Journal of Insect Physiology 119 (2019) 103965

Contents lists available at ScienceDirect

Journal of Insect Physiology

journal homepage: www.elsevier.com/locate/jinsphys

Proteases and nucleases across midgut tissues of *Nezara viridula* (Hemiptera:Pentatomidae) display distinct activity profiles that are conserved through life stages

Pablo Emiliano Cantón, Bryony C. Bonning*

Department of Entomology and Nematology, University of Florida, PO Box 110620, Gainesville, FL 32611, USA

ARTICLE INFO

Keywords:
Stink bug
Digestive protease
Nuclease
Midgut
Transcriptome

ABSTRACT

The southern green stink bug, *Nezara viridula* is a polyphagous pest of commercially important crops during both nymph and adult stages. This insect has recently transitioned from a secondary agricultural pest to one of primary concern. Novel management solutions are needed due to the limited effectiveness of current control strategies. We performed biochemical and transcriptomic analyses to characterize digestive enzymes in the salivary glands and along midgut tissues of *N. viridula* nymphs and adults fed on sweet corn. The digestive profiles were more distinct between midgut regions (M1 to M3) than between life stages. Aminopeptidase and chymotrypsin activities declined from the M1 (anterior) toward the M3 midgut region. Cysteine protease activity was higher in the M2 and M3 regions than in M1. Differences in sensitivity to chymotrypsin inhibitors between midgut regions suggest that distinct genes or isoforms are expressed in different regions of the gut. In nymphs, DNA and RNA degradation was higher in M1 than in M3. Adult nuclease activity was low across all midgut regions, but high in salivary glands. The differences in protease activities are reflected by transcriptomic data and functional enrichment of GO terms. Together, our results show that different regions of the digestive tract of *N. viridula* have specific and distinct digestive properties, and increase our understanding of the physiology of this organism.

1. Introduction

Although considered secondary pests in some crops of economic importance such as corn, soybean and cotton, the deleterious impact of stink bugs of the pentatomid complex has increased in recent years (Koch and Pahs, 2014). Stinkbug damage alters the appearance or taste of fruiting bodies, or induces premature maturation. Economic losses in 2008 from stinkbug damage to U.S. cotton alone were estimated at US $31 million (Pilkay et al., 2015). Current stink bug control strategies are limited to the spraying of chemical insecticides with inconsistent and diminishing effect, or the use of biological control (Jones, 1988). The increased pest status of stink bugs has been attributed to reduced insecticide use due to expanded adoption of transgenic crops (Greene et al., 1999), variable sensitivity to insecticides (Willrich et al., 2003), and reduced competition from controlled primary pest species (Zeilinger et al., 2016). Stink bugs rely on overlapping periods of flowering and fruiting to produce several generations of offspring in a year, with the distribution of life stages varying across species, cultivar, and seasons (Jones and Sullivan, 1982; Schumann and Todd, 1982; Tillman, 2010), which presents an additional challenge to their control. The southern green stinkbug, *Nezara viridula,* an important member of this pentatomid complex, is polyphagous and affects crops year round. *N. viridula* is widely distributed and is found in the Americas extending from northern Argentina to the eastern and southeastern United States (Panizzi et al., 2000). Effective control of *N. viridula* and related pest species will require novel and alternative methods, protein- (e.g. Bt toxin) and nucleic acid- (e.g. RNAi interference) based methods in particular, that should be tailored to the feeding habits and physiology of stink bugs (Chougule and Bonning, 2012).

N. viridula feeds by a piercing-sucking mechanism, whereby the

Abbreviations: BApNA, Nα-Benzoyl-D,L-arginine 4-nitroanilide hydrochloride; SAAPFpNA, N-Succinyl-Ala-Ala-Pro-Phe p-nitroanilide; LpNA, L-Leucine p-nitroanilide; pGFLpNA, pGlu-Phe-Leu p-nitroanilide; zRRpNA, Z-Arg-Arg p-nitroanilide; PMSF, Phenylmethylsulfonyl fluoride; TLCK, N_{α}-Tosyl-L-lysine chloromethyl ketone hydrochloride; TPCK, N_{α}-Tosyl-L-phenylalaninechloromethyl ketone; EDTA, Ethylenedintrilotetraecetic acid; DMSO, dimethyl sulfoxide; GO, Gene ontology; PCA, Principal component analysis

* Corresponding author.
E-mail address: bbonning@ufl.edu (B.C. Bonning).

https://doi.org/10.1016/j.jinsphys.2019.103965
Received 15 July 2019; Received in revised form 7 October 2019; Accepted 8 October 2019
Available online 11 October 2019

insect injects saliva through its mouthparts into the plant to liquefy the tissue, and then ingests the pre-degraded material to complete digestion in the gut. Thorough degradation and absorption of the dietary components requires the action of a combination of enzymes produced by the digestive organs. Despite its economic importance, biochemical studies on the digestive physiology of this species are not as numerous. The overall digestive profiles of saliva, salivary glands, and guts of adult *N. viridula* (Lomate and Bonning, 2016) and of the invasive species *Halyomorpha halys* (Lomate and Bonning, 2018), have been characterized. Complementary profiles of nucleases and proteases with different pH optima and specific activity in salivary glands and gut were found. The dominance of cysteine protease activity in the gut (primarily cathepsins), is consistent with that of the Heteroptera infraorder of Hemiptera (Terra and Ferreira, 2012; Terra and Ferreira, 1994).

The *N. viridula* midgut is anatomically divided into 4 separate ventricles or regions (Hirose et al., 2006). The anatomical separation of organs in the digestive tract is often indicative of distinct physiological roles, as shown in heteropteran species such as *Dysdercus peruvianus* and *Rhodnius prolixus* (Terra and Ferreira, 2012). Despite different feeding habits, the ultrastructure of midgut tissues among Heteroptera is shared (Santos et al., 2017). Different regions of the midgut vary in the length of microvilli, number of mitochondria, and glycogen granule content for example, pointing to specific roles of each region in digestion and absorption (Fialho et al., 2009; Meguid et al., 2013). Differences in general cysteine protease activity were noted between *N. viridula* midgut compartments during a study of the potential role of gut microbiota in digestive capabilities (Medina et al., 2018). Slight differences were also observed for protease activities in whole gut extracts between adult and fifth instar nymphs of the predatory stink bug, *Podisus maculoventris* (Bell et al., 2005). Whether the digestive profile is conserved across *N. viridula* development (from nymph to adult) remains to be addressed. As feeding on plant material may start as early as the first instar (Esquivel and Medrano, 2014), knowledge of digestive profiles throughout development is important toward management of both nymph and adult *N. viridula* that both contribute to crop losses.

Building on our previous work, we sought to address the lack of information on nymphal gut physiology and the differential roles of *N. viridula* midgut regions in digestion. To increase understanding of the digestive profiles, we combined enzymatic assays and transcriptomic analysis to identify the activity of proteases and nucleases in the M1, M2 and M3 midgut regions in adults and nymphs of *N. viridula*, along with those of the salivary gland. By use of these two approaches, we discovered differences in the digestive enzymes of the gut tissues that are supported by changes in transcript abundance and are consistent with a distinct tissue pH in each midgut region.

2. Materials and methods

2.1. Reagents

Azocasein, Nα-Benzoyl-D,L-arginine 4-nitroanilide hydrochloride (BApNA), N-Succinyl-Ala-Ala-Pro-Phe p-nitroanilide (SAAPFpNA), L-Leucine p-nitroanilide (LpNA), pGlu-Phe-Leu p-nitroanilide (pGFLpNA), Phenylmethylsulfonyl fluoride (PMSF), N_{α}-Tosyl-L-lysine chloromethyl ketone hydrochloride (TLCK), N_{α}-Tosyl-L-phenylalanine-chloromethyl ketone (TPCK),E-64,Ethylenedintrilotetraecetic acid (EDTA), and Calf thymus DNA were purchased from Sigma-Aldrich (St. Louis, MO, USA). Baker's Yeast RNA was purchased from Fisher Scientific/Alfa Aesar (Haverhill, MA, USA). Z-Arg-Arg p-nitroanilide (zRRpNA) was purchased from Bachem (Bubendorf, Switzerland). RNAlater stabilization solution was purchased from Invitrogen (Carlsbad, CA, USA).

2.2. Rearing and dissection of Nezara viridula

The *N. viridula* colony was maintained on a diet of thoroughly washed fresh sweet corn cobs, changed twice weekly, in a Percival incubator at 65% relative humidity, 28 °C, and a 16:8 hr light/dark photoperiod. Although this species is polyphagous and is routinely reared on mixed diet in the laboratory setting, we used a diet comprised of a single plant species to reduce variation that could result from ingestion of differing plant type. Eggs were hatched in sealed petri dishes with a moist cotton plug, and nymphs transferred to diet upon molt to second instar. For dissection, third instar nymphs were selected with a length of ~5 mm (approximately 23 days after egg laying). Upon molting, adults were transferred to fresh diet and allowed to feed for 24 h before dissection. Individuals were anaesthetized on ice before dissecting in a glass petri dish with phosphate buffered saline (PBS), pH 7.4 (25.6 g Na_2HPO_4, 80 g NaCl, 2 g KCl, 2 g KH_2PO_4 per liter of 10X stock). Salivary glands, M1, M2, and M3 midgut sections from adults and nymphs were isolated, and either flash frozen in liquid nitrogen and stored at −80 °C until processed, or submerged in RNAlater solution for future RNA extraction, and stored at −20 °C until processed. The M4 region, which has been reported to house bacterial symbionts in stinkbugs, was also dissected from adults (Bansal et al., 2014; Kikuchi et al., 2008; Prado et al., 2006).

