Expressed sequences in Three Leafhoppers Vectors of Pierce's Disease

Wayne B. Hunter[1], K.S. Shelby[2], L.E. Hunnicutt[3], R.F. Mizell[4], C. Tipping[5].

1.USDA, ARS, U.S. Horticultural Research Lab, Ft. Pierce, FL 34945. 2. USDA, ARS, Columbia, MO, 3. North Carolina State University, Genomic Sciences, Raleigh, NC 27695,

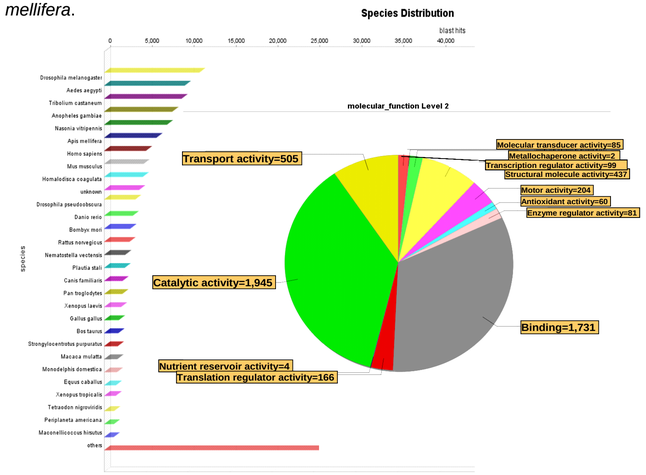
4. University of Florida, IFAS, NFREC, Quincy, FL 32351, 5.Delaware Valley College, Doylestown, PA 18901. Wayne.hunter@ars.usda.gov

July 2009

ABSTRACT- Leafhoppers are considered the second most important vector of agricultural diseases. We examined the gene expression across three leafhopper leafhoppers, *Homalodisca vitripennis, Graphocephala atropunctata*, and *Oncometopia nigricans,* which are vectors of the plant-infecting bacterium, *Xylella fastidiosa*, which causes Pierce's disease of grapes. The use of genomic data is providing new information on the biology and relatedness of these and other leafhoppers. Using a genomics approach has also advanced the understanding of leafhopper immunity, pathology, and development. As new developments in genomics and RNAi methodologies emerge, researchers will be able to use this genetic information to design highly specific and effective management tools to reduce either leafhopper populations, and/or leafhopper-transmitted diseases. The importance of these leafhoppers as the vectors of Pierce's disease, the abundance of ESTs produced for each, and their differences in host plant preferences, provide an excellent opportunity to conduct comparative examination of these leafhopper's. Several cDNA libraries which had been made from adult GWSS, BGSS, and BWSS, plus nymphs, and tissues, provided a resource totaling ~50,000 ESTs. When assembled we obtained ~5,000 specific transcripts for each species for comparison. This is approximately one-third of all predicted active genes available, as other insect genomes have demonstrated ~15,000 total genes. These were used for analyses between these species as well as for larger analysis to known genomes. Further analyses were conducted *in silico* using software programs available online Internet Resources, NCBI, EXPASY, and others to compare assembled data, predict proteins and compare them to the broader scope of insect genomes.

Many other genes of interest which have various functions in leafhopper biology and physiology have also been identified but are not reported herein. The EST sequences reported in this study have been deposited in GenBank's dbEST. (see references: Hunter 2005, 2006, 2007).

Fig. 1. Composite figure showing distribution of *Homalodisca vitripennis* transcripts across other species (along left), with the top 6 species homologies being in these insects whose genomes have been completed: *Drosophila melanogaster*, *Aedes aegyptii*, *Tribolium castaneum, Anopheles gambiae, Nasonia vitripennis, and Apis mellifera*.



