Transcriptome analysis of the salivary glands of potato leafhopper, *Empoasca fabae*

Bridget DeLay, Praveen Mamidala, Asela Wijeratne, Saranga Wijeratne, Omprakash Mittapalli, Jian Wang, William Lamp

Abstract

The potato leafhopper, *Empoasca fabae*, is a pest of economic crops in the United States and Canada, where it causes damage known as hopperburn. Saliva, along with mechanical injury, leads to decreases in gas exchange rates, stunning and chlorosis. Although *E. fabae* saliva is known to induce plant responses, little knowledge exists of saliva composition at the molecular level. We subjected the salivary glands of *E. fabae* to Roche 454-pyrosequencing which resulted in significant number (30,893) of expressed sequence tags including 2805 contigs and 28,088 singletons. A high number of sequences (78%) showed similarity to other insect species in GenBank, including *Trypanosoma brucei* and *Acrythosiphon pisum*. KEGG analysis predicted the presence of pathways for purine and thiamine metabolism, biosynthesis of secondary metabolites, drug metabolism, and lysine degradation. Pfam analysis showed a high number of cellulase and carboxylesterase protein domains. Expression analysis of candidate genes (alpha amylase, lipase, pectin lyase, etc.) among different tissues revealed tissue-specific expression of digestive enzymes in *E. fabae*. This is the first study to characterize the sialotranscriptome of *E. fabae* and the first for any species in the family of Cicadellidae. Due to the status of these insects as economic pests, knowledge of which genes are active in the salivary glands is important for understanding their impact on host plants.

1. Introduction

Insect saliva plays an important role in the ingestion of food and in the interaction between an insect and its host. Labial salivary glands are the most common type of insect salivary gland, followed by hypopharyngeal and mandibular salivary glands (Poiani and Da Cruz-Landim, 2010). The paired secretory lobes of the labial salivary glands of hemipterans are located in the head, and are connected to the labium by excretory ducts (Tsai and Perrier, 1996). Hemipterans produce two different types of saliva: sheath saliva and watery saliva (Miles, 1998). Sheath saliva hardens upon contact with air, and helps to stabilize the mouthparts of the insect as it probes into a host plant and prevents plant wound response to components in the watery saliva (Miles, 1964; Will and van Bel, 2006). Watery saliva contains a mixture of amino acids, proteins and digestive enzymes, and is thought to lubricate the stylets inside of the salivary sheath, aid in the digestion of plant material and prevent plant wound response (Carolan et al., 2009; Harmel et al., 2008; De Vos and Jander, 2009).

Leafhoppers are sap-feeding insects in the hemipteran family Cicadellidae and have long been recognized as significant pests of agricultural crops (Nault and Rodriguez, 1985; Poos, 1943). Species are usually classified as either sheath feeders or cell rupture feeders (Miles, 1972). Sheath feeders secrete saliva that hardens into a sheath surrounding their stylets as they feed from a single phloem cell (Hollebone et al., 1966). Cell rupture feeders lacerate multiple cells with their stylets and ingest the phloem that leaks out of the wounded cells while secreting watery saliva to prevent plant wound response (Backus and Hunter, 1989). Leafhopper feeding can cause a generalized plant response known as hopperburn, which is characterized by leaf chlorosis, stunted growth and reduced yield (Backus et al., 2005). In spite of their agricultural importance, sialotranscriptomes are unknown for any species in the family Cicadellidae.

