison against Arabidopsis sequences using the TAIR7 database and MapMan ontology. Major categorical assignments are shown in Figure 1. More detailed annotations will be made of specific differentially expressed genes based on the statistical analysis of microarray and Taqman qPCR results.

Objective 2a. Construction and analysis of walnut gene chips to compare expression levels between uninfected and infected plants.

Twelve samples from plants grown in the greenhouse in 2006 were used to test the NimbleGen walnut arrays: 3 replicates each of roots from uninfected plants, leaves from uninfected plants, roots from infected plants, and leaves from infected plants. The statistical analysis yielded lists of differentially expressed genes, which are shown in tables 2 and 3.

The results from the first microarrays demonstrated two things: first, the arrays were statistically validated on both the probe-level and the gene-level, which will allow them to be used for comprehensive comparisons involving more samples. Second, although a particular walnut gene may be differentially expressed, it is often not possible to assign even putative annotations. Of the nine genes listed in Table 2, four have no significant BLAST hits and a fifth is homologous only to sequenced genomic DNA. Additionally, the consensus sequences present on the array likely represent no more than 30% of the walnut transcriptome, leaving undiscovered many potentially important genes that could be involved in the response to nematode infection. Therefore, we looked for a model plant to enhance our research and further inform on the effects of nematode parasitism.

After evaluating several plants, *Medicago truncatula* was chosen for use as a model in this research. In vitro, it supports feeding and reproduction of *P. vulnus*, allowing the progress of infection to be easily viewed microscopically. It can be propagated both vegetatively and by seed, permitting experiments to be conducted with numerous plants. It grows very quickly from seed in the greenhouse, and is not deciduous, as walnut is. *M. trunacatula* is a model legume (and, phylogenetically, the closest model plant to walnut), with a nearly-sequenced genome and an Affymetrix microarray, which is available at comparable costs to the NimbleGen walnut arrays.

Although it is relatively simple to generate a large number of biological replicates by germinating seeds of homozygous inbred plant lines, such as the above-described *M. truncatula* A17, it is often difficult to achieve the same goal with vegetatively propagated clonal plants, such as walnut. Walnut cultivars are highly heterozygous, since they are derived from seedlings of outcrossed plants. Therefore, biological replicates of an interspecific hybrid clone (such as Paradox line Px1), must be derived via vegetative propagation. Additionally, the deciduous nature of walnut and its requirement for vernalization (cold-treatment), make it difficult to predict the survival rate of individual clones. Due to these constraints, the experiment for this

milestone was begun over a year early to ensure that it could be replicated, if necessary, during the time allotted for this project.

Lists of differentially expressed genes are being generated from the statistical analysis of each of the three sets of microarray experiments. Changes in expression levels for the specific genes will be confirmed by Taqman® real-time quantitative RT-PCR performed at the UC Davis Lucy Whittier Molecular and Diagnostic Core Facility.

Objective 2c. Functional analysis of nematode genes using RNA interference in vitro.

The cellular process by which double-stranded RNA (dsRNA) activates sequence-specific degradation of mRNA transcripts is believed to have evolved as a viral defense mechanism. RNA interference (RNAi) is triggered by dsRNA and certain ssRNA (for which the complementary strand is synthesized by an RNA-dependent RNA polymerase). The dsRNA is cleaved by a specific RNase III-type endonuclease (Dicer) into fragments (approximately 21 nt long with 2 nt 3′ overhangs) termed small interfering RNA (siRNA). The siRNA duplexes are subsequently incorporated into ribonuclotide-protein complexes (siRNPs), where the strands are separated. In an ATP-dependent process an siRNP is activated to form the RNA-induced silencing complex (RISC). There, the single-stranded siRNA guides the RISC to cleave complementary mRNAs, resulting in silencing of gene expression (Meister and Tuschl, 2004; Dykxhoorn *et al*., 2003).

