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 may remain dormant deep in the soil profile for several years, beyond reach of traditional pesticides. Nematicides are often very toxic to beneficial organisms and detrimental to the environment; several of the most effective are being phased out. Developing novel methods of diagnosis and control requires characterizing pathogens and their interactions with host and non-host plants through genomic and functional genetic analyses. This project focuses on interactions between walnut roots and *Pratylenchus vulnus*, an important nematode parasite. The primary goal is to develop a knowledge base of genes involved in interactions between the nematode and the walnut root on which it feeds. Genomics tools such as DNA sequencing and microarrays are being used 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 expression of specific genes in both organisms and will allow development of diagnostic tests which identify the presence and abundance of specific parasites within the plant's rhizosphere. In addition, functional genomic techniques are being used to suppress expression of specific nematode genes to identify potential targets for novel pesticides.

PROJECT OBJECTIVES

1. Survey the pattern of genes expressed in walnut roots and in feeding nematodes that determines 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 these 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. 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 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 *Medicago* root assays.

2008 was the third year of this four-year project. Objective 1 was completed in 2006-7, and was described in the January 2007 and January 2008 research reports. During 2008, objectives 2a, 2b, and 2c were primarily addressed.

PROCEDURES

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

Microarray analysis used the NimbleGen NimbleChip 4-plex gene expression array (Roche NimbleGen Inc., Madison, WI). A probe set for a NimbleChip 4-plex walnut array was developed from 8457 walnut genes. The arrays include sequences obtained in this project and others from *J. regia* and *J. hindsii* (parental species of Paradox) already available in the public database. Twelve samples were initially processed for array validation. Leaf and root RNAs from three infected and three uninfected plants 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 tissue (leaf or root) of infected and uninfected Px1 plants, with one plant representing a single biological replicate. Subsequently, an additional 58 samples were processed. These were from 28 walnut plants grown in the greenhouse in 2007, half of which were infected with nematodes. In this longitudinal study, whole plants were harvested and frozen on the day of inoculation and at one, two, four, or eight weeks post inoculation (wpi). Leaf and root RNA was extracted from each sample for microarray hybridization.

A similar longitudinal study was performed in 2007 using *Medicago truncatula*, a model plant that is also a host for *P. vulnus*. Total RNA was isolated from roots of ten biological replicates (five infected and five controls) 16 days after inoculation with *P. vulnus.* This RNA was processed into labeled cRNA using the Affymetrix one-cycle labeling kit and subsequently hybridized to Affymetrix Medicago genome arrays.

Gene expression data generated from the microarrays were analyzed using the R package Limma. A one-way ANOVA model was fit for each gene and the standard errors were adjusted by an empirical Bayes method. All *p*-values were adjusted for false discovery rate (FDR) in multiple hypotheses testing. Genes with FDR-adjusted *p*-values < 0.2 and log fold changes > 1.5 were considered differentially expressed among infected and control samples. These statistical comparisons were made for infected vs. uninfected plants for both roots and leaves at each time point of the Paradox longitudinal study. The resulting gene lists are being further analyzed to add annotation and generate clustered expression patterns over the course of infection. In addition, the *Medicago* and walnut arrays are being compared to better understand the pathways influenced by nematode infection.

Objective 2b. Confirmation of differentially expressed walnut genes by Taqman® real-time quantitative RT-PCR.

Seven *Medicago* genes were chosen for validation using Taqman real time quantitative RT-PCR. An additional 25 walnut genes that were significantly differentially expressed in infected and

uninfected plants over time were also validated. This design and validation is performed at the UC Davis Lucy Whittier Molecular and Diagnostic Core Facility. Analysis of these validation experiments and comparison with microarray data is currently ongoing.

Objective 2c. Functional analysis of nematode genes using RNA interference *in vitro*.
RNA interference (RNAi) is being used as a functional genomics tool to suppress expression of specific *P. vulnus* genes. As detailed in previous research reports, eight *P. vulnus* genes were selected for experiments to determine whether their expression can be suppressed by RNAi. 2000 individuals per replication are used for each gene. Nematodes were soaked for 24 hours in a solution containing *in vitro*-transcribed dsRNA, then frozen for transcript analysis by Taqman® real-time quantitative RT-PCR to determine if expression of the targeted gene has been suppressed in comparison with that of untreated nematodes. Between 2006 and 2008, multiple RNAi experiments were conducted for each of the eight *P. vulnus* genes. Statistical analyses were performed using the ΔΔCT method (ABI User Bulletin #2) to determine the relative transcription between control and treated samples. *P*-values were calculated using ANOVA on both normalized CT values and the percent transcription.