2.3. Determination of tissue pH

Tissue pH was analyzed as reported for other Hemiptera (Cristofoletti, et al., 2003; Deraison, et al., 2004). Fresh salivary glands, M1, M2 and M3 were dissected from adults and dipped in universal pH indicator solution (Fisher Scientific, Haverhill, MA, USA) for 5 min. Tissues were removed, and the colored tissue was evaluated against the color scoring card. The mean and dispersion of pH values was calculated for each tissue from 15 individual dissections.

2.4. Preparation of tissue extracts

Dissected tissues were homogenized with a Polytron 2500E device (Kinematica, Luzern, Switzerland) set at 10,000 rpm for 30 sec on ice in 1.5 ml microcentrifuge tubes, using a 3:1 v/w ratio of PBS to tissue. Tubes were centrifuged twice at 10,000*g*, 4 °C for 10 min and the supernatant transferred to a fresh 1.5 ml tube. Protein concentration in final extracts was determined by the Bradford method (Bio-Rad, Hercules, CA, USA) using BSA as standard.

2.5. Degradation of azocasein substrate

Proteolytic assays were performed as reported (Lomate and Bonning, 2016) with some modifications. Extract content, pH, and time where first optimized for enzymatic assays to ensure that product formation was linearly associated with enzyme concentration (data not shown). For total proteolytic activity, 10 µl of protein extract was mixed with 200 µl of a 1% w/v solution of azocasein in 0.1 M sodium acetate buffer, pH 5.0 for midgut extracts, or 0.1 M Tris-HCl buffer, pH 8.5 for salivary glands. Cysteine (3 mM) from freshly prepared stocks was added to reaction buffer before incubation for some assays. Under some circumstances, the presence of a reducing agent is required for full activity of cysteine proteases. The tubes were incubated for 2 h at 37 °C, and the reaction was stopped by addition of 300 µl of chilled 5% trichloroacetic acid. To remove precipitated protein, tubes were centrifuged for 10 min at 10,000*g*, 4 °C, and 150 µl of supernatant transferred to a 96-well clear bottom plate. To neutralize, 150 µl of 1 M NaOH was added to each well, and absorbance read at 450 nm in an iD3 SpectraMax plate reader (Molecular Devices, San Jose, CA, USA). Absorbance values were normalized for protein content of extracts and time. One activity unit is defined as a change in 0.1 absorbance per min per mg protein.

For inhibition assays, a reaction of 10 µl was prepared in 0.1 M acetate buffer, pH 5.0 with 50 µg (M1) or 30 µg (M2 and M3) of tissue extract in the presence or absence of the following concentrations of

protease inhibitors: 10 mM EDTA, 10 μM E-64, 100 μM TLCK, 100 μM TPCK, or 5 mM PMSF. Reactions were incubated for 30 min at 37 °C and afterwards 200 μl of a 1% azocasein solution in 0.1 M acetate buffer, pH 5.0 were added. Reactions were incubated at 37 °C for 2 h (M2 and M3) or overnight (M1), and then processed and analyzed as above. Activity values for all biological replicates were tested for normality by a Shapiro-Wilks test. Significant differences for inhibitor data compared to control without inhibitor were then determined by a Dunnett's test.

2.6. Degradation of protease class-specific substrates

Solutions of 1 mM synthetic substrate in 0.1 M acetate buffer pH 5.0 were prepared by first dissolving powder in DMSO and then slowly adding buffer, for a final concentration of 10% DMSO (except pGFLpNA at 30%). Reactions of 20 μl were prepared in acetate buffer with 50 μg (M1) or 30 μg (M2 and M3) of midgut extract, with or without protease inhibitors in the concentrations described in Section 2.4 above, and incubated for 30 min at 37 °C. Then, 100 μl of substrate solution was added to each reaction and incubated overnight at 37 °C. Reactions were stopped with 100 μl of 30% acetic acid, 200 μl from each reaction was placed into 96-well plates, and absorbance measured at 410 nm. One unit of activity is defined as the change in 0.1 absorbance per min per mg of protein. Activity values of biological replicates were tested for normality and differences within substrate assays determined by t-test with Bonferroni correction for multiple testing.

2.7. Nucleic acid degradation assays

Calf thymus DNA or Baker's yeast RNA was dissolved in nuclease free water, quantified by Nanodrop at 260 nm (ThermoFisher), and solutions of 0.1 mg/ml were prepared in 20 mM Tris-HCl pH 8.0 nuclease-free buffer containing 25 mM NaCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ (buffer A). For dsRNA substrate, the pGlo plasmid was used to amplify a 502 bp PCR GFP product with primers including the T7 promoter sequence (Table S1). This PCR product was used as template for *in vitro* synthesis of dsRNA using the TrascriptAid T7 High Yield Transcription Kit (ThermoFisher, Waltham, MA, USA) following the manufacturer's instructions. The dsRNA was purified using the PureLink RNA Mini Kit (Ambion, Foster City, CA, USA) and quantified by Nanodrop. Dilutions of dsRNA were prepared in buffer A.

Release of free nucleotides into solution from nucleic acid polymers was measured by an increase in absorbance at 260 nm (Fraser, 1980). For DNA and RNA, 200 μl of 0.1 mg/ml substrate were added to a 10 μl reaction mixture of buffer A and 10 μg of tissue extract. As a negative control, only 10 μl of buffer was added to substrate. For the positive control, reaction mixtures were prepared with 1 μl of 1 U/μl DNase I or 10 μg/μl RNase A (ThermoFisher) and 9 μl of buffer A. To measure background absorbance, 10 μg of tissue extract or 1 μl of control enzyme was added to buffer A, for a total volume of 210 μl. Reactions were prepared in triplicate and incubated for 30 min at 37 °C. Reactions were stopped by addition of 300 μl of chilled 10% trichloroacetic acid in nuclease-free water, with 20 mM sodium pyrophosphate, and then incubated on ice for 1 hr. For dsRNA degradation, reactions were prepared with 2 μg of dsRNA and 10 μg of tissue extract or RNase A in a final volume of 20 μl with buffer A. Control and background reactions were also prepared in 20 μl. Reactions were stopped with 30 μl of 10% trichloroacetic acid with 20 mM sodium pyrophosphate before incubation on ice. After incubation on ice, all tubes were centrifuged at 10,000*g*, for 10 min at 4 °C. The supernatant was recovered, and absorbance values quantified at 260 nm by Nanodrop. The average background of corresponding reactions was subtracted from readings, and then divided by 30 min. One unit was defined as the increase in 0.01 of absorbance at 260 nm.

For gel visualization, 10 μg of tissue extract was mixed with 200 ng, 1 μg, or 200 ng of DNA, RNA or dsRNA substrate respectively, in a final volume of 10 μl of buffer A. Reactions were incubated for 5 or 30 min at 37 °C. Control reactions were prepared without enzymes or with 1 μl of DNase I 1U/μl or RNase A 10 μg/μl. Reactions were run on an agarose gel stained with Gel Red (Biotium, Fremont, CA, USA). Additionally, 10 μg of tissue extract without substrate were run to account for background staining.

2.8. Extraction, purification, and sequencing of mRNA

Purified total RNA was obtained by use of the Ambion PureLink RNA Mini kit (Invitrogen). Tissues stored in RNAlater were resuspended in 600 μl of lysis buffer with β-mercaptoethanol after removing all traces of liquid from samples. Tissues were homogenized by passing several times through a 25G needle and syringe. To each tube, 600 μl 70% ethanol in nuclease-free water were added and mixed thoroughly. Samples were then transferred to spin columns and processed according to the manufacturer's instructions. Eluted RNA in 30 μl nuclease free water was quantified by Nanodrop at 260 nm, treated with DNase I (ThermoFisher), and cleaned up by overnight precipitation at −20 °C in 100 μl of isopropanol, 5 μl of sodium acetate 3 M pH 5.2, and 2.5 μl of RNA grade glycogen, followed by centrifugation at 12,000*g* at 4 °C, and washing of the resulting pellet with 70% ethanol in nuclease-free water at 7600 g for 5 min at 4 °C. Libraries of mRNA were prepared and sequenced by GeneWiz, Inc. (New Jersey, USA) in paired-end format for 150 cycles on an Illumina HiSeq 3000 chip.

2.9. Read processing, assembly, and annotation of de novo transcriptome

Tile filtering and adapter removal of Illumina reads was performed with, respectively, FilterByTile (BBMapSuite (Bushnell, 2015), parameters: d = 0.75, qd = 1,ed = 1, va = 0.5,qa = 0.5,ea = 0.5) and TrimGalore v 0.4.4((Krueger, 2015) parameters: –length 36, -q 5, –stringency 1, -e 0.1). Assembly of reads was performed with Trinity 2.8.3 ((Grabherr et al., 2011), –normalize_max_read_cov 200 and –min_kmer_cov 2). Trinity was also used to evaluate general assembly metrics, while BUSCO v3.01 (Waterhouse et al., 2017) was used to assess completeness for the arthropod ortholog dataset. Transcriptome abundance estimation and distribution was obtained by mapping sample reads to transcriptome assembly using Salmon v 0.12 (Patro et al., 2017) with the –gcBiasflag to correct for sample sequencing bias. Scripts from the Trinity package were used to obtain ExN50 values and Transcripts per Kilobase Million (TPM) values from Salmon abundance files. Annotation of the transcriptome was performed using Trinotate v3.0.1 (Haas et al., 2013) by following the developer's online vignette (Haas, 2018)

2.10. Differential expression analysis for assembled transcripts and functional enrichment

Analysis of transcripts was performed with packages from BioconductoR. The data in the Salmon quant.sf files were imported and adjusted to count tables with TxImport 1.8.0 (Soneson et al., 2015) using the lengthScaledTPM argument, and the counts were retained for individual transcripts rather than summarized to gene values by using the txOut = TRUE and dropInfReps = TRUE arguments. Statistical analysis of transcript expression was performed with DESeq2 1.20.0 (Love et al., 2014). Shrinkage of fold changes through the apeglm algorithm (Zhu et al., 2018) and a fold change threshold above 1 or below −1 were incorporated for statistical testing. An s-value of 0.005 was used as cutoff for significance. PCA plots were created with functions from the DESeq2 and ggplot packages. The IDs from lists of differentially expressed transcripts were used to recover annotation information in the Trinotate report for the *N. viridula* assembly. The number of transcripts with a BLASTp, BLASTx, or PFAM hit was counted for each list.