The potato leafhopper, *Empoasca fabae* (Harris), feeds and reproduces on over 220 species of plants in 26 families and is an especially important agricultural pest of legumes (Lamp et al., 1994). On alfalfa, *Medicago sativa*, *E. fabae* adults feed on the vascular tissues of leaves and stems of plants, while nymphs preferentially...
feed on leaves (Lamp et al., 2004). *E. fabae* are dynamic feeders, using both cell rupture and modified sheath feeding methods (Backus et al., 2005). As a cell rupture feeder, *E. fabae* mechanically injures phloem and parenchyma cells while injecting saliva to feed on the leaking cell contents. In addition, *E. fabae* can feed directly from vascular tissue, making a short-lived partial salivary sheath to stabilize its stylets during feeding (Zhou and Backus, 1999). Within 24 h of feeding on alfalfa stems, plants display reduced rates of photosynthesis and transpiration, disruption in the transport of photoassimilates, and accumulation of starch in the leaves (Pirone et al., 2005; Lamp et al., 2004; Nielsen et al., 1990). Subsequent to feeding, the generalized wound response caused by the feeding injury leads to the production of hopperburn associated characteristics in alfalfa (Pirone et al., 2005). While past studies have focused on mechanical injury by the mouthparts, saliva plays a role in the response (Ecale and Backus, 1995; DeLay and Lamp, unpublished data), yet the constituents of saliva involved in the plant response are unknown.

To date much of the sialotranscriptomes (salivary gland transcriptomes) have been deciphered in blood feeding insects viz., *Ixodes scapularis* (Nielsen et al., 1990; Valenzuela et al., 2002; Francischetti et al., 2005), *Anopheles gambiae* (Ribeiro et al., 2006; Arca et al., 2006; Calvo et al., 2006; Neira et al., 2009), *Dermacentor andersoni* (Das et al., 2010), *Triatoma brasiliensis* (Alarcon-Chaidez et al., 2007), *Ixodes ricinus* (Santos et al., 2007), *Triatoma infestans* (Chmela et al., 2008), *Glossina morsitans* (Assumpcao et al., 2008) and *Amblyomma variegatum* (Alves-silva et al., 2010), with little studies on phytophagous insects (Ribeiro et al., 2011). Roche® 454 pyrosequencing has in the recent past revolutionized functional genomic studies in non-model organisms, particularly in insects wherein little to no genetic information is available (Francischetti et al., 2007; Morozova and Marra, 2008; Margulies et al., 2005; Vera et al., 2008; Pauchet et al., 2009; Mittapalli et al., 2010; Bai et al., 2011). The developed transcriptomic database can subsequently be used as a reference for future functional studies like RNA seq and to mine for candidate targets for RNAi experiments. The primary goal of this study is to develop a sialotranscriptomic database for *E. fabae* (Expressed Sequence Tags, ESTs of saliva) using 454 pyrosequencing. Results obtained from this study provide insight into potential salivary components that play significant role(s) in the host response subsequent to *E. fabae* feeding injury.

2. Materials and methods

2.1. Insect samples

Potato leafhoppers were collected on alfalfa from the Western Maryland Research and Education Center in Keedysville, Maryland on the morning of June 30, 2009. Leafhoppers were caught with sweep nets, and individual adult *E. fabae* were aspirated into cages containing excised alfalfa sprouts. The cages were then brought back to the laboratory for salivary gland dissection in the afternoon.

2.2. Dissection of leafhopper salivary glands

Leafhopper adults were anesthetized by carbon dioxide and placed in a Petri dish that was kept cold on ice. Salivary glands (Fig. 1) were dissected in a microplate-well with a drop of the sterilized 1× Phosphate Buffered Saline (1× PBS) solution using fine-tipped forceps. This was accomplished by first pulling the head from the thorax with forceps, then carefully removing the salivary glands that emerged from the distal region of the severed head. A total of 200 salivary glands were dissected and directly dipped into 200 μl Trizol solution (Invitrogen, CA) for RNA preparation.