To investigate high variation between experiments, further assays were conducted to determine whether there was seasonal variation in base metabolism in *P. vulnus.* Carrot disks were inoculated with *P. vulnus* in December 2007 and incubated for two months at five different temperatures (10°C, 20°C, 25°C, 28°C and "room temperature"). The experiment was repeated six months later, with carrot disks inoculated in June 2008. After two months, nematodes were removed from each biological replicate (six per temperature) and counted. 2000 individual nematodes per biological replicate were used for Taqman analysis of the reference gene PVU26S. Statistical analyses were performed to determine differences in transcription of this basal metabolic gene.

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 January 2006 report)

Objective 1b. Construction of a database of nematode and walnut genomic information compiled from the ESTs and made available to all researchers.
Status: completed (see January 2007 report)

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

As described in our 2007 research report, initial analysis of the walnut consensus sequences was performed by a BLAST comparison against *Arabidopsis* sequences in the TAIR7 database. A more comprehensive analysis was also performed to generate annotations, including Gene Ontology and KEGG classifications. Annotations of the *P. vulnus* consensus sequences based on TBLASTX against multiple databases (including NCBI and Wormbase) were made to derive putative annotations, including Gene Ontology categories and KEGG pathways (E.C. numbers). This annotation is being used to interpret results from the microarray experiments.

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

As discussed in the 2007 research report, twelve initial samples from plants grown in the greenhouse in 2006 were used to test the NimbleGen walnut arrays: three replicates each of roots or leaves from infected or uninfected plants. Analysis of the resulting data validated the array design. Subsequently, 58 samples from the 2007 longitudinal study were processed for microarray analysis. Lists of differentially expressed genes are being generated from the statistical analysis of each of the three microarray experiments. Based on clustered expression patterns, 25 walnut genes from two clusters (one root cluster and one leaf cluster) were selected for validation with Taqman® real-time quantitative RT-PCR (Tables 1 and 2). The results of the Taqman validation experiments are currently being analyzed.

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

Ten RNA samples from Medicago roots (five infected, five uninfected) taken 16 days after inoculation with *P. vulnus* were processed and hybridized to Affymetrix arrays. Statistical analysis of the data yielded 14 genes that were significantly differentially expressed. These genes and their annotations are shown (Table 3).

Seven genes were chosen for validation using Taqman real time quantitative RT-PCR. A statistical comparison of the Taqman and microarray data from these seven genes is shown (Table 4). The corresponding expression patterns of the genes were consistent between methods, demonstrating the validity of the micaroarray statistical analysis. A further comparison of the *Medicago* and walnut experiments is underway.

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

RNAi is a cellular process by which double-stranded RNA (dsRNA) activates sequence-specific degradation of mRNA transcripts and is believed to have evolved as a viral defense mechanism. As described in previous research reports, eight *P. vulnus* genes were tested to determine if their expression can be suppressed by RNAi. Although many individual experiments showed reduced transcription, high variation among experiments made it difficult to confirm statistically that expression of these genes was actually being suppressed.

Because the variation appeared seasonal, additional experiments were conducted to determine whether the reference gene PVU26S varied seasonally or by temperature. CT values varied significantly among incubation temperatures and by season (Figure 1). CT values vary inversely with transcript abundance and are Log2 values due to the geometric amplification of PCR. Therefore, a sample with a CT value of 13 has only half the number of transcript molecules as a

sample with a CT value of 12. In winter, there was a very large difference in transcript abundance between 20-25°C, but in summer the differences were smaller, with only one set of samples (20°C 6/3) significantly different from the other sets that were incubated at 20-25°C.

This seasonal variation, independent of temperature, helps to explain why individual RNAi experiments could show suppression (since all samples of a particular experiment were incubated at the same time and at the same temperature), yet normalization of aggregated results was more difficult due to high variation between experiments conducted at different times. Statistical analysis of subsets of the experimental data led to the conclusion that the efficacy of RNAi gene suppression has a significant seasonal variation independent of temperature. When calibrated (Log fold change) values are compared from experiments conducted during the months of March and April vs. all other months, 78% of all experiments in March and April demonstrate increased transcription of the genes studied. However, during the rest of the year (May through February), 77% of all experiments resulted in decreased transcription of the same genes. This is consistent over the three-year period that experiments were run. These differences between seasons were consistent and statistically very significant ($p = 5.0 \times 10^{-15}$).

