CHANGES IN GENE EXPRESSION DURING RENIFORM NEMATODE INFESTATION OF COTTON SEEDLINGS AS DETERMINED BY DEEP SEQUENCING OF CDNA LIBRARIES

Ruijuan Li
Shankar Pant
Hongtao Hu
Narendra K. Singh
Robert Locy
Department of Biological Sciences
Auburn University, Alabama
Roelof B. Sikkens
David B. Weaver
Department of Agronomy and Soils
Auburn University, Alabama

Abstract

Six cotton varieties and genotypes (Delta Pearl, Suregrow 747, FiblerMax 966, Paymaster 1218, TX245 and TX1419) were infested with reniform nematode, and root samples were collected at 0 days (uninfected), 1 day, 3 day and 5 days post infestation. Total RNA was extracted from these 6 genotypes. Samples from the 4 cultivars at 0 day were pooled as the uninfested control, and the 1, 3, & 5 days post-infestation samples were pooled from all cultivars as the infested plants. Pooled samples from the 2 accessions corresponding to each of the time points were also prepared, giving 4 treatments. cDNA library were prepared and submitted for 454 Flex MPSS sequencing. Expression levels of genes that are up or down regulated during reniform nematode infestation were analyzed. The results indicated that specific genes up or down regulated are related to different physiological processes. This overview of gene expression during plant-nematode interaction will provide insights into indentifying key processes involved in reniform nematode infestation.

Introduction

Cotton (*Gossypium hirsutum L.*), an important textile crop worldwide, is susceptible to multiple species of plant parasitic nematodes. Currently the most damaging species of plant-parasitic nematodes affecting cotton in the US is reniform nematode (*Rotylenchulus reniformis*) (Robinson, 2007). During nematode infestation, specific plant genes may be differentially regulated, and a large number of genes that are differentially expressed during nematode infestation are likely to contribute to establishing the parasitic interaction (Williamson and Gleason, 2003; Gheysen and Fenoll, 2002). Our goal is to develop a system for understanding those genes that are involved in reniform nematode infestation of the cotton crop, and to genetically manipulate more reniform nematode resistant cotton plants using this information. For example, breeding increased resistance to reniform nematode is an important improvement objective at the present time, and in 2007, two cotton breeding lines resistant to *R. reniformis*, LONREN-1 and LONREN-2 where released by the USDA (Starr et al., 2007), but the implementation of these material in commercially useful germplasm has proven problematic.

In this study we have identified those genes that demonstrate expression changes during reniform nematode infestation (up-or downregulated), and here we report our progress on the use of next generation DNA sequencing to establish reniform nematode-infestation-mediated changes in the expression of various genes in six varieties and genotypes of cotton.

Materials and Methods

Plant Material and Stress Treatment

Six genotypes of cotton were selected, four are commercial cultivars susceptible to reniform nematode, they are Delta Pearl, SG747, FM966, and PM1218. Two additional genotypes that are USDA germplasm accessions, TX245 and TX1419, previously identified as more RN-resistance based on greenhouse studies were also employed. These cotton samples, both the S-cultivars and the USDA accessions (R-accessions) were infested with reniform nematodes and root samples were collected at 0, 1, 3 & 5 days post infestation. Samples at 0 day were taken as the uninfested controls, while samples from 1, 3 and 5 days post infestation were pooled and considered the reniform nematode infested samples.

RNA Extraction
Total RNA was extracted from these 4 samples using the hot borate method (Wan and Wilkins, 1994).

cDNA Library Preparation & Deep Sequencing
cDNA was prepared from the total RNA from each of these 4 treatments as described above, and submitted for sequencing utilizing Roche 454 Flex technology at the University of South Carolina Environmental Genomics facility (ENGENCORE).

Bioinformatic Analysis
Figure 1 describes the pipeline we have constructed to analyze sequence reads and determine biological functions of genes in cotton during *R. renifomis* infestation. Raw sequence reads (390,501 reads) coming straight from the MPSS sequencing, were processed using CLC high throughput sequencing tools. This step removed the 3' and 5' sequencing adaptor sequences from the raw reads and removed all sequenced inserts smaller than 75 nucleotides from the dataset. This yielded 320,054 adaptor-trimmed total reads (see Table 1).

