Gene expression in the Asian Citrus Psyllid: Vector of citrus greening (Hemiptera: Psyllidae)

Wayne Hunter, CS. Katsar, CL. McKenzie, RG. Shatters, Jr,

AA. Weathersbee,III, D. Hall

USDA, ARS, U.S. Horticultural Research Lab, 2001 South Rock Rd, Ft. Pierce, FL, USA 34945. Whunter@ushrl.ars.usda.gov

Agricultural Research Service
United States Department of Agriculture

07-24-2006

Introduction:

To better understand the biology of the *D. citri* and particularly those genes functioning in: feeding, disease transmission, and the development of insecticide resistance, we undertook a large-scale 5′ end sequencing project of cDNA clones from adult psyllids. Similar large-scale expressed sequence tag (EST) sequencing projects from other insects have provided the vehicle for answering important biological questions. Although there is a growing database in GenBank of ESTs from other insects, most of the available gene sequences are from the dipterans, *Drosophila* species and *Anopheles gambiae*; the lepidopteran - *Bombyx mori*; the hymenopteran *Apis mellifera*; and the coleopteran *Tribolium castaneum*. Relatively few genes have been specifically isolated from psyllids. In this paper, we describe the first public data set of ESTs from the psyllid *D. citri*. Over 5,906 cDNA clones were sequenced, resulting in 4,445 high-quality *D. citri* ESTs. Sequence alignment of the cDNAs resulted in 2,123 total assembled sequences, including both contiguous sequences and singlets. The putative protein transcript of each assembled sequence was annotated based on the biochemical function of matching gene sequences using BLASTX, TBLASTX, and BLASTN analyses, GenBank, nr ESTdb. The subsequent unigene set produced 517 sequences which had significant identities with homologous genes in the GenBank's database. The remaining 63% of the cDNA's showed 'no significant match' in either the non-redundant protein or nucleic acid databases, demonstrating that *de novo* EST sequencing projects can still provide new information to the scientific community. The *D. citri* gene expression data set advances current research efforts in the identification of genes and physiological processes of psyllids, of which a number of species are of economic importance. Knowledge of these genes and proteins will enable functional genomic studies for the development of new management strategies against psyllids and other hemipterans.

Importance of Expressed Sequence tags (ESTs):

By identifying these genes and the proteins they make, researchers will be able to develop the molecular tools needed to conduct functional genomic studies. By examining the genes that are being actively expressed under set conditions, the components of important pathways can be identified. The development and use of cDNA expression libraries have been proven to open new avenues for researchers to address difficult biological problems. The data from the Psyllid library will further accelerate research on the Asian Citrus Psyllid and Citrus Greening. The research that grows out of this and other research efforts will lay the groundwork to develop environmentally friendly solutions to address important scientific questions related to the transmission of diseases by psyllids and other insect pests.

Figure 1. Proportion of assembled sequences from the Asian Citrus Psyllid EST dataset **sorted by using Gene Ontology**. Of the 2,123 unigene set, 1,123 (53%) of the cDNA's had 'No Significant Homology' to any known protein or nucleotides in the GenBank non-redundant database, BLASTX analysis matches had an *E*-value of ≤-5. (NCBI , National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/).

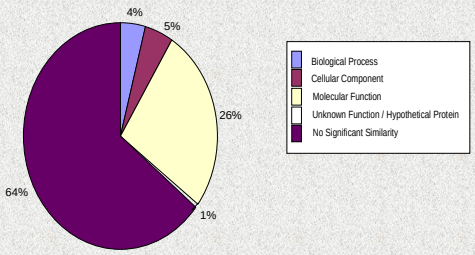

Table 1. The twenty most frequently represented sequences in the Asian Citrus Psyllid, *Diaphorina citri*, cDNA library.