TopGO (Alexa and Rahnenfuhrer, 2018) was used to determine overrepresented GO terms in the lists of differentially transcribed

transcripts. A custom GO annotation reference was prepared from the Trinotate results where each row contains a transcript ID, and all GO terms from the unique ontology terms were deduced by Trinotate from the BLAST and PFAM hits. In TopGO, information from the annotation file and an annotated list of differentially expressed transcripts was used to build the GO graph object. Only terms associated with node size of at least 5 transcripts were used to test for overrepresentation in sample lists. The 'weight' algorithm was used to evaluate significance of "Molecular Function" terms in the GO topology (Alexa et al., 2006). An exact Fisher test was used and a p-value of 0.05 established as cutoff for reported terms.

3. Results

3.1. Different parts of the digestive tract have different general digestive properties

We sought to determine whether the digestive profiles of different gut regions of *N. viridula* differed and whether these profiles varied between life stages. We selected nymphs that were halfway through third instar for dissection because they were actively feeding on the plant material, and because earlier stage nymphs were too small for adequate acquisition of tissues by dissection. Newly molted adults were transferred to fresh diet for a further 24 h before dissection to allow insects to proceed past sclerotization and initiate feeding behavior.

Tissues dissected from adults were stained with universal pH indicator solution and an average of pH according to the color scale of the indicator was obtained from 15 tissues of each region (Fig. S1). The pH of salivary glands ranged from 6.5 to 7.5, while that of M1 ranged from 6.5 to 7. The pH of M2 ranged from 5 to 6.5 (with a slight gradient visible along the tissue) The M3 region was also acidic, with a mean pH of 5.6.

We investigated the overall proteolytic activity of extracts from the dissected midgut regions (M1 to M4). M2 and M3 had the highest proteolytic activity (Fig. 1). The M1 region was lower in activity than these two, but higher than M4, which had negligible proteolytic activity in these assays. The salivary gland extracts also showed no degradation of the azocasein substrate (less than 10^{-6} units). Addition of a reducing agent (cysteine) to reactions with selected tissue extracts did not increase proteolytic activity (Fig. S2).

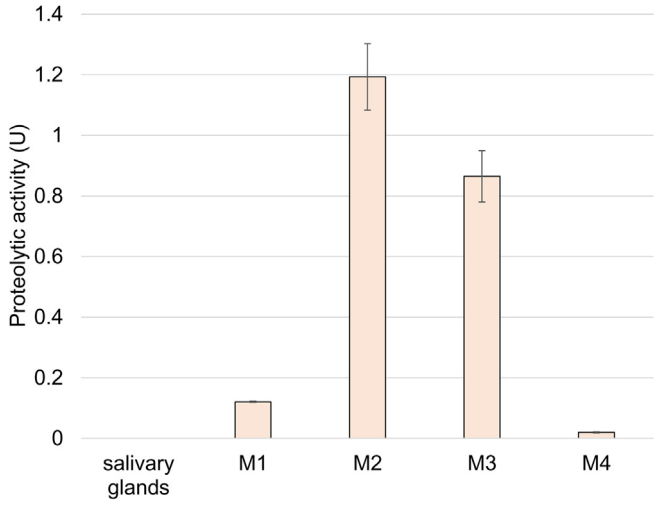

Fig. 1. Total proteolytic activity in the *N. viridula* digestive tract is highest in the M2 and M3 regions. Absorbance was measured on degradation of 1% azocasein substrate. A. One unit represents a change of 0.1 in absorbance per min per mg of protein extract. Error bars show standard deviation.

3.2. Midgut regions of N. viridula are not equally sensitive to protease inhibition

We evaluated the proteolytic activity of extracts from M1, M2 and M3 of adults and nymphs in the presence of inhibitors for different classes of proteases. We observed higher variation in adult M2 than other extracts. The serine protease inhibitor PMSF had little to no effect in reducing the degradation of azocasein by the extracts from the midgut tissues (Fig. 2A, B), as noted previously (Lomate and Bonning, 2016). For comparison, the trypsin and chymotrypsin inhibitors, TPCK and TLCK respectively, reduced activity in M2 and M3 extracts from adults but only significantly so in M3. Enzyme activity in M1 tissue from adults was not significantly inhibited by TPCK but was by TLCK. The cysteine protease inhibitor E-64 was also effective in reducing degradation of azocasein in all adult tissues, significantly so in M1 and M3. Comparatively, the metalloprotease inhibitor EDTA showed lower inhibition in all adult midgut tissues than other inhibitors except for PMSF. These patterns were similar to those of M2 and M3 from nymphs (Fig. 2B). Meanwhile, M1 from nymphs appeared to be less sensitive to all inhibitors in this assay, especially compared to E-64 and TLCK in the adult. To expand on this, we compared the inhibition of each compound between corresponding tissues in the adult and the nymph (Fig. S3). No significant differences were found.

3.3. Specific protease activities show differences between midgut tissues

The azocasein degradation assays showed differences among the midgut regions of *N. viridula*. We investigated further by measuring proteolytic activity in a series of assays of substrates specific to each protease class. With synthetic peptides we were able to individually measure proteolytic activity with reduced crosstalk or interference between proteases with distinct specificities (Fig. 3A, B). The activity of cysteine proteases, on the cathepsin B substrate zRRpNA in particular, was higher than trypsin or chymotrypsin activities for all midgut extracts in the adult. Noticeably, M2 shows a higher activity for all cysteine proteases than on the cathepsin B substrate, which suggests that additional enzymes are contributing to total class activity (Fig. 3A). This difference was not observed in M1 or M3. A considerable drop in aminopeptidase activity occurs from the anterior M1 region of the midgut towards the posterior M3 region. A similar but smaller drop was seen in chymotrypsin activity. Both of these changes are highly significant. This reduction shows a shift in the prevalent enzymatic activity that occurs in concert with the shift to lower pH (Fig. S1). These activity profiles were mirrored in the nymph midgut extracts (Fig. 3B). However, some differences were detected between the life stages (Fig. S4). Nymph M3 showed higher cysteine protease class activity. Nymphs also showed higher chymotrypsin and aminopeptidase activity in M2 and M3, with reduced activity along the length of the gut as seen for adults.

The degradation of class specific substrates was also measured in the presence of protease inhibitors. In general, the inhibition of proteolytic activity confirmed the selectivity of the protease class for each substrate. Interestingly, the amount of chymotrypsin inhibition by TPCK decreased from M1 to M2 and M3 (Fig. S5). This may indicate that the remaining chymotrypsin activity in the posterior midgut is due to chymotrypsin or chymotrypsin-like enzymes that do not share the same inhibitor sensitivity.

3.4. Nuclease activity is low across all midgut regions

When we analyzed the release of free nucleotides by protein extracts, we observed high degradation of all three types of nucleic acids in salivary gland samples from both adults and nymphs (Fig. 4). In contrast, none of the adult midgut extracts had considerable nuclease activity for DNA, RNA or dsRNA. Interestingly, the nymph M1 extracts showed very high DNA degradation, and both M1 and M2 showed

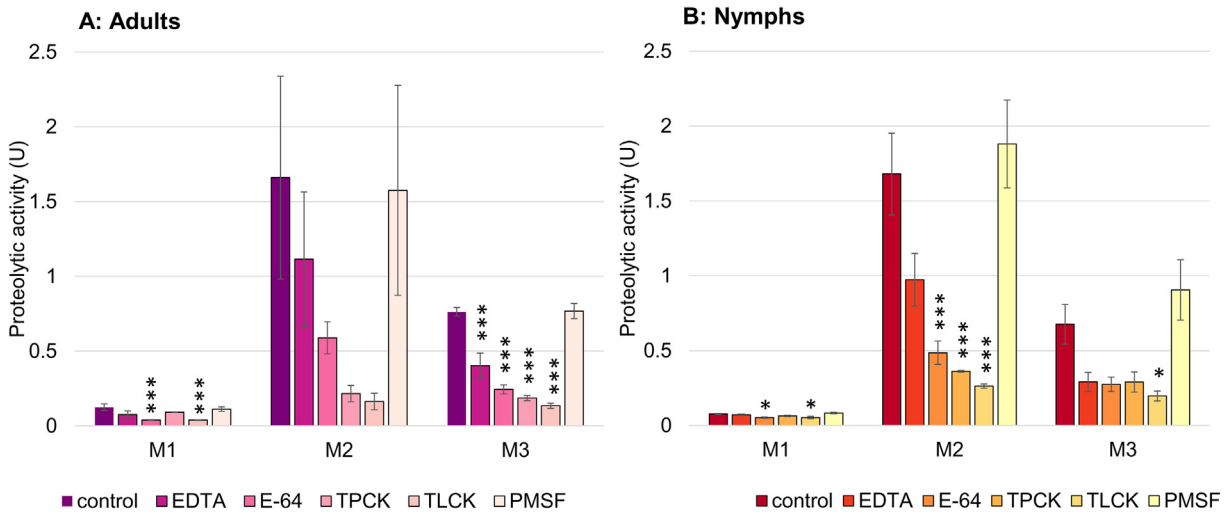

Fig. 2. Midgut regions in adults and nymphs differ in protease inhibitor sensitivity. Degradation of 1% azocasein was performed with tissue extracts in the presence of inhibitors for different protease classes: metalloproteases (EDTA), cysteine proteases (E-64), chymotrypsin (TPCK), trypsin (TLCK), PMSF (serine proteases). Residual activity values are expressed in comparison to the control without inhibitors. Error bars represent SEM from 3 individual experiments. Asterisks represent significance in a Dunnett's test of inhibitor treatments to control for each tissue: p-value * < 0.05, *** < 0.001.

moderate RNA degradation. The remaining nymph extract results show no significant difference to those of adult extracts.