2.3. RNA isolation and 454 pyrosequencing

Two hundred pairs of salivary glands of *E. fabae* were used for total RNA isolation using TRIzol® (Invitrogen). QC of the total RNA was assessed with RNA 6000 Nanochip. The library preparation and pyrosequencing was done at Purdue Genomics Core Facility, West Lafayette, IN as per Mittapalli et al., 2010. In brief, a SMART cDNA library construction kit (Clonetech, Mountain View, CA) was used following manufacturer’s instructions followed by shearing and nebulization of cDNA with subsequent extraction. The isolated DNA was blunt ended, ligated to adapters and immobilized on beads. Single stranded DNA was later isolated from these beads and subjected to QC using RNA 6000 (Agilent Technologies). The emPCR reactions were performed to amplify a single template onto a single sequencing bead. One-quarter of a pico-titer plate was sequenced at the Purdue Genomics Core Facility (West Lafayette, IN) using the GS FLX Titanium chemistry (Roche Diagnostics, Indianapolis, IN).

2.4. Bioinformatic data analysis

The 454 transcriptomic reads were assembled (after removal of adapters and low quality regions) using Newbler program (Roche) by the Purdue University Genomics Core facility. Initial annotation of assembled sequences, namely isoforms and contigs and the sequences that were not assembled into contigs or isoforms was done using Blast2Go software suite (Conesa et al., 2005; Gotz et al., 2008). Briefly, sequences were searched against GenBank non-redundant database with using BLASTx algorithm (Altschul et al., 1990) with *E* value cutoff of 10^-6_. The blast results were mapped to gene ontology terms and annotation was carried out using default annotation parameters in the Blast2Go software suit (Conesa et al., 2005; Gotz et al., 2008). For further functional annotation, the Kyoto Encyclopedia of Genes and Genomes (KEGG) mapping was carried out in Blast2Go. To obtain, species distribution, top blast hits were exported from the Blast2Go project and a modified
version of a python script available in NCBI taxonomy tree made easy repository (https://github.com/jhcepas.ncbi_taxonomy) was used to retrieve lineage from NCBI taxonomy database. For comparative genomics the sequences were compared to the protein sequences of *Drosophila melanogaster*, *Tribolium castaneum* and *Acrythosiphon pism* using BLASTx program with expect value of 10^-3. Protein domains were identified using the HMMER v3 program (Eddy, 1998) by importing the Blast2Go project generated by the Purdue Genomics Facility to a local server.

### 2.5. Quantitative real-time PCR

Quantitative real-time PCR was performed on total RNA extracted from the salivary glands, midguts and hind femurs of *E. fabae* collected on *M. sativa* at the University of Maryland’s Western Maryland Education and Research Center. Tissues were dissected from individual leafhoppers and pooled by tissue type before total RNA extraction using a Qiagen RNeasy mini RNA extraction kit. cDNA was synthesized with a Transcriptor First Strand cDNA Synthesis Kit (Roche), following the provided manufacturer’s protocol. This cDNA was used as the template for the qRT-PCR reactions, which were performed using a LightCycler 480 SYBR Green I Master Kit (Roche) on a LightCycler 480 qRT-PCR system (Roche), with the cycling parameters of 95 °C for 5 min followed by 45 cycles of 95 °C for 10 s, 60 °C for 10 s and 72 °C for 10 s. Primers used in the study (Table S1) were designed using IDT SciTools RealTime PCR software (Integrated DNA Technologies).

Relative expression analysis was performed with *E. fabae* specific ribosomal protein 15 as the internal control which was demonstrated to be a suitable internal control in other insects (Mittapalli et al., 2010b). Gene expression levels for each of the three tissue types were determined with the Relative Standard Curve method (Mittapalli et al., 2010), using threshold cycle (Ct) values, as detailed in the LightCycler 480 Instrument Operations Manual Version 1.0 (Roche). Relative expression values (REV) were calculated by dividing the quantity of mRNA detected in the target sample by the quantity of mRNA detected in the ribosomal protein sample. Analysis of Variance (ANOVA) was used to analyze the REVs of each target gene using the PROC MIXED procedure in SAS (SAS Institute Inc., 2011). For each sample, two biological replicates and two technical replicates were used for the statistical analysis, with the biological replicates used as a random effect in the model. Relative fold changes in tissue gene expression were calculated by setting the tissue type with the lowest REV for the gene (calibrator) at 1 ×. The calculated standard error shows the variance in the two biological replicates, each of which contained two technical replicates.

