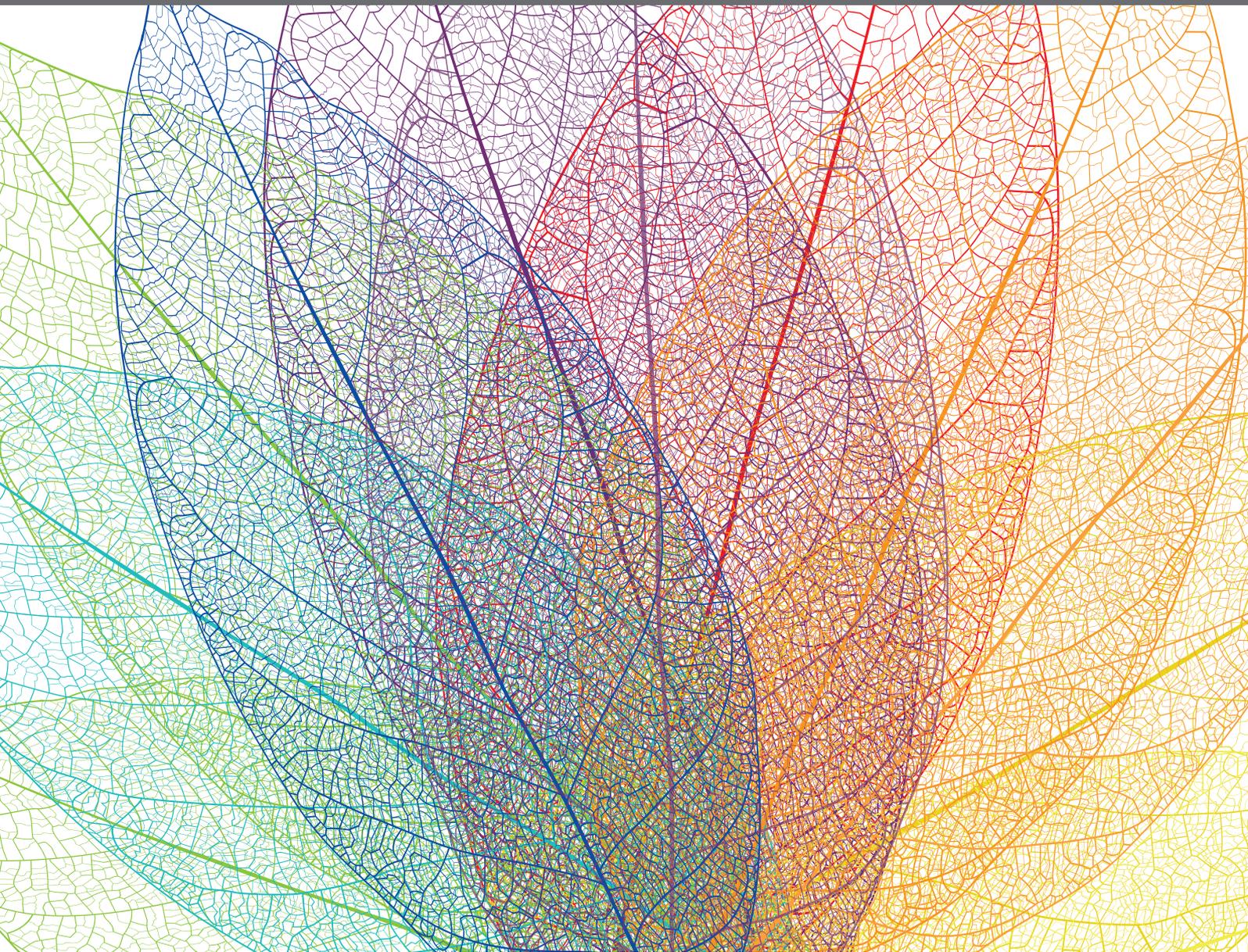


# PLANT-ARTHROPOD INTERACTIONS: EFFECTORS AND ELICITORS OF ARTHROPODS AND THEIR ASSOCIATED MICROBES

EDITED BY: Gary W. Felton, Giron David, Akiko Sugio and Isgouhi Kaloshian  
PUBLISHED IN: *Frontiers in Plant Science* and *Frontiers in Physiology*





# frontiers

## Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714  
ISBN 978-2-88966-305-7  
DOI 10.3389/978-2-88966-305-7

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)

# PLANT-ARTHROPOD INTERACTIONS: EFFECTORS AND ELICITORS OF ARTHROPODS AND THEIR ASSOCIATED MICROBES

Topic Editors:

**Gary W. Felton**, Pennsylvania State University (PSU), United States

**Giron David**, Centre National de la Recherche Scientifique (CNRS), France

**Akiko Sugio**, Institut de Génétique, Environnement et Protection des Plantes, France

**Isgouhi Kaloshian**, University of California, Riverside, United States

**Citation:** Felton, G. W., David, G., Sugio, A., Kaloshian, I., eds. (2020). Plant-Arthropod Interactions: Effectors and Elicitors of Arthropods and Their Associated Microbes. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-305-7

# Table of Contents

- 04** *Editorial: Plant-Arthropod Interactions: Effectors and Elicitors of Arthropods and Their Associated Microbes*  
Akiko Sugio, Gary W. Felton, David Giron and Isgouhi Kaloshian
- 06** *The Arabidopsis Lectin Receptor Kinase LecRK-I.8 is Involved in Insect Egg Perception*  
Caroline Gouhier-Darimont, Elia Stahl, Gaetan Glauser and Philippe Reymond
- 16** *The Principal Salivary Gland is the Primary Source of Digestive Enzymes in the Saliva of the Brown Marmorated Stink Bug, Halyomorpha halys*  
Sijun Liu and Bryony C. Bonning
- 33** *Differential Expression of Candidate Salivary Effector Genes in Pea Aphid Biotypes With Distinct Host Plant Specificity*  
Hélène Boulain, Fabrice Legeai, Julie Jaquiéry, Endrick Guy, Stéphanie Morlière, Jean-Christophe Simon and Akiko Sugio
- 45** *Transcription and Activity of Digestive Enzymes of Nezara viridula Maintained on Different Plant Diets*  
Pablo Emiliano Cantón and Bryony C. Bonning
- 56** *Chemical, Physiological and Molecular Responses of Host Plants to Lepidopteran Egg-Laying*  
Cinzia Margherita Berteà, Luca Pietro Casacci, Simona Bonelli, Arianna Zampollo and Francesca Barbero
- 63** *AcDCXR is a Cowpea Aphid Effector With Putative Roles in Altering Host Immunity and Physiology*  
Jacob R. MacWilliams, Stephanie Dingwall, Quentin Chesnais, Akiko Sugio and Isgouhi Kaloshian
- 80** *BSA-Seq Discovery and Functional Analysis of Candidate Hessian Fly (Mayetiola destructor) Avirulence Genes*  
Lucio Navarro-Escalante, Chaoyang Zhao, Richard Shukle and Jeffrey Stuart
- 96** *Juvenile Spider Mites Induce Salicylate Defenses, but Not Jasmonate Defenses, Unlike Adults*  
Jie Liu, Saioa Legarrea, Juan M. Alba, Lin Dong, Rachid Chafi, Steph B. J. Menken and Merijn R. Kant
- 108** *Big Genes, Small Effectors: Pea Aphid Cassette Effector Families Composed From Miniature Exons*  
Matthew Dommel, Jonghee Oh, Jose Carlos Huguet-Tapia, Endrick Guy, Hélène Boulain, Akiko Sugio, Marimuthu Murugan, Fabrice Legeai, Michelle Heck, C. Michael Smith and Frank F. White



# Editorial: Plant-Arthropod Interactions: Effectors and Elicitors of Arthropods and Their Associated Microbes

Akiko Sugio<sup>1\*</sup>, Gary W. Felton<sup>2</sup>, David Giron<sup>3</sup> and Isgouhi Kaloshian<sup>4</sup>

<sup>1</sup> INRAE, UMR1349, Institute of Genetics, Environment and Plant Protection, Le Rheu, France, <sup>2</sup> Department of Entomology, Pennsylvania State University, University Park, PA, United States, <sup>3</sup> Centre National de la Recherche Scientifique, Tours, France, <sup>4</sup> Department of Nematology, Institute for Integrative Genome Biology, University of California, Riverside, Riverside, CA, United States

**Keywords:** plant, insect, arthropod, herbivore, elicitor, effector

## Editorial on the Research Topic

### Plant-Arthropod Interactions: Effectors and Elicitors of Arthropods and Their Associated Microbes

With the advent of omics technologies, sequencing of genomes and transcriptomes of a number of arthropods have been accomplished. These achievements have brought about a renaissance in the study of host plant and herbivorous arthropod interactions. Using these approaches, intricate interactions have been revealed. Secretions from arthropods presumably delivered into the host plant and containing proteinaceous effectors and elicitors of both arthropod and microbial origins, were shown to modulate plant immunity and metabolism acting as inducers or suppressors of physiological responses. In this Research Topic, we aimed to gather research articles and reviews that describe the identification and characterization of the effectors and the elicitors involved in the interactions with host plants.

For many herbivores, the first contact with their host plant is during egg deposition or oviposition. Plants have evolved sensing mechanisms to recognize the mechanical and chemical cues associated with oviposition. In their review, Berteaux et al. provide an excellent overview of the responses of host plants to oviposition by Lepidoptera (i.e., moths and butterflies). Egg-induced defenses can directly impair or kill eggs through localized necrosis, neoplasm formation, and/or the direct production of ovicidal compounds. Plants also produce oviposition-induced plant volatiles which attract parasitoids that eventually kill the eggs or larvae. They argue that progress in understanding the specificity of these responses requires further characterization of egg-associated elicitors and the plant receptors that recognize these chemical cues. Gouhier-Darimont et al. contribute an important paper in understanding plant perception of oviposition. In *Arabidopsis*, eggs of the specialist butterfly, *Pieris brassicae* elicit a burst of reactive oxygen species and salicylic acid, and downstream defense gene expression and localized necrosis. Oviposition and egg cues trigger the localized expression of an L-type lectin receptor kinase LecRK-I.8. Using an *Arabidopsis* knock-out mutant *lecrk-I.8*, they found that the plant defense responses to these egg cues were significantly impaired in this mutant. Their results demonstrate that LecRK-I.8 is an early component of egg perception.

After the eggs hatch, herbivorous arthropods start to feed on their host plants and direct interactions between the animal and the plant begin. Tomato responses to two spider mite species, *Tetranychus urticae* and *T. evansi* were examined in detail by analyzing expression patterns of

## OPEN ACCESS

### Edited and reviewed by:

Michele Perazzolli,  
University of Trento, Italy

### \*Correspondence:

Akiko Sugio  
akiko.sugio@inrae.fr

### Specialty section:

This article was submitted to  
Plant Pathogen Interactions,  
a section of the journal  
Frontiers in Plant Science

**Received:** 25 September 2020

**Accepted:** 12 October 2020

**Published:** 04 November 2020

### Citation:

Sugio A, Felton GW, Giron D and  
Kaloshian I (2020) Editorial:  
Plant-Arthropod Interactions: Effectors  
and Elicitors of Arthropods and Their  
Associated Microbes.  
*Front. Plant Sci.* 11:610160.  
doi: 10.3389/fpls.2020.610160

marker genes for jasmonic and salicylic acids defense hormones (Liu et al.). In this analysis, they compared cumulative effect of mite life stages and effect of feeding by male and female adult mites. They also examined salivary effector expression patterns in similar cohorts of mites. Their study shows complex interactions of spider mites and their host and demonstrates fine-tuned regulation of salivary effector expressions in the two mite species.

While feeding, herbivorous arthropods secrete proteinaceous saliva. Various salivary components are identified and analyzed. Liu and Bonning took advantage of the enhanced availability of genomic resources for stink bugs to explore the repertoire of digestive enzymes through a tissue-specific transcriptome analysis. Their work provides evidence for the principal salivary gland being the primary source of proteases and nucleases used for efficient digestion of plant materials. They also show that *Halyomorpha halys* and *Nezara viridula* have a similar digestive biochemical arsenal and propose that the large diversity of salivary enzymes may mediate the ability of stink bugs to feed on multiple hosts. The ability of stink bugs to feed on diverse crop systems is further explored by Cantón and Bonning. They demonstrate that protease and nuclease activity of *N. viridula* maintained on different plant diets are similar. Conversely, their work shows that specific transcripts of the digestive enzymes are different. How diet could change the digestive physiology may help understand polyphagy and could open new avenues for the development of innovative control strategies of pests. Nevertheless, the study is limited in its finding because of inadequate genomic resources.

A thorough comparative analysis of salivary gene expression patterns in *Acyrtosiphon pisum* biotypes, which show distinct host plant specificity, reveal that the majority of the genes encoding candidate salivary effectors are expressed in two biotypes compared, and that there are small subsets of genes that are differentially expressed in a biotype-specific manner (Boulain et al.). As those subsets are enriched with duplicated and aphid-lineage-specific genes, the authors propose a scenario that biotype-specific salivary effectors have evolved recently and diversified through duplication events. Further, two candidate salivary effector families are reported in *A. pisum* (Dommel et al.). The members of these gene families encode highly conserved secretory signal peptides and divergent mature proteins derived from miniature exons. The family members are scattered throughout *A. pisum* genome and encoded in unusually large genomic regions. The authors propose a model that the gene families expanded in *A. pisum* through combinatorial assemblies of a common secretory signal cassette and novel coding regions, and hypothesis that the gene families facilitate the adaptation of the aphid to new hosts. MacWilliams et al. profile the salivary proteome of the cowpea aphid, *Aphis craccivora*. Their work

identifies a novel effector, AcDCXR, a member of short-chain dehydrogenases/reductases. They show that the recombinant AcDCXR protein has the predicted enzymatic activity in carbohydrate and dicarbonyl metabolisms with putative ability to enhance nutrition to the aphid as well as alter plant defense responses. Consistently, they show that transient expression of AcDCXR enhances the fecundity of the aphid. Their work also provides evidence for the existence of a novel pest defense metabolite, methylglyoxal, known for its role in abiotic stress.

Effectors are recognized by plant resistance (R) proteins and a way to overcome this resistance is the ability of the pest to mutate the effector to evade the recognition by the cognate R protein. Navarro-Escalante et al. describe the use of bulked-segregant analysis and whole genome sequence to identify virulent effectors from the Hessian flies (*Mayetiola destructor*) that have overcome single gene resistances in wheat. Their work confirms the identity of a previously identified virulence effector *vH6*, as well as identifies a second virulence effector *vHdic*. Using heterologous expression system, they show the ability of these two virulence effectors to suppress plant immune responses providing direct evidence for the role of effectors in pest virulence.

Taken together, these articles demonstrate that new technologies clearly expanded the opportunities to study a wide range of arthropod-plant interactions. These resources enabled identification of numerous effector/elicitor candidates, description of their expression patterns and their receptors. Yet, functional characterization of effectors/elicitors remains a big challenge. Model systems (e.g., Arabidopsis) could advance the field rapidly, but arthropod host specificity limits the use of model systems. In some cases, heterologous systems can be employed to overcome these difficulties. Further development of research tools is needed to understand functions of effectors, perception mechanisms of elicitors and how these activities are translated into the interactions between plants and herbivores.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Sugio, Felton, Giron and Kaloshian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The *Arabidopsis* Lectin Receptor Kinase LecRK-I.8 Is Involved in Insect Egg Perception

Caroline Gouhier-Darimont<sup>1</sup>, Elia Stahl<sup>1</sup>, Gaetan Glauser<sup>2</sup> and Philippe Reymond<sup>1\*</sup>

<sup>1</sup> Department of Plant Molecular Biology, University of Lausanne, Lausanne, Switzerland, <sup>2</sup> Neuchâtel Platform of Analytical Chemistry, University of Neuchâtel, Neuchâtel, Switzerland

Plants induce defense responses after insect egg deposition, but very little is known about the perception mechanisms. In *Arabidopsis thaliana*, eggs of the specialist insect *Pieris brassicae* trigger accumulation of reactive oxygen species (ROS) and salicylic acid (SA), followed by induction of defense genes and localized necrosis. Here, the involvement of the clade I L-type lectin receptor kinase LecRK-I.8 in these responses was studied. Expression of *LecRK-I.8* was upregulated at the site of *P. brassicae* oviposition and egg extract (EE) treatment. ROS, SA, cell death, and expression of *PR1* were substantially reduced in the *Arabidopsis* knock-out mutant *lecrk-I.8* after EE treatment. In addition, EE-induced systemic resistance against *Pseudomonas syringae* was abolished in *lecrk-I.8*. Expression of ten clade I homologs of *LecRK-I.8* was also induced by EE treatment, but single mutants displayed only weak alteration of EE-induced *PR1* expression. These results demonstrate that LecRK-I.8 is an early component of egg perception.

## OPEN ACCESS

### Edited by:

Akiko Sugio,  
INRA UMR1349 Institut de Génétique,  
Environnement et Protection des  
Plantes, France

### Reviewed by:

Zhonglin Mou,  
University of Florida, United States  
Milton Brian Traw,  
Nanjing University, China

### \*Correspondence:

Philippe Reymond  
Philippe.Reymond@unil.ch

### Specialty section:

This article was submitted to  
Plant Microbe Interactions,  
a section of the journal  
Frontiers in Plant Science

**Received:** 14 February 2019

**Accepted:** 26 April 2019

**Published:** 10 May 2019

### Citation:

Gouhier-Darimont C, Stahl E,  
Glauser G and Reymond P (2019) The  
*Arabidopsis* Lectin Receptor Kinase  
*LecRK-I.8* Is Involved in Insect Egg  
Perception. *Front. Plant Sci.* 10:623.  
doi: 10.3389/fpls.2019.00623

**Keywords:** *Arabidopsis thaliana*, lectin-like receptor kinase, oviposition, *Pieris brassicae*, *PR1* expression, herbivory

## INTRODUCTION

Herbivorous insects often deposit eggs on leaves and these seemingly inert structures have been shown to induce defense responses in different plant species (Reymond, 2013; Hilker and Fatouros, 2015). For example, direct defenses include localized hypersensitive response (HR)-like necrosis (Shapiro and DeVay, 1987; Balbyshev and Lorenzen, 1997; Fatouros et al., 2014; Griese et al., 2017), neoplasm formation (Doss et al., 2000; Petzold-Maxwell et al., 2011), production of ovicidal substances (Seino et al., 1996; Geuss et al., 2017), or tissue crushing (Desurmont et al., 2011), which all impair egg attachment or survival. In addition, oviposition-induced production of volatiles provides indirect defense by attracting egg parasitoids (Hilker et al., 2002; Fatouros et al., 2008; Büchel et al., 2011; Tamiru et al., 2011). Besides impacting egg survival, induced responses may also affect future success of hatching larvae. Indeed, reduced performance of larvae feeding on oviposited plants has been observed in pine (Beyaert et al., 2012), elm (Austel et al., 2016), *Nicotiana attenuate* (Bandoly et al., 2015, 2016), and Brassicaceae species (Pashalidou et al., 2012; Geiselhardt et al., 2013; Bonnet et al., 2017; Lortzing et al., 2019). However, this effect was not found with all tested insects and even an increased performance of a generalist insect feeding was reported in *Arabidopsis* (Bruessow et al., 2010; Pashalidou et al., 2012; Bandoly et al., 2016). Also, oviposition diminishes infection by bacterial pathogens, presumably for the benefit of hatching larvae (Hilfiker et al., 2014).

Although it is now clearly established that plants respond to oviposition, information on the nature of egg-associated cues that trigger the observed changes is scarce (Hilker and Fatouros, 2015). Bruchins are long-chain  $\alpha,\omega$ -diols purified from female bruchid beetles. They stimulate neoplasm formation on pea pods (Doss et al., 2000). Extracts from the female planthopper *Sogatella furcifera* contain various phospholipids that induce production of the ovidal substance benzyl benzoate in Japonica rice varieties (Seino et al., 1996; Yang et al., 2013). Benzyl cyanide is found in accessory reproductive glands from *Pieris brassicae* and induces leaf chemical changes that arrest an egg parasitoid on *Brassica oleracea* (Fatouros et al., 2008). Unknown proteins from oviduct secretions of the elm leaf beetle and the pine sawfly are responsible for egg-induced volatile emission (Meiners and Hilker, 2000; Hilker et al., 2005). Besides elicitors in secretions that are probably coating the egg surface, active molecules are also present within the egg. Crushed egg extract (EE) triggers neoplasm formation in pea (Doss et al., 1995) and arrest of parasitoids in maize (Salerno et al., 2013). EE from *P. brassicae* induces HR-like and expression of defense genes in *Arabidopsis* and *Brassica nigra* (Little et al., 2007; Bonnet et al., 2017). The activity is not proteinaceous and is enriched in the lipid fraction but a precise chemical characterization is still lacking (Bruessow et al., 2010; Gouhier-Darimont et al., 2013). Data thus indicate that various external and internal egg compounds activate defenses but how they reach a putative plant perception machinery is currently unknown.

The signal transduction pathway that links oviposition to downstream defense responses is starting to be unveiled. Reactive oxygen species (ROS) can be detected in oviposited or EE-treated plants, at the site of treatment (Little et al., 2007; Gouhier-Darimont et al., 2013; Bittner et al., 2017; Geuss et al., 2017). Salicylic acid (SA), a known signaling molecule in defense against biotroph pathogens, accumulates to high levels in response to insect eggs or EE in different plants, suggesting that the SA pathway is involved (Bruessow et al., 2010; Bonnet et al., 2017; Geuss et al., 2017; Lortzing et al., 2019). Indeed, the SA-responsive gene *PR1* is induced by oviposition (Little et al., 2007; Fatouros et al., 2014; Geuss et al., 2017) and its expression is abolished in SA-signaling *Arabidopsis* mutants *eds1-2*, *sid2-1*, and *npr1-1* (Gouhier-Darimont et al., 2013). EE-triggered *PR1* induction also depends on ROS accumulation but the nature of the ROS-generating process is still unknown, since *PR1* induction is still observed in mutants of NADPH oxidases (*rbohD/F*) that participate in pathogen-induced ROS production (Gouhier-Darimont et al., 2013). Ultimately, oviposition triggers a transcriptome signature that involves expression of many stress- and defense-related genes, and which is similar to SA-related transcriptomic responses to pathogens (Little et al., 2007; Fatouros et al., 2008; Büchel et al., 2011; Geuss et al., 2017; Drok et al., 2018). Furthermore, eggs from distantly related insect species induce the same defense genes, suggesting a common signaling pathway (Bruessow et al., 2010). Collectively, these findings are strikingly similar to the detection of pathogen-associated molecular patterns (PAMPs) by the plant innate immune

system, a process called pattern-triggered immunity (PTI) (Boller and Felix, 2009).

During plant pathogenesis, bacterial or fungal PAMPs are recognized by cell-surface pattern recognition receptors (PRRs) that constitute a large group of conserved proteins. These PRRs are receptor-like proteins (RLPs) or receptor-like kinases (RLKs) that share a transmembrane domain and a highly variable extracellular domain responsible for the specific binding of PAMPs. In addition, RLKs possess a cytosolic kinase domain (Boutrot and Zipfel, 2017). In *Arabidopsis*, hundreds of genes encode RLKs, and RLPs (Shiu et al., 2004), but only a handful of PRRs have been characterized, including the well-known flagellin and chitin receptors FLS2 and CERK1, respectively (Boutrot and Zipfel, 2017). To date, no PRR for an egg-associated elicitor has been identified. Previously, searching for RLKs that may be related to egg recognition in *Arabidopsis*, we discovered that a lectin receptor kinase, LecRK-I.8, was involved in the response to *P. brassicae* EE. *LecRK-I.8* was upregulated by oviposition and EE-treatment, and a T-DNA knock-out line exhibited a drastic reduction of EE-induced *PR1* expression (Little et al., 2007; Gouhier-Darimont et al., 2013). LecRK-I.8 is a L- (legume) type LecRK, whose family members have been associated with plant immunity (Singh and Zimmerli, 2013; Wang and Bouwmeester, 2017), and belongs to a subclade of eleven closely related members (Bellande et al., 2017). Here, we further investigated the role of *LecRK-I.8* and its homologs in *Arabidopsis* responses to *P. brassicae* eggs.

## MATERIALS AND METHODS

### Plant and Insect Material, Pathogens, and Growth Conditions

*Arabidopsis thaliana* Col-0 and mutant plants were grown in a growth chamber (Reymond et al., 2004) and were 4–5 week-old at the time of treatments. The *lecrk-I.8* T-DNA (SALK\_066416) mutant was described in Gouhier-Darimont et al. (2013). For other *lecrk* mutants, T-DNA insertion lines were obtained from the Nottingham *Arabidopsis* Stock Center. Specific forward and reverse primers were designed with SIGnAL T-DNA verification tool for all lines<sup>1</sup>. T-DNA lines and primers are listed in **Supplementary Table S1**.

A colony of *P. brassicae* was reared on *B. oleracea* var. *gemmifera* in a greenhouse (Bonnet et al., 2017). *Spodoptera littoralis* eggs were obtained from Syngenta (Stein, Switzerland).

### Cloning and Plant Transformation

For pLecRK-I.8::NLS-GFP-GUS reporter line, the *LecRK-I.8* promoter (795 bp) was amplified with Phusion enzyme (New England Biolabs) using specific primers (**Supplementary Table S1**) and cloned into pDONRP4-P1r (Thermo Fisher Scientific) to produce the Entry clone. Using the LR Clonase II (Thermo Fisher Scientific), the entry clone was cloned in the destination vector pMK7S\*NfM14GW,0 (Karimi et al., 2007). Plants were transformed using the floral-dip method

<sup>1</sup><http://signal.salk.edu/tdnaprimers.2.html>

(Clough and Bent, 1998) and selected on  $1/2$  MS agar containing 50  $\mu\text{g/ml}$  Kanamycin.

For complementation of *lecrk-I.8*, the *LecRK-I.8* promoter and coding sequence was amplified with Phusion enzyme (New England Biolabs) using specific primers (**Supplementary Table S1**). The *LecRK-I.8* amplicon (2769 bp) was cloned into a pGreenII0229-mVENUS plasmid containing the 3' OCS terminator. Transformants were selected on  $1/2$  MS agar containing 40  $\mu\text{g/ml}$  BASTA.

## Treatments

Egg extract preparation and application has been described previously (Bruessow et al., 2010; Gouhier-Darimont et al., 2013). In brief, *P. brassicae* eggs were crushed with a pestle in Eppendorf tubes. After centrifugation (15000  $g$  for 3 min), the supernatant (EE) was stored at  $-20^{\circ}\text{C}$ . Solid-phase extraction (SPE) was done as reported previously (Gouhier-Darimont et al., 2013). Total lipids were extracted with  $\text{CHCl}_3/\text{EtOH}$  (1:1, v/v), the solution evaporated in a speedvac, and the dried material resuspended in 10% dimethylsulphoxide (DMSO). Lipids were then loaded on a Sep-Pak C18-reverse phase cartridge (Waters AG, Baden, Switzerland) and eluted with 50% MeOH, followed by 80% MeOH, and 100% MeOH. The 100% MeOH fraction (SPE-F) was dried under a nitrogen flux, and resuspended at a concentration of 5  $\mu\text{g}/\mu\text{l}$  in 1% DMSO. For all experiments (except EE-induced SAR, see below), 2  $\mu\text{l}$  of EE (equivalent to one egg batch of 20–30 eggs), or SPE fraction was deposited on the abaxial side of fully developed leaves. For flagellin treatment, a solution of 100 nM flg22 (Peptron.com) was infiltrated in three leaves of each of three plants and leaves were collected after 20 h. Water infiltration was used as control. For natural egg deposition, plants were placed in a tent containing *P. brassicae* butterflies for 2–4 h. Oviposited plants were then transferred to a growth chamber for 96 h.