Digestive Enzymes: Aminopeptidases, several cathepsin L–like cysteine proteases, and other proteases have been identified in these leafhoppers which are also in other piercing-sucking feeding insects (Foissac et al., 2002, Wright et al., 2006, Zhu et al., 2003). In aphids, a cathepsin B protease has been shown to be constitutively expressed in all aphid individuals, suggesting gene duplication and evolution of a novel biological function of cathepsin B in the aphid lineage (Houseman and Downe 1983). Cathepsin B proteases were also identified in these leafhoppers and may show similar duplication. Cleavage of food proteins into peptides and amino acids is an important process for which an array of proteases of different substrate specificity and enzymatic activities are produced in the alimentary tract and are involved in protein digestion Annotation of these data advances current understanding of leafhopper biological pathways while providing clues to the genetic basis of such processes in insect-pathogen, and insect-plant interactions(Fig. 2, 3). The availability of genomic data for these leafhoppers continues to increase, thus uses of the current data provides a solid foundation for future studies in leafhopper functional genomics.

CONCLUSIONS - The information gained from this study provides the first investigation using comparative genomics of the transcriptomes from three leafhopper vectors of Pierce's disease of grapes: *H. vitripennis, G. atropunctata*, and *O. nigricans*. Amino acid sequence comparisons BLASTX, BLASTP with other known proteins relies on conserved motifs of specific domain(s), NCBI GenBank. *In silico* comparative analyses continues to be a valuable tool to identify proteins and their functional domains as widely accepted method to address many biological questions.

Many of the discoveries made in other insects, such as *Drosophila*, and Honey Bee, can be applied when the same genetic transcripts can be identified. For example we increased our understanding of the roles and pathways of heat shock proteins in leafhoppers by examining the data completed in Locusts, Flies, and Nematodes. The same is true for digestive enzymes. The increasing application of transcriptional data is leading the way in the development of new strategies to reduce plant diseases and their insect vectors. Application of RNAi against a wide range of insect species from spruce budworm to whiteflies are viewed as the future in insect pest control, and many new methods which incorporate the use of native endophytic bacteria and/or viruses as the mechanism for delivery or expression of dsRNA within plants are being widely evaluated. The main advantages of applying genomic data in this manner to solve agricultural problems is that the plants are not 'transformed', thus the quality of the crop is not altered, saving time, money, and reducing the effort needed to find solutions to many emerging devastating agricultural problems.