### 2.6. Data deposition

The Roche 454 reads of *E. fabae* were deposited to the NCBI Sequence Read Archive under the accession number SRA037848.1 and assembled isogroups which are above 200 nucleotides were deposited in Transcriptome Shotgun Assembly (TSA) under accession number, 175548 TSA.

### 3. Results and discussion

#### 3.1. 454 Pyrosequencing

The 454 pyrosequencing of *E. fabae* sialotranscriptome resulted in 255,491 transcriptomic reads (102,069,574 bp) from which 86.51% and 85.45% were aligned respectively with an inferred read error of 1.77%. These reads were further assembled (after removal of adapters and low quality regions) using Newbler program version 2.5 (Roche). The post assembly of the sequences resulted in 30,893 high quality ESTs including 2805 contigs and 28,088 singletons. The contigs ranged between 60 and 6199 bp with an average length of 1093 bp and totaled to 3031,962 bp. Singletons ranged from 50 to 919 bp with an average length of 396 bp totaled to 11,133,319 bp (Fig. S1; Table S2). To date, no ESTs are available for *E. fabae* in GenBank and to our knowledge this is the first comprehensive study on sialotranscriptome for this species. A sequence similarity search was done using BLASTx algorithm, which revealed high similarity of *E. fabae* (for 11,322 sequences with taxonomy information) with other insect species (83%) (Fig. 2). However, a portion of the sequences showed similarity to other eukaryotes (12%), bacteria (4%), fungi (0.4%) and virus (0.3%) as observed in other insect transcriptomic studies (Fig. 2; Table S3) (Pauchet et al., 2009; Mittapalli et al., 2010).

#### 3.2. Comparative analysis

The derived sequences of *E. fabae* were compared to the proteins of model insect species including the fruit fly (*D. melanogaster* Meigen) of Diptera, red flour beetle (*T. castaneum* Hebh) of Coleoptera and pea aphid (*A. pism* Harris) of Homiptera whose genomes are available (Karatolos et al., 2011; Adams et al., 2000; Richards et al., 2008). The majority of the *E. fabae* sequences showed similarity with those of *A. pism* (33.24%) followed by *T. castaneum* (33.12%) and *D. melanogaster* (30.75%) (Fig. 3). Similar observations were reported in a recent study on the brown plant-hopper (*Nilaparvata lugens* Stal) wherein *N. lugens* sequences shared a higher similarity with *T. castaneum* than with *A. pism* (Xue et al., 2010). At the current time, it is difficult to explain the similarity of *E. fabae* with *T. castaneum*, but future genomic studies may reveal the evolutionary relationship among these two species. A high percentage (62.58%) of sequences was unique to *E. fabae*, i.e., no significant similarity with the sequences of GenBank non-redundant database. This might be due to novel genes of *E. fabae* or untranscribed regions of 5′ and 3′ transcripts and/or assembly errors as observed in other transcriptomic studies (Pauchet et al., 2009; Mittapalli et al., 2010). We have also compared the *E. fabae* salivary transcripts with recent transcriptomic data of *N. lugens*, which resulted in 11.76% similarity (Xue et al., 2010). The lesser percentage of similarity with *N. lugens* might be due to the comparison with transcriptome data, instead of a fully sequenced genome (Table S4). Comparative genomics using the ESTs obtained from these studies will potentially reveal putative function of novel genes (Mittapalli et al., 2010). Moreover, ESTs retrieved from such studies could serve as a useful resource for rapid identification of transcripts involved in a particular biological process (Xue et al., 2010).
3.3. Gene ontology

Gene Ontology (GO) terms were assigned to a total of 18,027 transcriptomic sequences based upon their homology to GenBank protein sequences (Table S5). The GO terms were grouped into three main divisions: biological processes, molecular function and cellular components. The majority of the biological processes (Fig. 4) in the sialotranscriptome represented metabolic processes (4487 sequences) and cellular process (3597 sequences). These results indicate that the cells in the salivary glands are metabolically active, which correlates well with the biological function of the tissue of interest. In addition, the sialotranscriptome contained many sequences involved in biological regulation (1836 sequences), localization (1343) stimulus response (736 sequences). Therefore,
it is possible that the sequences coding for these physiological processes are involved in variation in the saliva composition (Guo et al., 2010). However, further functional studies need to be performed to validate these hypotheses.