## Histochemical Staining and SA Measurements

Reactive oxygen species visualization and quantification was done as in Gouhier-Darimont et al. (2013). GUS staining was done as in Little et al. (2007). Two leaves of each of six plants were treated with EE and 10–12 leaves were harvested after 72 h for ROS analysis. SA analysis was performed by ultra-high performance liquid chromatography-tandem mass spectrometry (UHPLC-MS/MS) as reported previously (Bruessow et al., 2010; Glauser et al., 2014). Three leaves of each of six plants were treated with EE. After 0, 48 and 96 h, 15 leaf discs of 10 mm diameter (ca. 100 mg FW) were collected, ground in liquid nitrogen, spiked with 10  $\mu\text{L}$  of a 100 ng/mL solution of SA-d4 as internal standard, and extracted twice with a mixture of ethylacetate:formic acid (99.5:0.5, v/v). After evaporation, the dried residues were reconstituted in 100  $\mu\text{L}$  of methanol 70%. An aliquot of 5  $\mu\text{L}$  was injected in the UHPLC-MS/MS system (a 4000 QTRAP from ABSciex coupled to an Ultimate 3000 RS from Dionex). The mass spectrometer was operated in negative electrospray with the transitions  $m/z$  137>93 and 141>97 for SA and SA-d4, respectively. Free SA quantification was achieved

by internal calibration using 5 calibration points containing all SA-d4 at 10 ng/mL.

## Gene Expression Analysis

Two leaves of each of four plants were treated with EE. After 72 h, EE was carefully removed and leaf discs of 5 mm diameter were collected at the site of treatment. For each genotype, 6 leaf discs were used for RNA extraction and Quantitative RT-PCR analysis. Expression analysis of selected genes was described previously (Bruessow et al., 2010; Gouhier-Darimont et al., 2013). *SAND* (At2g28390) was used as a reference gene. The list of gene-specific primers can be found in **Supplementary Table S1**.

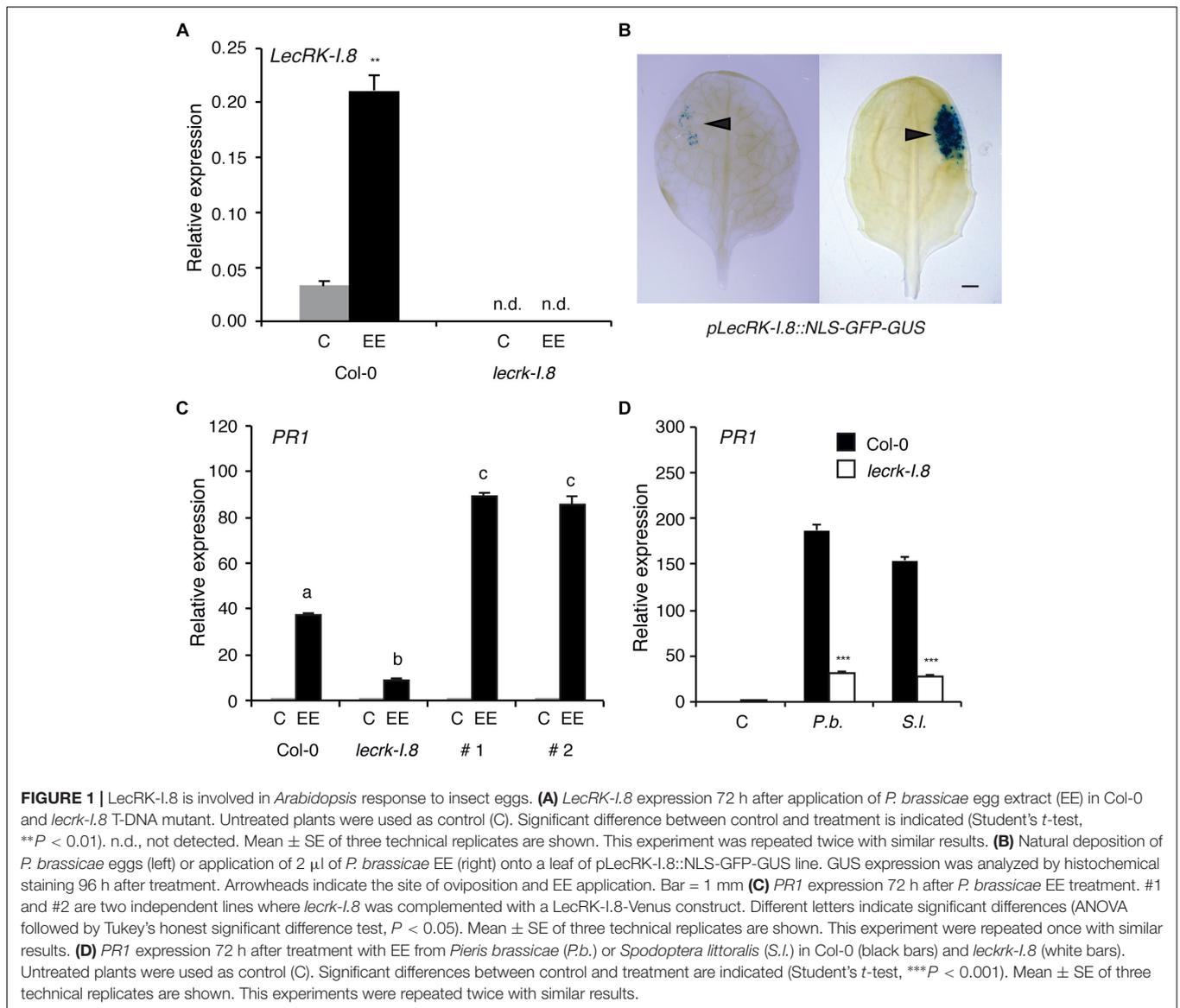
## EE-Induced SAR

SAR assay was performed as described previously (Hilfiker et al., 2014). *Pseudomonas syringae* pv. tomato DC3000 (*Pst*) was grown in King's B medium containing 50  $\mu\text{g/ml}$  rifampicin at  $28^{\circ}\text{C}$ . Overnight log phase cultures were washed three times with 10 mM  $\text{MgCl}_2$  and diluted to  $\text{OD}_{600}$  of 0.0005 for leaf inoculation. To induce SAR, three fully developed leaves of each of six Col-0 and *lecrk-I.8* plants were treated with 2  $\mu\text{l} \times 2 \mu\text{l}$  of EE from the abaxial side of the leaf. Five days after the treatment, EE was carefully removed with a brush and three untreated leaves distal to the site of EE treatment were inoculated with a suspension of *Pst* at  $\text{OD}_{600}$  0.0005 in 10 mM  $\text{MgCl}_2$  from the abaxial side with a 1 ml needleless syringe. The same amount of untreated plants was inoculated with *Pst* and served as controls. Growth of *Pst* in inoculated leaves was measured 48 h later by serial dilutions on LB plates.

## RESULTS

### Insect Eggs Trigger Local Expression of *LecRK-I.8*

Expression of *LecRK-I.8* (At5g60280) in response to *P. brassicae* EE treatment was monitored by QPCR and showed a more than fourfold increase 72 h after application (**Figure 1A**). A T-DNA knock-out line (*lecrk-I.8*, SALK\_066416) had no detectable *LecRK-I.8* expression in presence or absence of EE, confirming the KO nature of this mutant (**Figure 1A**). Using a promoter-NLS-GFP-GUS reporter line, we observed a strong activation of *LecRK-I.8* expression at the site of natural *P. brassicae* oviposition or at the site of EE treatment, indicating a precisely localized activation of this RLK (**Figure 1B**). As reported previously (Gouhier-Darimont et al., 2013), *P. brassicae* EE treatment triggered a substantial induction of the SA-marker gene *PR1*, and this response was significantly, although not fully, reduced in the *lecrk-I.8* mutant (**Figure 1C**). Similarly, induction of egg-responsive *CHIT*, *TI*, and *SAG13* (Little et al., 2007) was lower in *lecrk-I.8* (**Supplementary Figure S1**). To demonstrate that *LecRK-I.8* was directly responsible for the reduced expression of *PR1*, we generated *Arabidopsis* transgenic lines where *lecrk-I.8* was complemented with the *LecRK-I.8* gene under the control of its own promoter. In two independent lines, EE-dependent *PR1* induction was restored to even higher



levels than WT plants (Figure 1C). Finally, *PR1* induction in response to EE from *P. brassicae* or *S. littoralis* was similarly, reduced in *lecrk-I.8*, indicating that perception of eggs from two widely divergent herbivore species may depend on the same RLK (Figure 1D).

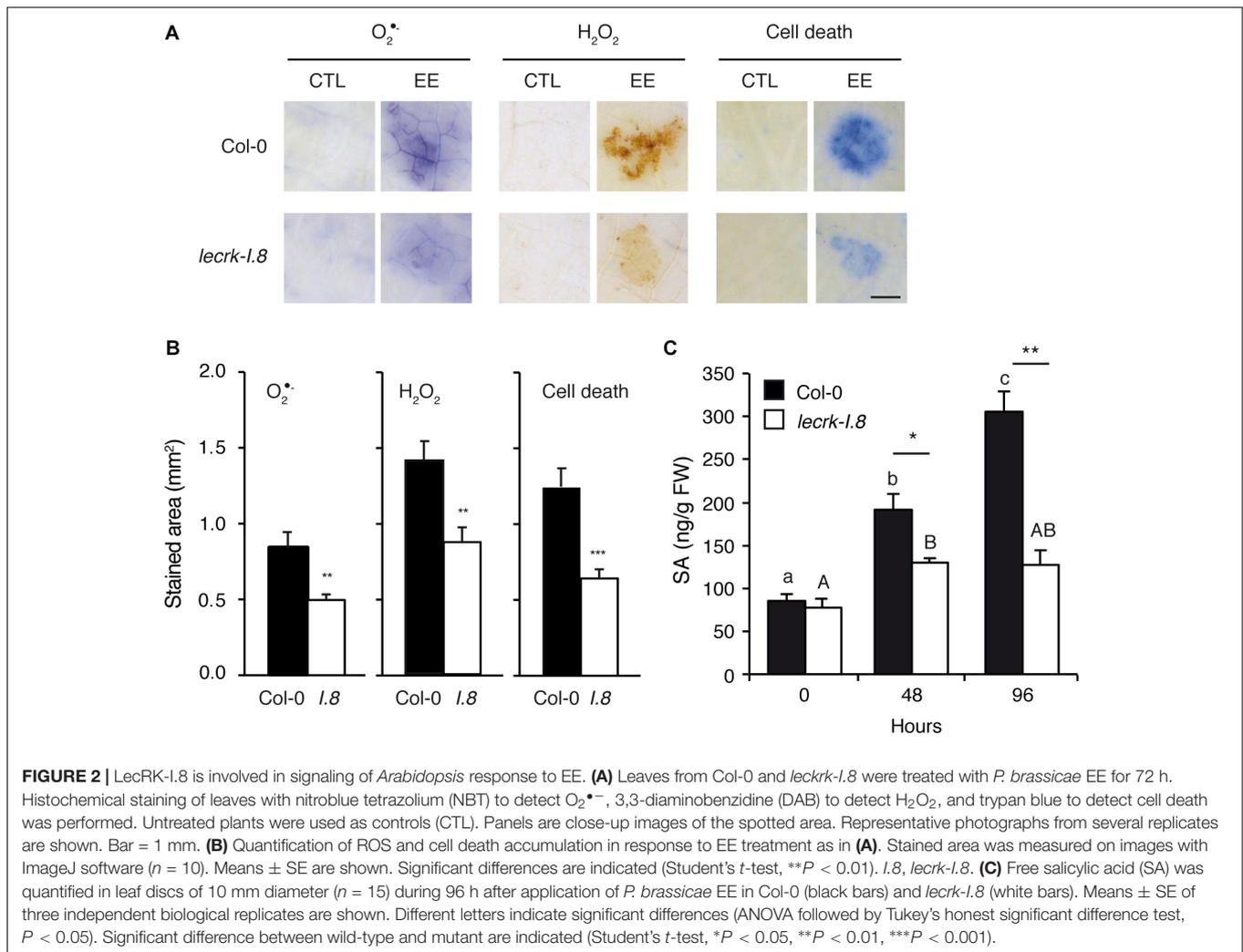
## LecRK-I.8 Modulates EE-Induced ROS and Cell Death

Oviposition triggers local ROS accumulation and cell death that depend on an intact SA pathway (Little et al., 2007; Gouhier-Darimont et al., 2013). We quantified  $O_2^{\bullet-}$  and  $H_2O_2$ , as well as cell death, in plants treated with EE for 72 h. Local accumulation of ROS and cell death was significantly reduced in *lecrk-I.8* compared to Col-0, implying that LecRK-I.8 plays an important role in this response (Figures 2A,B). However, the mutant exhibited ca. 50% of the wild-type response to EE

treatment, suggesting that other factors participate in ROS or cell death accumulation.

*Pieris brassicae* eggs or EE treatment induce a strong SA accumulation (Bruessow et al., 2010). We monitored free SA levels in Col-0 and *lecrk-I.8* from 0 to 4 days after EE treatment. At the start of the treatment, both genotypes had similar constitutive SA levels. However, the gradual EE-dependent increase of SA found in Col-0 was severely impaired in the mutant, although levels after 2 days of EE treatment were significantly higher than at time 0, indicating that a residual amount of SA can still accumulate in *lecrk-I.8* (Figure 2C). These results show that LecRK-I.8 is the main component controlling EE-induced SA accumulation.

We showed previously that total *P. brassicae* egg lipids and a lipidic fraction eluted with 100% MeOH from a SPE strongly activated *PR1* expression (Gouhier-Darimont et al., 2013). To test the specificity of LecRK-I.8 in response to active egg components,



we monitored cell death in naturally oviposited leaves and in leaves treated with EE or with the SPE fraction. Localized cell death was triggered by all treatments and significantly reduced in *lecrk-1.8* compared to Col-0 (Figure 3).

Because responses triggered by insect eggs resemble those induced during PTI, we assessed the role of LecRK-I.8 in PAMP-induced gene expression. After infiltration of the known PAMP flagellin (*flg22*), expression of *PR1*, *CHIT*, and *SAG13* was significantly induced in Col-0 but also to a similar extent in *lecrk-1.8*, suggesting that LecRK-I.8 is not required for flagellin perception but plays a specific role in egg perception (Figure 4).

### EE-Induced SAR Depends on LecRK-I.8

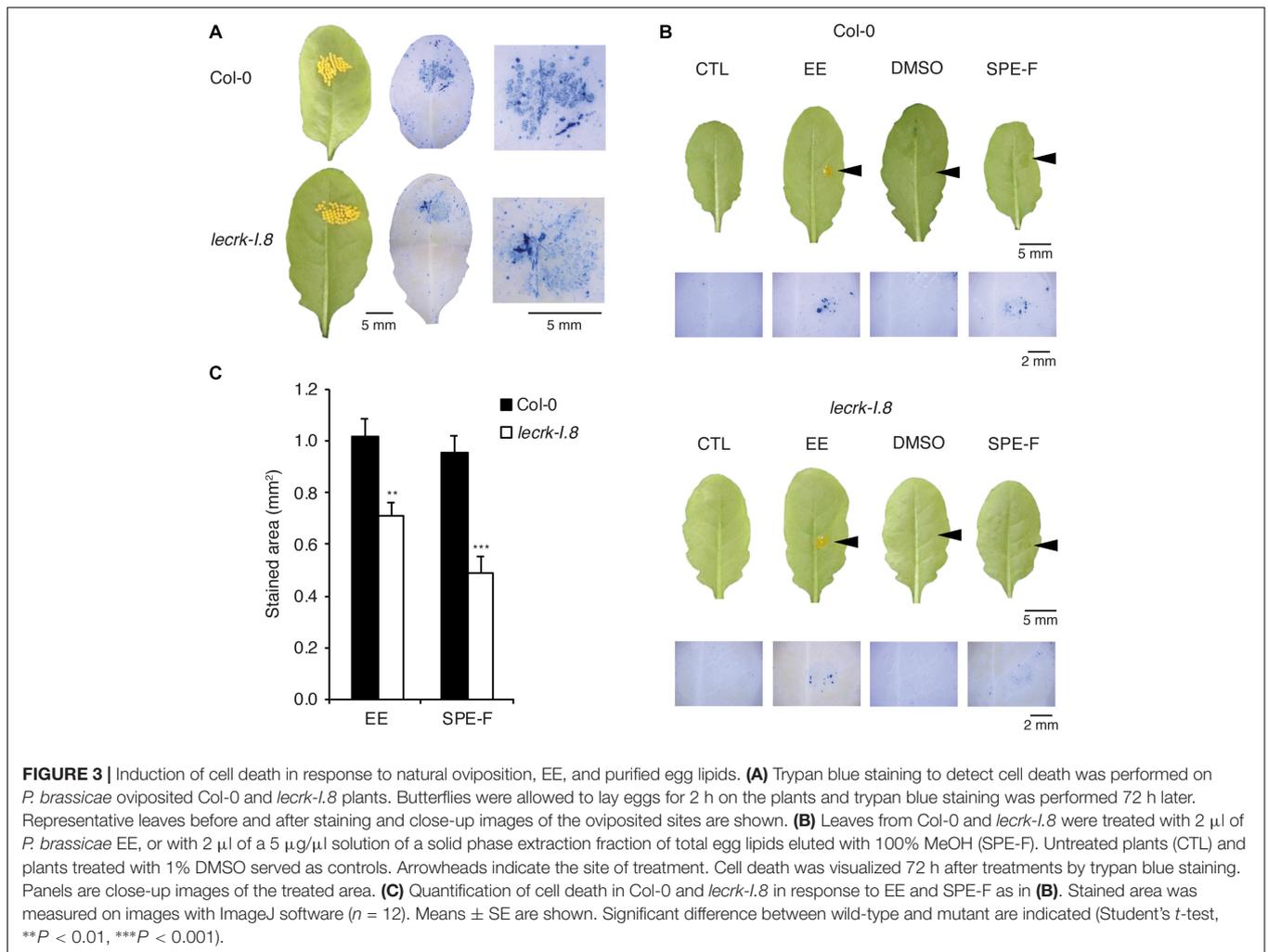
We previously found that oviposition by *P. brassicae* triggers a systemic acquired resistance (SAR) against the hemibiotrophic bacterial pathogen *P. syringae* (Hilfiker et al., 2014). To investigate the role of LecRK-I.8 in egg-induced SAR, we pretreated three *Arabidopsis* leaves with *P. brassicae* EE, and after 5 days three distal leaves were inoculated with *P. syringae* pv. *tomato* DC3000 (*Pst*). After 2 days, bacterial growth was monitored, and compared to control plants not treated

with EE. As reported previously, EE-pretreatment led to a significant inhibition of *Pst* growth in systemic leaves. Strikingly, this EE-induced SAR was abolished in *lecrk-1.8*, indicating that LecRK-I.8 is crucial for the establishment of EE-induced SAR (Figure 5).

### Role of LecRK-I.8 Homologs

LecRK-I.8 belongs to a subclade of 11 L-type LecRKs (Bellande et al., 2017). Since responses to EE tested in this study were not fully abolished in *lecrk-1.8*, we reasoned that this may be explained by some level of functional redundancy. We first assessed the expression of the 11 *LecRK-Is* in response to EE treatment. Like *LecRK-I.8*, all *LecRK-Is* genes were strongly up-regulated after 72 h of EE treatment (Figure 6A).

To investigate the role of each LecRK-Is in EE-induced gene expression, we obtained T-DNA mutants for all members, and quantitated *PR1* expression after EE treatment. Overall, none of the mutant except *lecrk-1.8* displayed a significantly altered *PR1* induction compared to Col-0, although there was a trend for reduced *PR1* induction in *lecrk-1.1* and *lecrk-1.4* (Figure 6B).

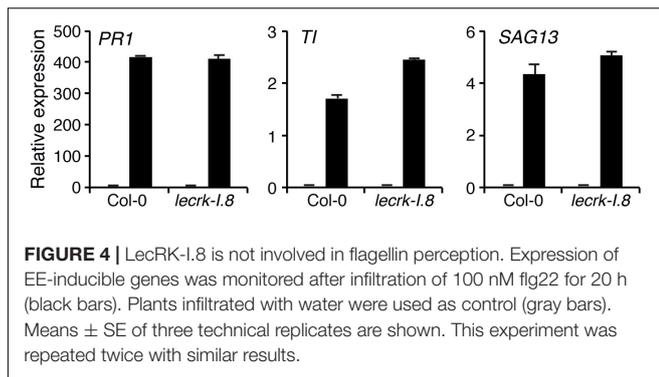


## DISCUSSION

Plants are equipped with a perception system to detect the presence of insect eggs and induce the accumulation of diverse signaling molecules including ROS and SA, followed by the activation of defense genes and localized cell death. Currently, very few insect-derived cues have been characterized and no plant receptor is known. We show here that a knock-out of the L-type lectin receptor kinase LecRK-I.8 is impaired in *Arabidopsis* responses to insect eggs. Indeed, EE-induced accumulation of the early signals  $O_2^-$  and  $H_2O_2$ , and of SA are significantly reduced in *lecrk-I.8*. In addition, expression of EE-inducible genes and localized cell death are also inhibited. These results indicate that LecRK-I.8 acts upstream of a signaling cascade that controls responses to oviposition. LecRK-I.8 is a plasma-membrane localized receptor kinase (Wang et al., 2017) and, as such, may well constitute a PRR for yet unknown egg-associated molecular patterns (EAMPs). Indeed, we show that a lipidic fraction from *P. brassicae* eggs triggers localized cell death and that this response is significantly attenuated in *lecrk-I.8*, suggesting that LecRK-I.8 is involved in the sensing

of an egg-derived lipidic compound. Testing this hypothesis will require the chemical identification of *P. brassicae* EAMPs and binding studies with LecRK-I.8 produced in heterologous systems. Alternatively, LecRK-I.8 may function as a co-receptor to modulate the activity of EAMP potential PRR(s). Searching for LecRK-I.8 interacting partners may help answering this question. Furthermore, although *Arabidopsis* response to insect eggs share similarities with PTI, the finding that flg22-induced *PR1* expression is not affected in *lecrk-I.8* suggests that LecRK-I.8 plays a specific role and further supports the idea that it is involved in EAMP perception.

Interestingly, expression of *LecRK-I.8* and its homologs is induced by EE treatment and experiments with the LecRK-I.8::NLS-GFP-GUS reporter line indicate that this activation is highly localized, at the site of egg deposition or EE treatment. Induced expression of PRR genes in response to PAMP treatment has been previously observed (Zipfel et al., 2006) and could represent a way to enhance the plant's ability to detect and respond to incoming pathogens. Here, the presence of eggs may as well stimulate a forward loop to increase the amount or number of potential LecRK receptors.

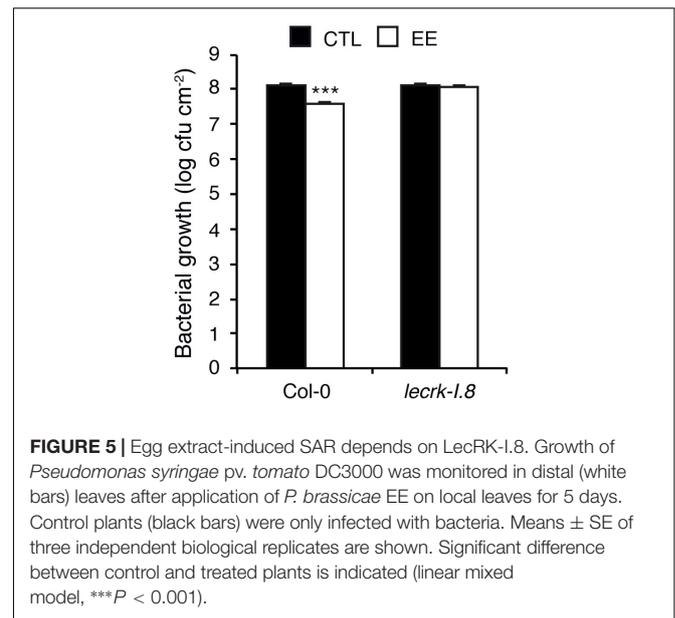


**FIGURE 4** | LecRK-I.8 is not involved in flagellin perception. Expression of EE-inducible genes was monitored after infiltration of 100 nM flg22 for 20 h (black bars). Plants infiltrated with water were used as control (gray bars). Means  $\pm$  SE of three technical replicates are shown. This experiment was repeated twice with similar results.

Generally, responses to oviposition in *Arabidopsis* have also been observed with EE treatment. Indeed, similar effects have been reported with both natural oviposition and EE treatment for defense gene expression, ROS production, cell death, SA accumulation, and EE-induced SAR (Little et al., 2007; Bruessow et al., 2010; Gouhier-Darimont et al., 2013; Hilfiker et al., 2014), strongly suggesting that EE treatment reflects natural oviposition. However, we cannot formally rule out that, in addition, intact eggs actively secrete elicitors or effectors that affect processes that have not yet been discovered. Capturing such molecules might be a challenge since eggs are firmly glued to the leaf surface. Current data indicate that passive diffusion of egg elicitors out of the egg into the leaf is the most parsimonious explanation for the observed responses. Once the exact chemical nature of the elicitor(s) will be obtained, further research should aim at understanding how they reach potential cell surface receptors.

Besides activating a signaling pathway that ultimately provokes an HR-like response and the expression of numerous defense genes, we previously reported that oviposition triggers a SAR that restricts bacterial growth in systemic leaves (Hilfiker et al., 2014). This phenomenon depends on a functional SA pathway and may constitute a strategy evolved by butterflies to protect the host on which eggs are deposited and will hatch (Hilfiker et al., 2014). Strikingly, we found here that EE-induced SAR is abolished in *lecrk-I.8*, in line with the lack of SA induction in the mutant. It thus appears that LecRK-I.8 is necessary for distinct responses to oviposition, confirming an involvement at the early phase of egg perception. Furthermore, the observation that the response to EE from two widely divergent insect species, *P. brassicae* and *S. littoralis*, is similarly impaired in *lecrk-I.8* strongly supports the notion that a generic EAMP is perceived by *Arabidopsis* and that this requires LecRK-I.8.