Patterns of staining of the incubated adult and nymph samples in an agarose gel for detection of nuclease activity were mostly consistent with the absorbance assays; no bands were apparent in lanes of samples incubated with salivary gland extracts indicative of complete degradation of nucleic acids (Fig. S6). M1, M2 and M3 extracts resulted in some DNA degradation as indicated by smearing, but not to the extent of digestion by salivary glands. These partially degraded molecules precipitate out and are therefore not recorded in the absorbance assays. Even after 30 min of incubation, dsRNA showed very little degradation by the adult midgut extracts. RNA degradation was harder to visualize because the size of RNA substrate was masked by background staining of the tissue extracts. Nucleases in nymph extracts were somewhat more active than those of adults and led to partial degradation on all the nucleic acid substrates

3.5. Differences in transcript abundance between tissues of the digestive tract

We obtained transcription profiles by sequencing mRNA libraries from adult salivary glands, M1, M2 and M3 in triplicate. Our biochemical assays indicated that the digestive profiles of the different tissues were generally conserved between nymph and adult stages. We chose to use adults to reduce the number of dissected individuals required, and thus to limit the genetic diversity of individuals that could increase the complexity of read assembly. We performed differential expression analysis with a common set of assembled sequences. Reads from this study and those from other *N. viridula* digestive tissue samples from insects maintained on green bean diet (Cantón and Bonning, unpublished) were pooled to enable *de novo* assembly of a common transcriptome reference. Two of the four datasets used for assembly were derived from insects fed on corn or bean-based diet. The resulting

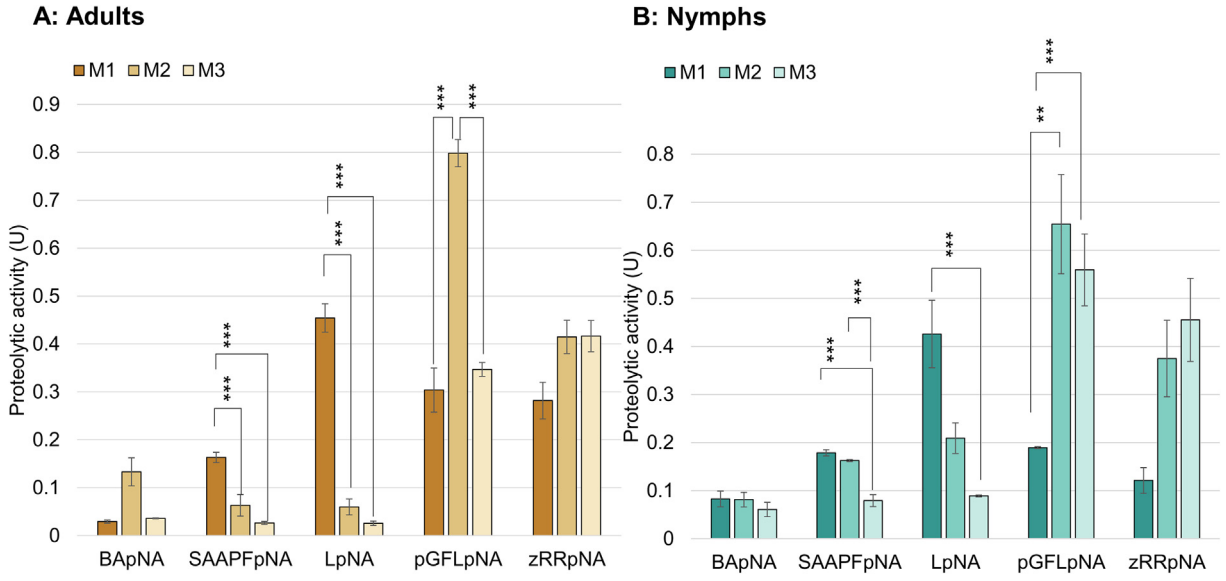

Fig. 3. Degradation of protease class-specific substrates indicates differences in prevalent digestive enzymes between midgut regions of adults and nymphs. Specificity of substrates: BApNA (trypsin), SAAPFpNA (chymotrypsin), LpNA (aminopeptidase-N), zRRpNA (cathepsin B), pGFLpNA (cysteine proteases). Error bars are SEM of 3 independent replicates. Statistical significance of a two-tailed *t*-test with Bonferroni correction of $\alpha = 0.05$ between tissues for each assay: p-value ** < 0.01, *** < 0,001.

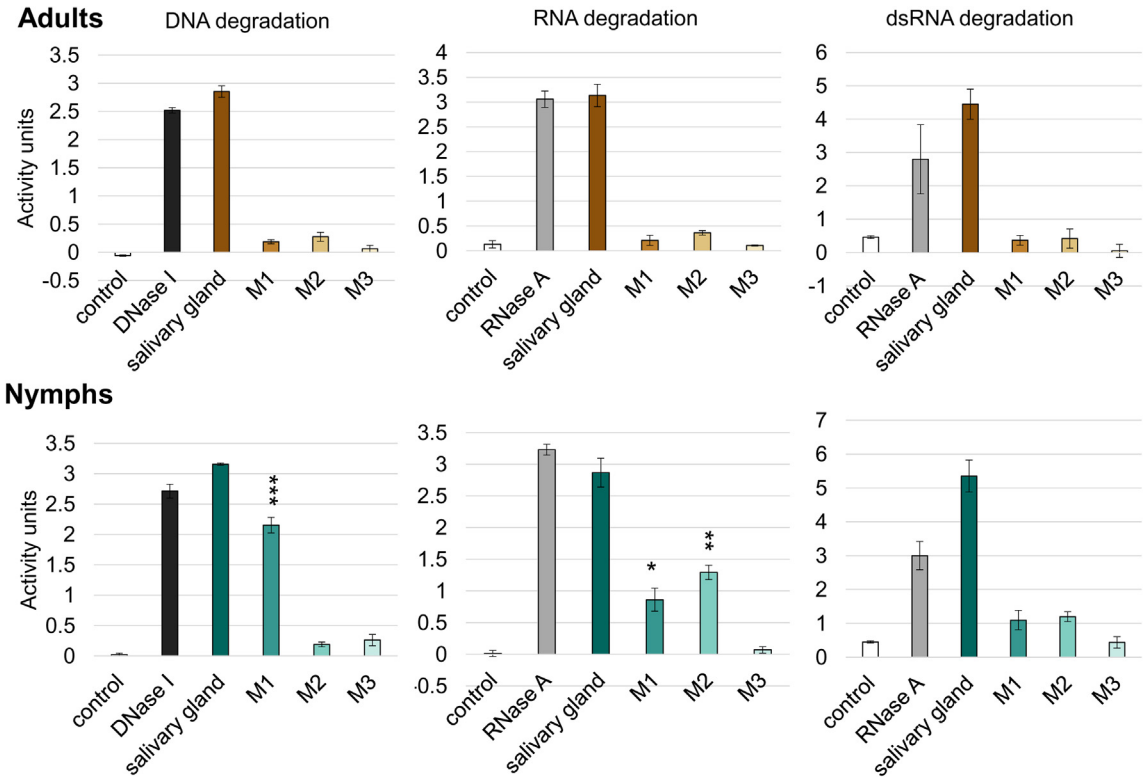

Fig. 4. Nuclease activity is generally high in salivary glands and low across midgut regions in both adults and nymphs. An activity unit is defined as a change of 0.01 in absorbance at λ = 260 nm in 30 min. Error bars represent SEM of three independent replicates. Asterisks indicate statistical difference in a two-tailed t-test p-value * < 0.05, ** < 0.01, *** < 0,001 for comparison of corresponding tissue and assay between adults and nymphs.

Table 1
Quality and completeness metrics of the Trinity assembly and annotation through Trinotate.

Metric	*N. viridula* assembly
Total assembled transcripts	298,322
N50 (nt)	1895
Median contig length (nt)	401
Average contig (nt)	899.09
Complete BUSCOs found	99.1%
Fragmented BUSCOs	0.3%
Missing BUSCOs	0.6%
Isoforms with 95% read data	19,727
Ex N50 of 95% read data (nt)	3,210
Isoforms with BLASTp, BLASTx or PFAM matches	48,330
Unique predicted ORFs	45,359
Unique 'genes' with BLASTp, BLASTX, or PFAM matches	20,320

de novo transcriptome had a large number of isoforms, but 95% of the read data mapped to only 19,727 transcript isoforms with an ExN50 of 3.2 kb (Table 1). BUSCO indicated that 99.1% of the single copy orthologs from the arthropod dataset were found in the transcriptome, indicating a high completeness of the assembly (Table 1). The vast majority of assembled transcripts were short and had low read support, most likely originating from pervasive transcription of non-coding or intergenic DNA. Following annotation through Trinotate, 45,359 transcripts had a predicted ORF, and 48,330 isoforms found a homology match through BLASTx, BLASTp, or PFAM domain searches, which were used to assign GO ontology terms where possible. These isoforms correspond to 20,320 unique Trinity genes. Overall, only 730 Trinity genes had more than 10 isoforms associated with them (Fig. S7).

Abundance data for the corn diet dataset were imported into DEseq2 through the TxImport package to create count data from the mapping results from Salmon. These length-adjusted data were then used to test for differential expression of isoforms between tissues. Simultaneous implementation of apeglm shrinkage and threshold testing in the current and future versions of DESeq2 yield s-values instead of p-values (Stephens, 2016; Zhu et al., 2018). The suggested cutoff for s-values is stricter than p-values at 0.005.