Because the susceptibility to RNAi by soaking differs between species of nematodes, it is necessary to optimize the soaking conditions for each species. One compound previously shown to induce the uptake of dsRNA in root knot nematodes (Rosso et al., 2005) is resorcinol. However, it appears that, in *P. vulnus*, gene suppression can be induced without the need for this chemical. Figure 2 shows the results of these experiments for four *P. vulnus* genes. In each case, gene expression is reduced solely with the addition of dsRNA (comparison of the first and second bars for each gene). Overall gene expression is also generally reduced by inclusion of resorcinol (most likely due to the toxic effects of this compound). However, the combination of resorcinol and dsRNA may not enhance suppression of target genes. Additional replicates for each gene are being performed so that the effects of dsRNA on gene expression can be confirmed statistically.

REFERENCES

Dykxhoorn, D.M., Novina, C.D. and Sharp, P.A. 2003. Killing the Messenger: Short RNAs That Silence Gene Expression. *Nature Reviews Molecular Cell Biology*. 4:457-467.

Meister, G. and Tuschl, T. 2004. Mechanisms of gene silencing by double-stranded RNA. *Nature*. 431:343-349.

Rosso, M.-N., Dubrana, M.P., Cimbolini, N., Jaubert, S. and Abad, P. 2005. Application of RNA Interference to Root-Knot Nematode Genes Encoding Esophageal Gland Proteins. *MPMI.* 18:615-620.

Thimm, O., Blasing, O., Gibon, Y., Nagel, A., Meyer, S., Kruger, P. Selbig., J., Muller, L.A., Rhee, S.Y., and Stitt, M. 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal* 37:914-939.

Usadel, B., Nagel, A., Thimm, O., Redestig, H., Blaesing, O.E., Palacios-Rojas, N., Selbig, J., Hannemann, J., Piques, M.C., Steinhauser, D., Scheible, W.-R., Gibon, Y., Morcuende, R., Weicht., D., Meyer, S., and Stitt, M. 2005. Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. *Plant Physiology*. 138:1195-1204.

Wilkins, T. and Smart, L.B. 1996. Isolation of RNA From Plant Tissue. In: *A Laboratory Guide to RNA: Isolation, Analysis, and Synthesis*. P.A. Krieg, ed. Wiley-Liss, Inc., New York. pp. 21-41.

Table 1. cDNA Libraries from *P. vulnus* and Paradox Walnut Rootstock (clone Px1)

Library	Year	Organism/Tissue	Source/Treatment	ESTs
PV-01	2004	*P. vulnus*	In vitro: On germinated Px1 somatic embryos	2485
PV-02	2006	*P. vulnus*	In vitro: On germinated Px1 somatic embryos	3327
WRO-01	2004	Paradox roots	In vitro: germinated somatic embryos	683
WRN-01	2004	Paradox roots	In vitro: germinated somatic embryos infected with *P. vulnus*	599
WRO-02	2006	Paradox roots	Plants in greenhouse	3250
WLO-02	2006	Paradox leaves	Plants in greenhouse	2715
WRN-02	2006	Paradox roots	Plants in greenhouse infected with *P. vulnus*	2554
WLN-02	2006	Paradox leaves	Plants in greenhouse infected with *P. vulnus*	2996

Table 2. Differentially Expressed Genes in Walnut Roots (Infected vs. Uninfected Controls)

Juglans Gene ID	Change in Expression	FDR Adjusted p-value	Annotation	Expect value
EL898453	up in infected plants	0.00085	No significant hits	>1.00E-05
EL899374	up in infected plants	0.00551	Transcription factor	2.00E-97
EL892579	up in infected plants	0.01200	No significant hits	>1.00E-05
WRN_2002785_C03	up in infected plants	0.02202	No significant hits	>1.00E-05
EL902546	up in infected plants	0.03480	Ubiquitin-like protein	2.00E-14
05apr07juglans1571	up in infected plants	0.03480	Polyphenol oxidase	5.00E-58
05apr07juglans2289	up in infected plants	0.04436	Similar to *Vitis vinifera* genomic DNA	9.00E-10
EL901628	up in infected plants	0.04854	No significant hits	>1.00E-05
EL896173	down in infected plants	0.04854	Raffinose synthase	4.00E-25

Table 3. Differentially Expressed Genes in Walnut Leaves (Infected vs. Uninfected Controls)