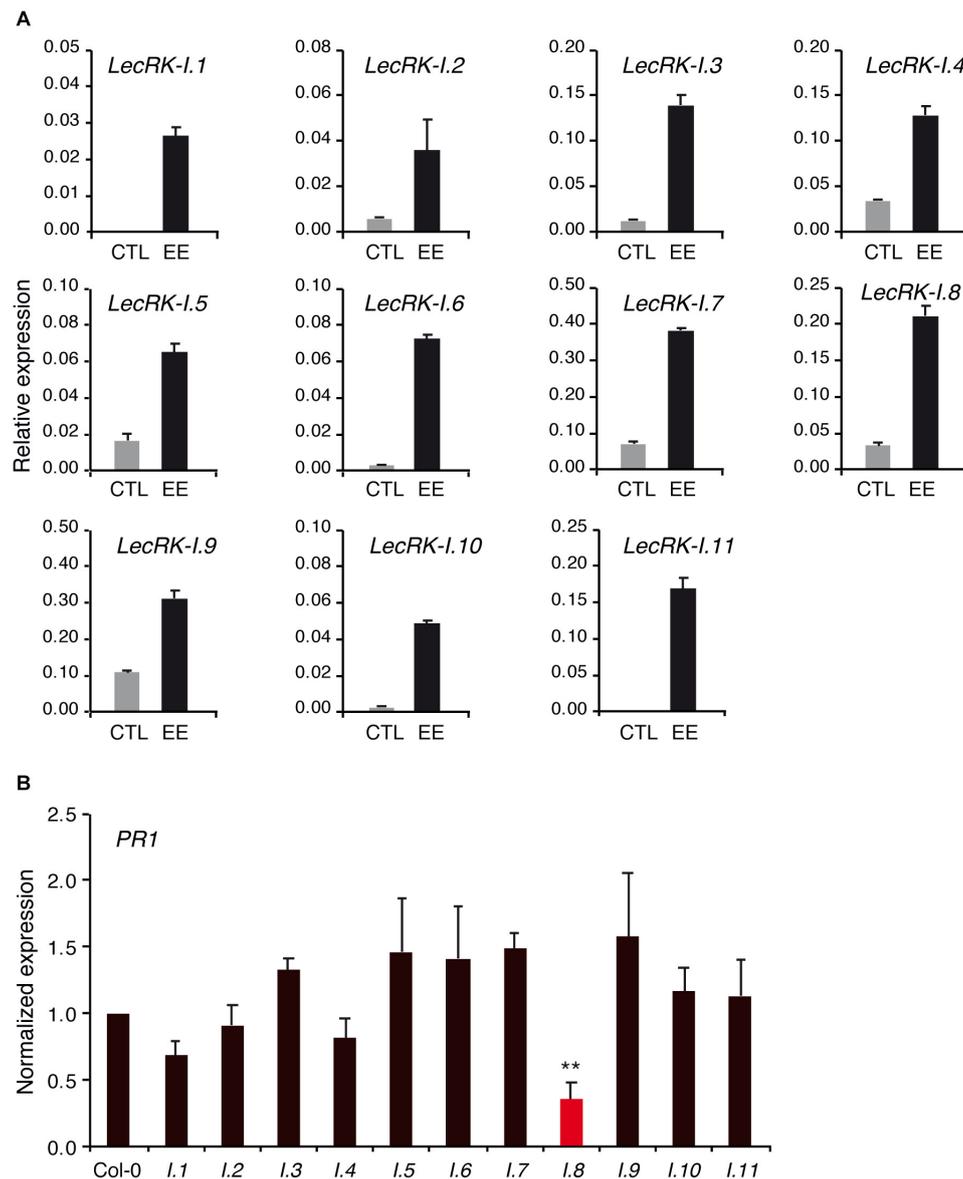
Although we demonstrate that LecRK-I.8 plays a significant role in *Arabidopsis* responses to eggs, expression of EE-inducible genes as well as ROS, SA, and cell death accumulation were not completely abolished in *lecrk-I.8*. At least two non-excluding hypotheses can explain these observations. First, plants contain a myriad of PRRs and specifically perceive different PAMPs from the same pathogen (Boutrot and Zipfel, 2017). It is conceivable that insect eggs release several EAMPs and that LecRK-I.8 is only perceiving one of them. As we are currently lacking a purified EAMP from *P. brassicae* eggs, we use a crude EE that may contain more than one active molecules. Second, all closely related



**FIGURE 5** | Egg extract-induced SAR depends on LecRK-I.8. Growth of *Pseudomonas syringae* pv. *tomato* DC3000 was monitored in distal (white bars) leaves after application of *P. brassicae* EE on local leaves for 5 days. Control plants (black bars) were only infected with bacteria. Means  $\pm$  SE of three independent biological replicates are shown. Significant difference between control and treated plants is indicated (linear mixed model, \*\*\* $P < 0.001$ ).

homologs of *LecRK-I.8* were induced by EE treatment, implying a role in perception. Although single mutants, except *lecrk-I.8*, are barely affected in EE-induced *PR1* expression, we cannot exclude some level of redundancy that may contribute to the residual responses in *lecrk-I.8*. Unfortunately, *LecRK-I.8* homologs are clustered in two loci of the *Arabidopsis* genome (Supplementary Figure S2), rendering the generation of higher order mutants by crossing difficult. Generating large deletions of *LecRK-Is* clusters by CRISPR-Cas9 technology may represent a useful strategy to test the role of these receptors in the responses to insect eggs.

Intriguingly, LecRK-I.8 was recently identified as a potential sensor for extracellular NAD<sup>+</sup> in *Arabidopsis* (Wang et al., 2017). Besides its role as an intracellular redox carrier that controls multiple metabolic reactions, including some defenses processes (Pétriacq et al., 2016), NAD(P) can be found in extracellular spaces after wounding or during pathogenesis (Zhang and Mou, 2009). Furthermore, exogenous application of NAD(P) triggers the expression of defense genes, including *PR1*, suggesting that perception of this extracellular signal could reinforce plant defenses (Zhang and Mou, 2009). Indeed, there is growing evidence that passive release of metabolites upon cell damage modulates innate immunity (Gust et al., 2017). Although the concentration of exogenous NAD<sup>+</sup> needed to trigger responses (millimolar range) is much higher than the binding affinity of LecRK-I.8 to NAD<sup>+</sup> (nanomolar range) (Wang et al., 2017), this finding raises the question of whether NAD<sup>+</sup> is involved in insect egg perception. Preliminary purification of *P. brassicae* EE has indicated that the active EAMP is present in a lipidic fraction that is unlikely to contain NAD<sup>+</sup> (Bruessow et al., 2010; Gouhier-Darimont et al., 2013). In addition, we show here that LecRK-I.8 is involved in the response to this lipidic fraction. Egg EAMP(s) could however trigger the release of extracellular NAD<sup>+</sup>, which would then be perceived by LecRK-I.8. Alternatively, we cannot formally exclude that LecRK-I.8 binds two different ligands. Future experiments should aim at clarifying these open questions.



**FIGURE 6 |** Role of *LecRK-I.8* homologs. **(A)** Expression of *LecRK-I.8* homologs 72 h after application of *P. brassicae* EE (black bars). Untreated plants were used as control (gray bars). Means  $\pm$  SE of three technical replicates are shown. This experiment was repeated twice with similar results **(B)** *PR1* expression in *leck* mutants 72 h after application of *P. brassicae* EE. Values were normalized to Col-0. Means  $\pm$  SE of three independent biological replicates are shown. Significant difference between Col-0 and each mutant is indicated (Student's *t*-test,  $**P < 0.01$ ).

Recent years have seen an emergence of studies implying LecRKs in plant immunity (Singh and Zimmerli, 2013; Wang and Bouwmeester, 2017). For instance, the closely related LecRK-I.9 mediates resistance to *Phytophthora brassicae* and *P. syringae* (Bouwmeester et al., 2011; Balagué et al., 2016). Interestingly, LecRK-I.9 was shown to bind extracellular ATP, in analogy with the NAD-binding property of LecRK-I.8 (Choi et al., 2014). Other members of clade I are also involved in defense against *Phytophthora* sp. or *Alternaria brassicicola* (Wang et al., 2014). LecRK-V.2, -V.5, -VI.2, -VII.1, and -IX.2 modulate PTI responses (Desclos-Theveniau et al., 2012;

Singh et al., 2012; Luo et al., 2017; Yekondi et al., 2018). In rice, a cluster of three G-type LecRKs confers resistance to the phloem-sucking brown planthopper (Liu et al., 2015). The *Arabidopsis* B-type LecRK LORE recognizes a bacterial PAMP lipopolysaccharide (Ranf et al., 2015). However, information about how LecRKs function at the molecular level and whether they act as PRRs or modulators of PRR signaling complexes is still lacking.

In conclusion, we have identified an important component of *Arabidopsis* perception system for insect eggs. LecRK-I.8 plays a role in early signal transduction steps and controls several

responses to *P. brassicae* eggs. Future studies should focus on identifying potential egg-derived ligands for LecRK-I.8 and investigating the occurrence of such ligand-receptor pair in other plant species, as well as in the context of different egg-plant interactions.

## DATA AVAILABILITY

All datasets for this study are included in the manuscript and/or the **Supplementary Files**.

## AUTHOR CONTRIBUTIONS

CG-D, ES, and PR designed and carried-out the experiments. GG quantified the salicylic acid. PR wrote the manuscript with the help of all authors.

## FUNDING

This work was supported by SNSF grant number 31003A\_169278.

## REFERENCES

- Austel, N., Eilers, E. J., Meiners, T., and Hilker, M. (2016). Elm leaves “warned” by insect egg deposition reduce survival of hatching larvae by a shift in their quantitative leaf metabolite pattern. *Plant Cell Environ.* 39, 366–376. doi: 10.1111/pce.12619
- Balagué, C., Gouget, A., Bouchez, O., Souriac, C., Haget, N., Boutet-Mercey, S., et al. (2016). The *Arabidopsis thaliana* lectin receptor kinase LecRK-I.9 is required for full resistance to *Pseudomonas syringae* and affects jasmonate signalling. *Mol. Plant Pathol.* 18, 937–948. doi: 10.1111/mpp.12457
- Balbyshev, N., and Lorenzen, J. (1997). Hypersensitivity and egg drop: a novel mechanism of host plant resistance to Colorado potato beetle (Coleoptera: Chrysomelidae). *J. Econ. Entomol.* 90, 652–657. doi: 10.1093/jee/90.2.652
- Bandoly, M., Grichnik, R., Hilker, M., and Steppuhn, A. (2016). Priming of anti-herbivore defence in *Nicotiana attenuata* by insect oviposition: herbivore-specific effects. *Plant Cell Environ.* 39, 848–859. doi: 10.1111/pce.12677
- Bandoly, M., Hilker, M., and Steppuhn, A. (2015). Oviposition by *Spodoptera exigua* on *Nicotiana attenuata* primes induced plant defence against larval herbivory. *Plant J.* 83, 661–672. doi: 10.1111/tj.12918
- Bellande, K., Bono, J.-J., Savelli, B., Jamet, E., and Canut, H. (2017). Plant lectins and lectin receptor-like kinases: how do they sense the outside? *Int. J. Mol. Sci.* 18:E1164. doi: 10.3390/ijms18061164
- Beyaert, I., Köpke, D., Stiller, J., Hammerbacher, A., Yoneya, K., Schmidt, A., et al. (2012). Can insect egg deposition “warn” a plant of future feeding damage by herbivorous larvae? *Proc. Biol. Sci.* 279, 101–108. doi: 10.1098/rspb.2011.0468
- Bittner, N., Trauer-Kizilelma, U., and Hilker, M. (2017). Early plant defence against insect attack: involvement of reactive oxygen species in plant responses to insect egg deposition. *Planta* 60, 1–15. doi: 10.1007/s00425-017-2654-3
- Boller, T., and Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu. Rev. Plant Biol.* 60, 379–406. doi: 10.1146/annurev.arplant.57.032905.105346
- Bonnet, C., Lassueur, S., Ponzio, C., Gols, R., Dicke, M., and Reymond, P. (2017). Combined biotic stresses trigger similar transcriptomic responses but contrasting resistance against a chewing herbivore in *Brassica nigra*. *BMC Plant Biol.* 17:127. doi: 10.1186/s12870-017-1074-7
- Boutrot, F., and Zipfel, C. (2017). Function, discovery, and exploitation of plant pattern recognition receptors for broad-spectrum disease resistance. *Annu. Rev. Phytopathol.* 55, 257–286. doi: 10.1146/annurev-phyto-080614-120106

## ACKNOWLEDGMENTS

We thank Oliver Kindler (Syngenta, Stein, Switzerland) for providing *S. littoralis* eggs, Blaise Tissot (Lausanne University) for help in growing plants, and Niko Geldner (Lausanne University) for the pGreenII plasmid.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2019.00623/full#supplementary-material>

**FIGURE S1** | Expression of EE-inducible genes in *lecrk-I.8*. Expression of *TI* (At1g73260), *CHIT* (At2g43570), and *SAG13* (At2g29350) was measured 72 h after application of *P. brassicae* EE (black bars). Untreated plants were used as control (gray bars). Means  $\pm$  SE of three technical replicates are shown. Significant difference between wild-type and mutant are indicated (Student's *t*-test, \*\*\**P* < 0.001). This experiment was repeated twice with similar results.

**FIGURE S2** | Position of *LecRK-I.8* homologs on *Arabidopsis* chromosomes 3 and 5.

**TABLE S1** | List of primers used in this study.

- Bouwmeester, K., de Sain, M., Weide, R., Gouget, A., Klammer, S., Canut, H., et al. (2011). The lectin receptor kinase LecRK-I.9 is a novel *Phytophthora* resistance component and a potential host target for a RXLR effector. *PLoS Pathog.* 7:e1001327. doi: 10.1371/journal.ppat.1001327
- Bruessow, F., Gouhier-Darimont, C., Buchala, A., Mettraux, J.-P., and Reymond, P. (2010). Insect eggs suppress plant defence against chewing herbivores. *Plant J.* 62, 876–885. doi: 10.1111/j.1365-313X.2010.04200.x
- Büchel, K., Malskies, S., Mayer, M., Fenning, T. M., Gershenzon, J., Hilker, M., et al. (2011). How plants give early herbivore alert: volatile terpenoids attract parasitoids to egg-infested elms. *Basic Appl. Ecol.* 12, 403–412. doi: 10.1016/j.baae.2011.06.002
- Choi, J., Tanaka, K., Cao, Y., Qi, Y., Qiu, J., Liang, Y., et al. (2014). Identification of a plant receptor for extracellular ATP. *Science* 343, 290–294. doi: 10.1126/science.343.6168.290
- Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x
- Desclous-Theveniau, M., Arnaud, D., Huang, T.-Y., Lin, G. J.-C., Chen, W.-Y., Lin, Y.-C., et al. (2012). The *Arabidopsis* lectin receptor kinase LecRK-V.5 represses stomatal immunity induced by *Pseudomonas syringae* pv. tomato DC3000. *PLoS Pathog.* 8:e1002513. doi: 10.1371/journal.ppat.1002513
- Desurmont, G. A., Donoghue, M. J., Clement, W. L., and Agrawal, A. A. (2011). Evolutionary history predicts plant defense against an invasive pest. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7070–7074. doi: 10.1073/pnas.1102891108
- Doss, R. P., Oliver, J. E., Proebsting, W. M., Potter, S. W., Kuy, S., Clement, S. L., et al. (2000). Bruchins: insect-derived plant regulators that stimulate neoplasm formation. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6218–6223. doi: 10.1073/pnas.110054697
- Doss, R. P., Proebsting, W. M., Potter, S. W., and Clement, S. L. (1995). Response of Np mutant of pea (*Pisum sativum* L.) to pea weevil (*Bruchus pisorum* L.) oviposition and extracts. *J. Chem. Ecol.* 21, 97–106. doi: 10.1007/BF02033665
- Drok, S., Bandoly, M., Stelzer, S., Lortzing, T., and Steppuhn, A. (2018). Moth oviposition shapes the species-specific transcriptional and phytohormonal response of *Nicotiana attenuata* to larval feeding. *Sci. Rep.* 8:10249. doi: 10.1038/s41598-018-28233-z
- Fatouros, N. E., Broekgaarden, C., Bukovinszkiné, G., van Loon, J. J. A., Mumm, R., Huijgens, M. E., et al. (2008). Male-derived butterfly anti-aphrodisiac mediates induced indirect plant defense. *Proc. Natl. Acad. Sci. U.S.A.* 105, 10033–10038. doi: 10.1073/pnas.0707809105

- Fatouros, N. E., Pineda, A., Huigens, M. E., Broekgaarden, C., Shimwela, M. M., Figueroa Candia, I. A., et al. (2014). Synergistic effects of direct and indirect defences on herbivore egg survival in a wild crucifer. *Proc. Biol. Sci.* 281:20141254. doi: 10.1098/rspb.2014.1254
- Geiselhardt, S., Yoneya, K., Blenn, B., Drechsler, N., Gershenzon, J., Kunze, R., et al. (2013). Egg laying of cabbage white butterfly (*Pieris brassicae*) on *Arabidopsis thaliana* affects subsequent performance of the larvae. *PLoS One* 8:e59661. doi: 10.1371/journal.pone.0059661
- Geuss, D., Stelzer, S., Lortzing, T., and Steppuhn, A. (2017). *Solanum dulcamara*'s response to eggs of an insect herbivore comprises ovidical hydrogen peroxide production. *Plant Cell Environ.* 40, 2663–2677. doi: 10.1111/pce.13015
- Glauser, G., Vallat, A., and Balmer, D. (2014). Hormone profiling. *Methods Mol. Biol.* 1062, 597–608. doi: 10.1007/978-1-62703-580-4\_31
- Gouhier-Darimont, C., Schmiesing, A., Bonnet, C., Lassueur, S., and Reymond, P. (2013). Signalling of *Arabidopsis thaliana* response to *Pieris brassicae* eggs shares similarities with PAMP-triggered immunity. *J. Exp. Bot.* 64, 665–674. doi: 10.1093/jxb/ers362
- Griese, E., Dicke, M., Hilker, M., and Fatouros, N. E. (2017). Plant response to butterfly eggs: inducibility, severity and success of egg-killing leaf necrosis depends on plant genotype and egg clustering. *Sci. Rep.* 7:7316. doi: 10.1038/s41598-017-06704-z
- Gust, A. A., Pruitt, R., and Nürnberger, T. (2017). Sensing danger: key to activating plant immunity. *Trends Plant Sci.* 22, 779–791. doi: 10.1016/j.tplants.2017.07.005
- Hilfiker, O., Groux, R., Bruessow, F., Kiefer, K., Zeier, J., and Reymond, P. (2014). Insect eggs induce a systemic acquired resistance in *Arabidopsis*. *Plant J.* 80, 1085–1094. doi: 10.1111/tplj.12707
- Hilker, M., and Fatouros, N. E. (2015). Plant responses to insect egg deposition. *Annu. Rev. Entomol.* 60, 493–515. doi: 10.1146/annurev-ento-010814-020620
- Hilker, M., Kobs, C., Varama, M., and Schrank, K. (2002). Insect egg deposition induces *Pinus sylvestris* to attract egg parasitoids. *J. Exp. Biol.* 205, 455–461. doi: 10.1242/jeb.01578
- Hilker, M., Stein, C., Schröder, R., Varama, M., and Mumm, R. (2005). Insect egg deposition induces defence responses in *Pinus sylvestris*: characterisation of the elicitor. *J. Exp. Biol.* 208(Pt 10), 1849–1854. doi: 10.1242/jeb.01578
- Karimi, M., Bleys, A., Vanderhaeghen, R., and Hilson, P. (2007). Building blocks for plant gene assembly. *Plant Physiol.* 145, 1183–1191. doi: 10.1104/pp.107.110411
- Little, D., Gouhier-Darimont, C., Bruessow, F., and Reymond, P. (2007). Oviposition by pierid butterflies triggers defense responses in *Arabidopsis*. *Plant Physiol.* 143, 784–800. doi: 10.1104/pp.106.090837
- Liu, Y., Wu, H., Chen, H., Liu, Y., He, J., Kang, H., et al. (2015). A gene cluster encoding lectin receptor kinases confers broad-spectrum and durable insect resistance in rice. *Nat. Biotechnol.* 33, 301–305. doi: 10.1038/nbt.3069
- Lortzing, V., Oberländer, J., Lortzing, T., Tohge, T., Steppuhn, A., Kunze, R., et al. (2019). Insect egg deposition renders plant defence against hatching larvae more effective in a salicylic acid-dependent manner. *Plant Cell Environ.* 42, 1019–1032. doi: 10.1111/pce.13447
- Luo, X., Xu, N., Huang, J., Gao, F., Zou, H., Boudsocq, M., et al. (2017). A lectin receptor-like kinase mediates pattern-triggered salicylic acid signaling. *Plant Physiol.* 174, 2501–2514. doi: 10.1104/pp.17.00404
- Meiners, T., and Hilker, M. (2000). Induction of plant synomones by oviposition of a phytophagous insect. *J. Chem. Ecol.* 26, 221–232. doi: 10.1023/A:100545383
- Pashalidou, F. G., Lucas-Barbosa, D., van Loon, J., Dicke, M., and Fatouros, N. (2012). Phenotypic plasticity of plant response to herbivore eggs: effects on resistance to caterpillars and plant development. *Ecology* 94, 702–713. doi: 10.1890/12-1561.1
- Pétriacq, P., Ton, J., Patrit, O., Tcherkez, G., and Gakière, B. (2016). NAD acts as an integral regulator of multiple defense layers. *Plant Physiol.* 172, 1465–1479. doi: 10.1104/pp.16.00780
- Petzold-Maxwell, J., Wong, S., Arellano, C., and Gould, F. (2011). Host plant direct defence against eggs of its specialist herbivore, *Heliothis subflexa*. *Ecol. Entomol.* 36, 700–708. doi: 10.1111/j.1365-2311.2011.01315.x
- Ranf, S., Gisch, N., Schäffer, M., Illig, T., Westphal, L., Knirel, Y. A., et al. (2015). A lectin S-domain receptor kinase mediates lipopolysaccharide sensing in *Arabidopsis thaliana*. *Nat. Immunol.* 16, 426–433. doi: 10.1038/ni.3124
- Reymond, P. (2013). Perception, signaling and molecular basis of oviposition-mediated plant responses. *Planta* 238, 247–258. doi: 10.1007/s00425-013-1908-y
- Reymond, P., Bodenhausen, N., Van Poecke, R. M. P., Krishnamurthy, V., Dicke, M., and Farmer, E. E. (2004). A conserved transcript pattern in response to a specialist and a generalist herbivore. *Plant Cell* 16, 3132–3147. doi: 10.1105/tpc.104.026120
- Salerno, G., De Santis, F., Iacovone, A., Bin, F., and Conti, E. (2013). Short-range cues mediate parasitoid searching behavior on maize: the role of oviposition-induced plant synomones. *Biol. Control* 64, 247–254. doi: 10.1016/j.biocontrol.2012.12.004
- Seino, Y., Suzuki, Y., and Sogawa, K. (1996). An ovidical substance produced by rice plants in response to oviposition by the whitebacked planthopper, *Sogatella furcifera* (HORVATH) (Homoptera: Delphacidae). *Appl. Entomol. Zool.* 31, 467–473. doi: 10.1303/aez.31.467
- Shapiro, A. M., and DeVay, J. E. (1987). Hypersensitivity reaction of *Brassica nigra* L. (Cruciferae) kills eggs of *Pieris* butterflies (Lepidoptera: Pieridae). *Oecologia* 71, 631–632. doi: 10.1007/bf00379310
- Shiu, S.-H., Karlowski, W. M., Pan, R., Tzeng, Y.-H., Mayer, K. F. X., and Li, W.-H. (2004). Comparative analysis of the receptor-like kinase family in *Arabidopsis* and rice. *Plant Cell* 16, 1220–1234. doi: 10.1105/tpc.020834
- Singh, P., Kuo, Y.-C., Mishra, S., Tsai, C.-H., Chien, C.-C., Chen, C.-W., et al. (2012). The lectin receptor kinase-VI.2 is required for priming and positively regulates *Arabidopsis* pattern-triggered immunity. *Plant Cell* 24, 1256–1270. doi: 10.1105/tpc.112.095778
- Singh, P., and Zimmerli, L. (2013). Lectin receptor kinases in plant innate immunity. *Front. Plant Sci.* 4:124. doi: 10.3389/fpls.2013.00124
- Tamiru, A., Bruce, T. J. A., Woodcock, C. M., Caulfield, J. C., Midega, C. A. O., Ogot, C. K. P. O., et al. (2011). Maize landraces recruit egg and larval parasitoids in response to egg deposition by a herbivore. *Ecol. Lett.* 14, 1075–1083. doi: 10.1111/j.1461-0248.2011.01674.x
- Wang, C., Zhou, M., Zhang, X., Yao, J., Zhang, Y., and Mou, Z. (2017). A lectin receptor kinase as a potential sensor for extracellular nicotinamide adenine dinucleotide in *Arabidopsis thaliana*. *eLife* 6:e25474. doi: 10.7554/eLife.25474
- Wang, Y., and Bouwmeester, K. (2017). L-type lectin receptor kinases: new forces in plant immunity. *PLoS Pathog.* 13:e1006433. doi: 10.1371/journal.ppat.1006433
- Wang, Y., Bouwmeester, K., Beseh, P., Shan, W., and Govers, F. (2014). Phenotypic analyses of *Arabidopsis* T-DNA insertion lines and expression profiling reveal that multiple L-type lectin receptor kinases are involved in plant immunity. *Mol. Plant Microbe Interact.* 27, 1390–1402. doi: 10.1094/MPMI-06-14-0191-R
- Yang, J.-O., Nakayama, N., Toda, K., Tebayashi, S., and Kim, C.-S. (2013). Elicitor(s) in *Sogatella furcifera* (Horváth) causing the Japanese rice plant (*Oryza sativa* L.) to induce the ovidical substance, benzyl benzoate. *Biosci. Biotechnol. Biochem.* 77, 1258–1261. doi: 10.1271/bbb.130055
- Yekondi, S., Liang, F.-C., Okuma, E., Radziejewski, A., Mai, H.-W., Swain, S., et al. (2018). Nonredundant functions of *Arabidopsis* LecRK-V.2 and LecRK-VII.1 in controlling stomatal immunity and jasmonate-mediated stomatal closure. *New Phytol.* 218, 253–268. doi: 10.1111/nph.14953
- Zhang, X., and Mou, Z. (2009). Extracellular pyridine nucleotides induce PR gene expression and disease resistance in *Arabidopsis*. *Plant J.* 57, 302–312. doi: 10.1111/j.1365-313X.2008.03687.x
- Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J. D. G., Boller, T., et al. (2006). Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. *Cell* 125, 749–760. doi: 10.1016/j.cell.2006.03.037

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Gouhier-Darimont, Stahl, Glauser and Reymond. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The Principal Salivary Gland Is the Primary Source of Digestive Enzymes in the Saliva of the Brown Marmorated Stink Bug, *Halyomorpha halys*

Sijun Liu<sup>1</sup> and Bryony C. Bonning<sup>2\*</sup>

<sup>1</sup> Department of Entomology, Iowa State University, Ames, IA, United States, <sup>2</sup> Department of Entomology and Nematology, University of Florida, Gainesville, FL, United States

## OPEN ACCESS

### Edited by:

Isgouhi Kaloshian,  
University of California, Riverside,  
United States

### Reviewed by:

Kelli Hoover,  
Pennsylvania State University,  
United States  
Jose Eduardo Serrão,  
Universidade Federal de Viçosa, Brazil

### \*Correspondence:

Bryony C. Bonning  
bbonning@ufl.edu

### Specialty section:

This article was submitted to  
Invertebrate Physiology,  
a section of the journal  
Frontiers in Physiology

**Received:** 19 August 2019

**Accepted:** 17 September 2019

**Published:** 11 October 2019

### Citation:

Liu S and Bonning BC (2019) The  
Principal Salivary Gland Is the Primary  
Source of Digestive Enzymes  
in the Saliva of the Brown  
Marmorated Stink Bug, *Halyomorpha*  
*halys*. *Front. Physiol.* 10:1255.  
doi: 10.3389/fphys.2019.01255

The brown marmorated stink bug, *Halyomorpha halys*, is an invasive, phytophagous stink bug of global importance for agriculture. Tissue-specific transcriptomic analysis of the accessory salivary gland, principal salivary gland (PSG) and gut resulted in identification of 234 putative protease and 166 putative nuclease sequences. By mapping the previously reported proteomes of *H. halys* watery saliva (WS) and sheath saliva to protein sequences translated from the assembled transcripts, 22 proteases and two nucleases in the saliva were identified. Of these, 19 proteases and both nucleases were present in the WS. The majority of proteases and nucleases found in WS were derived from the PSG, in line with ultrastructural observations, which suggest active protein synthesis and secretion by this tissue. The highly transcribed digestive proteases and nucleases of *H. halys* were similar to those of the southern green stink bug, *Nezara viridula*, indicating that these pentatomid stink bugs utilize a similar suite of proteases and nucleases for digestion of plant material. The comprehensive data set for the *H. halys* salivary glands and gut generated by this study provides an additional resource for further understanding of the biology of this pestiferous species.