Principal component analysis showed clustering of the tissue triplicates and a strong separation of the salivary gland samples from those of the midgut regions. Within the midgut, separation of the clusters follows the anatomy of *N. viridula*, with M1 and M3 being the most distant from each other and M2 located between these two tissues (Fig. 5). Differentially expressed transcripts for pairwise comparisons were retrieved from the DESeq2 test and separated into sets of transcripts with significantly higher or lower transcription according to their fold change sign and s-value threshold. Consistent with the PCA

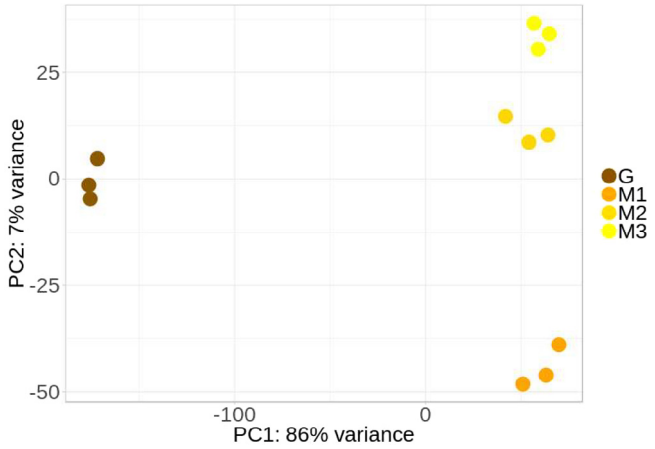

Fig. 5. Differences in transcript abundance reflect anatomical distinctions in the digestive tract of *N. viridula*. The PCA plot is shown for triplicate tissue samples from adult *N. viridula*. G: salivary glands.

Table 2
Number of transcripts with significant fold changes between tissue samples of *N. viridula*. S-values represent the probability of sign error, or the probability that the fold change observed is actually in the opposite direction and would have been found with more sampling. Transcripts with zero counts in all the samples included in a comparison were excluded from the total analyzed. A total of 251,288 transcripts were analyzed for all pairwise comparisons.

Test vs reference[a]	Transcription level	Significant[b]	With annotation[c]	Without annotation
M1 vs gland	higher	2869	1816	1053
	lower	1795	853	942
M2 vs gland	higher	2391	1573	818
	lower	1944	955	989
M3 vs gland	higher	2108	1366	742
	lower	1876	855	1021
M2 vs M1	higher	193	112	81
	lower	39	22	17
M3 vs M1	higher	371	216	155
	lower	799	481	318
M3 vs M2	higher	53	30	23
	lower	41	35	6

[a] Level of transcription is in reference to the second condition in the comparison.
[b] Differences were tested for |fold change| > 1 at an s-value < 0.005.
[c] Transcripts are considered annotated if they have at least one BLASTp, BLASTx, or PFAM match in the Trinotate annotation report.

plots, the salivary glands and midgut tissues were significantly different for a large number of transcripts (Table 2). As expected, there were more significant differences between M3 and M1 than between M2 and M1 or M2 and M3. In all sets of sequences with significant differences in transcription, a large proportion of these sequences did not return hits through BLASTx, BLASTp or PFAM searches, and remain without functional annotation. These transcript results further reinforce the biochemical differences of the midgut regions, particularly between M1 and M2/3.

3.6. Functional exploration of differentially expressed transcripts

The TopGO package was used to investigate functional enrichment in the significantly different subsets of transcripts in each of the tissue pairwise comparisons. (See Supplementary File 1 for the full list of molecular function GO terms). The analysis was limited by the fact that a high percentage of transcripts with significant differences could not be included due to lack of annotation. Nevertheless, the results of the comparisons of tissue samples correlated with the biochemical results, and agreed with previous studies of *N. viridula* (Liu et al., 2018; Lomate and Bonning, 2016).

In comparison to all three midgut tissues, the salivary glands showed enrichment for serine protease-related transcripts, yet also for serine protease inhibitors (Fig. 6). The highest number of other enriched transcripts were related to mannose binding. All three midgut tissues showed enrichment for transcripts for cysteine proteases relative to salivary glands (Fig. 7). In these sets, the most abundant cysteine protease transcripts are those annotated as cathepsin L followed by cathepsin B (Table 3). The number of transcripts with these annotations diminish from M1 to M3. Interestingly, M1 was the only midgut tissue that was also enriched in transcripts for cysteine protease inhibitors relative to salivary glands. While transcripts for the aspartic protease cathepsin D, which plays an important role in digestion in other Hemiptera (Terra and Ferreira, 2012), were similarly represented in M1 and M2 (Table 3), the GO term for aspartic peptidases was not significantly enriched. M1, unlike M2 or M3, showed enrichment for serine protease-related transcripts. Between the different midgut tissues, M2 had 16 cysteine protease-associated transcripts with higher transcription relative to M1. In contrast, M1 had only 2 with higher transcription relative to M2. M3 had no enrichment in this category for sequences with higher transcription over M2 or M1 (Fig. 8). The term for the aminopeptidase, a metalloprotease, was enriched in all tissues, but the number of associated transcripts was small. This number of transcripts also showed a slight decrease from the salivary glands towards M3, similar to the trend seen in the enzymatic assays. Besides protease related terms, all three midgut tissues showed high enrichment over salivary glands for transcripts related to detoxification enzymes (i.e. monooxygenases and carboxylesterases; Fig. 7). These terms were also those where every comparison between midgut tissues showed at least a moderate enrichment, and the one with the highest number of transcripts of M3 with higher levels over M2 and M1 (Fig. 8).

Nuclease-related terms were not overly enriched in any of the midgut tissues (Figs. 7 and 8), with a small number of associated transcripts in M1, M2 and M3. This result indicates no difference in nuclease transcription between these tissues. The T2 ribonuclease term was statistically enriched and had a few highly expressed transcripts in all tissues. However, the associated transcripts were not the same: isoforms of TRINITY_DN27_c0_g2_i9 were highly transcribed in salivary glands while TRINITY_DN2192_c0_g1_i7 and isoforms of TRINITY_DN36_c1_g1 were highly transcribed in midgut tissue.

To take a closer look at transcript abundance for the protease inhibitors that varied across tissues, the Salmon abundance estimated TPM values were retrieved for the corresponding transcripts in the tissue samples. While the TPM for some of these transcripts was high, other transcripts had low TPM despite a statistically significant fold-change. For example, TRINITY_DN824_c29_g1_i4, encoding a serpin, had an average TPM of 1764 in salivary glands but only around 1.5 in all midgut tissues (Fig. S8A). The other transcripts for serine protease inhibitors with notably higher TPM in salivary glands were isoforms of this gene, but also TRINITY_DN1754_c0_g1_i1. The TRINITY_DN57_c0_g1_i2 and TRINITY_DN57_c0_g1_i1 isoforms correspond to cysteine protease inhibitors and were highly transcribed in the salivary glands (Fig. S8B). For the midgut tissues, the transcript abundance for TRINITY_DN1283_c0_g1_i4 and TRINITY_DN1283_c0_g1_i7, encoding a homolog of onchocystatin, was high in both M2 and M3, but not in M1 (Fig. S9B). In M1, TRINITY_DN2863_c0_g1_i2 had the largest TPM of the transcripts for the cysteine protease inhibitors exclusive to that tissue, while TRINITY_DN951_c0_g1_i4 had the highest TPM overall. These transcripts could have roles in the regulation of digestive activity, as M1 had lower cysteine protease activity than M2 or M3. The serine protease inhibitor transcripts contributing to the enrichment of this term in the midgut tissues had significantly higher transcription but have a low TPM overall (Fig. S9A).

4. Discussion

The goal of this study was to establish whether the anatomically distinct sections of the *N. viridula* gut had specialized enzymatic functions in digestion, and to address whether the digestive enzyme profiles are conserved across stink bug development. Insights into stink bug digestive physiology may in the future facilitate the development of targeted control measures to limit crop damage caused by both adults and nymphs.

From the results of this study, we conclude that the M1, M2 and M3 regions of the southern green stink bug midgut differ in their biochemical and digestive enzyme profiles, and likely function sequentially to achieve complete degradation of ingested plant material. Prior analysis of the midgut as a single organ masked the specific properties of each region. For example, we observed differences in pH range in each of the tissues analyzed, consistent with previous reports for other members of the Pentatomorpha (Hirose et al., 2006; Terra and Ferreira, 1994), but opposite to those of other hemipterans such as aphids (Cristofoletti et al., 2003). The overall proteolytic activity differed between midgut regions and was marginal for adult M4. This result reinforces the role of M4 as a symbiont-housing organ that has minimal or no direct role in digestive processes, with gut contents accumulating in

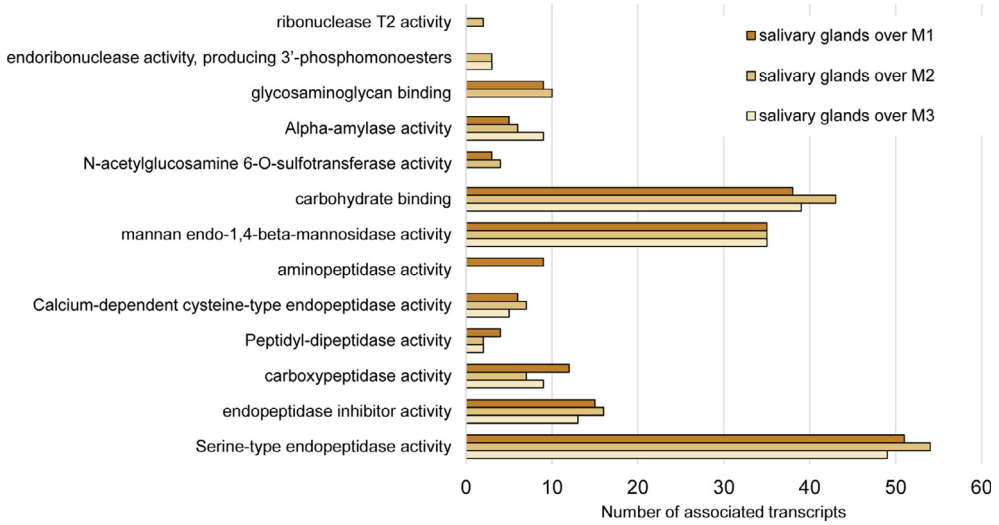

Fig. 6. Selected GO terms enriched for transcripts with lower transcription level in midgut tissues relative to salivary glands. Salivary glands were enriched for terms for carbohydrate binding, serine peptidases, and peptidase inhibitors, mainly for serine peptidases.