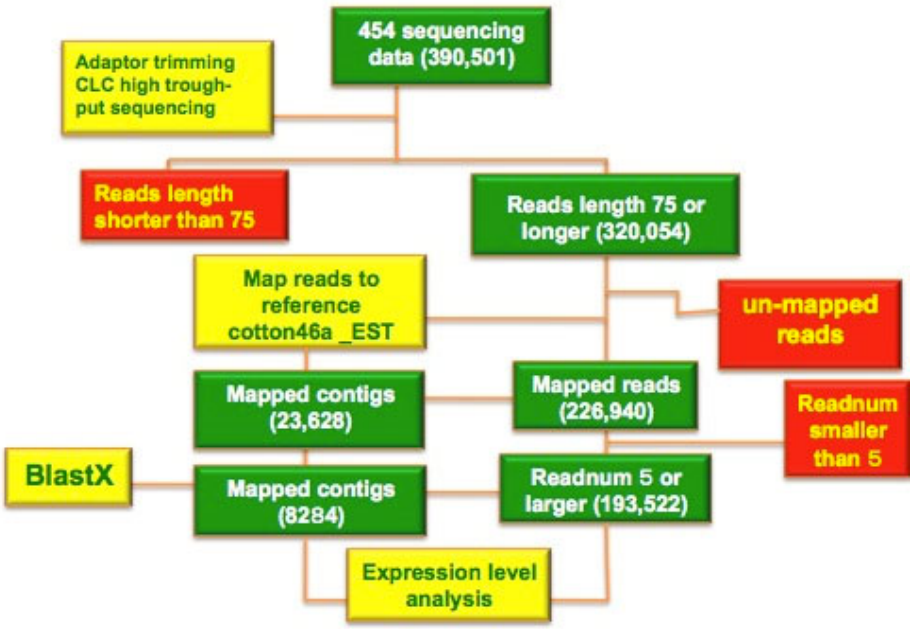

Figure 1. Flow diagram showing the processing of the raw sequence data reads from 454 Flex MPSS sequencing. Unique read abundance is utilized to reveal gene expression level for each expressed sequence. Biological functions of cotton genes were also bioinformatically predicted from BLAST search comparisons to Arabidopsis protein sequences.

These reads were then mapped to cotton 46a EST database obtained from the Comparative Evolutionary Genomics of Cotton webpage, and the number of mapped reads were determined for each of the 4 datasets and for the total dataset. Low abundance reads (readnum =< 5) were then filtered out yielding 193,522 reads that match to 8284 contigs in cotton 46a EST dataset. Those sequences with a match to contigs in the cotton 46a EST database were further analyzed to examine differential gene expression levels.

BlastX searches of the expressed sequences against NCBI Arabidopsis protein database was performed to predict biological functions of cotton genes differentially expressed during *R. renifomis* infestation.

Results and Discussion

Sequencing Results
A total of 390,501 sequence reads (Table 1) were obtained from 454 sequencing. After trimming low quality sequences and adaptor removal, sequences shorter than 75 nucleotides were discarded, and the remaining 320,054 reads were mapped to the cotton 46a EST database obtained from the Comparative Evolutionary Genomics of

Cotton webpage. 222,600 reads in four of the treatment sequencing databases show greater than 85% homology to sequences in the 46a EST database. Subsequently all sequence reads with total reads number less than 5 were filtered obtaining 8284 contigs matching to 193,522 reads in four of the databases. There is significant variation in the number of reads in each dataset that generally correlates with number of assembled, mapped contigs in each dataset.

Table 1. Results of the sequencing of the 4 cDNA libraries.