**Keywords:** *Halyomorpha halys*, protease, nuclease, stink bug, salivary gland, gut, sheath saliva, watery saliva

## INTRODUCTION

The family Pentatomidae is comprised of 896 genera and 4,722 species of stink bugs (Rider, 2011), and includes multiple species that are significant pests of agriculture on a global scale (Panizzi et al., 2000). The pestiferous species include the brown marmorated stink bug, *Halyomorpha halys*, the southern green stink bug (SGSB), *Nezara viridula*, the green stink bug, *Acrosternum hilare*, and the brown stink bug, *Euschistus servus*. Management challenges are posed by their high reproductive capacity and by the development of resistance to the classical chemical insecticides used for suppression of stink bug populations (Leskey et al., 2012). Further complications are caused by the wide host range of many

**Abbreviations:** ASG, accessory salivary gland; BMSB, brown marmorated stink bug; PSG, principal salivary gland; SS, sheath saliva; WS, watery saliva.

stink bug species with damage resulting from the feeding of both nymphs and adults (Bergmann et al., 2013; Panizzi, 2015).

*H. halys*, is an East Asian species that spread into Europe and North America. First detected in the United States in the 1990s (Lee et al., 2013), *H. halys* has spread to most states, and is a serious pest in agriculture in addition to being a nuisance when overwintering inside homes and businesses (Biddinger et al., 2014; Leskey and Nielsen, 2018). *H. halys* can feed on more than 120 host plants (Bergmann et al., 2013; Haye et al., 2015), with the ability to feed on multiple plants important for development and survival. *H. halys* has caused dramatic losses in apple, peach, corn, peppers, tomatoes, and soybean (Biddinger et al., 2014). Management is primarily via chemical control (Kuhar and Kamminga, 2017) and pheromone-based attractants show promise (Weber et al., 2014; Rice et al., 2018).

Stink bugs feed by inserting their piercing-sucking mouthparts (stylets) into plant tissues (phloem or xylem) either by salivary sheath feeding or by physically rupturing cells (Backus et al., 2005; Lucini and Panizzi, 2018a,b). For salivary sheath feeding on phloem or xylem vessels, stink bugs secrete gelling or SS to form a flange at the site of penetration into the plant and a stabilizing sheath around the stylets (Lucini and Panizzi, 2018b). For both feeding strategies, WS is released to digest cell contents, and the predigested plant material is subsequently ingested. Further digestion occurs within the gut. The complementary digestive enzymes in the saliva and gut tissues result in efficient metabolic use of ingested plant material by the stink bug (Lomate and Bonning, 2016, 2018; Liu et al., 2018).

The WS produced by hemipteran insects was hypothesized to contain enzymes required for the digestion of plant proteins (Miles, 1964, 1972; Moreno et al., 2011). The *H. halys* WS and SS proteomes revealed distinct protein compositions (Peiffer and Felton, 2014), but few proteases and nucleases were identified from this study as genomic resources for *H. halys* were limited at the time. Since then, genomic resources for *H. halys* have significantly improved (The i5k Initiative, 2017). Two transcriptome studies of *H. halys* that characterized transcriptomes of whole insects at various developmental stages using different bioinformatics tools have been reported (Ioannidis et al., 2014; Sparks et al., 2014). In addition, we characterized the digestive proteases and nucleases of the southern green stink bug, *N. viridula*, at the biochemical, transcriptomic, and proteomic levels with a focus on the salivary gland (ASG and PSG) and gut tissues (Lomate and Bonning, 2016; Liu et al., 2018). The annotated *H. halys* genes provided a blueprint for our *N. viridula* transcriptomic and proteomic analyses. We also conducted a biochemical analysis of digestive enzymes in the same tissues of *H. halys* (Lomate and Bonning, 2018). These studies reinforced the complementary roles of the gut and salivary glands in producing different sets of enzymes for efficient digestion of plant materials by stink bugs.

The goals of this study were to (1) assess whether common digestive enzymes are used by different phytophagous stink bugs, and (2) determine the relative roles of the ASG and PSG in production of salivary enzymes. To this end, we conducted transcriptomic analysis of the *H. halys* salivary gland (ASG and PSG) and gut tissues. Transcripts for putative digestive

proteases and nucleases were identified and relative transcription levels determined. Transcripts for digestive enzymes were then translated, and *H. halys* and *N. viridula* proteomes mapped to the translated sequence dataset. This analysis allowed for further identification of secreted proteins including proteases and nucleases in the WS and SS. In addition to providing for comprehensive characterization of *H. halys* digestive enzymes, this study also allowed for comparison of enzyme types and transcription levels by tissue with those of *N. viridula*.

## MATERIALS AND METHODS

### Tissue Collection and RNA Isolation

The ASG, PSG, and gut tissues were dissected from one hundred *H. halys* adults. Tissues of each type were pooled and directly homogenized in Trizol reagent (Invitrogen, Carlsbad, CA, United States). Total RNA was isolated from the tissues according to the manufacturer's directions. The quality and integrity of the RNA samples was determined using a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, United States) and agarose gel electrophoresis.

### Preparation of cDNA Libraries and Illumina Sequencing

Three mRNA-Seq libraries derived from ASG, PSG, and gut were prepared by using the TruSeq RNA kit (Illumina Inc., San Diego, CA, United States) according to the manufacturer's instructions. Single-end sequencing was performed using the Illumina HiSeq2500™ (Illumina Inc., San Diego, CA, United States) to generate 100 base reads. Construction of the mRNA-Seq libraries and sequencing were performed by the DNA Facility at Iowa State University using standard procedures.

### Sequence Assembly, Data Analysis, and Bioinformatics

The quality of the raw sequence reads was examined using FASTQC<sup>1</sup> (Wingett and Andrews, 2018). Low quality reads and bases were trimmed using the FASTQ Quality Filter of the FASTx-toolkit<sup>2</sup>. Transcripts were *de novo* assembled using Trinity assembler (v2.1.1) (Haas et al., 2013). Reads per kilobase million (RPKM) were estimated using the "align\_and\_estimate\_abundance.pl" of Trinity software with RSEM (RNA-Seq by Expectation-Maximization) methods (Li and Dewey, 2011). Contigs of  $\geq 200$  nt were selected for further analysis. Sequence annotation for the assembled transcripts was performed using the BLASTx search engine against the NCBI non-redundant (nr) protein database. Gene ontology (GO) annotation of transcripts was achieved by use of the BLAST2GO software<sup>3</sup> (Conesa et al., 2005). Protease and nuclease transcripts were identified based on the identity of top hits from BLASTx analysis. Transcripts ( $\geq 300$  nt) with top hits of protease, proteinase,

<sup>1</sup>FASTQC <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

<sup>2</sup>FASTx-toolkit [http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)

<sup>3</sup>BLAST2GO <http://www.geneontology.org>

peptidase, or nuclease from the BLAST search were selected for further analysis.

The transcripts of putative protease and nuclease enzymes were further verified by BLASTp annotation. RNA and protein sequence alignments and other analyses such as sequence similarity and identity, were performed either by use of the multiple sequence alignment tool (Clustal Omega<sup>4</sup>) (Sievers and Higgins, 2018) or by use of BioEdit<sup>5</sup>.

Identification of conserved domains and putative function associated with the enzymes was conducted using the BLAST domain search. Putative enzymes with functions in the mitochondrion or with tRNA activity were excluded from analysis. The sequences of the selected transcripts were checked individually, and unique transcripts, including those with incomplete sequences, were determined by sequence analysis. The presence of a potential signal peptide encoded by full-length protease sequences was predicted using the web-based SignalIP 4.1 server<sup>6</sup> (Nielsen, 2017; Almagro Armenteros et al., 2019).

Raw sequence data were submitted to NCBI Sequence Read Archive (SRA BioProject: PRJNA560285).

## Mapping of Putative Protein Sequences to Proteomic Profiles Derived From *H. halys* and *N. viridula*

To identify putative proteases and nucleases expressed in the ASG, PSG and gut of *H. halys*, the putative protein sequences of  $\geq 100$  amino acids (aa) were translated using TransDecoder software<sup>7</sup>. Acquisition of proteomic data for *H. halys* WS and SS has been described previously (Peiffer and Felton, 2014) and these data were kindly provided for the current study by Drs. Michelle Peiffer and Gary Felton, Department of Entomology, Pennsylvania State University, United States. Methods for mapping the *H. halys* proteomics data and previously published *N. viridula* gut and salivary gland proteomics data (Lomate and Bonning, 2016) to putative protein sequences translated from assembled *H. halys* transcripts were adapted from Liu et al. (2018).

## Construction of Phylogenomic Trees

Proteases and nucleases identified from the salivary proteomes of *H. halys* were aligned to the NCBI nr database by BLASTp. For gene hits derived from insects and other arthropod species, full-length protein sequences were selected for investigation of their phylogenomic relationships. Protein sequences were aligned by MAFFT software (Kato et al., 2017). The resulting aligned sequences were entered into IQ-TREE version 1.6.7.1 (Nguyen et al., 2015) for construction of phylogenomic trees with maximum likelihood (ML) algorithms and 10,000 Ultrafast bootstrap approximation (Minh et al., 2013; Hoang et al., 2018). The best fit model for the ML tree was determined using the Bayesian information criterion by ModelFinder implemented in IQ-TREE (Kalyaanamoorthy et al., 2017). The resulting ML tree

files were uploaded to iTOL (Letunic and Bork, 2019) for editing. Trees were presented as mid-point rooted trees.

## RESULTS

### Assembly and Annotation of the *H. halys* Tissue Transcriptomes

Deep sequencing of the transcriptomes isolated from ASG, PSG, and gut of *H. halys* resulted in generation of 66.5 (PSG), and 81.7 (ASG) million single-end reads. The raw reads were trimmed and the resulting high-quality reads were used for assembly of transcripts. Transcripts from the ASG, PSG, and gut were assembled separately. The numbers of transcripts assembled (contigs) for each sample are shown in **Table 1**. The numbers of contigs encoding putative peptides of  $\geq 100$  aa were 22,185 ( $\sim 30\%$  of total ASG contigs of  $> 200$  nt), 16,745 (42% of PSG contigs), and 20,240 (36% of gut-derived contigs). A summary of statistics for assembly of the transcriptomes is provided in **Table 1**.

Initial annotation of the assembled transcripts was performed by BLASTx search against the NCBI nr database at an E-value of  $1 \times 10^{-3}$ . The numbers of annotated contigs were 31,523 (42%) for ASG, 23,528 (59%) for PSG and 28,234 (50%) for the gut. The top hit sequences were derived from 746 species for ASG, 450 species for PSG, and 630 species for the gut transcriptome. As expected, the majority ( $> 74\%$ ) of the transcripts hit predicted genes of *H. halys* (**Figure 1**). The E-values for *H. halys* hits were  $< 1 \times 10^{-20}$  (data not shown). The proportion of transcripts for the top 10 species hit by *H. halys* transcripts are shown in **Figure 1**. For all three tissues, the organism with the second highest number of hits (5.7–7.6% of transcripts) was *Nosema*, a symbiont commonly associated with stink bugs (Sparks et al., 2014; Hajek et al., 2017). Transcripts that hit sequences of *Candidatus Pantoea carbekii*, a primary gut symbiont of *H. halys*, were only found in the gut transcriptome. Approximately 4% of the transcripts were derived from a *Candidatus* species (**Figure 1**). The Gene Ontology (GO) annotation, with transcripts grouped by functions of “Biological process,” “Cellular component,” and “Molecular function” is

**TABLE 1** | Summary of *H. halys* ASG, PSG, and gut transcriptome statistics.

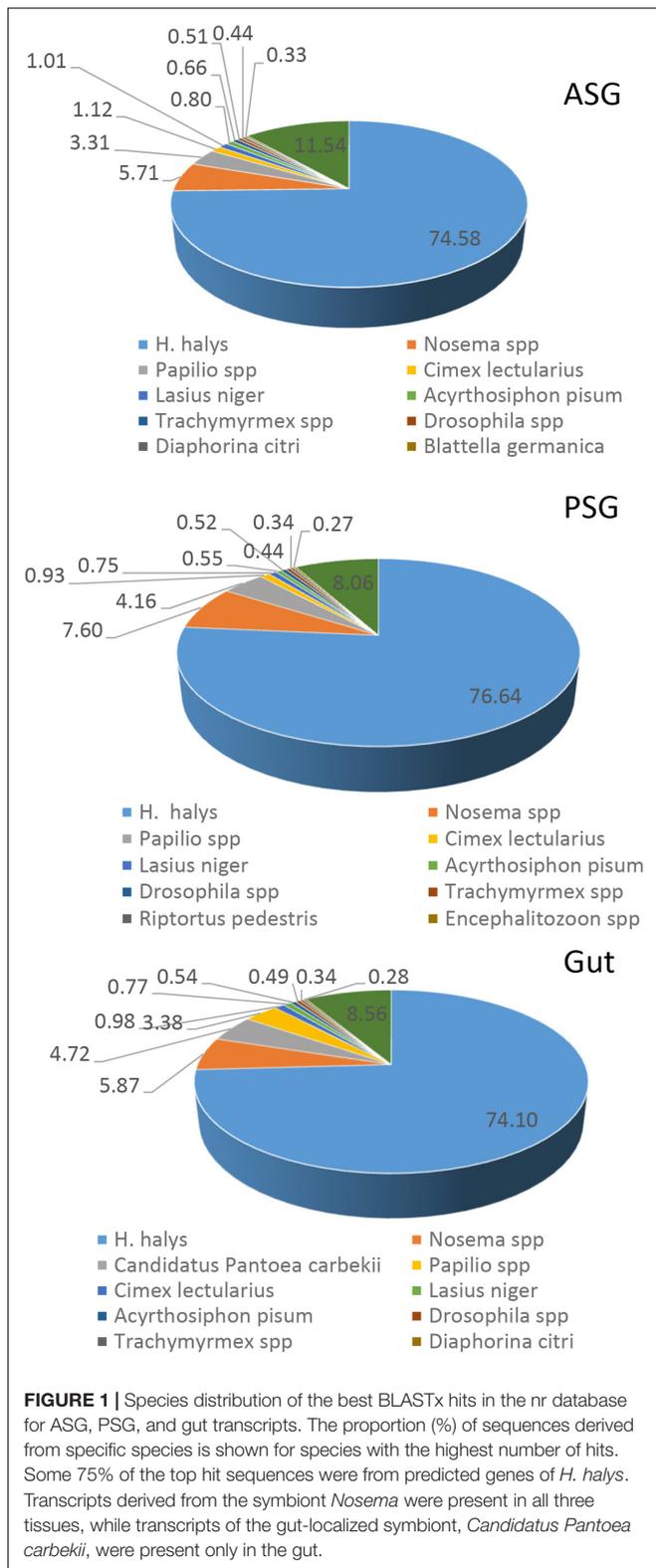
Tissue	ASG	PSG	Gut
Total raw reads (million)	81.7	66.5	77.1
Reads after trimming (million)	72.3	61.5	71.7
Total no. of contigs ( $\geq 100$ nt)	74,632	39,684	55,967
Total length (nt)	54,253,793	26,160,907	40,490,240
Mean length (nt)	654	659	723
N50	1,114	916	1106
No. annotated contigs	31,523	23,528	28,234
% annotated contigs	42	60	50
No. contigs encoding potential proteins ( $\geq 100$ aa)	22,185	16,745	20,240
No of translated peptide sequences ( $\geq 100$ aa)	28,724	20,466	25,226

<sup>4</sup>Clustal Omega; <http://www.ebi.ac.uk/Tools/msa/clustalo/>

<sup>5</sup>BioEdit <https://softfamously.com/bioedit/>

<sup>6</sup>SignalIP 4.1 server <http://www.cbs.dtu.dk/services/SignalP/>

<sup>7</sup>TransDecoder <https://transdecoder.github.io>



summarized in **Figure 2**. The annotated transcripts derived from the three tissues were comprised of similar numbers of GO terms in each functional category.

## Identification of *H. halys* Protease and Nuclease Transcripts

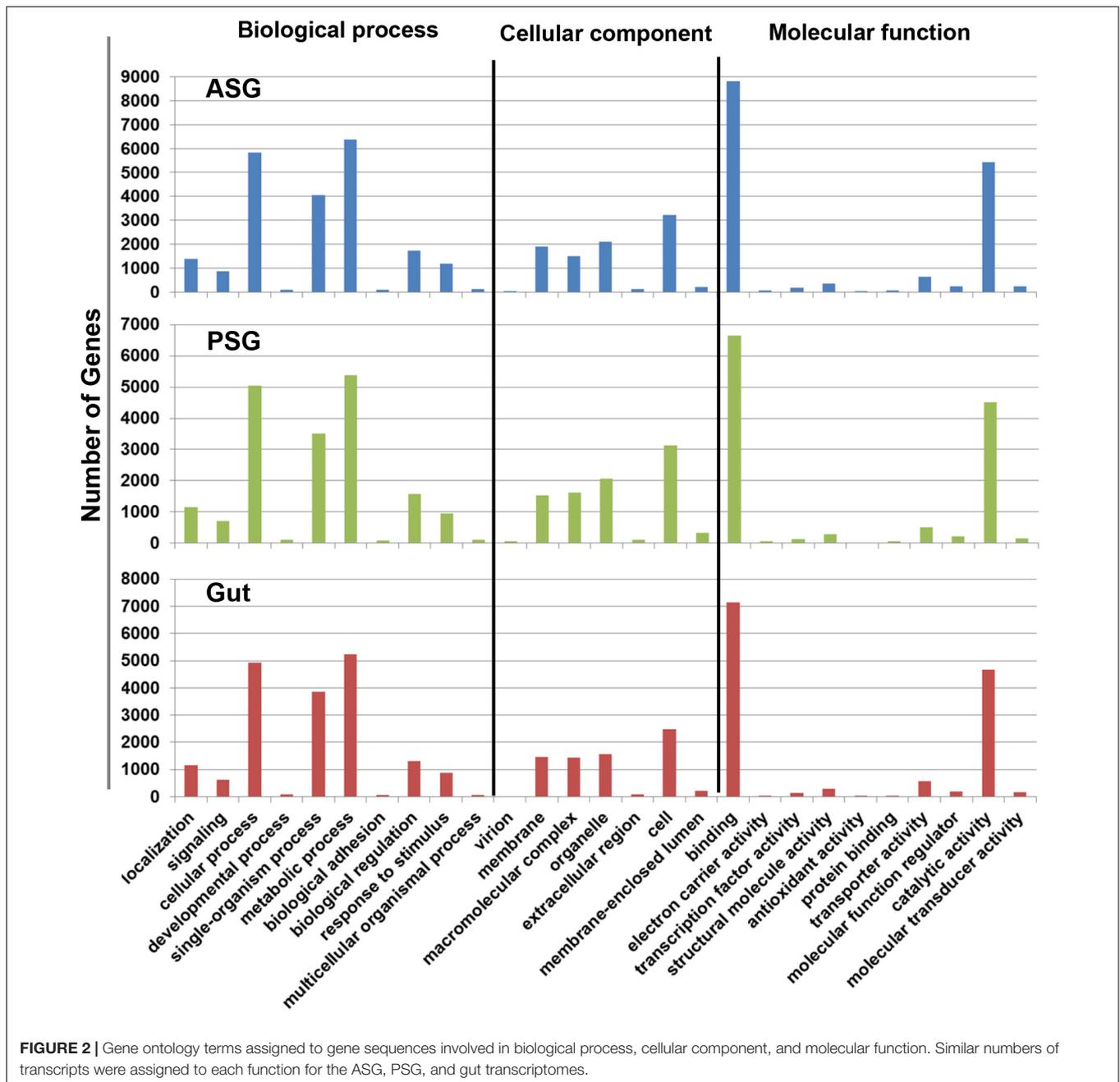
From the BLAST annotation results and sequence analysis we were able to identify unique transcripts of 234 putative proteases and 166 putative nucleases. The majority of the protease and nuclease transcripts identified were full- or near full-length. The proteases and nucleases identified are listed in **Supplementary Tables S1, S2** respectively, along with relative levels of transcription (RPKM). A summary of the different categories of protease and nuclease transcripts identified in the *H. halys* tissues is presented in **Table 2**. Among the proteases, contigs of 44 aminopeptidases, 55 peptidases, 59 cathepsin-like/cysteine protease, and 48 trypsin-like/serine proteases were identified. In addition to the 211 protease sequences derived from *H. halys*, 23 proteases were apparently derived from symbionts of *Nosema* or *C. Pantoea* (including six genes that hit *Papilio xuthus*) (**Supplementary Table S1**).

One hundred sixty-six putative nucleases were identified from the three transcriptomes (**Supplementary Table S2**). Of these, 113 of the contigs hit *H. halys* genes, 50 hit nucleases of symbionts, bacteria or microsporidia and three were from other insects. Fewer full-length sequences were acquired for putative nucleases from the transcriptomes, likely due to the lower levels of transcription relative to protease enzymes (**Supplementary Table S2**). Remarkably, 41% of the unique nuclease sequences appeared to be derived from symbionts, in contrast to 8.3% of the protease sequences that hit symbiont genes (**Table 2**).

## Mapping of *N. viridula* Proteomes to Predicted *H. halys* Protein Sequences

The three sets of assembled *H. halys* tissue-derived transcripts were translated and the resulting protein sequences ( $\geq 100$  aa) were used for mapping of proteome-derived peptide sequences. Proteomics libraries derived from the salivary gland (SG) and gut of *N. viridula* (Liu et al., 2018) were used for mapping. The proteomics profiles of *N. viridula* were useful for identification of *H. halys* proteins based on the high protein sequence identities between these two species. Peptide mapping results for the *N. viridula* proteomes are shown in **Figure 3**. From 8 to 12% of the *H. halys* predicted protein sequences were mapped by peptides derived from the *N. viridula* salivary gland (SG) proteome, while only 3% of the *H. halys* protein sequences were mapped by *N. viridula* gut proteins.

A total of 113 WS and 92 SS proteins mapped to the predicted protein sequences derived from the assembled *H. halys* transcripts, although some of the mapping results had low sequence coverage (**Supplementary Tables S3, S4**). Comparison of the proteins mapped by WS and SS peptides revealed that only 24 proteins were common to both WS and SS, with 89 and 68 proteins unique for WS and SS, respectively. The differences in the primary components of WS and SS likely reflect the respective biological functions of the WS and SS. The functions of 24 WS proteins and 22 SS proteins were unknown with either no hits or hits to uncharacterized *H. halys* proteins. The proteins common to the two salivary proteomes, many of which are involved in digestive processes, included amylases, carbonic anhydrases,



chitinases, glycosidase, lectins, lipases, proteases, and nucleases. WS proteins included two proteins derived from *C. Pantoea*.

## Proteases and Nucleases Identified From *H. halys* Watery Saliva and Sheath Saliva Proteomes

Proteomics libraries derived WS and SS of *H. halys* (Peiffer and Felton, 2014) were next mapped to the *H. halys* predicted protein sequences. Putative proteases and nucleases identified from mapping of WS and SS peptides to predicted protein sequences derived from the assembled *H. halys* transcripts are

listed in **Table 3**. In total, 22 proteases, one ribonuclease, and one potential nuclease were identified from the saliva of *H. halys*. Notably, no aminopeptidases were identified from either the WS or SS protein profiles. The proteases found in WS were peptidases (two carboxypeptidase B-like), cathepsin-like (two cathepsin L1-like), chymotrypsins (three) and trypsin-like serine proteases (14), while no chymotrypsin-like proteases were identified from SS. Only four (peptidase-5, trypsin-42, -45, -50) were found in both WS and SS, with 15 and three proteases being unique to WS and SS, respectively. Signal peptides were predicted for all of the proteases identified with complete N-terminal sequences, confirming secretion of these

**TABLE 2** | Proteases and nucleases identified from *H. halys* transcriptomes.

Enzyme	No. of top hits			Unique hits	
	Total	<i>H. halys</i>	Symbiont	<i>H. halys</i>	Symbiont
<b>Protease</b>					
Aminopeptidase	44	31	13*	28	10
Peptidase	55	50	5	46	3
Cathepsin-like/cysteine protease	59	59	0	54	0
Chymotrypsin	5	5	0	5	0
Trypsin-like/serine protease	48	48	0	45	0
Other putative protease	25	22	5	20	5
Total proteases	239	215	23	198	18
<b>Nuclease</b>					
Endonuclease	24	10	14	7	14
Exonuclease	20	11	9	8	9
Nucleotidase	20	18	1	10	1
Nuclease	21	16	5	8	5
Endoribonuclease	9	9	0	3	0
Exoribonuclease	19	11	8	8	8
Ribonuclease	53	38	13	25	13
Total nucleases	166	113	50	69	50

\*Hits assigned to *Papilio xuthus* appear to be derived from *Nosema* spp.

proteases from the salivary gland into saliva (Table 3). The RPKM values indicate that most of the enzyme transcripts with RPKM of >1000 were produced by the PSG, the exceptions being cathepsin-25 in the gut, and trypsin-48 in the ASG. Three proteases were transcribed at very high levels with RPKM > 10,000 with two (cathepsin-25 and trypsin-44) located in WS, and one (trypsin-45) in SS (Table 3). Surprisingly, only one nuclease (ribonuclease-31) and two uncharacterized nucleases (uncharacterized nuclease\_f410 and uncharacterized nuclease\_f435, which hit LOC106684787 LOC106684787/venom nuclease-like protein 1) (Supplementary Table S4), were identified from the WS of *H. halys* (Table 3).

## The Transcripts of Proteases and Nucleases Identified in the WS and SS Proteomes Were Highly Expressed

To determine relative transcript levels, the RPKM distributions of transcripts from each tissue (ASG, PSG, and gut) were determined. Similar RPKM distribution patterns were observed in the ASG, PSG, and gut transcriptomes: The RPKM values of ~50% of the transcripts were less than 1.5, and ~75% were less than 3, demonstrating that the majority of the transcripts were expressed at relatively low levels. In contrast, less than 2% of the transcripts had an RPKM of more than 100. Most of the proteases and nucleases had high RPKM in either PSG or ASG. The exceptions to this were cathepsin-25 and cathepsin-57, with high transcription levels in the gut (Table 3). Remarkably, the transcripts of 16 (67%) proteins had the highest RPKM in PSG. Seventy-five% of the identified putative enzymes had an RPKM of more than 500 in either PSG, ASG, or gut tissues. These results are consistent with our previous observation that proteins identified

in the *N. viridula* tissue proteomes were derived from genes with high transcription levels (Liu et al., 2018).