Among the molecular function GO terms identified the majority were predicted to be involved in binding (4687 sequences) and catalytic (4572 sequences) functions (Fig. 5). This number includes sequences annotated to be involved in protein, nucleic acid, ion, cofactor and enzyme binding. Hydrolase activities were also identified, wherein 656 *E. fabae* sequences showed homology to hydrolase sequences deposited in GenBank. Hydrolases are enzymes that catalyze chemical bond formation, and have been described in the saliva of the Russian wheat aphid (*Diuraphis noxia*), bird cherry-oat aphid (*Rhopalosiphum padi*), and the mosquito *Culex quinquefasciatus* (Bede et al., 2006; Ni et al., 2000).

The majority of cellular component (Fig. S2) GO terms showed homology with cell (4447) and organelle (2754) sequences in GenBank. In addition, 1711 sequences appeared to be involved in macromolecular complex, and 422 sequences had homology to sequences of membrane enclosed lumen.

### 3.4. KEGG analysis

The KEGG pathways predicted in the sialotranscriptome of *E. fabae* were purine metabolism (215), thiamine metabolism (36), drug metabolism (48), and lysine degradation (43) (Table 1; Table S6). Besides these, we also recovered transcripts involved in pathways of fatty acid and amino acid metabolism (63), starch and sucrose metabolism (50) and 20 sequences involved in metabolism of xenobiotics by cytochrome P450.

### 3.5. Protein domains

We identified 676 distinct domains in 3298 sialome transcripts of *E. fabae* using HMMER3 software (Table S7). Among the top Pfam domains, a cellulase (273) domain was the highest in occurrence (Table 2). Phytophagous insects feeding on plant biomass degrade cellulose to glucose and utilize the latter as an energy source. However, the mechanism of carbohydrate metabolism (breakdown) in insects is poorly understood (Ribeiro and Francischetti, 2001). Initially, it was thought that the source of cellulases in insects was from their endosymbionts such as bacteria, fungi and protozoa (Watanabe and Tokuda, 2010). However, Watanabe et al., 1997 described the first insect-origin cellulase gene from *Reticulitermes speratus* RsEG, which encodes a endo-β-1,4-glucanase. There is further evidence that insects from Dictyoptera, Orthoptera and Coleoptera produce their own cellulases in the foregut, midgut or salivary glands (Martin, 1983; Scharf et al., 2003; Sugimura et al., 2003). Insect cellulases which possess high relative activity at alkaline pH are believed to have potential applications in the biofuel industry (Martin, 1983; Wei et al., 2006).

Next to cellulase domains, we found a high number of carboxylesterases (COE) domains (=268). COEs are multigene families and are widely distributed among prokaryotes and eukaryotes. These
are primarily involved in detoxification, development and neurogenesis (Willis et al., 2011; Satoh and Hosokawa, 1998; Marshall et al., 2003; Ranson et al., 2002; Bornscheuer, 2002). The role of COEs in detoxification is well studied in insects and is thought to play important roles in defense against plant allelochemicals and various synthetic chemicals within the context of metabolic resistance (Mittapalli et al., 2005; Small and Hemingway, 2000). However, the function of salivary COEs in phytophagous insects remains elusive.