Accession	EC #	BLAST DeKegg Ids	GO TYPE	GO1	GO2	GO3	HC_ID	E-value (HC)	GA_ID	E value (GA)	ON_ID	-value (ON)
Q7QEH5		AgCP7442	Molecular Function	motor activity	structural molecu	structural r	WHMg 2758	0	WHGA067-6_	5.4086E-102	WHON042_E12	1.1E-134
Q86GF8		Hypothetic Citrate cycle (TCA cycl	Molecular Function	aconitate hydratase	mitochondrion	tricarboxyl	WHMg 2485	0	WHGA2354	7.01752E-63	WHON0097	0
Q6PPI6		Putative cytoplasmic actin A3a1	Molecular Function	motor activity	structural molecu	structural r	WHHC51531	0	WHGA2663	4.55082E-67	WHON0017	0
Q6RFY9	[3.6.5.3]	Putative elongation factor 1-alpha	Molecular Function	translation elongatio	GTP binding	cytoplasm	WHHC51398	0	WHGA0038	0	WHON0147	0
Q6PPI5		Putative muscle actin	Molecular Function	motor activity	structural molecu	structural r	WHHC51395	0	WHGA2643	1.92228E-73	WHON0021	0
Q7PMH3	1.6.5.3	ENSANGP Ubiquinone biosynthes	Molecular Function	iron ion binding	electron transpor	mitochond	WHHC2058	0	WHGA1900	8.2498E-107	WHON1268	0
Q7PPE7	2.7.1.40	ENSANGP Glycolysis / Gluconeog	Molecular Function	magnesium ion bind	pyruvate kinase	glycolysis	WHHC0110	0	WHGA2158	1.54504E-62	WHON0828	2.1E-138
Q6PPH2		Putative activated protein kinase	Molecular Function	receptor activity	kinase activity		WHHC2042	0	WHGA006-44	1.6435E-120	WHON1478	1.98E-48
Q6PPI6		Putative cytoplasmic actin A3a1	Molecular Function	motor activity	structural molecu	structural r	WHHC0062	0	WHGA2663	4.55082E-67	WHON0017	0
Q6PPH9	1.14.19.1	Putative delta-9 desaturase 1	Molecular Function	stearoyl-CoA 9-desc	iron ion binding	endoplasm	WHHC0255	0	WHGA0206	0	WHON0283	0
Q6RFY9	3.6.5.3	Putative elongation factor 1-alpha	Molecular Function	translation elongatio	GTP binding	cytoplasm	WHHC0034	0	WHGA0038	0	WHON0147	0
Q6PPI0	4.1.2.13	Putative fn Glycolysis / Gluconeog	Molecular Function	fructose-bisphospha	glycolysis	lyase activi	WHHC0177	0	WHGA1213	0	WHON0126	0
Q6PPI5		Putative muscle actin	Molecular Function	motor activity	structural molecu	structural r	WHHC0074	0	WHGA2643	1.92228E-73	WHON0021	0
Q6PPI1		Putative rhodopsin	Molecular Function	rhodopsin-like recep	receptor activity	G-protein r	WHHC0120	0	WHGA051-49_	3.1982E-122	WHON0281	0
Q6PPI3		Glyceraldel Glycolysis / Gluconeog	Molecular Function	glyceraldehyde-3-ph	glucose metabol	glycolysis	WHMg 1552	4.3383E-168	WHGA0261	1.9645E-169	WHON0243	4.9E-173
Q6PPH2		Putative activated protein kinase	Molecular Function	receptor activity	kinase activity		WHHC52008	4.6847E-166	WHGA2745	9.75547E-72	WHON1478	1.98E-48
Q6PPI0		Putative fr Glycolysis / Gluconeog	Molecular Function	fructose-bisphospha	glycolysis	lyase activi	WHMg 2297	2.39564E-147	WHGA1213	0	WHON013_H01	2.59E-19
Q17083		Vitellogenin precursor	Molecular Function	lipid transporter activ	lipid transport	nutrient res	WHHC0041	3.3648E-143	WHGA019-70_	2.47424E-39	WHON0059	1.1E-105
Q9VFF0		CG3731-PA	Molecular Function	metalloendopeptidas	proteolysis		WHHC52247	5.2528E-142	WHGA085-15_	4.12314E-65	WHON036_H10	5.27E-30
Q7PXA8		AgCP12715	Molecular Function	catalytic activity	GTP binding	tricarboxyl	WHHC1385	8.5228E-136	WHGA2584	2.53721E-45	WHON007_F09	2.43E-73
Q7PF83		ENSANGP00000024398	Molecular Function	electron transporter	iron ion binding	tricarboxyl	WHHC2462	1.1902E-129	WHGA1354	1.55212E-69	WHON030_F08	2.06E-64
Q7PPE7	[2.7.1.40]	ENSANGP00000021580	Molecular Function	magnesium ion bind	pyruvate kinase	glycolysis	WHSg044_B07	7.9038E-108	WHGA2158	1.54504E-62	WHON0828	2.1E-138
Q7PPE7		ENSANGP Glycolysis / Gluconeog	Molecular Function	magnesium ion bind	pyruvate kinase	glycolysis	WHMg 2868	2.4039E-107	WHGA2158	1.54504E-62	WHON0828	2.1E-138
Q9VFF0		CG3731-PA	Molecular Function	metalloendopeptidas	proteolysis		WHMg046_G12	6.9007E-106	WHGA085-15_	4.12314E-65	WHON036_H10	5.27E-30
Q9V4E3		CG1970-PA	Molecular Function	electron transporter	mitochondrion	electron tra	WHSg031_H09	6.3715E-104	WHGA2174	1.50812E-85	WHON032_H08	4.25E-96
Q17083		Vitellogenin precursor	Molecular Function	lipid transporter activ	lipid transport	nutrient res	WHHC0115	4.2689E-103	WHGA1453	4.34174E-57	WHON0059	1.1E-105
Q6PPI2		Putative ferritin GF2	Molecular Function	binding	iron ion transpor	iron ion hc	Contig 1413	3.50913E-95	WHGA0709	7.98921E-91	WHON0085	2.56E-92
Q7QB64		AgCP2476	Molecular Function	malic enzyme activit	malate metabolis	oxidoreduc	WHSg041_F10	3.9499E-95	WHGA1635	5.71834E-84	WHON034_G07	5.85E-93
Q6PPI2		Putative ferritin GF2	Molecular Function	binding	iron ion transpor	iron ion hc	WHMg 1739	4.47344E-95	WHGA0709	7.98921E-91	WHON0085	2.56E-92