## Comparison of Highly Transcribed Proteases and Nucleases in *N. viridula* and *H. halys*

*Halyomorpha halys* and *N. viridula* both belong to the family Pentatomidae and have highly homologous genes (Liu et al., 2018). To compare transcription of proteases and nucleases from these two species, we selected enzymes with the highest RPKM values of  $\geq 100$  for proteases and  $\geq 20$  for nucleases. In total 66 putative enzymes (54 proteases and 12 nucleases) were selected. The *N. viridula* counterparts of the selected *H. halys* proteins shared 60–97% sequence identities (Supplementary Table S5). The heat map of RPKM demonstrated that the vast majority of transcripts from the two stink bug species had similar transcription profiles (Supplementary Table S5). For example, peptidases, trypsin, and chymotrypsins were highly transcribed in the PSG, while cathepsins were primarily transcribed in the gut. Only a few chymotrypsins and trypsin showed moderate transcription levels in the gut for both *H. halys* and *N. viridula*. The only exception is *H. halys* cathepsin-53 which was highly expressed in ASG, but not in PSG or gut.

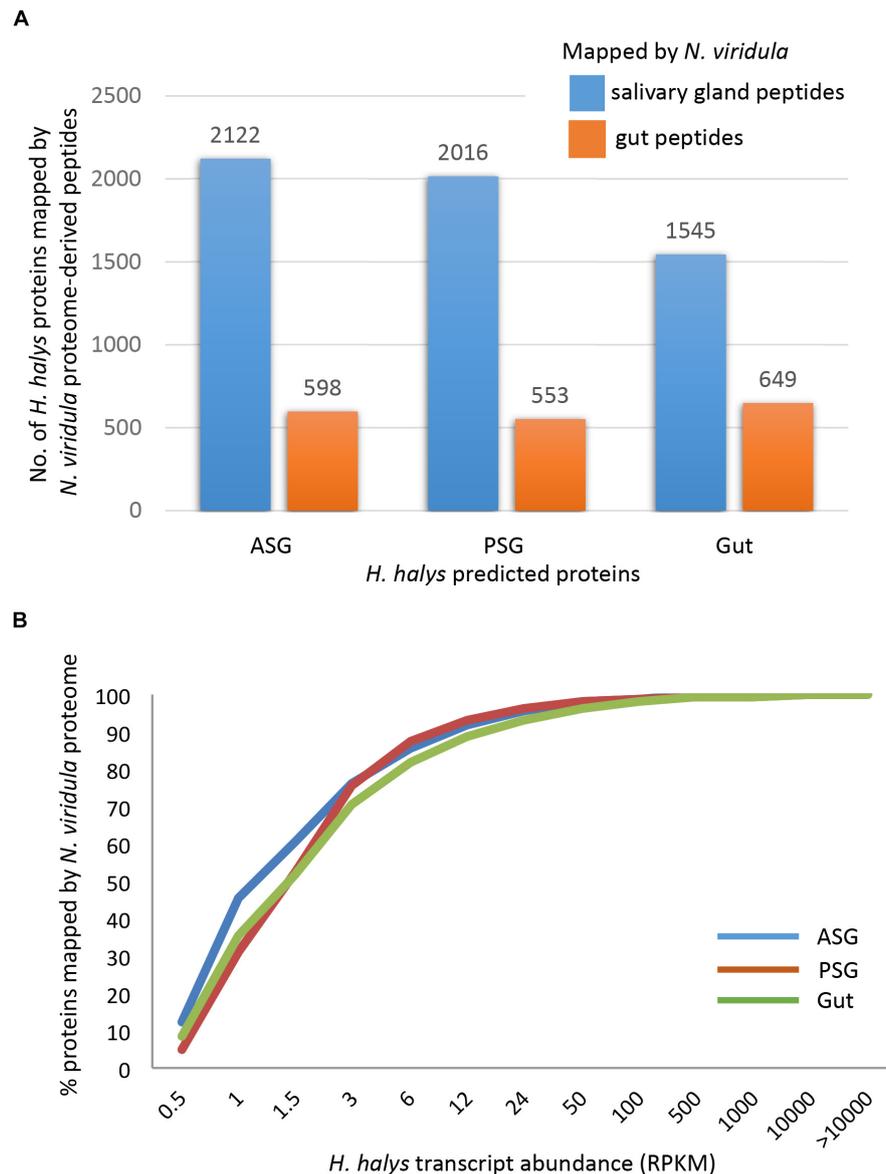
Nuclease transcription was lower overall than protease transcription levels, and nuclease transcription was generally higher in the ASG and gut tissues. Only ribonuclease-31 (ribonuclease Oy-like) was transcribed at a high level in the PSG and was detected in WS. Eight proteases and two nucleases identified in *H. halys* were not identified from *N. viridula* (Supplementary Table S5).

## Analysis of Proteases and Nucleases Identified From WS and SS Peptidases

The two peptidases, peptidase-5/XP\_014275318.2 and peptidase-9/XP\_014277059.1, are both carboxypeptidase B-like with 44% sequence identity. Similar conserved domains, e.g., *propep\_M14\_superfamily* and *Peptidase\_M14\_like\_superfamily* (domain accessions: *cd03860*, *smart00631*, *pfam00246*, *pfam002244*, and *COG2866*) were identified in these two peptidases. The transcription of peptidase-5 was nearly 10-fold higher than that of peptidase-9 in PSG (Table 3). Peptidase-5 was observed in both WS and SS, suggesting that this enzyme provides a primary function. Phylogenetically, the two peptidases group into the same clade, along with other peptidases from stink bugs and the bed bug (*Cimex lectularius*), and were distant from peptidases of other insects (Figure 4).

## Cathepsins

Two different types of cathepsin-like proteases were identified from the salivary proteomes. Cathepsin L1-like proteases (cathepsin-25/XP\_014281793.1 and cathepsin29/XP\_01492127.1) were identified from WS and cathepsin-57/XP\_014278765.1, a putative cysteine proteinase CG12163-like was found in SS. Similar conserved domains (accessions: *pfam00112*, *cd02248*, *smart00645*, *PTZ00203*,



**FIGURE 3 |** Mapping of peptides from the *N. viridula* gut and salivary gland proteomes to *H. halys* translated transcripts. *H. halys* transcripts were translated to protein sequences ( $\geq 100$  aa). Translated protein sequences derived from the ASG, PSG, and gut were used as targets for mapping of peptides identified from *N. viridula* SG and gut (Lomate and Bonning, 2016). **(A)** Some 8–12% of the *H. halys* predicted protein sequences were mapped by peptides derived from the *N. viridula* salivary gland proteomes, while only 3% of the *H. halys* protein sequences were mapped by *N. viridula* gut-derived peptides. **(B)** Proteins translated from highly abundant *H. halys* transcripts, were more likely to be mapped by the *N. viridula* proteome. Translated protein sequences derived from the ASG, PSG, and gut were used as targets for mapping of peptides identified from *N. viridula* SG and gut (Lomate and Bonning, 2016).

*smart00848*, *pfam08246*, and *GOG4870*) were identified in both cathepsin-25 and cathepsin-27 proteins. In contrast to cathepsin L1-like cysteine proteases, cathepsin-57 contains multiple domains of the *CY* superfamily (accessions: *smart00043*, *cd00042*, and *pfam00031*) in addition to the domains found in cathepsin-25 and cathepsin-27. A phylogenetic tree based on selected sequences of arthropods showed two large clusters (cysteine protease CG12163-like and cathepsin -L1 like). In the cathepsin L1-like group, cathepsin-57, cathepsin-25, and cathepsin-27 were located on two separate branches

(Figure 5). Cathepsin-25 was highly expressed in the gut, while cathepsin-27 was mainly expressed in ASG (Table 3). Differential expression of the two cathepsin L1-like proteases may reflect differences in function.

### Chymotrypsins and Trypsins

Among the three chymotrypsins and 14 trypsin identified from WS and SS, chymotrypsin-1 and chymotrypsin-3 were highly homologous (sequence identity of 92%), and proteome peptide mapping did not distinguish between them (data

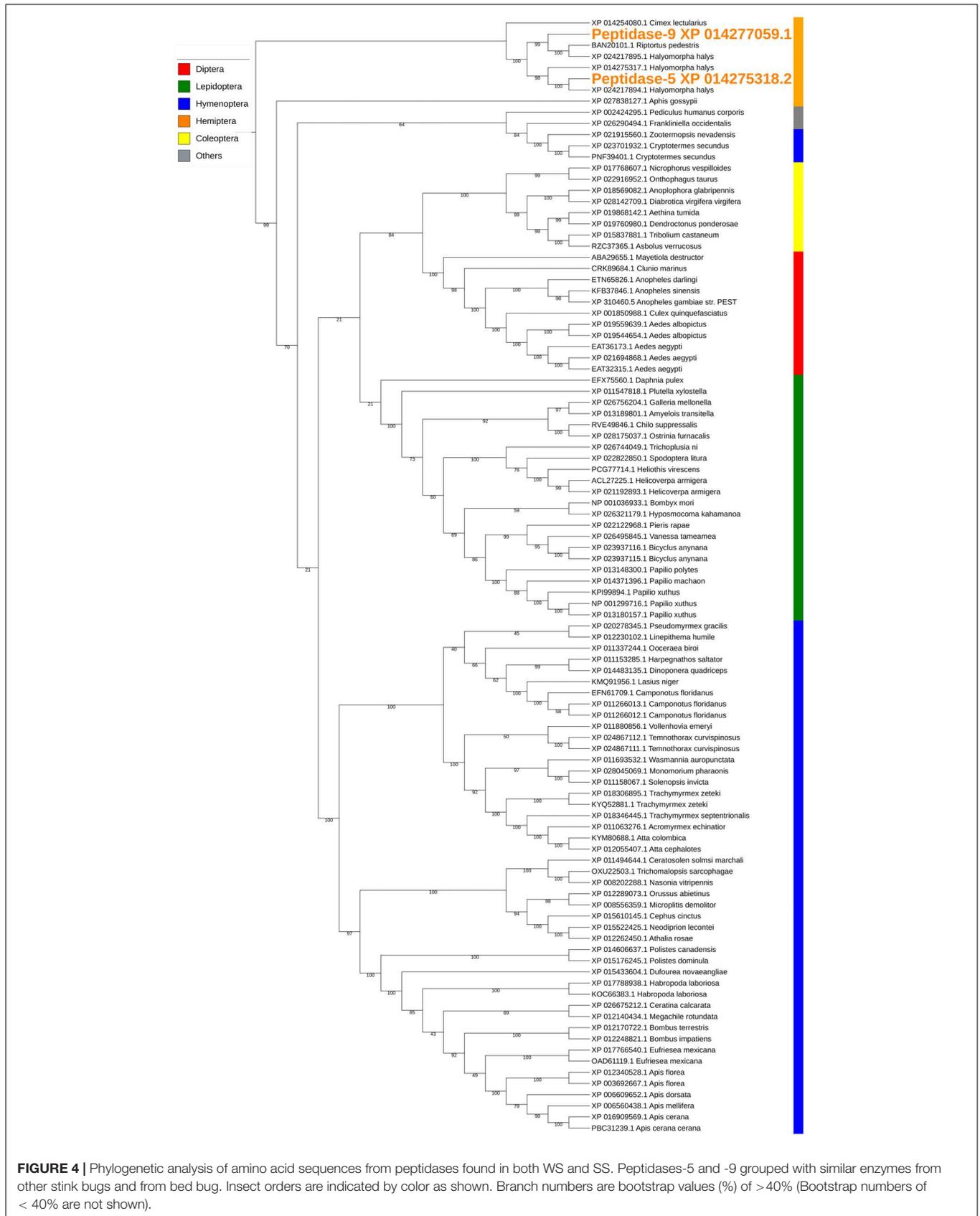
**TABLE 3** | Protease and nuclease transcripts identified by mapping of *H. halys* watery saliva and sheath saliva proteomes to translated sequences.

Protein ID	Enzyme	Hit accession	In proteome?		Signal peptide	RPKM		
			WS	SS		ASG	PSG	Gut
Peptidase -5	carboxypeptidase B-like	XP_014275318.1	Y	Y	Y	27.41	5081.11	4.62
Peptidase -9	carboxypeptidase B-like	XP_014277059.1	Y	N	Y	3.16	522.1	0.48
Cathepsin-25	cathepsin L1-like	XP_014281793.1	Y	N	Y	18.85	9.55	13,836.51
Cathepsin-27	cathepsin L1-like	XP_014292127.1	Y	N	Y	316.34	6.16	1.12
Cathepsin-57	putative cysteine proteinase CG12163 isoform X1	XP_014278765.1	N	Y	Y	71.12	20.7	446.78
Chymotrypsin-1	chymotrypsin-1-like	XP_014274701.1	Y	N	Y	36.78	7502.07	5.45
Chymotrypsin-2	chymotrypsin-1-like	XP_014272491.1	Y	N	Y	8.37	1586.36	1.17
Chymotrypsin-3	chymotrypsin-1-like	XP_014274702.1	Y	N	Y	25.01	1809.5	3.19
Trypsin-20	venom protease	XP_014271293.1	N	Y	Y	22.47	3.5	5.93
Trypsin-23	venom serine protease-like	XP_014286426.1 <sup>&amp;</sup>	Y	N	Y	7.11	912.22	1.46
Trypsin-24	venom serine protease-like	XP_014280682.1	Y	N	Y	4.42	605.69	0.8
Trypsin-36	venom protease-like isoform X2	XP_014285104.1	Y	N	U*	33.98	0.86	3.18
Trypsin-38	venom serine protease	XP_014289625.1	Y	N	Y	3.61	593.27	1.21
Trypsin-39	venom serine protease-like	XP_014293119.1	Y	N	Y	9.23	1762.07	1.51
Trypsin-40	venom serine protease-like	XP_014286910.1	N	Y	U	14.77	2812.85	2.95
Trypsin-41	venom serine protease-like	XP_014292325.1	Y	N	Y	538.65	11.87	0.82
Trypsin-42	venom serine protease-like	XP_014291671.1 <sup>&amp;</sup>	Y	Y	Y	19.03	3550.97	4.37
Trypsin-43	venom serine protease-like	XP_014289432.1	Y	N	Y	123.02	0.5	0.68
Trypsin-44	venom serine protease 34-like	XP_014277254.1	Y	N	Y	50.78	10,209.03	14.24
Trypsin-45	venom serine protease 34-like	XP_014272550.1	Y	Y	Y	65.08	13,245.27	12.1
Trypsin-47	venom serine protease 34-like	XP_014291670.1 <sup>&amp;</sup>	Y	N	U	101.21	166.27	0.16
Trypsin-48	venom serine protease	XP_014292641.1	Y	Y	Y	2266.41	17.2	1.92
Ribonuclease-31	ribonuclease Oy-like	XP_014273779.1	Y	N	N/A	20.56	3112.01	7.83
Uncharacterized protein (2 isoforms)	LOC106684787/venom nuclease-like protein 1(DNA/RNA non-specific endonuclease; pfam01223)	XP_024218583.1	Y	N	N/A	102.99	4141.56	4.71

The presence of each enzyme (protein) in watery saliva (WS) and sheath saliva (SS) is indicated, along with RPKM values shown in heat map format. RPKM values  $\geq 10,000$  are shown in orange; 1,000–9,999 in yellow; 10–999 in pale green; 0.1–9.9 in dark green. \*Undetermined, missing N-terminal sequences. <sup>&</sup>This record was removed from NCBI database as a result of standard genome annotation processing.

not shown). Interestingly, three trypsins (trypsin-23, trypsin-42, and trypsin-47) were homologs of previously predicted trypsin genes of *H. halys* (XP\_014286426.1, XP\_014291671.1, and XP\_014291670.1), but were later removed from the NCBI nr database. All identified chymotrypsins and trypsins contained

common *Tryp\_SpC* superfamily domains (accession: *cd00190*, *smart00020*, *pfam00089*, and *COG5640*). In addition, CLIP domain (*pfam1203* and *smart00680*), a regulatory domain in various trypsins, was present in two trypsins (trypsin-20/XP\_014271293.1 and trypsin-36/XP\_014285104.1), while

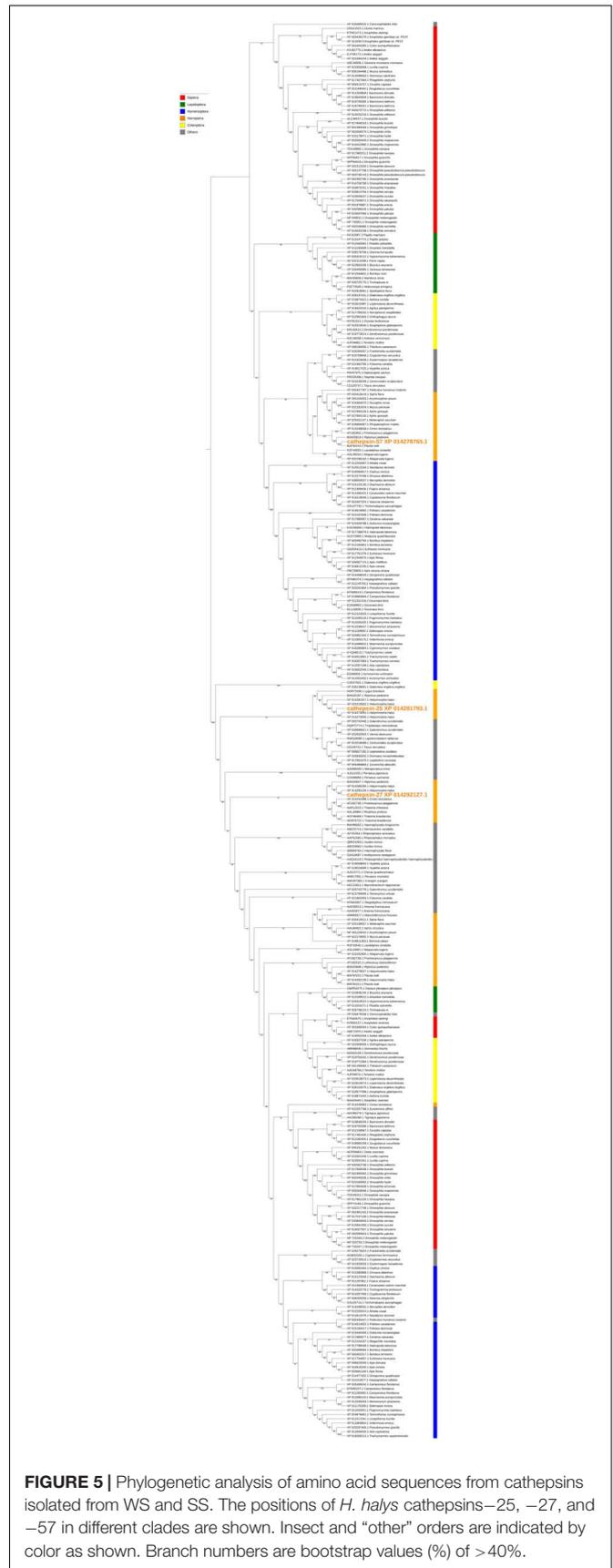


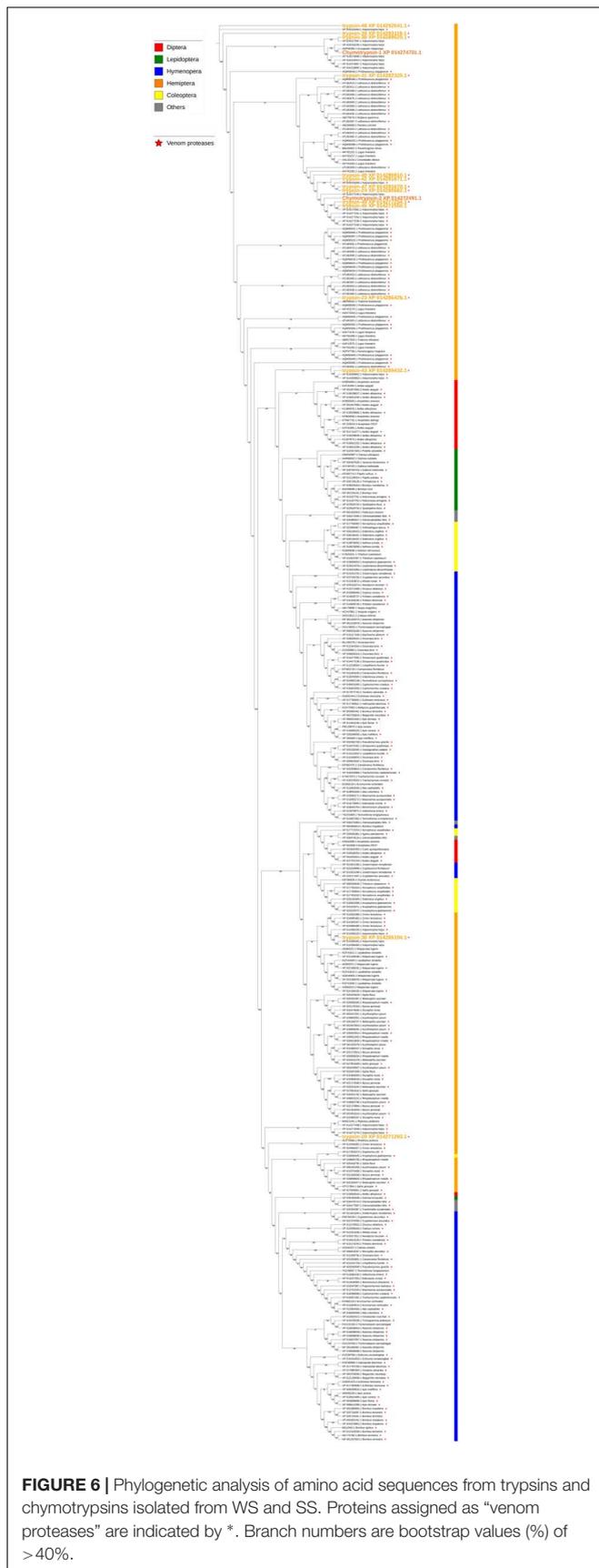
CUB domain (*cd00041* and *smart00042*), an extracellular domain, was identified in trypsin-41/XP\_014292325.1 and trypsin-43/XP\_014289432.1. A phylogenetic tree based on selected trypsin sequences grouped the chymotrypsins and trypsins of WS and SS into two major clades of trypsin-like proteases from hemipteran insects. The chymotrypsins and a majority of the trypsins found in the saliva of *H. halys* were located in the same clade, with the salivary trypsins grouped into various sub-clades (Figure 6). In contrast, the two trypsins with CLIP domain motifs (trypsin-20 and -36) were located in a distant hemipteran clade.

## Nucleases

Ribonuclease-31/XP\_014273779.1 is a homolog of our previously reported SGSB\_Ribonuclease-C20 (with 72% sequence identity) (Liu et al., 2018) and both genes were highly expressed in the PSG of *H. halys* and *N. viridula* (Supplementary Table S5). These ribonuclease Oy-like RNases contain *RNase\_T2* superfamily domains (*cd01961*, *pfam00445*, and *COG3719*), and may play an important role in host RNA degradation. The phylogenetic tree of ribonuclease-31 and homologous RNases of other insects showed that ribonuclease-31-like RNases group with similar RNases identified from stink bugs and the bed bug (Figure 7).

Two uncharacterized nucleases hit the LOC1064484787/XP\_024218583.1 gene of *H. halys*. These two transcripts encoded 410 and 436 aa, longer than the predicted XP\_024218583.1 protein (362 aa). Proteins with homology to XP\_024218583.1 have not previously been identified from *N. viridula* because the key word nuclease was missing from the BLAST annotation. On further analysis of *N. viridula* transcripts, we identified a homologous protein of XP\_024218583.1 from the *N. viridula* transcriptomes of 411 aa, only one aa shorter than the uncharacterized nuclease\_f410 and sharing 82% sequence identity. Protein sequence alignment of XP\_024218583.1 with the three homologous protein sequences suggests that the predicted XP\_024218583.1 was missing a 73 aa sequence (Figure 8). The sequence of the uncharacterized nuclease\_f435 is identical to the predicted XP\_024218583.1 sequence except for 73 aa missing from XP\_024218583. The N-terminal sequence of the uncharacterized nuclease\_f410 is similar to that of the XP\_024218583-homologous protein of *N. viridula*, but different from the N-terminal sequences of the uncharacterized nuclease\_f435 and XP\_024218583 (Figure 8). These results suggest that two isoforms of XP\_024218583 were transcribed in *H. halys*, and that the predicted XP\_024218583.1 could be incorrect. Analysis of transcript abundance of the *N. viridula* version of uncharacterized nuclease\_f410 (SGSB\_XP\_024218583\_like) indicated that, similar to uncharacterized nuclease\_f410, SGSB\_XP024218583\_like, it was highly expressed in PSG (RPKM: 3908 in PSG, 7.08 in ASG, and 2.2 in gut of *N. viridula*). The uncharacterized XP\_024218583 like nucleases contain *NUC\_superfamily* domains (*cd00091*, *pfam01223*, *smart00892*, *COG1864*, and *PTZ00259*), suggesting that they are DNA/RNA non-specific endonucleases that may function in digesting double- or single-stranded DNA and RNA. Similar to ribonuclease-31 (Figure 7), XP\_024218583\_like





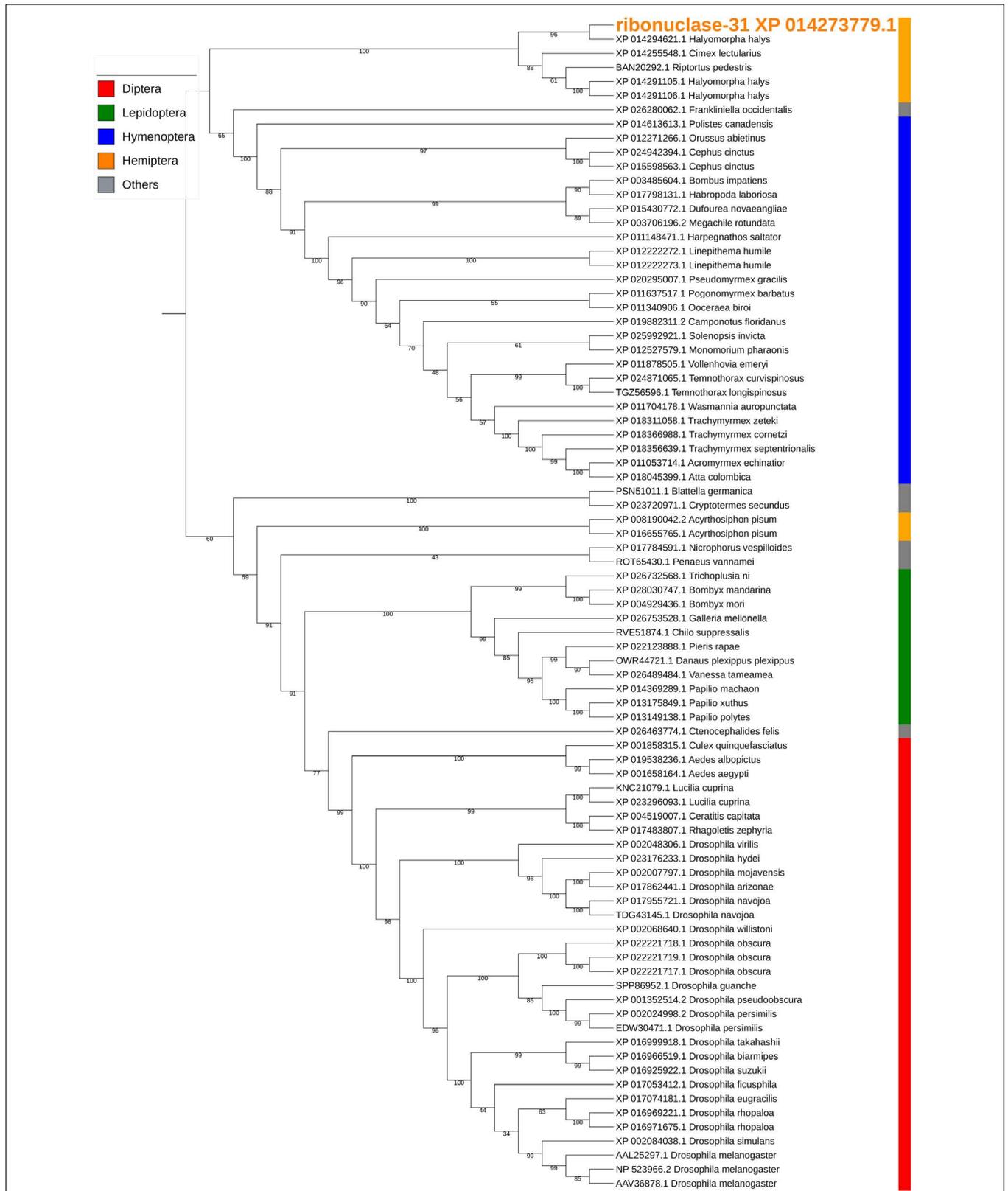
endonucleases of *H. halys* were closely related to those of other stink bugs, the bed bug and other hemipteran species (**Figure 9**).