M3 for absorption without proceeding to M4. Transport of nutrients into cells is an energy-intensive process, which correlates with the increased abundance of mitochondria reported for the posterior midgut in closely related species (Fialho et al., 2009; Meguid et al., 2013). Differences between tissues were similar in nymphs and adults, suggesting that digestive profiles are stable throughout growth, with physiological roles of each tissue established early in development.

Table 3
Number of differentially transcribed cathepsins. The numbers indicate the transcripts found to have significantly higher transcription in midgut tissues compared to salivary glands in pairwise comparisons. Cathepsin L and B are cysteine proteases, while cathepsin D is an aspartic protease.

	M1	M2	M3
Cathepsin L-like	137	83	59
Cathepsin B-like	19	21	8
Cathepsin D-like	15	15	2

4.1. Differences in nuclease activity between tissues and life stages

In this study we found that nuclease activity is highest for salivary glands in both adults and nymphs for all types of nucleic acid substrates, with relatively low levels in midgut tissues. This result is consistent with nuclease activity in other Hemiptera (Allen and Walker, 2012; Christiaens et al., 2014), Diptera, and Lepidoptera (Singh et al., 2017). Activity levels that were greater than those of pure enzymes used as positive controls may be attributed to the simultaneous action of a cocktail of nucleases present in the extract. Based on the RPKM values for nuclease transcripts, there was no differential regulation of nucleases between principal and accessory salivary glands and midgut in the previous transcriptome (Liu et al., 2018). Consistent with this, we found no notable differences in transcription of nucleases between adult midgut tissues. We did however detect moderate DNase and RNase activity in nymph M1 and M2 by absorbance, and some nuclease activity in adult M1 and nymph extracts by gel staining. In the light of the transcription data, this activity may have resulted from salivary nucleases released into the food and subsequently ingested, rather than

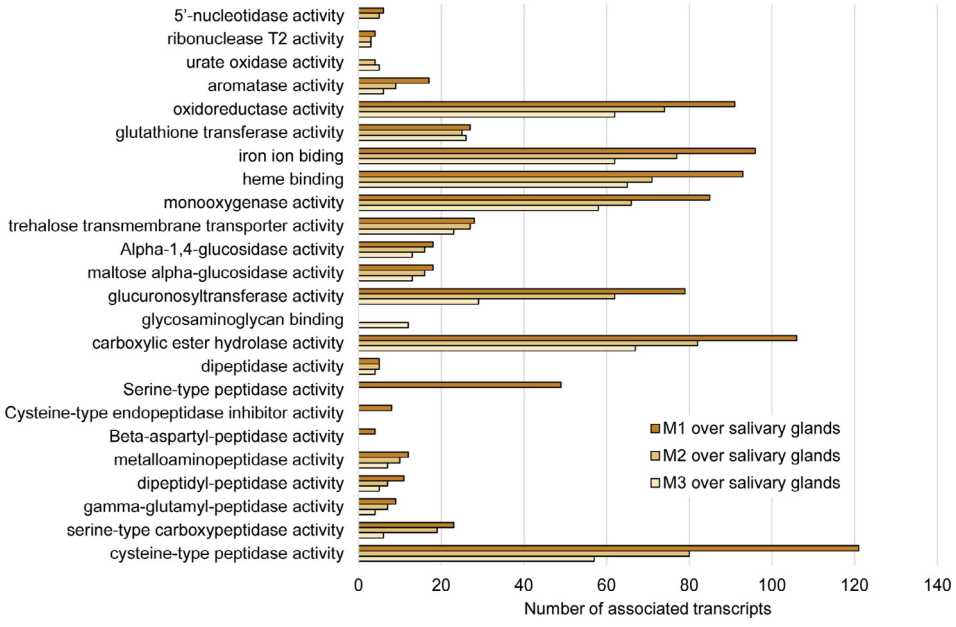

Fig. 7. Select GO terms enriched for transcripts with a higher transcription level in the different midgut regions over salivary gland samples. All midgut tissues were enriched for cysteine proteases, carboxylesterases and oxidoreductase enzymes. Additionally, M1 showed enrichment for serine peptidases and cysteine protease inhibitors.

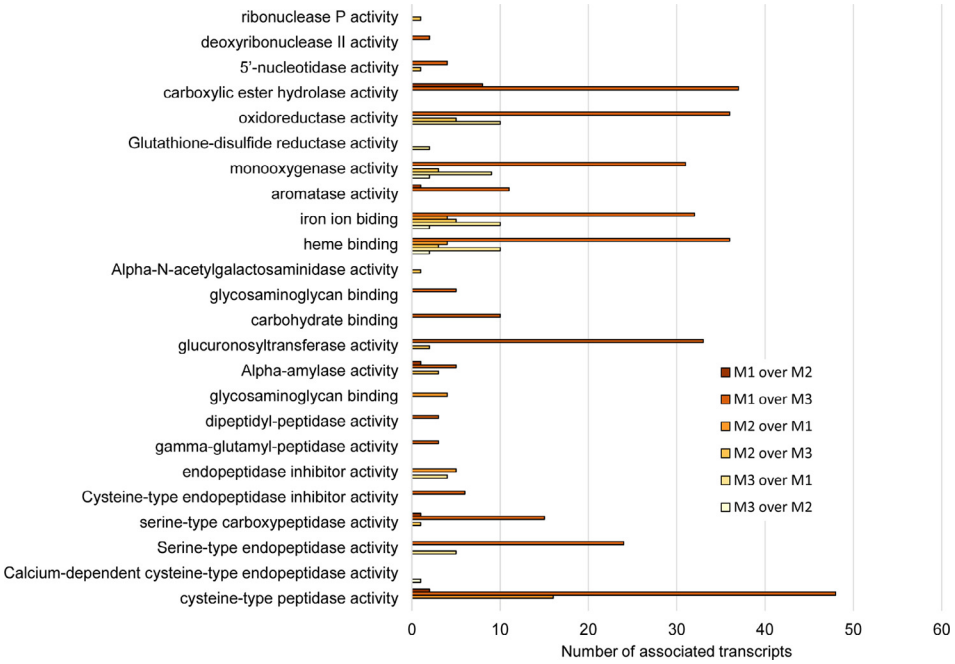

Fig. 8. Selected GO terms enriched in transcripts from comparisons between midgut tissues. Comparisons of M1 over M3 show the terms with the highest number of associated transcripts.

from enzymes produced endogenously by M1 and M2, which explains the progressive loss of activity toward the posterior end of the midgut.

Gene silencing by use of RNA interference (RNAi) is widely studied as a tool for disrupting gene function in insects. Although the effects of RNAi may be consistent in coleopterans (Baum et al., 2007), this is not the case for other insect orders (Joga et al., 2016). The high nuclease activity present in the saliva and salivary glands of hemipterans limits the efficacy of orally delivered dsRNA. In this study, we have shown relatively low nuclease activity in all three adult midgut tissues, but partial degradation of dsRNA in extracts from the nymphal midgut regions. These results highlight the need for a delivery vehicle that can protect the payload from salivary and foregut nucleases if RNAi is to be employed for suppression of genes for *N. viridula* research or management, particularly in the case of nymphs.

4.2. Protease activity, inhibition, and regulation in different tissues

Differences in enzyme class activity reflect the functional specialization of each gut region. For example, in both adults and nymphs we found high aminopeptidase (APN) activity in M1 but not M2 or M3. Additionally, through the transcriptome we saw that only a few transcripts for this function were associated in each tissue. These two facts suggest that only a few proteins contribute to APN activity in M1. Cysteine proteases, as well as aspartic proteases, are more active at acid pH than other enzymes such as trypsin. The evolutionary recruitment of lysosomal cathepsins enabled this proteolytic activity in the hemipteran gut (Terra and Ferreira, 1994). The shift to lower pH parallels the increase in cysteine protease activity along the midgut regions. The importance of cysteine proteases in M2 and M3 is supported both by high inhibition by E-64 and by their proportionally higher activity. Based on the number of differentially transcribed cathepsins in the midgut tissues, this activity may correspond to the movement of cathepsins L or B expressed by M1 into M2 and M3. When focused on M1 however, cysteine proteases might not be the most suitable to study due to the lower biochemical activity in that tissue. The substrate assays showed that general cysteine protease activity measured with pGFLpNA was highest in M2 (Fig. 2), while degradation of the zRRpNA substrate was similar in M2 and M3 (Fig. 3). Although this substrate is preferentially digested by cathepsin B-like enzymes, cathepsin L can also degrade this peptide. Hence, cathepsin L-like enzymes can contribute to the activities recorded. Differences in degradation of zRRpNA and pGFLpNA are likely to result from cathepsin L-like enzymes, particularly for M2 and in nymphal M3. Although we did not analyze other acid stable proteases, specifically cathepsin D, the presence of a few differentially expressed genes with this annotation in M1 and M2 of the midgut indicates a relevant role. It has been proposed that cathepsin D helps to degrade diet-derived cysteine protease inhibitors (Pimentel et al., 2017). The presence of cysteine protease inhibitors in the diet could explain the low cysteine protease activity in M1 despite the high transcription of these enzymes. The cysteine protease inhibitors transcribed in M1 could act on either endogenous or exogenous enzymes. Some plants induce proteases as a herbivore defense mechanism including a cysteine protease in corn (although not in seeds) (Pechan et al., 2000). The presence of cysteine protease inhibitors in M1 would neutralize such enzymes in the diet as they enter the midgut. Alternatively, these inhibitors could regulate cysteine proteases in M1, thereby reducing enzymatic activity of this class in the tissue. Identification of enzyme targets of the transcribed cystatins will improve understanding of their physiological role in M1.