A high number of lipase domains (84) were predicted in the current study. Lipases are multifunctional proteins involved in lipid acquisition, storage and mobilization besides their involvement in reproduction and development (Mamidala et al., 2011a; Horner et al., 2009). Lipases perform a diverse and unique array of functions in insects and are well documented in several blood feeding insects, however, these are poorly understood in phytophagous insects (Arrese and Soulages, 2010; Tunaz and Stanley, 2004; Anderson et al., 2006). In a recent study of *Mayetiotla destructor*, a lipase salivary transcript (MesL1) was shown to be involved in extra-oral digestion and host cell permeability (Anderson et al., 2006). The other Pfam domains of digestive enzymes in the current study include alpha amyrase (56), glycosyl hydrolase (48) and trypsin (44) domains. Alpha amyrase, glycosyl hydrolase and trypsin are known for their digestive role in insects and in recent studies have been reported to play an important role in salivary secretions of insects (Shukle et al., 2009; Hosseininaveh et al., 2009). Functional characterization of these digestive enzymes of *E. fabae* may shed light on their mode of extra-oral digestion.

Besides the above-mentioned digestive associated transcripts, we also found a high number of cytochrome P450 (39) and glutathione-S-transferase (GST) domains (22), which are often associated with detoxification of plant allelochemicals, insecticides and endogenous metabolites (Scharf et al., 2010; Small and Hemingway, 2000). The role of cytochrome P450s and GSTs in particular are well documented in insect midgut and fat body tissues (Pauchet et al., 2009). However, little is known on the role of cytochrome P450s in insect saliva. Other protein domains that were predicted from the sialotranscriptome of *E. fabae* were Zinc finger C4H2 (77), Protein kinase (68), RNA recognition motif (54), Hemocyanin (33), Ras family (28) and Proteasome (27) domains (Table 2).

### Table 2

<table>
<thead>
<tr>
<th>Pfam accession</th>
<th>Pfam domain</th>
<th>Pfam domain description</th>
<th># occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF00135.21</td>
<td>COE</td>
<td>Carboxylesterases</td>
<td>268</td>
</tr>
<tr>
<td>PF00150.11</td>
<td>Cellulase</td>
<td>Cellulase</td>
<td>273</td>
</tr>
<tr>
<td>PF00096.19</td>
<td>Zinc finger</td>
<td>Zinc finger C4H2</td>
<td>77</td>
</tr>
<tr>
<td>PF00151.12</td>
<td>Lipase</td>
<td>Lipase</td>
<td>80</td>
</tr>
<tr>
<td>PF00128.17</td>
<td>Alpha amyrase</td>
<td>Alpha amyrase</td>
<td>56</td>
</tr>
<tr>
<td>PF00794.21</td>
<td>Glyco hydro, 18</td>
<td>Glycosyl hydrolase</td>
<td>48</td>
</tr>
<tr>
<td>PF00069.18</td>
<td>Protein kinase</td>
<td>Protein kinase</td>
<td>68</td>
</tr>
<tr>
<td>PF00076.15</td>
<td>RRM,1</td>
<td>RNA recognition motif (a.k.a RPM, RBD, or RNP domain)</td>
<td>54</td>
</tr>
<tr>
<td>PF00089.19</td>
<td>Trypsin</td>
<td>Trypsin</td>
<td>44</td>
</tr>
<tr>
<td>PF01607.17</td>
<td>CBM,14</td>
<td>CBM,14</td>
<td>47</td>
</tr>
<tr>
<td>PF00067.15</td>
<td>p450</td>
<td>Cytochrome P450</td>
<td>39</td>
</tr>
<tr>
<td>PF03723.7</td>
<td>Hemocyanin</td>
<td>Hemocyanin</td>
<td>33</td>
</tr>
<tr>
<td>PF00071.15</td>
<td>Ras</td>
<td>Ras family</td>
<td>28</td>
</tr>
<tr>
<td>PF00227.19</td>
<td>Proteasome</td>
<td>Proteasome</td>
<td>27</td>
</tr>
<tr>
<td>PF02798.13</td>
<td>GST</td>
<td>Glutathione-S-transferase</td>
<td>22</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Candidate genes</th>
<th>Number of occurrences in salivary glands</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endo-beta-glucanase</td>
<td>58</td>
</tr>
<tr>
<td>Alpha-amyrase</td>
<td>36</td>
</tr>
<tr>
<td>Chitinase, acidic mammalian</td>
<td>14</td>
</tr>
<tr>
<td>Chitinase 10</td>
<td>4</td>
</tr>
<tr>
<td>Chitinase 1</td>
<td>2</td>
</tr>
<tr>
<td>Lipase</td>
<td>7</td>
</tr>
<tr>
<td>Pancreatic lipase</td>
<td>7</td>
</tr>
<tr>
<td>Carbonic anhydrase</td>
<td>6</td>
</tr>
<tr>
<td>Pectin lyase</td>
<td>3</td>
</tr>
<tr>
<td>Aldehyde dehydrogenase</td>
<td>2</td>
</tr>
<tr>
<td>Beta-glucosidase</td>
<td>2</td>
</tr>
<tr>
<td>Protein phosphatase</td>
<td>2</td>
</tr>
<tr>
<td>Superoxide dismutase</td>
<td>2</td>
</tr>
<tr>
<td>Alpha-mannosidase</td>
<td>1</td>
</tr>
<tr>
<td>Catalase</td>
<td>1</td>
</tr>
</tbody>
</table>