## DISCUSSION

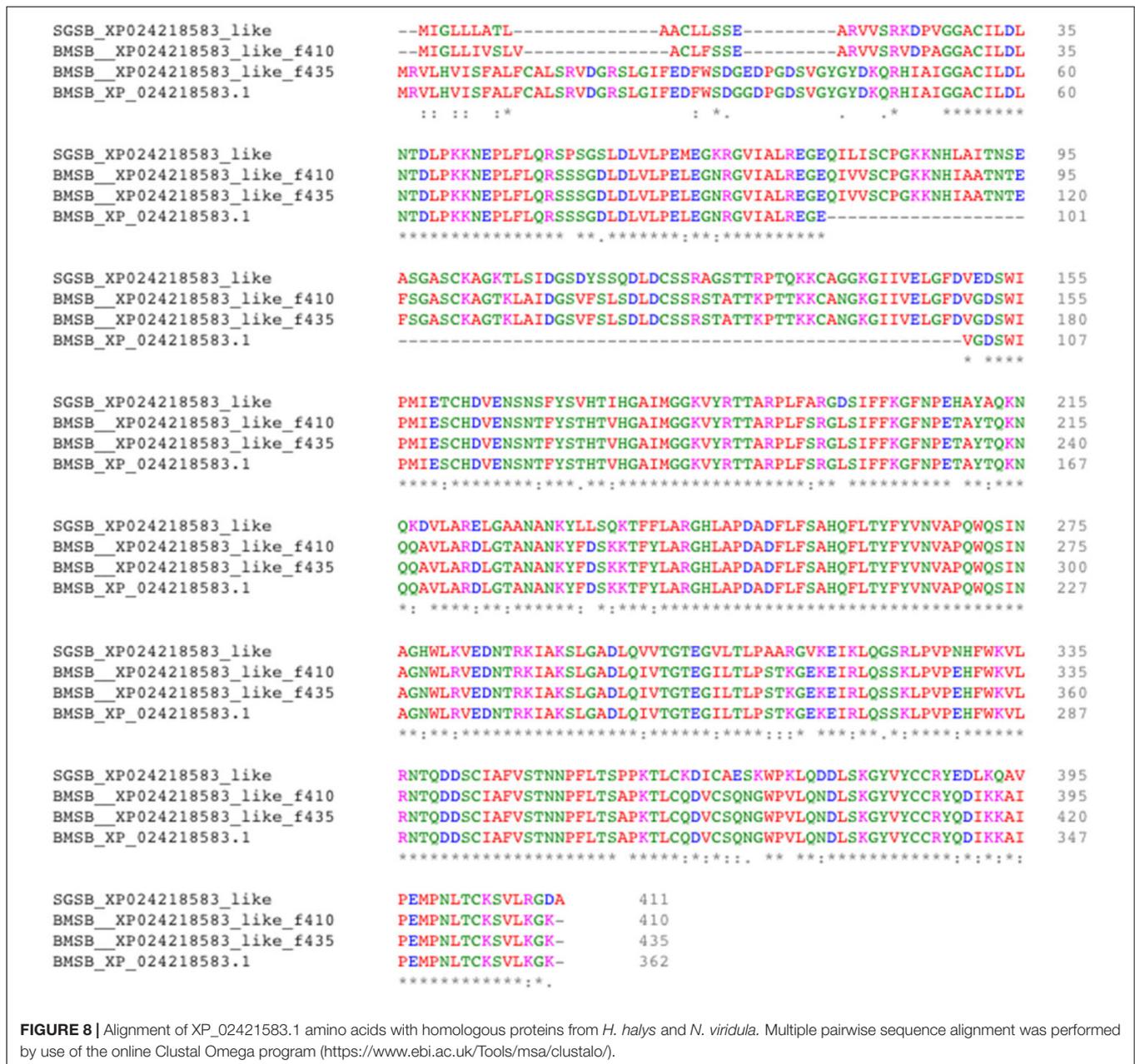
The goals for this study were investigation of whether different phytophagous stink bugs employ common digestive enzymes, and to assess the relative roles of the ASG and PSG in production of salivary enzymes. From this work, we can draw the following conclusions: (1) *H. halys* produces at least 400 putative digestive enzymes (234 proteases, 166 nucleases) identified from the assembled sequences of the ASG, PSG, and gut transcriptomes. (2) More than 20 proteases and nucleases were identified from WS and SS and analysis of both proteomic and transcriptomic datasets indicated that the majority of proteases in WS were derived from the PSG. (3) The majority of the highly transcribed proteases and nucleases of *H. halys* were similar to those of *N. viridula* (Liu et al., 2018), indicating that phytophagous stink bugs employ a similar suite of proteases and nucleases for extra-oral and gut-based digestion.

Analysis of the ASG and PSG transcriptomes allowed for the identification of additional proteins present in the previously described *H. halys* WS and SS proteomes (Peiffer and Felton, 2014). The majority of digestive enzymes identified were present in the WS (19 proteases, 2 nucleases), with only 7 proteases found in SS. Of these 7 proteases, three were not identified in WS. Identification of these enzymes in the SS and WS proteomes implies functionality in extra-oral digestion. Although 44 putative aminopeptidases were identified from the ASG, PSG, and gut transcriptomes, none were found in the WS or SS proteomes. Among the two cathepsins (cathepsin-25 and -27) found in WS and cathepsin-57 detected in SS, cathepsin-25 and -57 were highly expressed in gut, suggesting that *H. halys* gut cathepsins can be delivered into saliva. There is a precedent for this suggestion: First instar *Tuberaphis styraci* aphid soldiers inject midgut-expressed cathepsin B-like proteases through their stylets into enemies, resulting in paralysis and death of the victims (Kutsukake et al., 2004). Similarly, the serine proteases detected in *H. halys* saliva were likely produced in the midgut and transferred to the saliva. The aminopeptidases of pentatomid stink bugs are highly expressed in the gut (**Supplementary Table S1**) (Liu et al., 2018), similar to other hemipterans (Cristofoletti et al., 2006). The majority of aminopeptidases are membrane-associated, which may explain why no aminopeptidases were found in WS and SS of *H. halys*.

Hemipteran SS was originally presumed to be produced by the PSG, while digestive enzymes were assumed to be produced by the ASG (Miles, 1972). If correct, the ASG would be the primary source of WS secretions. However, analysis of the *H. halys* WS and SS proteomes, ASG and PSG transcriptomes, and comparison with those of *N. viridula* (Liu et al., 2018) support a primary role for the PSG in production of enzymes (proteases and one nuclease) destined for the WS. Of the *H. halys* proteases and nucleases with RPKM of >1,000 (**Table 3**), eleven were produced in the PSG and one in the ASG. Of the 11 proteases highly transcribed in the PSG, all but one (trypsin-40)



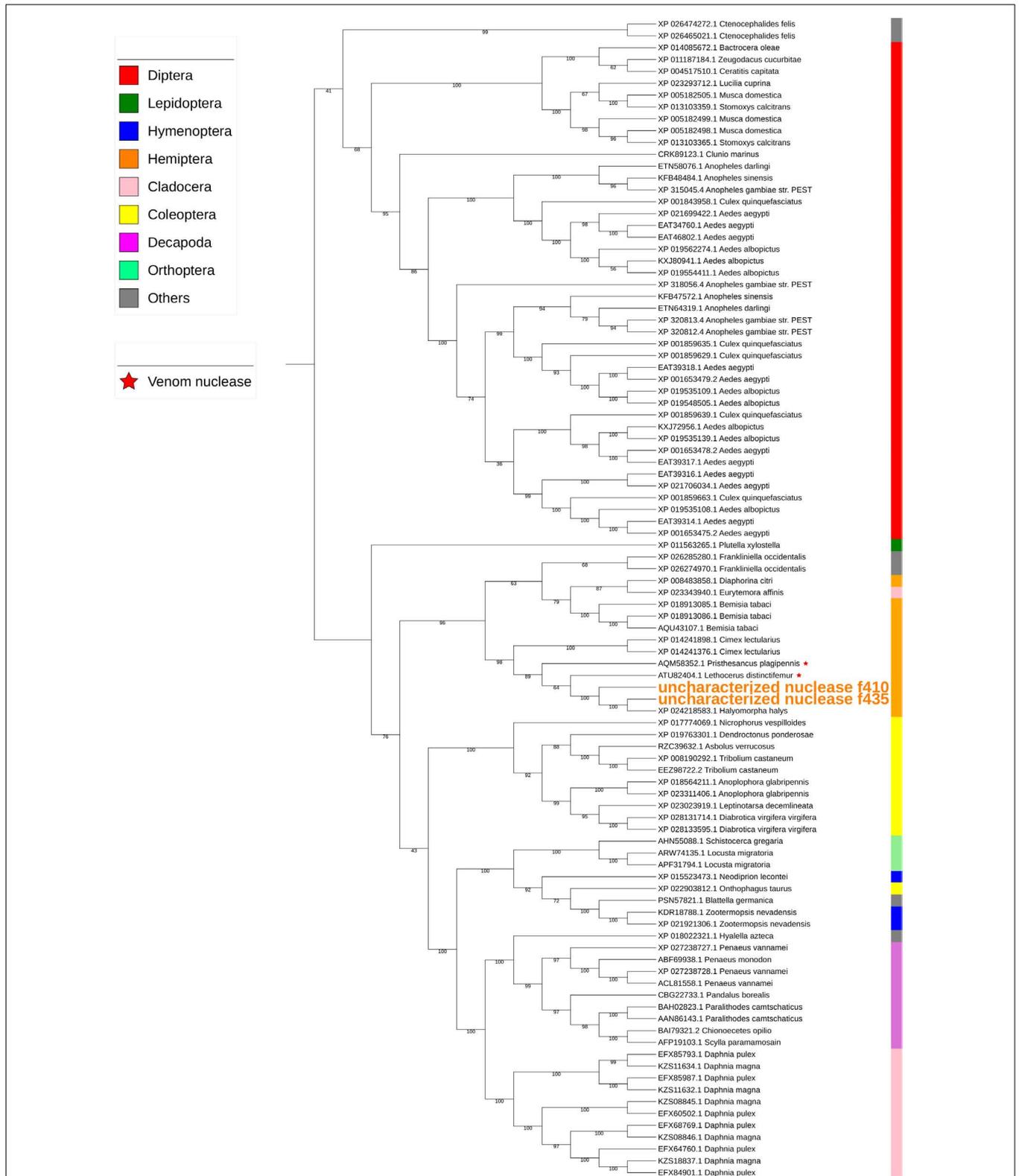
**FIGURE 7 |** Phylogenetic analysis of amino acid sequences from ribonuclease-31. Ribonuclease-31 grouped with similar enzymes from other stink bugs and from bed bug. Insect orders are indicated by color as shown. Branch numbers are bootstrap values (%) of >40%.



were found in WS and three were found in SS (trypsin-40, -42, -45). Trypsin-48, which was highly transcribed in the ASG, was detected in both WS and SS. These results indicate that enzymes produced by the PSG or ASG are not exclusively destined for the WS or SS, respectively. The different compositions of the SS and WS show that stink bugs are able to regulate the composition of their saliva.

Results from an ultrastructural analysis of the salivary glands of the Neotropical brown stink bug, *E. heros*, provide additional insight into the roles of the salivary gland tissues (Castellanos et al., 2017). Characteristics at the ultrastructural level suggest production of different compounds by the anterior and posterior glandular lobes of the PSG, muscle-mediated regulation of the

mixing of these compounds, and control of the amount of saliva released from ASG and PSG at any given point during development (Castellanos et al., 2017). The ultrastructure of the ASG implicates this tissue in water transport and secretion but with limited storage capacity, implying that proteins synthesized are likely to be transported to the lumen of the PSG (Castellanos et al., 2017). The appearance of the PSG is typical of a tissue active in protein synthesis and secretion. The authors suggest that the anterior lobe of the PSG produces proteins for extra-oral digestion, while the posterior lobe produces other salivary components such as carbohydrates, lipids and other proteins. It follows that the ability of the stink bug to modify the composition of salivary components produced by the



**FIGURE 9 |** Phylogenetic analysis of amino acid sequences from uncharacterized XP\_02421583.1-like nucleases. Uncharacterized nucleases f410 and f435 grouped with other hemipteran nucleases. Insect and crustacean orders are indicated by color as shown. Proteins assigned as “venom nucleases” are indicated by \*. Branch numbers are bootstrap values (%) of >40%.

different tissues in the salivary gland facilitates the polyphagous habit of these insects. As the anterior and posterior lobes of the PSG were not separated prior to RNA extraction in the present study, we are unable to determine whether the digestive enzymes produced by PSG are primarily produced by the anterior lobe.

Interestingly, two proteins derived from the gut symbiont *C. Pantoea carbekii* were identified in the *H. halys* WS proteome. It is conceivable that these proteins were transported from the gut to the ASG, which appear to function in the transport of proteins from the hemolymph (Castellanos et al., 2017), and subsequently into the WS.

We previously observed that nuclease enzymes were abundant in *H. halys* saliva and salivary gland (Lomate and Bonning, 2018). In RNA-seq and proteomic analyses, a ribonuclease Oy-like RNase was highly expressed in the salivary glands of both *H. halys* and *N. viridula* (Liu et al., 2018), and was also identified in the WS proteome of *H. halys*. Ribonuclease-Oy-like RNase is a member of the RNase T2 family. T2 family RNases catalyze cleavage of single-stranded RNA, are found in a wide array of organisms (including protozoans, plants, bacteria, animals, and viruses) and have a broad range of functions (Luhtala and Parker, 2010). The other putative nuclease found from *H. halys* WS was an uncharacterized protein (LOC106684787; XP\_024218583.1), which is an endonuclease\_NS-like DNA/RNA non-specific endonuclease. A polyA binding protein (XP\_013171827.1) was also detected in WS. As polyA binding protein is associated with mRNA turn-over (Mangus et al., 2003), this protein may also be involved in the degradation of host plant mRNA. It is hypothesized that nucleases secreted by stink bugs into the host plant function to degrade viral RNAs.

Many salivary proteases (trypsin-like) of *H. halys* hit proteases assigned as “venom proteases” by BLAST annotation. Phylogenetic analysis also indicated that salivary trypsins were related to “venom proteases” (Figure 6). Similarly, the uncharacterized nucleases-f410 and -f435 were closely related to two “venom nucleases” isolated from the assassin bug (a hemipteran predator), *Pisthesancus plagipennis*, and from the giant water bug or giant fishkiller, *Lethocerus distinctifemur* (Figure 9). Venoms from blood feeding insects and from insect predators, share features with the venoms of other organisms (Walker et al., 2016, 2017). While the composition of venom is complex, trypsin-like and chymotrypsin-like proteases are major venom components. Homologs of venom proteases are also found in plant-feeding hemipterans (Walker et al., 2016). It is unclear what role these “venom protease-like” trypsins and “venom nuclease-like” nucleases play following injection into plant hosts beyond potential functions in the degradation of host plant proteins and nucleotides.

## CONCLUSION

In conclusion, we have generated *H. halys* gut and salivary gland transcriptomes and identified the major proteases and

nucleases produced by the ASG, PSG, and gut, along with those present in the WS and SS. The proteases and nucleases of *H. halys*, together with our previous characterization of proteases and nucleases from *N. viridula*, show that these phytophagous stink bugs encode and express similar suites of proteases and nucleases for extraoral digestion and gut-based digestion. Based on ultrastructural analysis, the differential mixing and release of salivary components from the ASG and PSG (anterior and posterior lobes) may mediate the ability of stink bugs to feed on multiple host plants (Castellanos et al., 2017). The comprehensive analysis of stink bug digestive enzymes presented here may provide leads for novel control strategies targeting digestive enzymes for management of multiple stink bug species, and highlight the common enzymatic challenges faced by bioactives in development for stink bug control.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the NCBI Sequence Read Archive (SRA BioProject: PRJNA560285).

## AUTHOR CONTRIBUTIONS

SL conducted the bioinformatics analyses. BB conceived the study and contributed to the design of the experiments. Both authors contributed to the writing and review of the manuscript.

## FUNDING

This work was supported by the National Science Foundation I/UCRC, the Center for Arthropod Management Technologies (grant numbers IIP-1338775 and IIP-1821914), and by industry partners.

## ACKNOWLEDGMENTS

The authors thank Drs. Michelle Peiffer and Gary Felton, Department of Entomology, Pennsylvania State University, United States for provision of the WS and SS proteome data, Dr. Donald Weber, USDA, BARC-West Beltsville, MD, United States and Dr. Alberto Bressan at Bayer CropScience for provision of *H. halys* for use in this study, and Dr. Purushottam R. Lomate for conducting tissue dissections.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphys.2019.01255/full#supplementary-material>

## REFERENCES

- Almagro Armenteros, J. J., Tsirigos, K. D., Sonderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. doi: 10.1038/s41587-019-0036-z
- Backus, E. A., Serrano, M. S., and Ranger, C. M. (2005). Mechanisms of hopperburn: an overview of insect taxonomy, behavior, and physiology. *Annu. Rev. Entomol.* 50, 125–151. doi: 10.1146/annurev.ento.49.061802.123310
- Bergmann, E., Bernhard, K. M., Bernon, G., Bickerton, M., Gill, S., Gonzales, C., et al. (2013). *Host Plants of the Brown Marmorated Stink Bug in the U.S.* Available at: [http://www.stopbmsb.org/where-is-bmsb/host-plants/#host\\_plants\\_table](http://www.stopbmsb.org/where-is-bmsb/host-plants/#host_plants_table) (accessed July 12, 2016).
- Biddinger, D. J., Krawczyk, G., Tooker, J. F., Mitchell, W., Rice, K. B., Herbert, A., et al. (2014). Biology, ecology, and management of Brown Marmorated Stink Bug (Hemiptera: Pentatomidae). *J. Integ. Pest Manag.* 5, A1–A13. doi: 10.1603/ipm14002
- Castellanos, N., Martinez, L. C., Silva, E. H., Teodoro, A. V., Serrao, J. E., and Oliveira, E. E. (2017). Ultrastructural analysis of salivary glands in a phytophagous stink bug revealed the presence of unexpected muscles. *PLoS One* 12:e0179478. doi: 10.1371/journal.pone.0179478
- Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610
- Cristofolletti, P. T., de Sousa, F. A., Rahbe, Y., and Terra, W. R. (2006). Characterization of a membrane-bound aminopeptidase purified from *Acyrtosiphon pisum* midgut cells. A major binding site for toxic mannose lectins. *FEBS J.* 273, 5574–5588. doi: 10.1111/j.1742-4658.2006.05547.x
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi: 10.1038/nprot.2013.084
- Hajek, A. E., Solter, L. F., Maddox, J. V., Huang, W. F., Estep, A. S., Krawczyk, G., et al. (2017). *Nosema maddoxi* sp. nov. (Microsporidia, Nosematidae), a widespread pathogen of the Green Stink Bug *Chinavia hilaris* (Say) and the Brown Marmorated Stink Bug *Halyomorpha halys* (Stal). *J. Eukaryot. Microbiol.* 65, 315–330. doi: 10.1111/jeu.12475
- Haye, T., Garipey, T., Hoelmer, K., Rossi, J. P., Streito, J. C., Tassou, X., et al. (2015). Range expansion of the invasive brown marmorated stinkbug, *Halyomorpha halys*: an increasing threat to field, fruit and vegetable crops worldwide. *J. Pest Sci.* 88, 665–673. doi: 10.1007/s10340-015-0670-2
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35, 518–522. doi: 10.1093/molbev/msx281
- Ioannidis, P., Lu, Y., Kumar, N., Creasy, T., Daugherty, S., Chibucos, M. C., et al. (2014). Rapid transcriptome sequencing of an invasive pest, the brown marmorated stink bug *Halyomorpha halys*. *BMC Genomics* 15:738. doi: 10.1186/1471-2164-15-738
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285
- Katoh, K., Rozewicki, J., and Yamada, K. D. (2017). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* doi: 10.1093/bib/bbx108 [Epub ahead of print].
- Kuhar, T. P., and Kamminga, K. (2017). Review of the chemical control research on *Halyomorpha halys* in the USA. *J. Pest Sci.* 90, 1021–1031. doi: 10.1007/s10340-017-0859-7
- Kutsukake, M., Shibao, H., Nikoh, N., Morioka, M., Tamura, T., Hoshino, T., et al. (2004). Venomous protease of aphid soldier for colony defense. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11338–11343. doi: 10.1073/pnas.0402462101
- Lee, D. H., Short, B. D., Joseph, S. V., Bergh, J. C., and Leskey, T. C. (2013). Review of the biology, ecology, and management of *Halyomorpha halys* (Hemiptera: Pentatomidae) in China, Japan, and the Republic of Korea. *Environ. Entomol.* 42, 627–641. doi: 10.1603/EN13006
- Leskey, T. C., Lee, D. H., Short, B. D., and Wright, S. E. (2012). Impact of insecticides on the invasive *Halyomorpha halys* (Hemiptera: Pentatomidae): analysis of insecticide lethality. *J. Econ. Entomol.* 105, 1726–1735. doi: 10.1603/ec12096
- Leskey, T. C., and Nielsen, A. L. (2018). Impact of the invasive Brown Marmorated Stink Bug in North America and Europe: history, biology, ecology, and management. *Annu. Rev. Entomol.* 63, 599–618. doi: 10.1146/annurev-ento-020117-043226
- Letunic, I., and Bork, P. (2019). Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239
- Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* 12:323. doi: 10.1186/1471-2105-12-323
- Liu, S., Lomate, P. R., and 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. doi: 10.1016/j.ibmb.2018.10.003
- Lomate, P. R., and 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. doi: 10.1038/srep27587
- Lomate, P. R., and 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. doi: 10.1002/arch.21459
- Lucini, T., and Panizzi, A. R. (2018a). Electropenetrography (EPG): a breakthrough tool unveiling Stink Bug (Pentatomidae) feeding on plants. *Neotrop. Entomol.* 47, 6–18. doi: 10.1007/s13744-017-0574-3
- Lucini, T., and Panizzi, A. R. (2018b). Electropenetrography monitoring of the neotropical Brown-Stink Bug (Hemiptera: Pentatomidae) on soybean pods: an electrical penetration graph-histology analysis. *J. Insect Sci.* 18:5. doi: 10.1093/jisesa/iey108
- Luhtala, N., and Parker, R. (2010). T2 family ribonucleases: ancient enzymes with diverse roles. *Trends Biochem. Sci.* 35, 253–259. doi: 10.1016/j.tibs.2010.02.002
- Mangus, D. A., Evans, M. C., and Jacobson, A. (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.* 4:223. doi: 10.1186/gb-2003-4-7-223
- Miles, P. (1964). Studies on the salivary physiology of plant bugs: the chemistry of formation of the sheath material. *J. Insect Physiol.* 10, 147–160. doi: 10.1016/0022-1910(64)90103-9
- Miles, P. W. (1972). The saliva of Hemiptera. *Adv. Insect Physiol.* 9, 183–255. doi: 10.1016/s0065-2806(08)60277-5
- Minh, B. Q., Nguyen, M. A., and von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* 30, 1188–1195. doi: 10.1093/molbev/mst024
- Moreno, A., Garzo, E., Fernandez-Mata, G., Kassem, M., Aranda, M. A., and Fereres, A. (2011). Aphids secrete watery saliva into plant tissues from the onset of stylet penetration. *Entomol. Exp. Appl.* 139, 145–153. doi: 10.1111/j.1570-7458.2011.01117.x
- Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Nielsen, H. (2017). Predicting secretory proteins with signalp. *Methods Mol. Biol.* 1611, 59–73. doi: 10.1007/978-1-4939-7015-5\_6
- Panizzi, A. R. (2015). Growing problem with stink bugs (Hemiptera: Heteroptera: Pentatomidae): species invasive to the U.S. and potential neotropical invaders. *Am. Entomol.* 61, 223–233. doi: 10.1093/ae/tmv068
- Panizzi, A. R., McPherson, J. E., James, D. G., Javahery, M., and McPherson, R. M. (2000). “Economic importance of stink bugs (Pentatomidae),” in *Heteroptera of Economic Importance*, eds C. W. Schaefer, and A. R. Panizzi, (Boca Raton, FL: CRC Press), 421–474. doi: 10.1201/9781420041859.ch13
- Peiffer, M., and Felton, G. W. (2014). Insights into the saliva of the brown marmorated stink bug *Halyomorpha halys* (Hemiptera: Pentatomidae). *PLoS One* 9:e88483. doi: 10.1371/journal.pone.0088483
- Rice, K. B., Bedoukian, R. H., Hamilton, G. C., Jentsch, P., Khramian, A., MacLean, P., et al. (2018). Enhanced response of *Halyomorpha halys* (Hemiptera: Pentatomidae) to its aggregation pheromone with ethyl decatrienoate. *J. Econ. Entomol.* 111, 495–499. doi: 10.1093/jee/tox316
- Rider, D. A. (2011). *Number of Genera and Species of Pentatomidae*. Fargo, ND: North Dakota State University.