	Tissue and treatment				
Dataset	SI	SU	RI	RU	Total
Total Reads	101,380	79,775	153,745	55,601	390,501
After Trimming	80,117	61,226	130,823	47,888	320,054
Mapped Reads	57,436	42,682	91,723	35,099	226,940
Filtered Read	50,050	37,320	37,320	31,065	193,522
Assembled Contigs	7,139	6,274	7,901	5,836	8,284
Average Reads/Contig	7.01	5.95	10.06	5.32	23.9

R. reniformis **Infestation Regulates Gene Expression in Cotton Roots**

The 4 cDNA sequencing datasets were further analyzed by examining the number of reads found in various categories of gene expression. The ratio of expression of a sequence in infested tissue divided by the expression of a sequence in uninfested tissue was calculated for both the resistant accession and susceptible cultivar datasets, and for the sum of the 2 groups of pooled genotypes. Then these data were plotted on a scale that involves the logarithm in the base 2 of this ratio. The overall expression of EST sequences in cotton roots in response to reniform nematode infestation is shown in Figure 2.

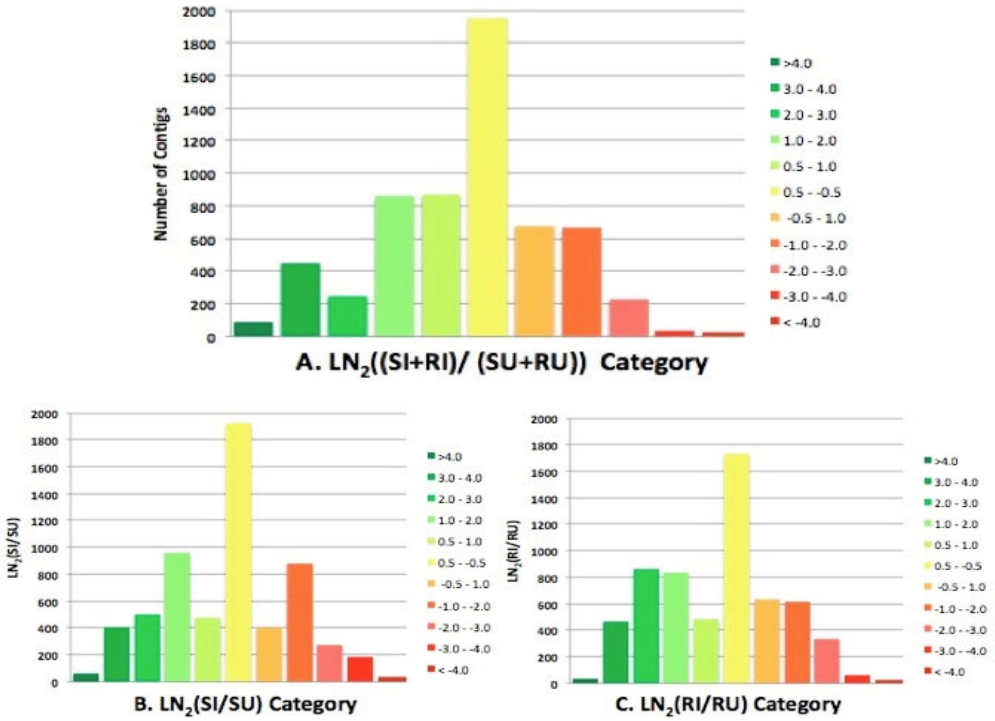

Figure 2. Panel A: The overall expression of EST sequences in sum of these two groups of genotypes of cotton roots in response to reniform nematode infestation is shown. Panel B: The overall expression of EST sequences in roots of susceptible genotypes of cotton in response to reniform nematode infestation is shown. Panel C: The overall expression of EST sequences in resistant genotype of cotton roots in response to reniform nematode infestation is shown.

Figure 3 shows the distribution of the expression levels of cotton root ESTs. The sequences were sorted based on expression level. It is noteworthy that a larger number of ESTs demonstrate up-regulation rather than down-regulation in both susceptible and resistant genotypes.

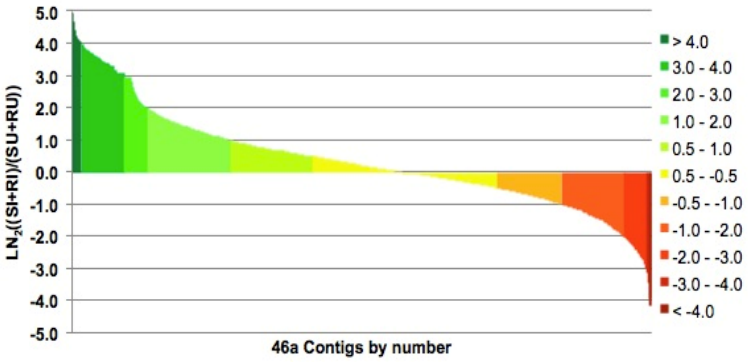

Figure 3. Distribution of the expression levels of cotton roots ESTs during *R. reniformis* infestation.