Recent transcriptomic analyses of *N. viridula* (Liu et al., 2018) and *H. halys* (Bansal and Michel, 2018; Sparks et al., 2014) have revealed a large number of protease genes, with particularly high numbers of serine proteases. Our *de novo* assembly and annotation of tissue specific transcriptomes verifies this finding. The unannotated transcripts with high TPM and significant differences should be explored further, perhaps through machine learning algorithms (Ofer and Linial, 2015), to search for digestive enzymes with biochemical properties of interest. The transcription of digestive enzymes has been studied in different tissues for other species of Hemiptera (Bansal and Michel, 2018; Ribeiro et al., 2014; Rispe et al., 2008), which may facilitate functional assignments of unannotated transcripts. The divergence of insect digestive enzymes compared to mammalian homologs can change substrate and inhibitor specificity (Deraison et al., 2004). Notably, PMSF, which is used to inhibit mammalian serine proteases, has little effect on insect digestive enzymes as seen here and in other work (Kuwar et al., 2015). We also observed high inhibition by the TPCK and TLCK inhibitors,

especially in M2 and M3, yet we found low trypsin and chymotrypsin activity in those tissues with synthetic peptides. One explanation is that the serine proteases sensitive to these inhibitors act on motifs present in the azocasein substrate that are absent in the synthetic peptide. Alternatively, TLCK and TPCK can also inhibit cathepsins, and this inhibition contributes to the high inhibition observed in M2 and M3. An additional consideration is that the BApNA substrate, typically used to determine trypsin activity, can be digested by some insect cathepsin-like enzymes (Gooding, 1966). However, degradation of the chymotrypsin substrate SAAPFpNA was on par with that of BApNA in all midgut homogenates analyzed, indicative of low serine protease activity in these tissues rather than crosstalk with cathepsin L. The sequences of proteases and associated isoforms reported here should facilitate phylogenetic analysis to determine insect -and potentially order- specific signatures of insect enzymes.

4.3. Additional enzymatic differences in tissues of the digestive tract.

While the focus for this study was differential protease and nuclease activities, it is apparent from the transcriptome data that carboxylesterases and oxidoreductases were also abundantly expressed along the midgut of *N. viridula*. These enzymes contribute to metabolic plasticity and differentiation, as terms in this category were enriched in all pairwise comparisons between midgut regions and salivary glands. Transcripts for these enzymes exhibited the most changes between midgut tissues. Cytochrome P450s are a large family of oxidoreductases involved in many metabolic, developmental, and detoxification processes in insects (Feyereisen, 2012). Although we did not test for P450 activity, this enzymatic activity is clearly more important for midgut physiology compared to that of the salivary glands, and perhaps especially for M3, where all remaining diet products, including the phytoesterols and flavonoids from plant tissues, must be degraded and absorbed. The expansion and overexpression of these gene families is often associated with resistance to insecticides (David et al., 2013; Puinean et al., 2010; Yang et al., 2013). Cytochrome P450s have also been implicated in the polyphagous nature of the closely related brown marmorated stink bug (Bansal and Michel, 2018). More detailed classification and analysis of these enzymes can provide important information for evaluating the potential for resistance alleles in the Pentatomid family. Given their role in digestion of potentially lethal plant defense compounds, these enzyme transcripts can prove ideal silencing targets for disruption of stinkbug midgut physiology. Finally, salivary glands showed a marked enrichment of carbohydrate binding relative to midgut tissues, particularly for mannose. Enzymes in this category may be necessary for degradation of plant cell polysaccharides. Some carbohydrate metabolism-related terms were also enriched in M1, though these appear to be related to transport and metabolism of simple sugars (maltose or trehalose). The production of carbohydrate binding enzymes in the salivary gland may have implications for the initial contact of the stink bug stylet with plant tissue and the sheath saliva seal produced before feeding, but also for the ability to degrade multiple dietary components with complex glycoprotein structures.

5. Conclusions

In this study we examined the digestive enzyme profiles of individual tissues of the digestive tract of *N. viridula* adults and nymphs, complementing and expanding previous findings. This more detailed description of enzyme activities across gut regions and life stages provides insight into the differing physiological roles of each gut compartment in nymph and adult *N. viridula* complementing ultrastructural observations. This work also highlights the tissue- and stage-specific enzymatic challenges associated with the use of RNAi or toxin-based management of *N. viridula*.

6. Data accessibility

Raw reads sequenced for this project available through NCBI SRA accession SRP193118. The transcript abundance estimates used for differential expression analysis are accessible through GEO with ID GSE130097.

Funding

This material is based upon work supported by the National Science Foundation I/UCRC, the Center for Arthropod Management Technologies, under Grant Nos. IIP-1338775 and 1821914, and by industry partners.

Author contributions

PEC performed the experiments and bioinformatics analysis, PEC and BCB devised the experiments and wrote the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Dr. Ke Wu (University of Florida) kindly provided the pGlo plasmid and primer sequences to prepare GFP dsRNA. The authors acknowledge University of Florida Research Computing for computational resources and support for this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jinsphys.2019.103965.

References

Alexa, A., Rahnenfuhrer, J., 2018. topGO: Enrichment analysis for gene ontology, R package version 2.34.0 ed.

Alexa, A., Rahnenfuhrer, J., Lengauer, T., 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 22, 1600–1607.

Allen, M.L., Walker 3rd, W.B., 2012. Saliva of *Lygus lineolaris* digests double stranded ribonucleic acids. J. Insect Physiol. 58, 391–396.

Bansal, R., Michel, A., 2018. Expansion of cytochrome P450 and cathepsin genes in the generalist herbivore brown marmorated stink bug. BMC Genom. 19, 60.

Bansal, R., Michel, A.P., Sabree, Z.L., 2014. The crypt-dwelling primary bacterial symbiont of the polyphagous pentatomid pest *Halyomorpha halys* (Hemiptera: Pentatomidae). Environ. Entomol. 43, 617–625.

Baum, J.A., Bogaert, T., Clinton, W., Heck, G.R., Feldmann, P., Ilagan, O., Johnson, S., Plaetinck, G., Munyikwa, T., Pleau, M., Vaughn, T., Roberts, J., 2007. Control of coleopteran insect pests through RNA interference. Nat. Biotechnol. 25, 1322–1326.

Bell, A.H., Down, E.R., Edwards, P.J., Gatehouse, A.J., Gatehouse, M.R.A., 2005. Digestive proteolytic activity in the gut and salivary glands of the predatory bug *Podisus maculiventris* (Heteroptera: Pentatomidae); effect of proteinase inhibitors. Eur. J. Entomol. 102, 139–145.

Bushnell, B., 2015. BBMap: A Fast Accurate Splice-Aware Aligner.

Chougule, N.P., Bonning, B.C., 2012. Toxins for transgenic resistance to hemipteran pests. Toxins (Basel) 4, 405–429.

Christiaens, O., Swevers, L., Smagghe, G., 2014. DsRNA degradation in the pea aphid (*Acyrthosiphon pisum*) associated with lack of response in RNAi feeding and injection assay. Peptides 53, 307–314.

Cristofoletti, P.T., Ribeiro, A.F., Deraison, C., Rahbe, Y., Terra, W.R., 2003. Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid *Acyrthosiphon pisum*. J. Insect Physiol. 49, 11–24.

David, J.P., Ismail, H.M., Chandor-Proust, A., Paine, M.J., 2013. Role of cytochrome P450s in insecticide resistance: impact on the control of mosquito-borne diseases and use of insecticides on Earth. Philos. Trans. Roy. Soc. London B Biol. Sci. 368, 20120429.

Deraison, C., Darboux, I., Duportets, L., Gorojankina, T., Rahbe, Y., Jouanin, L., 2004. Cloning and characterization of a gut-specific cathepsin L from the aphid *Aphis gossypii*. Insect Mol. Biol. 13, 165–177.