Fig. 2. Sequence Distribution: Molecular Functions. Categories had to have at least 50 members. Represents EST's from three cDNA libraries, Adults, 5th instar, and Midgut. *Homalodisca vitripennis*, (Blast2GO analysis). Highest Categories in descending order: Ribosome structure= 296, Calcium ion binding=218, ATPase activity=154, Actin binding= 134, Microfilament motor activity=180, Endopeptidase activity=114, Oxidoreductase activity= 109, Protein Kinase activity= 107.

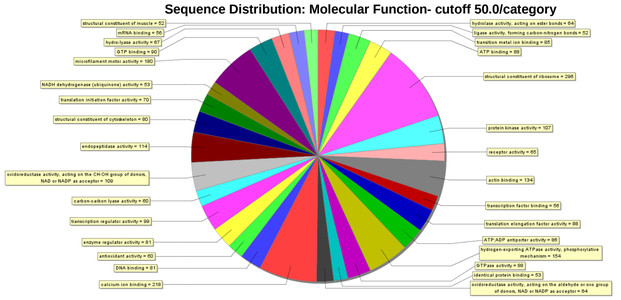
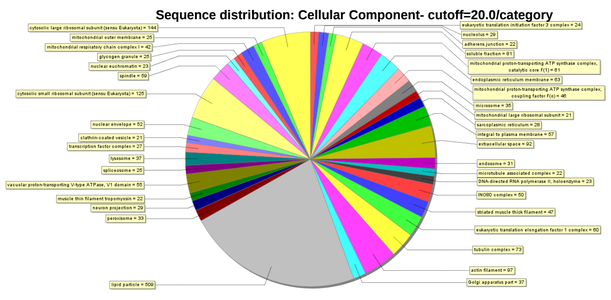


Fig. 3. Sequence Distribution: Cellular Component. Categories had to have at least 20 members. Represents EST's from three cDNA libraries, Adults, 5th instar, and Midgut. *Homalodisca vitripennis,* (Blast2GO analysis).



Hunter WB, Mizell III, RF, Tipping C, Dang PM, Hunnicutt LE. 2005. Adult sharpshooter leafhopper *Oncometopia nigricans*, (Hemiptera: Cicadellidae), DR755012-DR759538. National Center for Biotechnology Information, NCBI.

Hunter WB, Hunnicutt LE, Wistrom, CM, Purcell AH. 2006. Proteins expressed in the Blue-green sharpshooter, *Graphocephala atropunctata* (Hemiptera: Cicadellidae). 44 Proteins, DQ445499-DQ445542. NCBI.

Hunter WB, Hunnicutt LE, Wistrom CM, Purcell AH. 2007. Gene expression in adult blue-green sharpshooters, *Graphocephala atropunctata* (Signoret) (Hemiptera: Cicadellidae). EH655849-EH662332. NCBI.