Table 2 shows summarizes by category, the ESTs that are up or downregulated greater than 3 fold during reniform nematode infestation. Among the 1175 upregulated sequences, 5.9% or 488 sequences of the total 8284 sequences expressed in roots with or without nematode infestation were upregulated more than 3-fold in both groups of genotypes following nematode infestation. Only 1.9% of the 8284 sequences or 161 sequences were downregulated in both the S-cultivars and R-Accessions following reniform nematode infestation. By comparison only 2.6% or 212 of the total sequences were upregulated in S-cultivars while 471 or 5.7% of the total sequences were upregulated in the R-accessions. The numbers of downregulated sequences in both groups of genotypes were more nearly equal and fewer in number. Thus, of the 1175 upregulated sequences, nearly 60% were upregulated in the S-cultivars; while over 80% of these sequences were upregulated in the R-accessions.

It can be concluded that a greater number of genes respond to nematode infestation in the wild accessions than in the more highly bred cultivars. This however, does not appear to be of universal utility in creating nematode resistant materials although it does suggest that a more diverse gene pool is available in these materials that may be useful for generating low levels of horizontal resistance that may ultimately prove useful.

Table 2. Greater than 3 fold up- and downregulated sequences.

	Number of Contigs	% of 8294 Total
Upregulated > 3x Total	1175	
Upregulated > 3x in both	488	5.9%
Upregulated >3x only in “S” cultivars	212	2.6%
Upregulated >3x only in “R” accessions	471	5.7%
-3x < level < 3x	6459	78.0%
Downregulated < -3x Total	651	
Downregulated < -3x in both	161	1.9%
Downregulated < -3x only in “S” cultivars	261	3.2%
Downregulated < -3x only in “R” accessions	224	3.2%

Functional Annotation

BLASTX searches were conducted using the up- and down-regulated sequences obtained from our study. These sequences were “blasted” against the Arabidopsis Protein database at NCBI, and the results were used to develop an annotation table showing the best hit in the Arabidopsis genome for each of the sequences. Figure 4 shows functional categorization of contigs, which are up or down-regulated greater than 3-fold. Of the cotton sequences differentially expressed only a few sequences (those labeled “Not Available”), comprising 67 sequences upregulated by > 3x and 53 down-regulated sequences, appeared to have no comparable sequence in the Arabidopsis genome. The largest number of sequences found is listed in a category labeled function unknown. This group contains sequences labeled as hypothetical proteins, unknown function proteins, or unnamed proteins, as well as a few additional proteins for genes involved in developmental processes but that have no defined biochemical or cellular function. Because this result is a preliminary analysis, further effort is presently being placed on identifying additional categories of significance within this large group. Note that the categories labeled “Metabolic”, “Transcription Factors”, “Signal Transduction” represent genes whose products are involved in metabolic functions,

transcriptional regulation, and signal sensing and transduction respectively, are among the groups having significant numbers of up-regulated genes. Note also that to a lesser extent, genes involved in membrane and water transport, mitochondria, chloroplasts, oxidative stress, and the cytoskeleton are not abundantly differentially expressed.