Esquivel, J.F., Medrano, E.G., 2014. Ingestion of a marked bacterial pathogen of cotton conclusively demonstrates feeding by first instar southern green stink bug (Hemiptera: Pentatomidae). Environ. Entomol. 43, 110–115.
Feyereisen, R., 2012. Insect CYP genes and P450 enzymes. In: Gilbert, L.I. (Ed.), Insect Molecular Biology and Biochemistry. Academic Press, San Diego, pp. 236–316.
Fialho, M.D.C.Q., Zanuncio, J.C., Neves, C.A., Ramalho, F.S., Serrão, J.E., 2009. Ultrastructure of the digestive cells in the midgut of the predator *Brontocoris tabidus* (Heteroptera: Pentatomidae) after different feeding periods on prey and plants. Ann. Entomol. Soc. Am. 102, 119–127.
Fraser, M.J., 1980. Purification and properties of *Neurospora crassa* endo-exonuclease, an enzyme which can be converted to a single-strand specific endonuclease. Methods Enzymol. 65, 255–263.
Gooding, R.H., 1966. *In vitro* properties of proteinases in the midgut of adult *Aedes aegypti L.* and *Culex fatigans* (Wiedemann). Comp. Biochem. Physiol. 17, 115–127.
Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.
Greene, J.K., Turnipseed, S.G., Sullivan, M.J., Herzog, G.A., 1999. Boll damage by southern green stink bug (Hemiptera: Pentatomidae) and tarnished plant bug (Hemiptera: Miridae) caged on transgenic *Bacillus thuringiensis* cotton. J. Econ. Entomol. 92, 941–944.
Haas, B.J., 2018. Trinotate: transcriptome functional annotation and analysis. https://github.com/Trinotate/Trinotate.github.io/wiki. Accesed 28, Septemeber, 2018.
Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N., Regev, A., 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494–1512.
Hirose, E., Panizzi, A.R., Souza, J.T.D., Cattelan, A.J., Aldrich, J.R., 2006. Bacteria in the gut of southern green stink bug (Heteroptera: Pentatomidae). Ann. Entomol. Soc. Am. 99 (91–95), 95.
Joga, M.R., Zotti, M.J., Smagghe, G., Christiaens, O., 2016. RNAi efficiency, systemic properties, and novel delivery methods for pest insect control: what we know so far. Front. Physiol. 7, 553.
Jones, W.A., 1988. World review of the parasitoids of the southern green stink bug, *Nezara viridula (L.)* (Heteroptera: Pentatomidae). Ann. Entomol. Soc. Am. 81, 262–273.
Jones, W.A., Sullivan, M.J., 1982. Role of host plants in population-dynamics of stink bug (Hemiptera, Pentatomidae) pests of soybean in South Carolina. Environ. Entomol. 11, 867–875.
Kikuchi, Y., Hosokawa, T., Fukatsu, T., 2008. Diversity of bacterial symbiosis in stinkbugs. Microbial Ecology Research Trends. Nova Sicence Publishers, Inc.
Koch, R.L., Pahs, T., 2014. Species composition, abundance, and seasonal dynamics of stink bugs (Hemiptera: Pentatomidae) in Minnesota soybean fields. Environ. Entomol. 43, 883–888.
Krueger, F., 2015. A wrapper around Cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files, with extra functionality for RRBS data. http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
Kuwar, S.S., Pauchet, Y., Vogel, H., Heckel, D.G., 2015. Adaptive regulation of digestive serine proteases in the larval midgut of *Helicoverpa armigera* in response to a plant protease inhibitor. Insect Biochem. Mol. Biol. 59, 18–29.
Liu, S., Lomate, P.R., Bonning, B.C., 2018. Tissue-specific transcription of proteases and nucleases across the accessory salivary gland, principal salivary gland and gut of *Nezara viridula*. Insect Biochem. Mol. Biol. 103, 36–45.
Lomate, P.R., Bonning, B.C., 2016. Distinct properties of proteases and nucleases in the gut, salivary gland and saliva of southern green stink bug, *Nezara viridula*. Sci. Rep. 6, 27587.
Lomate, P.R., Bonning, B.C., 2018. Proteases and nucleases involved in the biphasic digestion process of the brown marmorated stink bug, *Halyomorpha halys* (Hemiptera: Pentatomidae). Arch. Insect Biochem. Physiol. 98, e21459.
Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550.
Medina, V., Sardoy, P.M., Soria, M., Vay, C.A., Gutkind, G.O., Zavala, J.A., 2018. Characterized non-transient microbiota from stinkbug (*Nezara viridula*) midgut deactivates soybean chemical defenses. PLoS One 13, e0200161.
Meguid, A.A., Awad, H.H., Omar, A.H., Elelimy, H.A.S., 2013. Ultrastructural study on the midgut regions of the milkweed bug, *Spilostethus pandurus* Scop., (Hemiptera:Lygaeidae). Asian J. Biol. Sci. 6, 54–66.
Ofer, D., Linial, M., 2015. ProFET: feature engineering captures high-level protein functions. Bioinformatics 31, 3429–3436.
Panizzi, A.R., McPherson, J.E., James, D.G., Javahery, M., McPherson, R.M., 2000. Stink bugs (Pentatomidae). In: Schaefer, C.W., Panizzi, A.N.R. (Eds.), Heteroptera of Economic Importance. CRC Press, Boca Raton, FL, pp. 421–474.
Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., Kingsford, C., 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat. Meth. 14, 417–419.
Pechan, T., Ye, L., Chang, Y., Mitra, A., Lin, L., Davis, F.M., Williams, W.P., Luthe, D.S., 2000. A unique 33-kD cysteine proteinase accumulates in response to larval feeding in maize genotypes resistant to fall armyworm and other Lepidoptera. Plant Cell 12, 1031–1040.
Pilkay, G.L., Reay-Jones, F.P., Toews, M.D., Greene, J.K., Bridges, W.C., 2015. Spatial and temporal dynamics of stink bugs in southeastern farmscapes. J. Insect Sci. 15. https://doi.org/10.1093/jisesa/iev006.
Pimentel, A.C., Fuzita, F.J., Palmisano, G., Ferreira, C., Terra, W.R., 2017. Role of cathepsins D in the midgut of *Dysdercus peruvianus*. Comp. Biochem. Physiol. B: Biochem. Mol. Biol. 204, 45–52.
Prado, S.S., Rubinoff, D., Almeida, R.P.P., 2006. Vertical transmission of a pentatomid caeca-associated symbiont. Ann. Entomol. Soc. Am. 99, 577–585.
Puinean, A.M., Foster, S.P., Oliphant, L., Denholm, I., Field, L.M., Millar, N.S., Williamson, M.S., Bass, C., 2010. Amplification of a Cytochrome P450 gene is associated with resistance to neonicotinoid insecticides in the aphid *Myzus persicae*. PLoS Genet. 6, e1000999.
Ribeiro, J.M., Genta, F.A., Sorgine, M.H., Logullo, R., Mesquita, R.D., Paiva-Silva, G.O., Majerowicz, D., Medeiros, M., Koerich, L., Terra, W.R., Ferreira, C., Pimentel, A.C., Bisch, P.M., Leite, D.C., Diniz, M.M., Junior, da, S.G.V., Da Silva, M.L., Araujo, R.N., Gandara, A.C., Brosson, S., Salmon, D., Bousbata, S., Gonzalez-Caballero, N., Silber, A.M., Alves-Bezerra, M., Gondim, K.C., Silva-Neto, M.A., Atella, G.C., Araujo, H., Dias, F.A., Polycarpo, C., Vionette-Amaral, R.J., Fampa, P., Melo, A.C., Tanaka, A.S., Balczun, C., Oliveira, J.H., Goncalves, R.L., Lazoski, C., Rivera-Pomar, R., Diambra, L., Schaub, G.A., Garcia, E.S., Azambuja, P., Braz, G.R., Oliveira, P.L., 2014. An insight into the transcriptome of the digestive tract of the bloodsucking bug, *Rhodnius prolixus*. PLoS Neglected Trop. Dis. 8 e2594.
Rispe, C., Kutsukake, M., Doublet, V., Hudaverdian, S., Legeai, F., Simon, J.C., Tagu, D., Fukatsu, T., 2008. Large gene family expansion and variable selective pressures for cathepsin B in aphids. Mol. Biol. Evol. 25, 5–17.
Santos, H.P., Rost-Roszkowska, M., Vilimova, J., Serrao, J.E., 2017. Ultrastructure of the midgut in Heteroptera (Hemiptera) with different feeding habits. Protoplasma 254, 1743–1753.
Schumann, F.W., Todd, J.W., 1982. Population-dynamics of the southern green stink bug (Heteroptera, Pentatomidae) in relation to soybean phenology. J. Econ. Entomol. 75, 748–753.
Singh, I.K., Singh, S., Mogilicherla, K., Shukla, J.N., Palli, S.R., 2017. Comparative analysis of double-stranded RNA degradation and processing in insects. Sci. Rep. 7, 17059.
Soneson, C., Love, M.I., Robinson, M.D., 2015. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res 4, 1521.
Sparks, M.E., Shelby, K.S., Kuhar, D., Gundersen-Rindal, D.E., 2014. Transcriptome of the invasive brown marmorated stink bug, *Halyomorpha halys* (Stal) (Heteroptera: Pentatomidae). PLoS One 9, e111646.
Stephens, M., 2016. False discovery rates: a new deal. Biostatistics 18, 275–294.
Terra, W., Ferreira, C., 2012. Biochemistry and molecular biology of digestion. In: Gilbert, L.I. (Ed.), Insect Molecular Biology and Biochemistry. Academic Press, San Diego, pp. 365–418.
Terra, W.R., Ferreira, C., 1994. Insect digestive enzymes – properties, compartmentalization and function. Comp. Biochem. Physiol. B 109, 1–62.
Tillman, P.G., 2010. Composition and abundance of stink bugs (Heteroptera: Pentatomidae) in corn. Environ. Entomol. 39, 1765–1774.
Waterhouse, R.M., Seppey, M., Simao, F.A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E.V., Zdobnov, E.M., 2017. BUSCO applications from quality assessments to gene prediction and phylogenomics. Mol. Biol. Evol. https://doi.org/10.1093/molbev/msx319.
Willrich, M.M., Leonard, B.R., Cook, D.R., 2003. Laboratory and field evaluations of insecticide toxicity to stink bugs. (Heteroptera: Pentatomidae) v, 7.
Yang, X., Xie, W., Wang, S.-L., Wu, Q.-J., Pan, H.-P., Li, R.-M., Yang, N.-N., Liu, B.-M., Xu, B.-Y., Zhou, X., Zhang, Y.-J., 2013. Two cytochrome P450 genes are involved in imidacloprid resistance in field populations of the whitefly, *Bemisia tabaci,* in China. Pestic. Biochem. Physiol. 107, 343–350.
Zeilinger, A.R., Olson, D.M., Andow, D.A., 2016. Competitive release and outbreaks of non-target pests associated with transgenic Bt cotton. Ecol. Appl. 26, 1047–1054.
Zhu, A., Ibrahim, J.G., Love, M.I., 2018. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics 35, 2084–2092.