- Sievers, F., and Higgins, D. G. (2018). Clustal omega for making accurate alignments of many protein sequences. *Protein Sci.* 27, 135–145. doi: 10.1002/pro.3290
- Sparks, M. E., Shelby, K. S., Kuhar, D., and Gundersen-Rindal, D. E. (2014). Transcriptome of the invasive brown marmorated stink bug, *Halyomorpha halys* (Stal) (Heteroptera: Pentatomidae). *PLoS One* 9:e111646. doi: 10.1371/journal.pone.0111646
- The i5k Initiative (2017). *Genome Assembly and Annotation Report: Halyomorpha halys (brown marmorated stink bug)*. Available at: [https://www.ncbi.nlm.nih.gov/genome/14101?genome\\_assembly\\_id=368443](https://www.ncbi.nlm.nih.gov/genome/14101?genome_assembly_id=368443) (accessed May 16, 2018).
- Walker, A. A., Madio, B., Jin, J., Undheim, E. A., Fry, B. G., and King, G. F. (2017). Melt with this kiss: paralyzing and liquefying venom of the Assassin Bug *Pristhesancus plagipennis* (Hemiptera: Reduviidae). *Mol. Cell Proteom.* 16, 552–566. doi: 10.1074/mcp.M116.063321
- Walker, A. A., Weirauch, C., Fry, B. G., and King, G. F. (2016). Venoms of heteropteran insects: a treasure trove of diverse pharmacological toolkits. *Toxins* 8:43. doi: 10.3390/toxins8020043
- Weber, D. C., Leskey, T. C., Walsh, G. C., and Khrimian, A. (2014). Synergy of aggregation pheromone with methyl (E,E,Z) -2,4,6-decatrienoate in attraction of *Halyomorpha halys* (Hemiptera: Pentatomidae). *J. Econ. Entomol.* 107, 1061–1068. doi: 10.1603/ec13502
- Wingett, S. W., and Andrews, S. (2018). FastQ screen: a tool for multi-genome mapping and quality control. *F1000Res* 7:1338. doi: 10.12688/f1000research.15931.2

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Liu and Bonning. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Differential Expression of Candidate Salivary Effector Genes in Pea Aphid Biotypes With Distinct Host Plant Specificity

Hélène Boulain<sup>1†</sup>, Fabrice Legeai<sup>1,2</sup>, Julie Jaquiéry<sup>1</sup>, Endrick Guy<sup>1</sup>, Stéphanie Morlière<sup>1</sup>, Jean-Christophe Simon<sup>1</sup> and Akiko Sugio<sup>1\*</sup>

<sup>1</sup> INRA, UMR1349, Institute of Genetics, Environment and Plant Protection, Le Rheu, France, <sup>2</sup> University of Rennes 1, Inria, CNRS, IRISA, Rennes, France

## OPEN ACCESS

### Edited by:

Brigitte Mauch-Mani,  
Université de Neuchâtel,  
Switzerland

### Reviewed by:

Owain Rhys Edwards,  
CSIRO Land and Water,  
Australia  
Eduard Venter,  
University of Johannesburg,  
South Africa

### \*Correspondence:

Akiko Sugio  
akiko.sugio@inra.fr

### †Present address:

Hélène Boulain,  
EAWAG, Swiss Federal Institute of  
Aquatic Science and Technology,  
Dübendorf, Switzerland

### Specialty section:

This article was submitted to  
Plant Microbe Interactions,  
a section of the journal  
Frontiers in Plant Science

Received: 09 July 2019

Accepted: 18 September 2019

Published: 22 October 2019

### Citation:

Boulain H, Legeai F, Jaquiéry J,  
Guy E, Morlière S, Simon J-C and  
Sugio A (2019) Differential Expression  
of Candidate Salivary Effector Genes  
in Pea Aphid Biotypes With Distinct  
Host Plant Specificity.  
Front. Plant Sci. 10:1301.  
doi: 10.3389/fpls.2019.01301

Effector proteins play crucial roles in determining the outcome of various plant-parasite interactions. Aphids inject salivary effector proteins into plants to facilitate phloem feeding, but some proteins might trigger defense responses in certain plants. The pea aphid, *Acyrtosiphon pisum*, forms multiple biotypes, and each biotype is specialized to feed on a small number of closely related legume species. Interestingly, all the previously identified biotypes can feed on *Vicia faba*; hence, it serves as a universal host plant of *A. pisum*. We hypothesized that the salivary effector proteins have a key role in determining the compatibility between specific host species and *A. pisum* biotypes and that each biotype produces saliva containing a specific mixture of effector proteins due to differential expression of encoding genes. As the first step to address these hypotheses, we conducted two sets of RNA-seq experiments. RNA-seq analysis of dissected salivary glands (SGs) from reference alfalfa- and pea-specialized *A. pisum* lines revealed common and line-specific repertoires of candidate salivary effector genes. Based on the results, we created an extended catalogue of *A. pisum* salivary effector candidates. Next, we used aphid head samples, which contain SGs, to examine biotype-specific expression patterns of candidate salivary genes. RNA-seq analysis of head samples of alfalfa- and pea-specialized biotypes, each represented by three genetically distinct aphid lines reared on either a universal or specific host plant, showed that a majority of the candidate salivary effector genes was expressed in both biotypes at a similar level. Nonetheless, we identified small sets of genes that were differentially regulated in a biotype-specific manner. Little host plant effect (universal vs. specific) was observed on the expression of candidate salivary genes. Analysis of previously obtained genome re-sequenced data of the two biotypes revealed the copy number variations that might explain the differential expression of some candidate salivary genes. In addition, at least four candidate effector genes that were present in the alfalfa biotype but might not be encoded in the pea biotype were identified. This work sets the stage for future functional characterization of candidate genes potentially involved in the determination of plant specificity of pea aphid biotypes.

**Keywords:** host-specificity, transcriptomics, effectors, copy number variation, phytophagous insects

## INTRODUCTION

A large majority of herbivorous insects feeds on specific host plant species (Forister et al., 2015). Host plants not only provide food sources but may also provide insect habitat and mating sites. Such continuous and intimate interactions with certain plant species are considered as major driving forces in insect evolution and specialization to host plants, potentially leading to new species through reduction of gene flow between plant-specialized populations and mechanisms reinforcing reproductive isolation (Butlin and Smadja, 2017). Understanding the adaptation mechanisms of insects to their host plants is of paramount importance to increase knowledge on the role of natural selection in species formation but also to contribute to applied issues, notably to respond to the increasing need to develop sustainable crop pest-management strategies. However, the molecular mechanisms of insect specialization to host plant species are little understood, and these mechanisms seem to vary between combinations of plant and insect species (Simon et al., 2015; Birnbaum and Abbot, 2018).

Aphids are major crop pests worldwide and have a very specialized feeding style. Most aphid species have a narrow range of host plants (Peccoud et al., 2010). Aphids feed on plant sap by using their needle-like mouthparts, called stylets. In the process of inserting the stylets into phloem sieve cells and establishing phloem feeding, aphids puncture various plant cells and secrete watery saliva that contains a battery of proteins, many of them expressed in salivary glands (SGs) (Moreno et al., 2011; Boulain et al., 2018). Several salivary proteins were shown to increase aphid fecundity when expressed in plants or to reduce aphid fecundity when their expression was silenced in aphids, providing evidence that these proteins function like effectors of microbial pathogens (Mutti et al., 2006; Mutti et al., 2008; Bos et al., 2010; Pitino et al., 2011; Atamian et al., 2012; Pitino and Hogenhout, 2013; Elzinga et al., 2014; Naessens et al., 2015; Wang et al., 2015; Kettles and Kaloshian, 2016; Guy et al., 2016). *In planta* expression of the salivary effectors C002, Mp1, and Mp2 from the generalist aphid *Myzus persicae* increases the fecundity of *M. persicae* on its host plants *Arabidopsis thaliana* and *Nicotiana benthamiana*, while expression of orthologous genes from a legume-specialist species (*Acyrtosiphon pisum*) in these plants has no effect on *M. persicae* fecundity, suggesting host-specific functions of some salivary proteins (Pitino and Hogenhout, 2013). On the other hand, *in planta* expression of aphid salivary proteins (e.g., Mp10 and Mp42 from *M. persicae*) reduces aphid fecundity, suggesting a possible property of salivary proteins as avirulence proteins, which are recognized by a plant and trigger plant defense reactions against aphids (Bos et al., 2010). These results indicate that a set of salivary effectors can determine the outcome of plant-aphid interactions.

*Acyrtosiphon pisum* is a model aphid species and is often regarded as a single insect species. However, the *A. pisum* complex actually encompasses at least 15 biotypes with differential fitness on specific host plants (Peccoud et al., 2009; Peccoud et al., 2015). Each biotype is specialized to one or a few legume species and cannot perform well on other plants (Peccoud et al., 2009). They have a similar but distinct genetic makeup; therefore, *A. pisum*

biotypes are an ideal system for studying the mechanisms of aphid specialization to host plants. Interestingly, all the 15 biotypes feed well on *Vicia faba*, which is considered as a universal host plant of *A. pisum*. Previous analysis of 390 microsatellite markers (Jaquiéry et al., 2012) and pool-seq analysis (Nouhaud et al., 2018) of three *A. pisum* biotypes represented by 60 individual aphids, both indicated that the genomic regions that are highly differentiated between the biotypes are significantly enriched in candidate salivary effector genes. In addition, gene expression analysis of six biotypes of *A. pisum* showed that a relatively high proportion of candidate salivary effector genes is differentially expressed (DE) between the biotypes (Eyes et al., 2016). These studies indicate potential involvement of the salivary effector genes in host plant specialization.

Previously, we conducted transcriptomics analysis and bioinformatics prediction of secreted proteins of the *A. pisum* reference line LSR1 (alfalfa biotype) and identified 3,603 SG-expressed candidate salivary effector genes (Boulain et al., 2018), of which, 740 were upregulated in the SGs compared to the alimentary tract (AT). Proteomics analysis of aphid-fed diet also identified 51 secreted proteins, all of them expressed in the SGs. A comparative genomic analysis using 17 arthropod genomes revealed that the SG-upregulated effector set contains a high proportion of aphid lineage-specific genes and tends to evolve faster. The study also revealed that the salivary effector set was enriched with members of gene families, some of which were expanded in the pea aphid genome compared to other aphid species (Boulain et al., 2018).

Based on the accumulated results of functional characterization of aphid salivary effector proteins and genome-wide analyses of *A. pisum*, we hypothesized that *A. pisum* biotypes express different salivary effector proteins and that biotype-specific mixture of salivary proteins might be required for host plant adaptation. To characterize biotype-specific differences in salivary effector composition and expression level, we conducted transcriptomic analysis of two *A. pisum* biotypes on both the universal (*V. faba*) and specific host plants.

We have chosen the pea biotype to compare with the alfalfa biotype, which includes the reference line LSR1, because they are closely related (limiting the chances to identify highly differentiated genes that are not involved in host specificity) (Peccoud et al., 2009), show distinct phenotypes on the two specific host plants, and various genetic resources and techniques are available in pea (*Pisum sativum*), which will facilitate the follow-up study of effector functions (Guy et al., 2016; Meziadi et al., 2016; Meziadi et al., 2017). As a first step to compare these two biotypes, we created a list of salivary genes using an *A. pisum* pea-adapted line because the previous candidate salivary gene list was created only for the alfalfa-adapted line, LSR1 (Boulain et al., 2018), and may not include the salivary genes that are specific to the pea biotype. To take into account aphid lineage-specific expression differences and to identify the genes that show biotype-specific differential expression patterns, we conducted a transcriptomic study using three genetically distinct aphid lines for each biotype. We also examined the effect of feeding plants (universal host *V. faba* vs. specific host) on the expression patterns of identified salivary genes. Due to the enormous task of

dissecting SGs to provide a sufficient amount of RNA for RNA-seq, we used aphid head samples to examine the transcriptome of three aphid lines *per* biotype and the effect of host plants. Nonetheless, we were able to successfully identify salivary genes that are DE in a biotype-specific manner and evaluate the impact of host plants on the expression pattern of salivary genes.

## MATERIALS AND METHODS

### Aphids, Plants, and Growth Conditions

To explore biotype effects, we studied six different lines of *A. pisum*, of which three lines represented each biotype (Supplementary Table S1). To avoid the potential influence of secondary symbionts on overall aphid fitness and plant exploitation mechanisms, we used aphid lines that were free of facultative symbiont. The six aphid lines used in this study, including the LSR1 line for which the genome is sequenced (The International Aphid Genomics Consortium, 2010) were maintained in a growth chamber at 18°C with a 16-h-day/8-h-night photoperiod on their universal host, the broad bean, *V. faba* (cv. Castel), at low density to avoid the production of winged individuals. All plants were grown in a growth chamber at 18°C with a 16-h-day/8-h-night photoperiod. Before installing the aphids for the experiments, *V. faba* and pea, *P. sativum* (cv. Baccara), were grown for 10 days whereas alfalfa, *Medicago sativa* (cv. Comète), was grown for 4 weeks.

### Aphid Performance Assays

Adult aphids from the six lines were installed on each tested plant (*V. faba*, *P. sativum*, *M. sativa*) so that the nymphs produced did not experience a switch of host plant species. One 1-day-old aphid nymph was installed on each test plant (with 12 test plants *per* condition), and their offspring were counted 18 days later. The experiment was conducted in a growth chamber at 18°C with a 16-h-day/8-h-night photoperiod.

Differences in numbers of offspring produced by each aphid line on the three tested plants were analyzed with a Kruskal-Wallis test performed in R (R Development Core Team, 2017).

### RNA Sequencing

To prepare RNA samples from SGs and ATs of the pea biotype, we used 9-day-old aphids from the line P123 reared at a low density of 10–15 aphids *per* *V. faba* plant. The aphids were dissected in saline solution. Dissected organs were soaked in RNA later (QIAGEN) to avoid RNA degradation and pooled in batches before RNA extraction (three replicates *per* line and *per* organ). On average, RNA samples from 200 pairs of SGs or 20 ATs that were dissected on the same day were pooled for one replicate of an RNA-seq experiment. Three biological replicates *per* condition were prepared.

To prepare RNA samples from heads of the three alfalfa biotype lines (LSR1, LL01, L84) and the three pea biotype lines (ArPo58, P123, S1PS02), we used 9-day-old aphids reared since birth on the universal host *V. faba* and on the specific hosts (*P. sativum* or *M. sativa*) at a density of 10 aphids *per* plant. Aphids were then collected, flash frozen in liquid nitrogen and heads

(in front of the first pair of legs) were cut by scalpel while whole aphid bodies were frozen. Three replicates *per* line and *per* plant were prepared. On average, 20 aphid heads harvested on the same day were pooled for one replicate. Three replicates *per* condition were prepared.

RNA from SGs, ATs, and heads were extracted by NucleoSpin RNA XS (Macherey-Nagel) and quantified. rRNA depletion, single stranded-RNA library preparation, multiplexing, and sequencing were performed by Genewiz (New Jersey, USA). Sequencing was performed on the Illumina HiSeq2500 platform, with a 2 × 125 bp paired-end (PE) configuration in the High Output mode (V4 chemistry). Each sample was sequenced on four different flowcell lanes to avoid lane effect. In total, 269,440,904 reads were obtained for the six SG samples, 257,744,832 reads for the three AT samples, and 1,678,378,894 reads for the 33 head samples. Raw data are available in NCBI Sequence Read Archive (<https://trace.ncbi.nlm.nih.gov/Traces/sra/>) with reference numbers shown in Supplementary Table S2.

### De Novo Assembly

Reads from the three SG samples from P123 (this study) and LSR1 (Boulain et al., 2018) were trimmed using trimmomatic (version 0.36, options ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36), and an assembly for each biotype was done using Trinity (v2.5.1) (Grabherr et al., 2011). Lowly expressed contigs were removed by applying a filter with RSEM (--fpkm\_cutoff 0.5, --isopt\_cutoff=15.0) (Li and Dewey, 2011). The remaining contigs were mapped on the LSR1 reference genome (The International Aphid Genomics Consortium, 2010) with gmap (version 2018-03-25) (Wu et al., 2016).

Unmapped contigs from each LSR1 and P123 SG library were searched against the nonredundant protein database using a blastx (BLAST+ v2.5.0, e-value = 1e-8) (Camacho et al., 2009) and P123 contigs were blasted against LSR1 contigs to identify those unmapped contigs that were similar between both aphid lines (blastn, e-value = 1e-8).

### Read Mapping and Gene Expression Analysis

The gene expression patterns of *A. pisum* SG, AT, and head samples were analyzed using the Acyr\_2.0 reference genome assembly (GCF\_000142985.2) with the NCBI *A. pisum* Annotation Release 102, both available at <ftp://ftp.ncbi.nlm.nih.gov/genomes>. The PE reads were mapped on the reference genome using STAR v2.5.2 (Dobin et al., 2013) with the following parameters: outFilterMultimapNmax=5, outFilterMismatchNmax=3, alignIntronMin=10, alignIntronMax=50,000, alignMatesGapMax=50,000. Subread featureCounts (Liao et al., 2014) was used to estimate fragment counts *per* gene using default parameters. Because some viruses might be associated with adaptation of the pea aphid to its host plants (Lu et al., 2019), reads were also mapped to the genomes of the eight known aphid viruses: the *Acyrtosiphon pisum* virus (NC\_003780.1), the *Rhopalosiphum padi* virus (NC\_001874.1), the *Brevicoryne brassicae* virus (NC\_009530.1), the rosy apple aphid virus (DQ286292.1), the *Aphis*

*glycines virus 2* (NC\_028381.1), the *Macrosiphum euphorbiae virus 1* (NC\_028137.1), the *Myzus persicae densovirus* (NC\_005040.1), and the *Dysaphis plantaginea densovirus* (NC\_034532.1).

Three gene expression analyses were conducted separately (with SGs and ATs only, with LSR1 and P123 heads only and finally with all heads) following previously described workflows (Chen et al., 2016; Lun et al., 2016). First, the raw fragment counts were converted to counts *per million* (CPM) using the edgeR (Robinson et al., 2010) R-implemented package (R Development Core Team, 2017). Expressed genes were filtered based on a CPM > 1 in at least three of the libraries incorporated in the analysis and then CPMs were normalized using the edgeR TMM method for Normalization Factor calculation (Robinson and Oshlack, 2010). The reproducibility of replicates was then assessed by multidimensional scaling (MDS) of distances between gene expression profiles based on filtered and normalized CPMs (Ritchie et al., 2015). Filtering, normalization, and clustering steps realized for the different analyses are presented in **Supplementary Figures S1, S2, and S3**. The MDS analysis revealed three head samples (two replicates of P123 and one replicate of S1PS02 both from pea plant condition) that did not cluster with other pea biotype samples. These three samples were removed before further analyses (**Supplementary Figure S3, Supplemental Table S2**). Based on the different analyses, we defined a set of 13,203 *A. pisum* genes that were expressed in at least one condition (CPM > 1) and considered as our working gene set.

## Differential Expression Analyses

The differential expression between samples was then explored with different functions implemented in edgeR that allowed us to i) estimate the common dispersion among the data, ii) fit a quasi-likelihood negative binomial generalized log-linear model to the data, and iii) perform empirical Bayes quasi-likelihood F-tests to determine DE genes (Lund et al., 2012). Statistical tests were taken into account only when expression level averages were above CPM >1 in at least one of the conditions that were compared; otherwise, comparisons were treated as nonsignificant. Fold changes (FCs) between conditions were calculated from average CPM and a FC threshold was fixed at 1.5 to be considered as a DE gene. P-values of the statistical tests were adjusted using the False Discovery Rate (FDR) (Benjamini and Hochberg, 1995). A first contrast matrix was designed to test for organ effect (SGs vs. ATs) in LSR1 and P123 lineages and therefore identify the genes that show upregulated expression in SGs compared to ATs. Then, contrast matrices were designed to analyze the plant (universal vs. specific), biotype (pea vs. alfalfa), and line effects among the head samples of the six *A. pisum* lines. The plant and line effects were tested within each biotype whereas the biotype effect was tested between biotypes. DE genes were retained based on the FC and FDR from edgeR as previously described, except for testing the biotype effect. As we noticed that some genes showing intra-biotype variability were still present in our biotype-DE set of genes, we applied a Student t-test after the edgeR statistical test (calculated on average CPM from each line) and filtered the DE genes based on a p-value < 0.05 in both methods.

## Secretion Prediction and Orthology Analysis

Signal peptides and nonclassical secretion signals of *A. pisum* proteins were identified using a combination of SignalP v3.0, v4.1 (Bendtsen et al., 2004b; Petersen et al., 2011) and SecretomeP v2.0 (Bendtsen et al., 2004a), as described by Boulain et al. (2018). Then, among these proteins that are predicted to be secreted, the ones containing membrane-inserted domains such as transmembrane domains (Krogh et al., 2001) or GPI anchors (Pierleoni et al., 2008) were removed as they are likely not secreted.

To assign an orthology level to each *A. pisum* gene, we determined groups of orthologs among 17 arthropod genomes (Boulain et al., 2018). The longest protein isoforms from each arthropod species were used to run OrthoDB\_soft\_1.6 (Kriventseva et al., 2015) and the levels of orthology were assigned referring to the species phylogeny established by Boulain et al. (2018). The differences in orthologous categories between salivary effector subsets and other genes were then analyzed using proportion tests implemented in R. The groups of orthologs generated by OrthoDB were also used to identify *A. pisum* unique (single copy) or duplicated (multiple copy) genes. Then, we examined whether the salivary effector sets contained more duplicated genes than expected by chance alone. A significant effect was demonstrated if the number of genes that were duplicated lay above the 95% confidence interval (CI) of the expectation. In addition, 95% CI was computed by randomly sampling the number of genes contained in each salivary effector subset (152, 103, and 3,291 for alfalfa-up, pea-up, and non-DE, respectively) from the list of 3,546 salivary effector genes and counting the number of duplicated genes in this random sample. This step was repeated 10,000 times.

## Copy Number Variation Analysis

Population genomic data from Nouhaud et al. (2018), that consisted of Illumina sequencing of two pools of 60 pea- or alfalfa-adapted genotypes (pool-seq) with coverage values >110X each, were used to evaluate copy number variation. The reads from pea and alfalfa biotype pools were mapped following the protocol described by Nouhaud et al. (2018), only primary alignments were kept, and low-quality mapping ( $q < 20$ ) and identically located reads resulting from PCR duplication were removed with MarkDuplicates from Picard tools (<http://broadinstitute.github.io/picard/>). A mean coverage for each exon for each biotype was then computed with Bedtools coverageBed (Quinlan, 2014) with the mean option. The mean coverage of each gene was then calculated by summing its exon coverages and dividing by its total exon size. Then, for the purpose of normalization, these coverages were divided by the average coverage of each gene calculated separately on each biotype pool. Finally, the ratio of coverages was computed for each gene using the normalized mean coverages obtained on both pools.

## RESULTS

### Plant Specialization of *A. Pisum* Lines That Belong to Alfalfa and Pea Biotypes

We selected three aphid lines (LSR1, LL01, and L84) identified as alfalfa biotypes and another three lines (P123, ArPo58, and S1PS02)

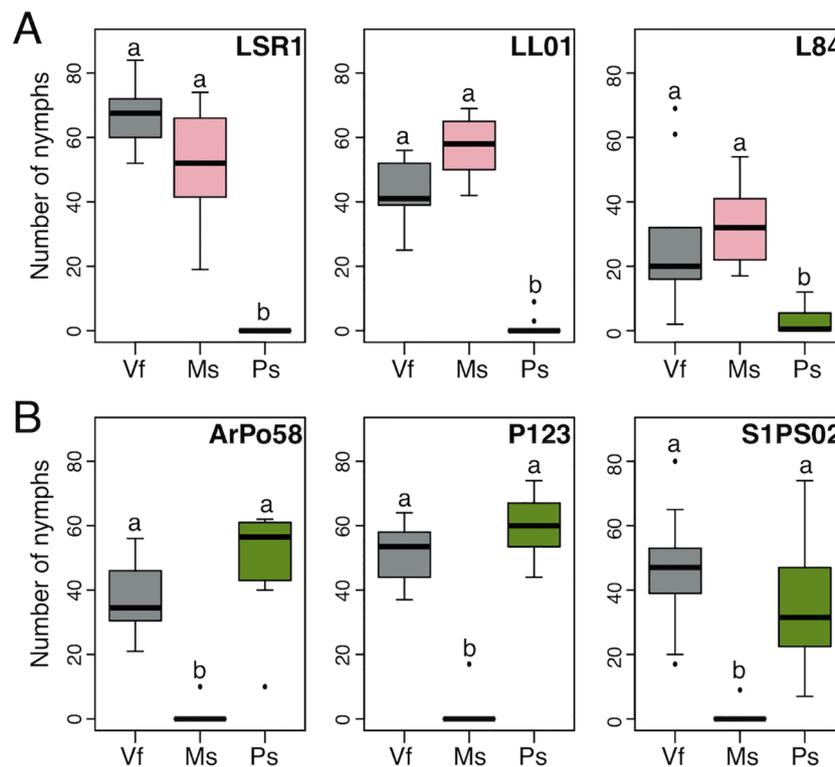
identified as pea biotypes based on the plant from which they were collected and their genetic profiles at several polymorphic microsatellite loci (Peccoud et al., 2009). These aphid lines were collected in different locations and maintained in our lab on the universal host of *A. pisum*, *V. faba* (faba bean) (Supplementary Table S1). To confirm the assigned biotypes, we examined their fecundity on *M. sativa* (alfalfa), *P. sativum* (pea), and *V. faba*. Although there was variation in total nymph production between the lines, both biotypes produced a large number of nymphs on *V. faba* and their respective specific host plant but not on the nonspecific host plant (Figure 1). Hence, these six aphid lines showed distinct host specificity and served as a model system to examine biotype-specific gene expression patterns. We also showed that all the lines performed equally well on *V. faba* and their specific hosts, confirming the “universal host status” of *V. faba*.

### Candidate Salivary Effector Genes Were Identified From Two *A. Pisum* Lines

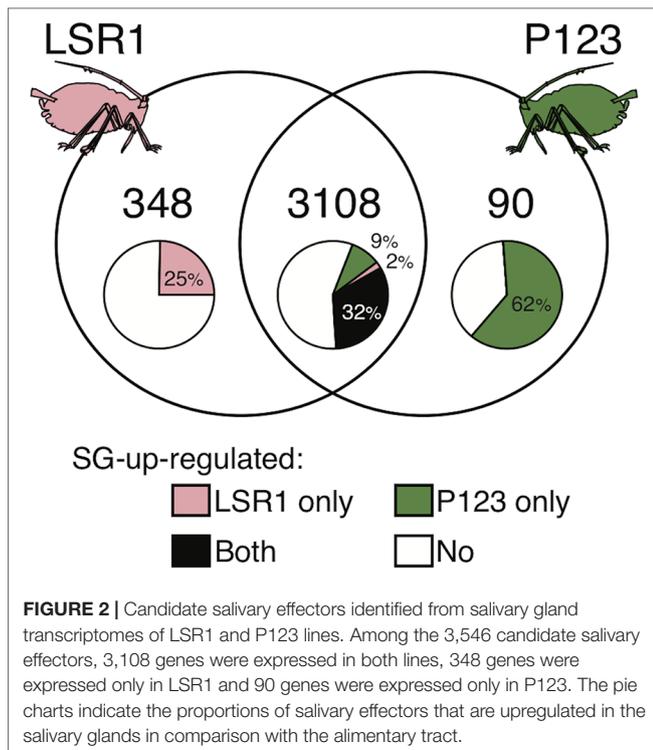
RNA-seq analysis of P123 (pea biotype) SG and AT samples along with LSR1 SG and AT samples (Boulain et al., 2018) retained 12,421 protein-coding genes for analysis and identified 3,546 genes that are expressed (CPM > 1) in SGs of at least one of the aphid lines and encoding proteins that are predicted to be secreted (Boulain et al., 2018). Out of the 3,546 candidate salivary effector genes, 3,108 genes were commonly expressed

in SGs of the two aphid lines, while 348 and 90 genes were specifically expressed in LSR1 and P123, respectively (Figure 2). The comparison between the SG and AT samples from each aphid line allowed us to identify SG-upregulated genes among salivary effector genes. Among the 3,108 common salivary effectors, 32% (1,007 genes) were SG-upregulated in both LSR1 and P123 lines, whereas 2% (63 genes) and 9% (273 genes) were SG-upregulated only in LSR1 and P123, respectively. Out of the LSR1-specific salivary genes, 25% (86 genes) were upregulated in LSR1 SGs, whereas 62% (56 genes) of the P123-specific salivary effectors were upregulated in P123 SGs. The overlap between the two lines led to a total set of 1,485 SG-upregulated effector candidates (Figure 2, Supplementary Table S3).