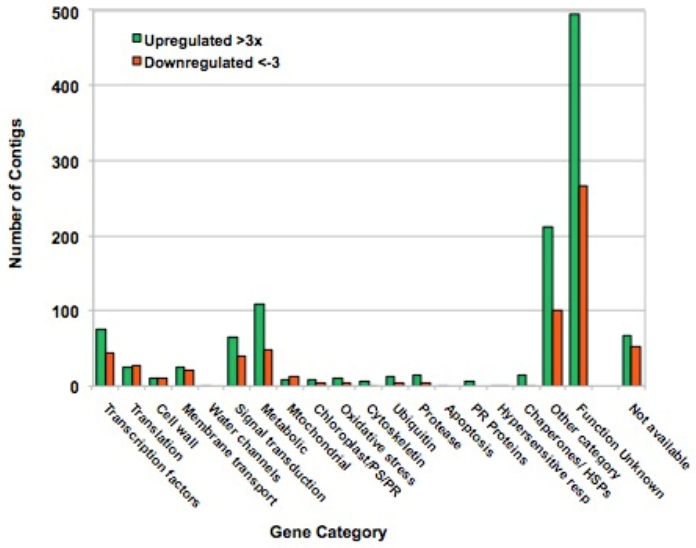

Figure 4. Categorization of contigs into various functional groups.

Table 3 shows the differential regulation for a set of gene categories that are most likely involved in host-pathogen interaction. These include genes involved in protein degradation (Ubiquitin-related and Proteases), Apoptosis-related genes, Pathogenesis-related proteins, Chaperones-heat shock proteins, and proteins involved in the hypersensitive response. A disproportionally greater number of these groups are up-regulated rather than down-regulated during nematode infestation.

Table 3. The number of genes in various categories that are up- or down-regulated greater than 3-fold are indicated.

	Up-regulated < 3x		Down-regulated < -3x	
	(Number)	(%)	(Number)	(%)
Transcription factors	75	6.4%	45	6.9%
Translation	26	2.2%	28	4.3%
Cell wall	11	0.9%	10	1.5%
Membrane transport	25	2.1%	22	3.4%
Water channels	3	0.3%		
Signal transduction	66	5.6%	40	6.1%
Metabolic	109	9.3%	48	7.4%
Mitochondrial	8	0.7%	13	2.0%
Chloroplast/PS/PR	8	0.7%	4	0.5%
Oxidative stress	11	0.9%	4	0.6%
Cytoskeleton	7	0.6%	3	0.5%
Ubiquitin	13	1.1%	4	0.6%
Protease	14	1.2%	4	0.6%
Apoptosis	2	0.2%		
Hypersensitive response	2	0.2%	2	0.3%
PR Proteins	7	0.6%		
Chaperones/ HSPs	14	1.2%	3	0.5%
Other category	212	18.0%	101	15.5%
Function unknown	495	42.1%	267	41.0%
Not available	67	5.7%	53	8.1%

Conclusions

1. Deep sequencing of RNA from cotton root during *R. reniformis* infestation has generated a set of data that demonstrate changes in gene expression that take place in a more susceptible group of cultivars and a less susceptible pair of USDA accessions.
2. Up-regulation of gene expression is more abundant during nematode infestation than down-regulation.
3. The putative "R-accessions" show a greater diversity of up-regulated sequences than do the "S-cultivars".
4. Sequences most likely involved in host-pathogen interactions, are among the sequences demonstrating differential expression, although we are unable to confirm that such sequences are causally involved in resistance.
5. Studies are underway to extend this gene-expression approach to more stably resistant genotypes and to demonstrate changes in gene expression in resistant and susceptible genotypes with the intent of establishing genes and alleles that are associated with resistance and susceptibility.

Acknowledgements

This work was supported by grants from the Alabama Cotton Commission, Cotton, Inc., and the Alabama Agricultural Experiment Station.

References

Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410.

Gheysen G and Fenoll C. 2002. Gene expression in nematode feeding sites. Annual Review of Phytopathology 40: 191-219.

Robinson AF. 2007. Reniform in the U.S. cotton: When, where, why, and some remedies. Annual Review of Phytopathology 45: 263-288.

Starr JL, Koenning SR, Kirkpatrick, TL, Robinson AF, Roberts PA, and Nichols RL. 2007. The future of nematode management in cotton. Journal of Nematology 39(4): 283- 294.

Wan CY and Wilkins TA. 1994. A Modified Hot Borate Method Significantly Enhances the Yield of High-Quality RNA from Cotton (*Gossypium hirsutum* L.). Analytical. Biochem. 223, 7-12.

Willamson VM and Gleason CA. 2003. Plant–nematode interactions. Current Opinion in Plant Biology 6:1–7.