### 3.6. Genes of interest

We screened the *E. fabae* sialotranscriptomic database for known insect saliva gene products. The ESTs encoding for endo-beta-glucanase (EBG), alpha-amyrase, chitinase, lipase etc. were found to be predominant among top ten potential salivary gland specific proteins (Table 3). Among the candidate genes listed in Table 3, we found high occurrence of EBGs (58) putatively involved in the break-down of cellulose, a major component of plant cell walls. The salivary gland specific expression of these EBGs is well documented in termites (*R. speratus*), the blister beetle (*Mylabris pustulata*), and various species of aphids (*Motoyama and Dauterman, 1980; Slaytor, 1992; Watanabe et al., 1997*). EBGs are also reported in the foregut and midgut of termites and cockroaches, where they are involved in breaking down ingested plant matter (*Motoyama and Dauterman, 1980*). Further, the EBGs of *Lygus* are well demonstrated to degrade the cell walls of alfalfa (*M. sativa*), damage which was previously thought to be caused by the insect’s mouthparts (*Will and van Bel, 2006*). Future studies on the function of *E. fabae* saliva may shed light on the role of EBGs and other cell wall degrading enzymes in the development of hopperburn associated symptoms (*Harmel et al., 2008*).
the role that alpha-amylase plays in the saliva of *E. fabae* is likely to be similar. The high EST occurrence (36) and higher mRNA levels of alpha-amylase better explain their putative role in digesting ingested plant compounds.

A high number of ESTs coding for chitinases (20) in *E. fabae* sialotranscriptome is intriguing (Table 3). Insects are known to produce chitinases for chitin degradation during molting (Feng et al., 1996). However, chitinases have also been reported in the saliva of ants that feed on fungi, where the enzyme is necessary to degrade the fungal cell wall (Merzendorfer and Zimoch, 2003). The occurrence of these chitinases in *E. fabae* is suggests that their role is to degrade polysaccharides found in plant cell walls, allowing the insect to feed more easily.

Among the identified lipases (14) in the current EST database, we detected transcript levels for one candidate lipase (Isotig 00445) which was specifically expressed in the salivary glands of *E. fabae* (Fig. 6A). Lipases have been reported in the salivary glands of multiple species of insects, including the Hessian fly (*M. destructor*) (Shukle et al., 2009), the mosquito (*Anopheles stephensi*) (Valenzuela et al., 2003), the froghopper (*Aeneolamia varia saccharind*) (Haygle, 1966), and the milkweed bug (*Oncopeltus fasciatus*) (Francischetti et al., 2007). In these insects, lipases are reported to break down lipids of the host cells, facilitating ingestion and digestion. Therefore, it is likely that the lipases found in the sialotranscriptome of *E. fabae* are also used to break down cell membranes in order to facilitate feeding. In the case of the froghopper (*A. varia saccharind*), salivary lipases cause the red streaking of leaf tissue associated with blight symptoms in sugar-cane (Haygle, 1966). Other studies have shown that phospholipases in saliva can induce plant response cascades that can cause symptoms of feeding damage (Munnik et al., 1995; Wang, 1999). These studies suggest that lipases are capable of causing wound response cascades that are similar to the hopperburn caused by *E. fabae* feeding. Therefore, the role that salivary lipases play in the plant response to *E. fabae* should be investigated further.