Juglans Gene ID	Change in Expression	FDR Adjusted p-value	Annotation	Expect value
05apr07juglans1898	up in infected plants	0.03511	Galactinol synthase 2 (drought inducible gene)	1.00E-174
CV197871	up in infected plants	0.03511	Galactinol synthase 1 (drought inducible gene)	3.00E-72
CV198475	up in infected plants	0.03511	Haloacid dehalogenase-like hydrolase	3.00E-72
EL891230	up in infected plants	0.03511	Haloacid dehalogenase-like hydrolase	6.00E-60
EL901737	up in infected plants	0.04171	No significant hits	>1.00E-05
EL896574	down in infected plants	0.00636	Heat shock protein	1.00E-33
05apr07juglans1243	down in infected plants	0.00636	Heat shock protein	1.00E-40
CV195370	down in infected plants	0.03511	No significant hits	>1.00E-05
EL896324	down in infected plants	0.03511	Heat shock protein	8.00E-44
05apr07juglans2002	down in infected plants	0.03511	Heat shock protein	5.00E-34
EL897323	down in infected plants	0.03962	Thioredoxin	2.00E-18
05apr07juglans1859	down in infected plants	0.03962	Heat shock protein	1.00E-138
EL893910	down in infected plants	0.04171	Peptidylprolyl isomerase (Rotamase)	5.00E-53

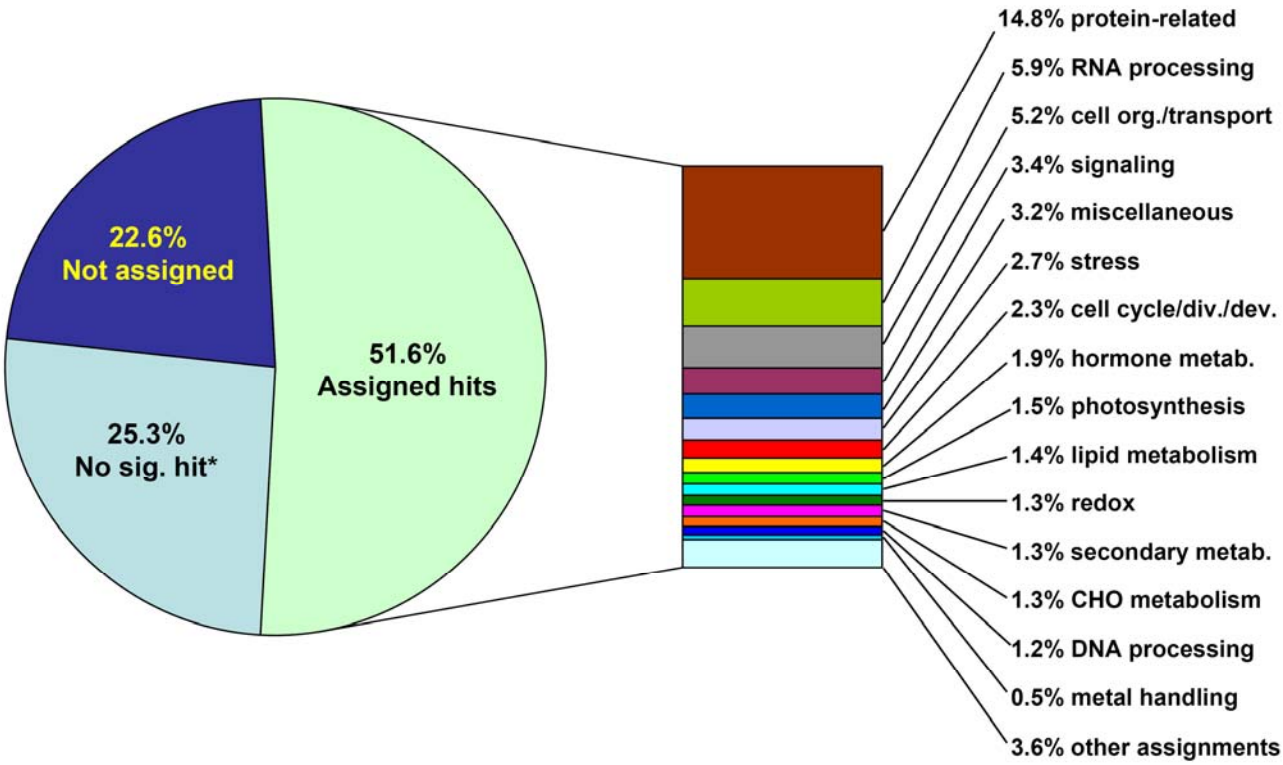

Figure 1. Assignment of Walnut Consensus Sequences to Metabolic Categories. The 8622 walnut consensus sequences were compared to *Arabidopsis thaliana* genes (TAIR7 database) using BLASTX with a significance (E-value) cutoff of 10^{-5} and were assigned to categories based on the MapMan ontology (Thimm et al., 2004; Usadel et al., 2005).