Sequence ID	# Sequences	Description[a]	Best Match[b]	E.C.#	E value
WHDc 0014	52	Cytochrome c oxidase subunit I	Q69HD6	1.9.3.1	1.5035E-150
WHDc 0030	45	ATP synthase F0 subunit 6	Q69HD3		2.70664E-32
WHDc 0080	41	Cytochrome c oxidase subunit III	Q69HD2	1.9.3.1	5.07862E-46
WHDc 0021	31	Putative ferritin GF2	Q6PPI2		7.47398E-63
WHDc 0035	25	Cytochrome b	Q69HC6		2.7731E-119
WHDc 0023	23	Cytochrome c oxidase subunit II	Q69HD5	1.9.3.1	9.85564E-77
WHDc 0065	18	myosin light chain	Q7PUV3		8.0000E-48
WHDc 0076	18	ATP synthase F0 subunit 6	Q69HD3		6.18037E-06
WHDc 0086	15	NADH dehydrogenase subunit 4	Q69HC9	1.6.5.3	2.8032E-99
WHDc 0592	14	NADH dehydrogenase subunit 2	Q69HC4	1.6.5.3	1.86387E-36
WHDc 0125	14	Ejaculatory bulb specific protein III precursor	Q9W1C9		7.48136E-32
WHDc 0187	14	Translationally controlled tumor protein	Q75VN3		1.85962E-80
WHDc 0020	10	NADH dehydrogenase subunit 5	Q69HD0	1.6.5.3	2.06965E-80
WHDc 1078	8	Ribosomal protein L32	Q6F449	3.6.5.3	2.37581E-49
WHDc 0049	8	S3e ribosomal protein	Q6EV05	3.6.5.3	2.6657E-100
WHDc 0175	8	NADH dehydrogenase subunit 1	Q68RL0		7.91356E-87
WHDc 0177	8	Thioredoxin-like protein	Q9U515		2.38394E-19
WHDc 0441	8	Ribosomal protein S27	Q962Q3	3.6.5.3	3.31106E-41
WHDc 0472	8	40S ribosomal protein S9	Q9VT06		5.0000e-86
WHDc 0502	8	Abnormal wing disc-like protein	Q8MUR5	2.7.4.6	9.19534E-73

aDescription derived from GenBank nr database Blastx search.
bEST match indicates the gi number of the most similar annotated cDNA in GenBank nr database, BLASTX analysis matches had an *E*-value of ≤-5.

Table 2. Enzymatic Classification of Psyllid Sequences.
Classification is hierarchical. All functional assignments of *Psyllid sequences* described here are "inferred from electronic evidence" using the top 5 BLASTX hits with an *E*-value of ≤-5 generated from NCBI's nr database. The definition term associated with each sequence was defined according to The International Union of Biochemistry and Molecular Biology's Enzyme classification system.

EC #	Class	Subclass	# sequences
1.1	Oxidoreductase	acting on the CH-OH group of donors	7
1.11	Oxidoreductase	Peroxidases	2
1.14	Oxidoreductase	Oxygen	4
1.3	Oxidoreductase	Acting on the CH-CH Group of Donors	3
1.6	Oxidoreductase	Acting on NADH or NADPH	11
1.7	Oxidoreductase	Acting on other Nitrogenous Compounds as Donors	1
1.9	Oxidoreductase	Acting on a Heme Group of Donors	8
2.1	Transferase	Transferring One-Carbon Groups	2
2.3	Transferase	Acyltransferases	1
2.4	Transferase	Glycosyltransferases	4
2.5	Transferase	Transferring alkyl or aryl groups, other than methyl groups	1
2.6	Transferase	Transferring nitrogenous groups	1
2.7	Transferase	Transferring phosphorus-containing groups	19
3.1	Hydrolases	Acting on ester bonds	11
3.2	Hydrolases	Glycosidases	11
3.3	Hydrolases	Acting on ether bonds	2
3.4	Hydrolases	Acting on peptide bonds (Peptidases)	13
3.5	Hydrolases	Acting on Carbon-Nitrogen Bonds, other than Peptide Bonds	1
3.6	Hydrolases	Acting on Acid Anhydrides	127
4.1	Lyases	Carbon-Carbon Lyases	4
4.2	Lyases	Carbon-Oxygen Lyases	4
4.3	Lyases	Carbon-Nitrogen Lyases	1
4.4	Lyases	Carbon-Sulfur Lyases	1
4.6	Lyases	Phosphorus-Oxygen Lyases	1
5.1	Isomerases	Racemases and Epimerases	1
5.2	Isomerases	*cis-trans*-Isomerases	2
5.3	Isomerases	Intramolecular Oxidoreductases	4
6.1	Ligases	Forming Carbon-Oxygen Bonds	4
6.3	Ligases	Forming Carbon-Nitrogen Bonds	6
6.4	Ligases	Forming Carbon-Carbon Bonds	1
		Total	258

Summary:

As each new set of Psyllid cDNA libraries are completed, each will be posted to GenBank for public access. The Scientific Impact: The USDA, ARS, U.S. Horticultural Research Laboratory effort is developing national collaborations to establish expression libraries from important insect pest and vectors of plant diseases (Leafhoppers, Aphids, Whiteflies, Psyllids, Mealybugs). The ability to identify and studying insect genes at the molecular level leads to analyses of functional genomics which will yield significant advances, including discoveries about psyllid biology, development, pathology, and on pathogen interactions. This new knowledge enables researchers to examine new ways to manage these economically important insect pests, and to reduce the spread of plant diseases.