This difference is even more pronounced in experiments in which no inducing compound was added. Over the course of the experiments, resorcinol (a neural toxin reported to induce uptake of dsRNA in other nematode species) was added to many experimental replicates. When only those samples without resorcinol were compared, 90% of experiments in March and April showed increased gene expression, while 100% of experiments during the rest of the year demonstrated decreased transcription. This is also statistically very significant ($p = 1.42 \times 10^{-15}$). The average expression of samples treated with dsRNA in May through February compared to untreated controls is shown (Figure 2). Expression of seven of the eight genes was statistically significantly decreased (FDR Adj. $p < 0.05$) when compared to untreated controls. The same comparison for experimental samples from March and April showed no gene with significantly decreased transcription and 5 of 7 genes showed increased transcription (Figure 3). The eighth gene had no replicates during those months.

This seasonal variation is not easy to explain, as it does not appear to derive from varying experimental conditions. It is possible that the carrot cultivar on which the nematodes were grown varied, influencing the nematodes' metabolism. However, this would be unusual as the phenomenon is only observed during a narrow window in late winter/early spring, but was seen in experiments conducted in three different years (2006, 2007, and 2008). In the last phase of this project, induction of *in planta* RNAi is being attempted using transgenic *Medicago* roots. Results of those experiments may shed additional light on this seasonal effect. Since all *Medicago* seedlings are of the same cultivar (line A17) and all roots in the experiments are clonal, there should be no variation in plant material between experiments conducted in different months.

Table 1. Differentially Expressed Genes in Leaf Cluster 3 Chosen for Taqman Validation

Gene ID	Annotation	EC
05apr07juglans0427	Formamide amidohydrolase (amino acid metabolic process)	3.5.1.49
05apr07juglans1846	Drought induced protein (response to water stress)	
05apr07juglans1971	Periplasmic protein transport	
CV195261	Wound-response family protein	
CV196886	Mapkkk wak 1 kinase (signal transduction, implicated in cell death)	2.7.11.0
CV197193	NAC domain-containing protein (flower development; implicated in response to stress and abiotic stimulus)	
CV197767	No annotation (homolog found in grape genomic sequence)	
EL891574	Translational repressor protein (RNA binding)	
EL891750	No annotation (no significant hits in database)	
EL894037	Two-component response regulator-like (protein-binding; response to abiotic stimulus)	
EL897875	Plant cell wall protein	

Table 2. Differentially Expressed Genes in Root Cluster 7 Chosen for Taqman Validation

Gene ID	Annotation	EC
05apr07juglans0610	No annotation (no significant hits in database)	
05apr07juglans1197	Ras-related protein (signal transduction; plasma membrane)	
05apr07juglans1852	Fimbrin-like protein (cytoskeleton; response to stress/abiotic stimulus)	
05apr07juglans2010	Pyrophosphatase	3.6.1.1
05apr07juglans2555	Citrulline-aspartate ligase (amino acid derivative)	6.3.4.5
05apr07juglans2646	Uncharacterized protein (secondary metabolic process)	
CV195647	Exocyst complex protein sec5 (exocytosis)	
EL893186	Opa3-like protein	
EL896318	Tomato: Meloidogyne-induced giant cell protein	
EL898292	Polygalacturonate galacturonosyltransferase (carbohydrate metabolism leading to pectin)	2.4.1.43
EL900372	Nuclear DNA-binding (“late elongated-hyptocotyl)	
EL900435	Lustrin a-like protein (matrix component)	
EL901455	Glycogenin glucosyltransferase (extracellular carbohydrate metabolism)	2.4.1.186
WRN_2002784_A05	No annotation (no significant hits in database)	

Table 3. Differentially Expressed Genes in *Medicago* Roots Using Affymetrix Microarrays
(Log fold change > 1.5, Adjusted *p*-value < 0.2)

Genes with increased expression in infected roots:

Medicago Gene	Annotation	LogFC	FDR Adj. *p*-value
Mtr.36013	Ribosomal RNA gene	4.614	0.001
Mtr.9366	Uncharacterized protein	3.010	0.013
Mtr.4639	Nodulin-like protein	1.909	0.113
Mtr.47907	Nodulin family protein	2.538	0.115
Mtr.16274	Proteinase inhibitor (extracellular)	1.894	0.119
Mtr.12647	Galactose-binding lectin precursor	2.628	0.127
Mtr.31817	Elongation factor (protein translation)	1.997	0.155