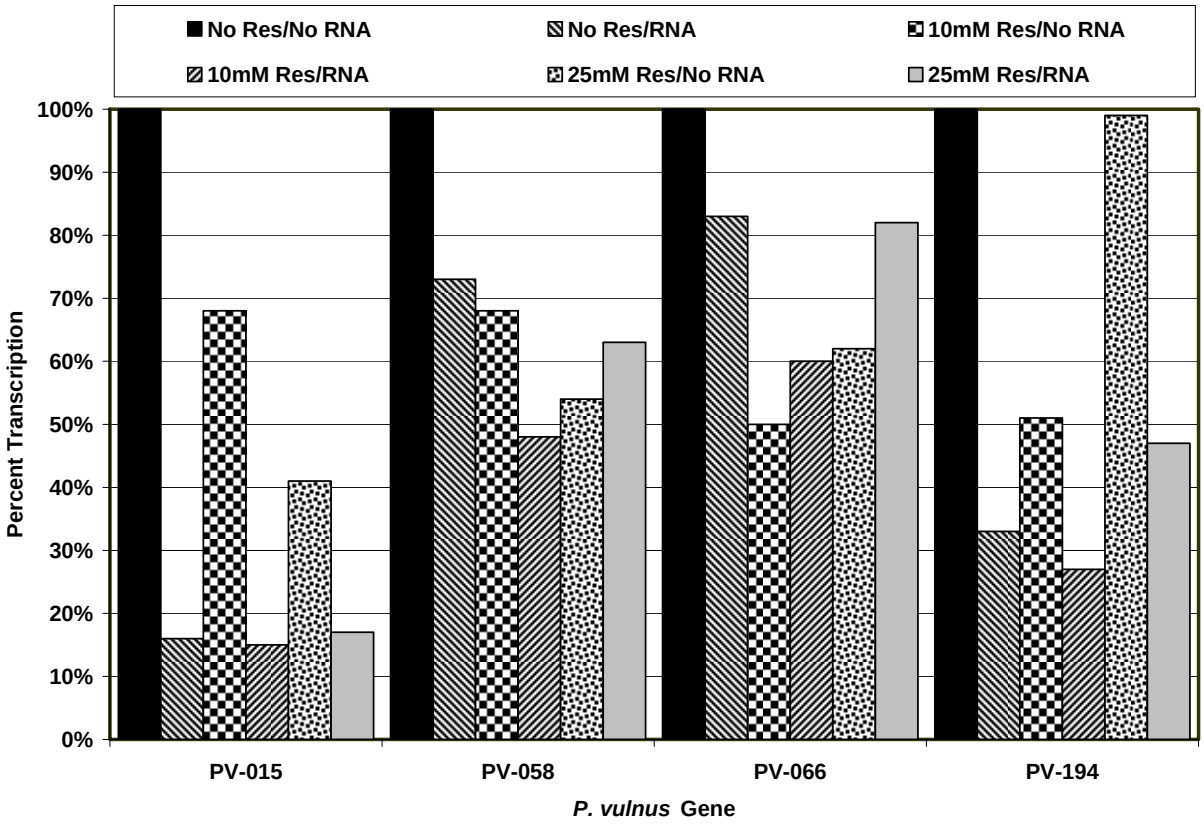

Figure 2. Suppression of *P. vulnus* Gene Expression by RNAi. Each bar represents an average of 2 or 4 samples (of approximately 2000 nematodes each) that have been soaked in a solution of dsRNA corresponding to a specific nematode gene (i.e., PV-015, PV-058, PV-066 or PV194) and resorcinol at concentrations of 0, 10 or 25mM. Percent transcription for each treatment is compared to negative controls with no resorcinol, which are shown as 100% transcription.