In addition to the above putative salivary proteins, we also examined mRNA transcript levels for alkaline phosphatase, pectin lyase, laccase and Wolbachia surface membrane protein putatively involved in digestion. Among these, alkaline phosphatase expression was detected only in the salivary glands (Fig. 5S), whereas the transcript levels of pectin lyase were found to be higher in the salivary glands (20×) followed by the midgut (16×) (Fig. 6B). While alkaline phosphatases have important functions in the salivary glands, development, cuticle sclerotization, and neural and renal function of insects, the pectin lyases are major insect salivary proteins involved in pectin (the major polysaccharide in plant cell walls) degradation and modification, facilitating digestion of ingested plant material (Harper and Armstrong, 1972; Chang et al., 1993; Yang et al., 2000; Funk, 2001). The peak mRNA transcript levels of laccase in midgut (650×) and salivary glands (450×) (Fig. 6C) suggests their induced expression for rapid oxidation of phytotoxic compounds, which is evident with other insect species (Hattori et al., 2005). The expression levels of Wolbachia membrane protein were exclusively seen in the midgut (Fig. S3). Wolbachia is an insect-associated bacterium found in the midgut of many other insect species, and can be transferred horizontally via saliva injection into an insect's host plant (Sintupachee et al., 2006).

This study focused on the sialotranscriptome of adult *E. fabae* that have fed on alfalfa. Analysis of gene expression in the salivary glands of *E. fabae* nymphs or individuals that have fed on different food sources may show different expression patterns. Plant-feeding insects are known to secrete different salivary components depending on the developmental stage of the insect (Gouinguene et al., 2003; Takabayashi et al., 1995), and we sought to minimize variance in gene expression by using only adult *E. fabae* in this study. In addition, phytophagous insects are known to vary their salivary protein composition in response to the nutritional content and defensive compounds of different host plants (Peiffer and Felton, 2005). Therefore, we collected leaftoppers from a single plot of alfalfa to minimize the potential effects of host plant composition on the sialotranscriptome of *E. fabae*. As adult *E. fabae* are polyphagous and capable of flight, the leaftoppers used in the study may have fed on host plants other than alfalfa. Therefore, we pooled the salivary glands of 200 adult *E. fabae* collected from the interior of an alfalfa field into one sample in order to minimize the genetic contribution of the few individuals that may have fed on an alternative host plant to the assembled sialotranscriptome data.

4. Conclusions

This is the first comprehensive study of the *E. fabae* sialotranscriptome, and the first of any cicadellid leaftopper species. The goal of this study was to understand which genes are active in the salivary glands of adult leaftoppers in an attempt to understand the components of the saliva produced by this economic pest. A number of known insect salivary enzymes were detected in the sialotranscriptome of *E. fabae*, as well as sequences involved in cellular metabolism and biological processes. Tissue-specific expression analysis suggests that some putative digestive enzymes identified in the sialotranscriptome are produced at higher levels in the salivary glands than in the midgut. These results, along with the sequences deposited in GenBank, provide insight into the functioning of the salivary glands of adult *E. fabae* that have fed on alfalfa.

Acknowledgments

The authors would like to thank Jianhua Huang for his help in the salivary gland dissections. Funding for this work was provided by the Maryland Agricultural Experiment Station Grant Program, 2010–11, and MAES Hatch Project MD-ENTM-1016.
Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jinsphys.2012.10.002.

References