Genes with decreased expression in infected roots:

Medicago Gene	Annotation	LogFC	FDR Adj. *p*-value
Mtr.43713	Chalcone synthase	-3.259	0.001
Mtr.1360	Hypothetical protein	-2.408	0.011
Mtr.41979	Nod19-like protein	-2.878	0.011
Mtr.17288	Ethylene-insensitive metal transporter (membrane)	-2.794	0.012
Msa.2831	Ferric reductase (plasma membrane transport)	-2.628	0.014
Mtr.11719	Ferric reductase (plasma membrane transport)	-2.215	0.099
Mtr.32081	Ribosomal protein	-1.881	0.115

Table 4. Comparison of *Medicago* Gene Expression Determined by Microarrays or Taqman qPCR in Infected or Uninfected Roots.

Medicago Gene	Array LogFC	Taqman LogFC	Taqman % Transcription*	Array Unadj. *p*-Value	Array FDR Adj. *p*-value	Taqman *p*-value
Mtr.36013	4.614	8.456	35112%	3.18E-07	0.0015	0.0200
Mtr.41979	-2.878	-3.523	9%	5.35E-06	0.0109	0.0328
Mtr.17288	-2.794	-2.455	18%	7.48E-06	0.1222	0.0723
Mtr.47907	2.538	3.610	1221%	0.0001	0.1146	0.0153
Mtr.33457	1.538	1.520	287%	0.0011	0.3933	0.0650
Mtr.18799	2.889	4.925	3038%	0.0012	0.4013	0.0935
Mtr.4242	0.831	0.005	100%	0.0622	0.9281	0.9919

*Transcription in infected roots relative to 100% transcription in uninfected roots

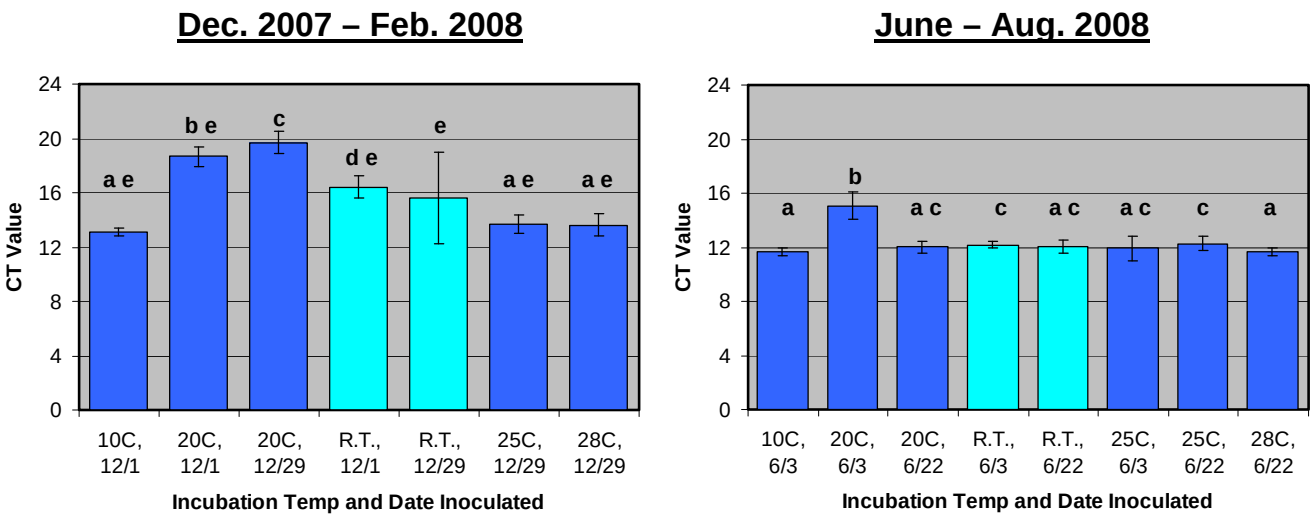

Figure 1. Seasonal variation in *P. vulnus* Reference Gene (PVU26S) CT expression. Carrot disks were inoculated with *P. vulnus* and incubated at various temperatures for two months. A replicate experiment was performed six months later. CT expression varied significantly both between and within temperature sets, with more variation observed during the winter months.

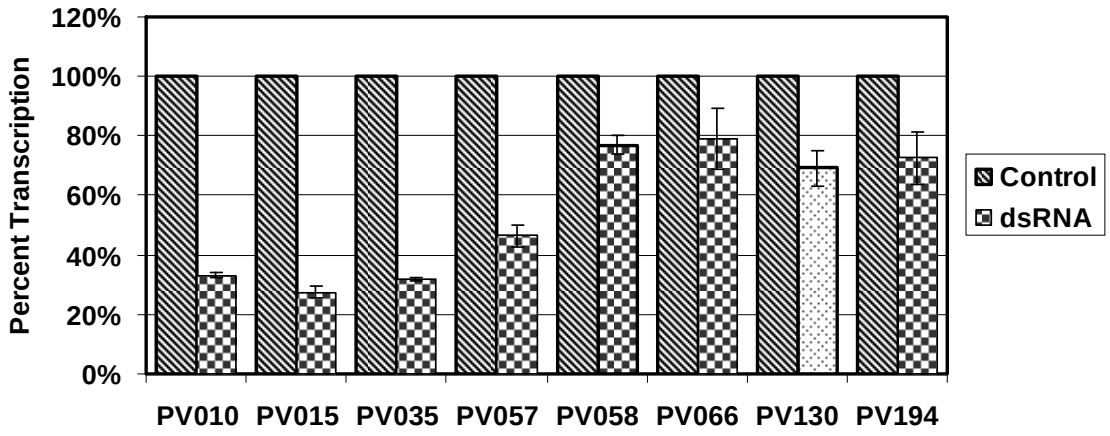

Figure 2. Suppression of *P. vulnus* Gene Expression by RNAi. Each bar represents the average transcription of several nematode samples that were soaked in a solution of dsRNA corresponding to a specific *P. vulnus* gene. All experiments conducted between May and February (2006, 2007, and 2008) with no added resorcinol are included. Controls are set at 100% transcription, and treated samples are shown as a percent of that transcription. All genes showed statistically significant reduced expression except PV130, for which there were only two replicates during the time period, preventing a more precise statistical analysis.

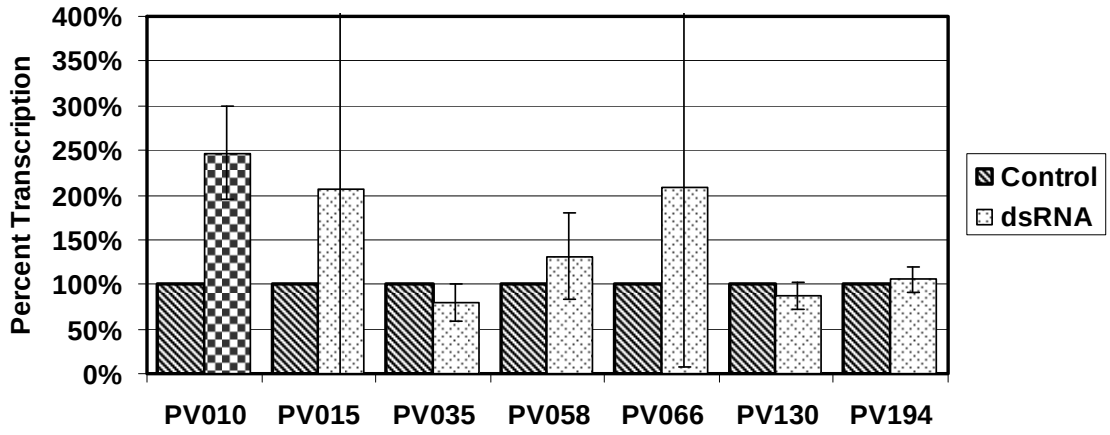

Figure 3. RNAi Experiments Conducted in March and April of 2006, 2007, and 2008. Each bar represents the average transcription of several nematode samples that were soaked in a solution of dsRNA corresponding to a specific *P. vulnus* gene. All experiments conducted in March or April with no added resorcinol are included. Controls are set at 100% transcription, and treated samples are shown as a percent of that transcription. Genes PV010, PV015, PV058, PV066, and PV194 show increased expression. Only PV010 is statistically significantly different than the controls; all other genes are not statistically significant. (There is no data for PV057 from this time period.)