There was a possibility of not detecting expression of P123-specific genes in this analysis because the LSR1 reference genome was used for mapping and counting the RNA-seq data. Therefore, we conducted *de novo* assembly of SG RNA samples for each aphid line and mapped them on the reference genomes (LSR1 and the obligate symbiont, *Buchnera aphidicola*). LSR1 SG RNA samples produced 565 unmapped contigs (mean length 454 bp, median 284 bp) while P123 SG RNA samples produced 566 unmapped contigs (mean length 453 bp, median 276 bp), out of which 108 showed high homology to unmapped *de novo* assembled LSR1 contigs. Unmapped contigs were BLASTed against NCBI nonredundant protein sequences (Supplementary Table S4). More than 360 contigs in each sample had no BLAST hit and



**FIGURE 1** | *Acyrthosiphon pisum* lines that belong to the alfalfa (A) or pea (B) biotype show distinct host specificity. The numbers of nymphs produced by single aphid mother during 18 days on the universal host plant *Vicia faba* (Vf), specific host plants *Medicago sativa* (Ms), and *Pisum sativum* (Ps) are shown. Letters above boxes indicate significant differences determined by multiple Kruskal-Wallis tests for each aphid line (in LSR1:  $H = 25.235$ , 2 d.f.,  $P < 0.001$ ; in LL01:  $H = 26.479$ , 2 d.f.,  $P < 0.001$ ; in L84:  $H = 21.499$ , 2 d.f.,  $P < 0.001$ , in ArPo58:  $H = 25.778$ , 2 d.f.,  $P < 0.001$ ; in P123:  $H = 23.416$ , 2 d.f.,  $P < 0.001$  and in S1PS02:  $H = 23.179$ , 2 d.f.,  $P < 0.001$ ).



more than 120 contigs of each sample showed similarity with hypothetical or uncharacterized proteins. Since these unmapped contigs from two aphid lines showed similar numbers, short length, and high rate of no BLAST hit, we concluded that use of the LSR1 reference genome for mapping and counting the SG RNA-seq data would not miss a large number of P123-specific salivary genes, if they exist, and continued to use the reference genome for further study.

## A Large Majority of Candidate Salivary Effector Genes Was Detected in Head Samples

We reasoned that examination of gene expression patterns in multiple aphid lines that belong to the same biotype would distinguish biotype-specific gene expression patterns from line-specific expression patterns. However, dissection of SGs is difficult and preparation of SG samples for six aphid lines was not realistic for us. Hence, we decided to use head samples, which are easier to prepare compared to SG samples, to examine the expression patterns of candidate salivary effector genes. We examined expression patterns of the 3,546 candidate salivary effector genes in the SG and head samples of LSR1 and P123 (reared on *V. faba*). In both sets of samples, gene expression levels in SGs and heads were well correlated (Supplementary Figure S4), and 3,165 (91.6%) and 3,107 (97.1%) of candidate salivary effector genes identified for each line were detected in head samples of LSR1 and P123, respectively. Hence, the aphid head samples provide approximate information on the expression levels of salivary genes and can be exploited to identify the candidate salivary genes that are expressed in a biotype-specific manner. Note that

none of the reads mapped to the eight aphid viral genomes; thus, no aphid line seemed to be infected by the viruses.

## Aphid Line and Biotype, But Not Host Plants, Had a Marked Effect on the Expression of Candidate Salivary Effector Genes

The six aphid lines were reared on either *V. faba* or on their specific host plant (*M. sativa* and *P. sativum*, respectively, for alfalfa and pea biotypes) for 9 days and RNA of heads was prepared and subjected to RNA-seq analysis. A distance-based clustering analysis of global expression patterns showed a strong effect of aphid lines and biotypes whereas the clustering was not influenced by host plant (Supplementary Figure S3C). We tested the effects of the three factors (line, biotype, and plant) and identified DE genes due to each factor (Table 1, Supplementary Table S3). Only six and 12 genes were DE depending on the host plants in the alfalfa biotype and the pea biotype respectively. Two genes were commonly downregulated in the two aphid biotypes feeding on *V. faba* compared to the specific plants (*M. sativa* or *P. sativum*) and encoded a linear gramicidin synthase subunit D and an unknown protein. Out of the 16 DE genes, four encoded candidate effectors and all of these were upregulated in the pea biotype when they were feeding on *V. faba* compared to *P. sativum*. These four candidate effector genes (predicted to encode an uncharacterized protein, a dnaJ homolog subfamily B member 11, a probable low-specificity L-threonine aldolase 2, and an endoplasmic reticulum resident protein 44), as well as the rest of the plant DE genes encoded seemingly unrelated proteins (Supplementary Table S3). Meanwhile 689 and 7,207 genes were DE depending on aphid biotype and line, respectively. More than one third (255) of the genes that showed biotype-specific differential expression patterns were candidate salivary effector genes (Table 1, Supplementary Table S3).

**TABLE 1 |** Differentially expressed genes in the head samples of the alfalfa and pea biotypes reared on the universal and specific host plants.

Contrast	# of DE genes <sup>a</sup>	# of DE effector genes <sup>b</sup>
<b>Plant effect</b> (universal vs. specific)		
• in alfalfa biotype	6	0
• in pea biotype	12	4
Total	16	4
<b>Biotype effect</b> (alfalfa vs. pea) <sup>c</sup>		
Total	689	255
<b>Line effect</b> (line vs. line) <sup>d</sup>		
• in alfalfa biotype	6583	1934
• in pea biotype	2165	697
Total	7207	2116

<sup>a</sup>Number of protein-coding genes that are differentially expressed with a FDR < 0.05 and a FC > 1.5 (from overall set of 18,601 protein coding genes existing in the NCBI *Acyrtosiphon pisum* Annotation Release 102).

<sup>b</sup>Number of differentially expressed genes that are considered as candidate salivary effectors (from overall set of 3,546 candidate salivary genes).

<sup>c</sup>In addition to FDR and FC filtering, a Student t-test was applied to exclude DE genes that showed high intra-biotype variability.

<sup>d</sup>The line effect is computed independently in each biotype and a gene is considered as DE when a FDR < 0.05 and a FC > 1.5 are observed between at least two of the three lines that constitute each biotype.

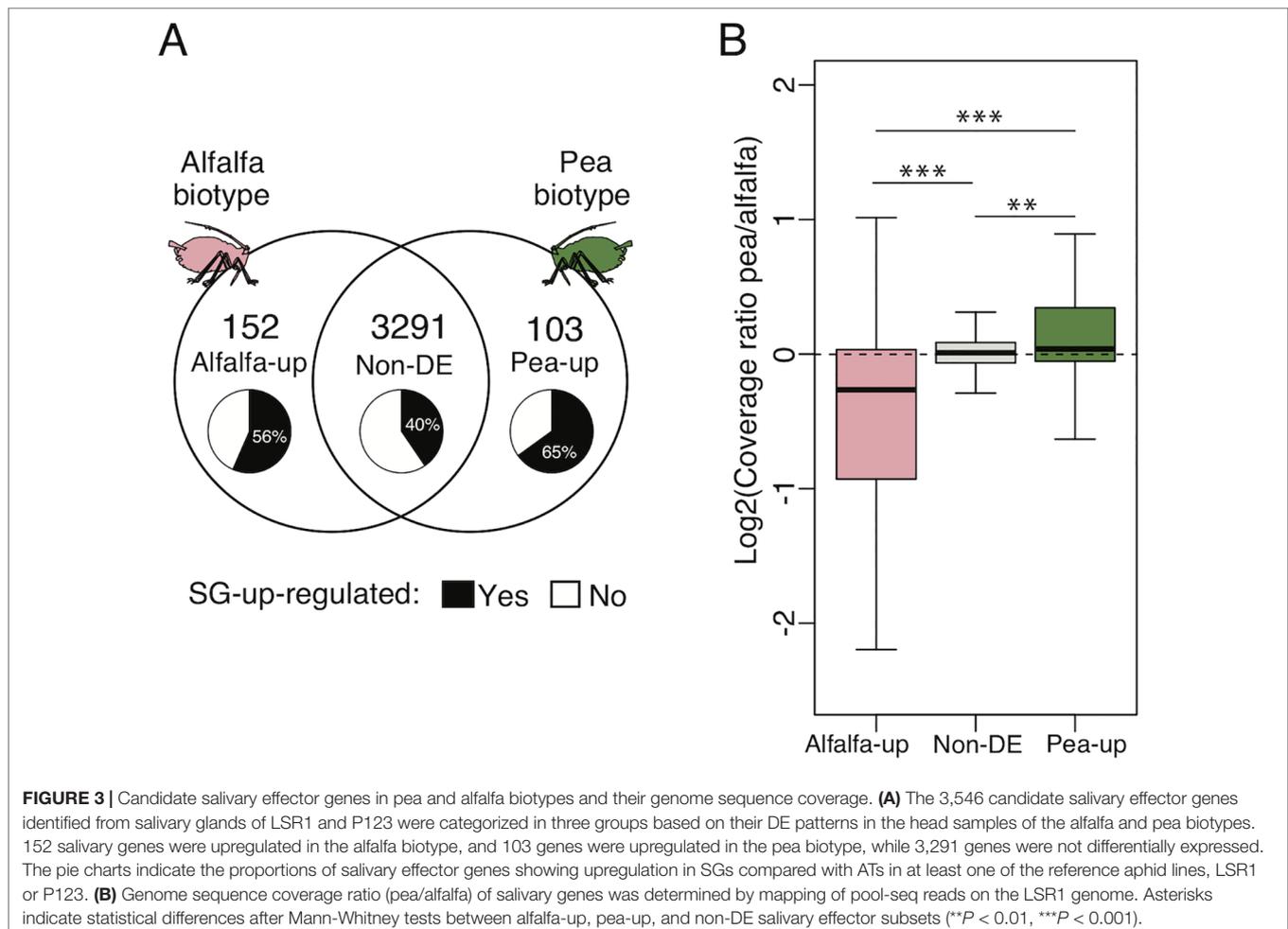
## Biotype-Specific DE Salivary Effector Sets Were Enriched With Duplicated and Aphid-Specific Genes

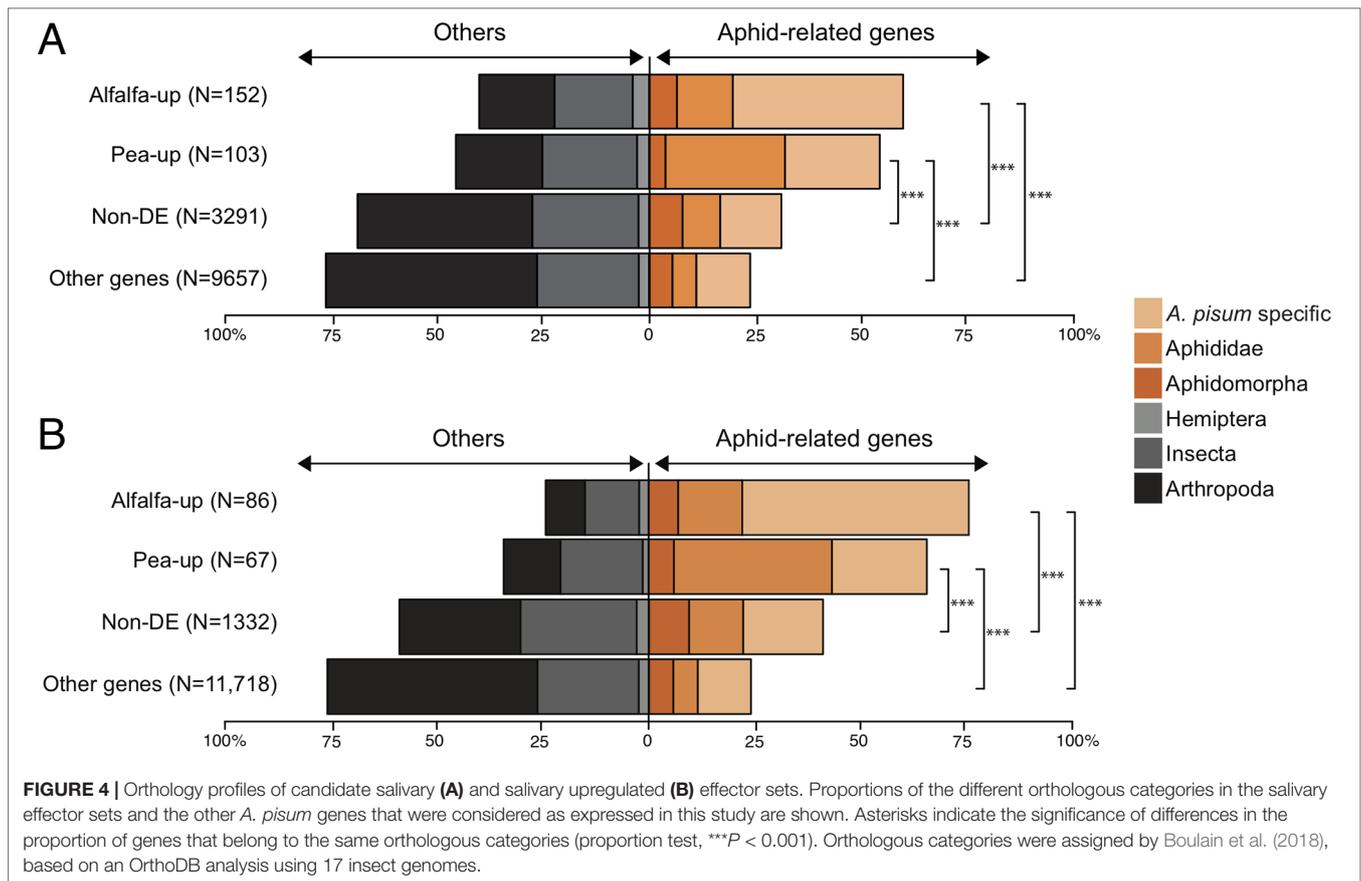
Out of the 3,546 candidate salivary effector genes identified from the SGs of LSR1 and P123 reference lines, 152 were significantly upregulated in the alfalfa biotype (alfalfa-up) compared to the pea biotype, and 103 were upregulated in the pea biotype (pea-up) compared to the alfalfa biotype (Figure 3A). The rest of the 3,291 genes were not significantly DE between the two biotypes (non-DE). Among these alfalfa-up, pea-up, and non-DE subsets, 86 (56%), 67 (65%), and 1,332 (40%) candidate salivary effectors, respectively, were upregulated in the SGs compared to the ATs in at least one of the two reference lines.

Orthology analysis showed that both alfalfa-up and pea-up salivary effector sets contained high proportions of aphid lineage-specific genes compared to the non-DE sets and the other genes that were not considered as candidate salivary effector genes (Figure 4A). The proportion of aphid lineage-specific genes was even higher (>60%) in the alfalfa-up and pea-up subsets when only SG-upregulated genes of each category were considered (Figure 4B). The alfalfa-up and pea-up sets contained 79 (52%) and 57 (55%) genes that encode uncharacterized proteins, respectively (Supplementary Table S3).

In our previous study, we found that *A. pisum* candidate salivary effector genes contained multiple members of multigene families (Boulain et al., 2018). Thus, we examined whether the alfalfa-up or pea-up subsets contained more duplicated genes than expected by chance alone (tested on genes having at least one paralogue). The observed numbers of duplicated genes in the two subsets always lay above the 95% CI, reflecting a higher number of duplicated genes than expected (alfalfa-up: 65 genes, 95% CI = [32, 52] and pea-up: 42 genes, 95% CI = [20, 38]). In contrast, the non-DE subset contained fewer duplicated genes than expected as the number of observed genes lay below the 95% CI (894 genes, CI = [915, 942]).

Among these duplicated genes, a subset of the *A. pisum*-expanded Aminopeptidase-N gene family showed a clear biotype-specific expression pattern (Figure 5). Out of the 27 Aminopeptidase-N proteins that are predicted to be effectors (Boulain et al., 2018), seven were included in the alfalfa-up set while the remaining 20 were included in the non-DE set. Moreover, these alfalfa-up Aminopeptidase-N genes were classified as either “clade 4”, in which episodic events of positive selection have been reported, or “no clade” due to their diversified sequences (Boulain et al., 2018). Many Aminopeptidase-N genes with no assigned clade were lowly expressed in both biotypes while more than half of the genes classified to other clades were highly expressed in both biotypes.





## Differential Expression of Candidate Salivary Genes Is Associated With Copy Number Variation Between the Two Biotypes

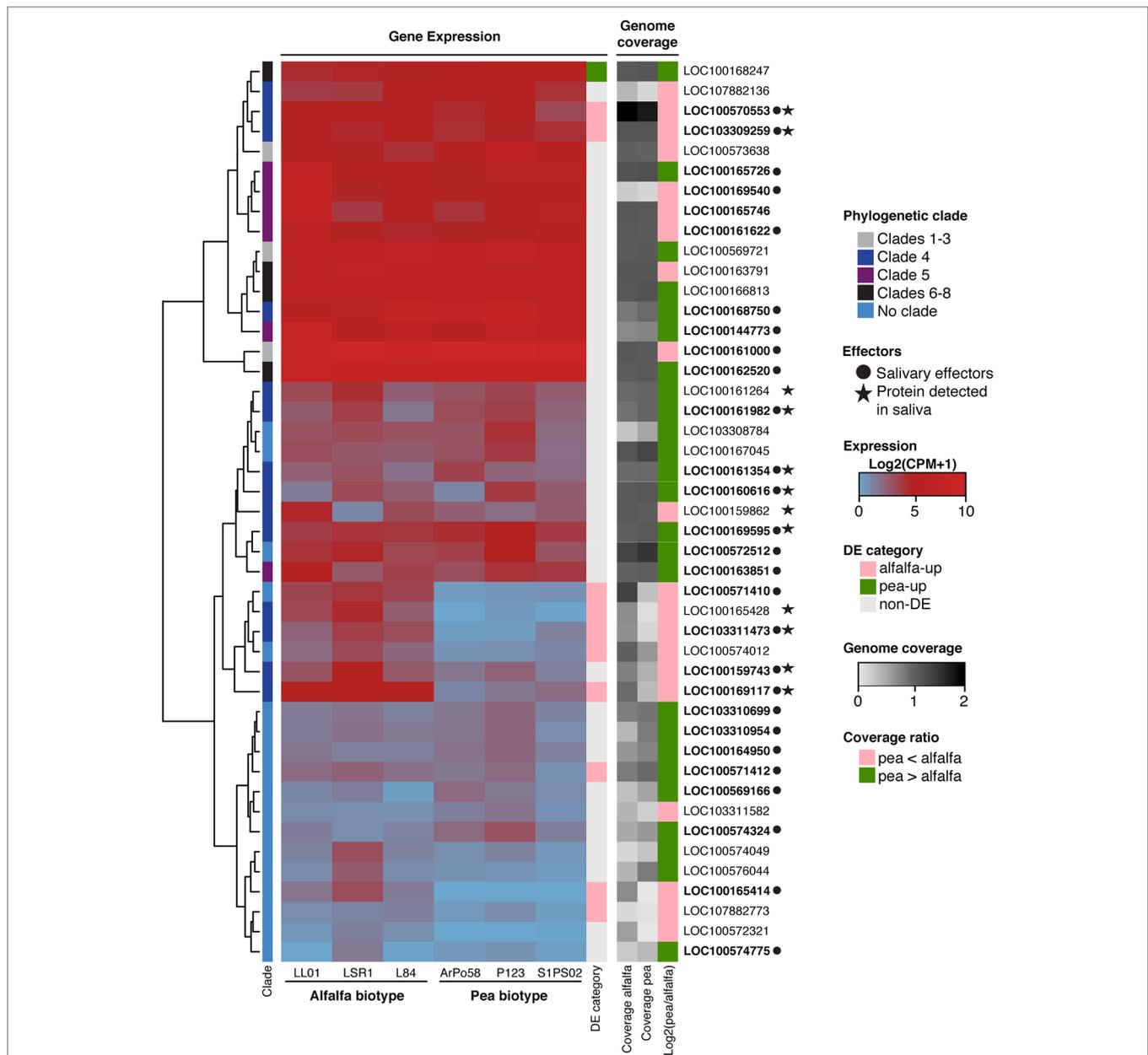
As the biotype-specific differential expression of salivary genes may result from copy number variation between the alfalfa and pea biotypes, we examined the sequence coverage of the genomes of the two biotypes using the genomic pool-seq data created previously (Nouhaud et al., 2018). Comparison of the sequencing coverage ratio between the two biotypes revealed copy number variation. Mean coverage ratio (pea/alfalfa) of the alfalfa-up set was significantly lower than the non-DE set and that of the pea-up set was significantly higher. This pattern was observed among the salivary effectors (Figure 3B) as well as in the SG-upregulated effectors (Supplementary Figure S5). The coverage of four alfalfa-up salivary effector genes was very low in the pool-seq of the pea biotype ( $<0.1$ ), and some of these genes were very lowly expressed in the SG of P123 and the head samples of the three pea biotype lines (CPM  $< 1$ ), while they were expressed (CPM  $> 1$ ) in the three lines of the alfalfa biotype. These genes were predicted to encode an Aminopeptidase-N-like protein, a fatty acid synthase-like protein, a ubiquitin-C-like protein, and an uncharacterized protein. Those genes may not be encoded in the genome of the pea biotype lines and be specific to the alfalfa biotype although their predicted functions do not seem to be related. No such gene (very low coverage and expression

value in the alfalfa biotype) was observed in the pea-up gene set (Supplementary Table S3).

## DISCUSSION

To understand the molecular basis of host plant adaptation in *A. pisum* biotypes, we created a comprehensive list of candidate salivary genes using two aphid lines that belong to the pea or alfalfa biotype and compared their expression patterns in the two biotypes, each represented by three genetically distinct aphid lines. Due to the difficulty of creating SG RNA samples, we used aphid head samples to examine biotype-specific expression patterns of candidate salivary genes and the effect of host plants. Comparison of gene expression levels in the head and SG samples showed that expression levels of the majority of genes were correlated between the two sample types with some exceptions. The head samples contain many organs (eyes, antennae, brain, etc.) in addition to SGs. Some of the SG-expressed genes might be expressed in other organs than the SGs and, in such cases, correlation between the expression values in the SGs and the heads is not expected. Nonetheless, this study presents one of the most thorough and comparative analyses of salivary gene expression in genetically related insect lines with clearly distinct host plant specificity.

The analyses of the head samples showed strong line and biotype effects on aphid gene expression and revealed a very weak effect of host plant. Our results are in line with the study



**FIGURE 5 |** Gene expression and genome sequence coverage of Aminopeptidase-N gene family. The tree on the left is a hierarchical clustering of expression values and squares on the right side of the tree correspond to the clades of Aminopeptidase-N members assigned by Boulain et al. (2018). A heatmap of expression levels of Aminopeptidase-N family members in various aphid lines is shown in the middle. DE categories and sequencing coverage of the genomes and NCBI accession numbers are shown on the right. The genes were considered as candidate salivary effector genes if they were present in the 3,546 set established by the transcriptomes of LSR1 and P123 salivary glands. Detection of proteins in aphid saliva was reported by Boulain et al. (2018) and Carolan et al. (2009).

of Eyres et al. (2016), which examined transcriptional patterns of six pea aphid biotypes reared on their specific and universal host plants and found little expression change caused by host plant type. Unlike the generalist aphid *M. persicae*, which shows large changes in gene expression to acclimatize to host plant (Mathers et al., 2017), *A. pisum* biotypes seem to make very little transcriptional adjustment to their host plants. This difference in transcriptional plasticity may explain the differences in host range of the two aphid species (generalist vs. specialist) although

further examination of multiple generalist and specialist aphids are required to link the transcriptional plasticity with host range.

We focused our analyses on the expression patterns of the candidate salivary gene sets created by LSR1 and P123 SG transcriptomes. Although the effect of aphid line on gene expression patterns was considerable, we were able to identify 153 and 103 candidate salivary genes that are upregulated in the alfalfa and the pea biotypes, respectively. Differential expression of salivary genes in six *A. pisum* biotypes was reported previously using a smaller list

of candidate salivary genes (307 genes) published at the time and by using multiple aphid lines as biological replicates of a biotype (Eyres et al., 2016). Our study refined the analysis by creating and compiling biotype-specific salivary gene sets for an alfalfa- and a pea-adapted *A. pisum* line, by expanding the candidate salivary genes list by more than 10 times and by including three biological replicates for each aphid line and for each condition.

The orthology analysis of candidate salivary genes revealed that the alfalfa-up and pea-up gene sets contain higher proportions of aphid lineage-specific genes and the proportion of those genes was even higher when only the SG-upregulated salivary genes were analysed. These alfalfa-up and pea-up gene sets also contain higher numbers of duplicated genes than expected. These observations support the scenario that biotype-specific salivary effectors may have evolved recently and diversified through duplication events, possibly in relation to the diversification of the pea aphid complex of biotypes (Peccoud et al., 2009). Under this scenario, certain gene duplicates would tend to be recruited differently in the pea and alfalfa biotypes to achieve better performance on each host plant while other copies would maintain basic functions and lie in the non-DE set. Analysis of the gene family of Aminopeptidase-N supports this scenario as it revealed a subset of genes that show high expression values in all the six aphid lines and another set of genes that show differential expression in a biotype specific manner.

Four alfalfa-up salivary genes showed virtually no expression values in both heads and SGs and low genome coverage in the pea biotype. These genes may not exist in the three lines of pea biotype studied here and be considered as alfalfa biotype-specific genes. Although the expression levels of those genes in the alfalfa biotype tend to be low, they may be required for efficient feeding on alfalfa plants or may trigger unwanted responses in pea plants. On the contrary, all the pea-up genes were highly expressed in the pea biotype and all of them seem to be encoded in the alfalfa biotype genome. Our analysis of *de novo* assembled transcripts showed very little difference between LSR1 and P123 lines. Although there is a possibility that we missed some genes that are specifically encoded in the pea biotype and absent in the alfalfa biotype by using the LSR1 genome as reference for the RNA-seq analyses, the number of such genes should be very small. Thus, except for a few biotype-specific genes, the repertoires of the salivary genes in the two biotypes were almost identical, and small sets of genes showed differential expression which might determine host plant specificity. The evolutionary history of specialization to *P. sativum* and *M. sativa* in *A. pisum* biotypes has not been elucidated, and we cannot speculate on the evolutionary process of the differential expression (gene loss vs. gain, induction vs. suppression) of these genes.

Although the effect of biotype on salivary effector expression was small, the host plants showed an even smaller effect on effector transcription. This suggests that the gene expression differences in candidate salivary effectors between the biotypes largely result from genomic variation and not from expression plasticity. This is supported by another result showing a low ratio of genome coverage (pea/alfalfa) for the alfalfa-up gene set and a higher ratio for the pea-up gene set: differential expression of these two sets of genes can be partly explained by copy number variation in the two biotypes. In addition to copy number variation, variation of coding sequences or promoter regions (small insertion/deletion/inversion, SNPs) and

gene rearrangements may also be the causes of differential expression of candidate salivary effectors between the biotypes. As genome sequences of different aphid lineages and a better assembly of the pea aphid genome are becoming available (Nouhaud et al., 2018; Li et al., 2019), dedicated studies are needed for a thorough investigation of biotype-specific amino acid sequence polymorphism of candidate effectors and potential causes of differential gene expression.

In conclusion, this study provides a comprehensive list of candidate salivary effectors and brings evidence that a subset of salivary genes that include a high proportion of aphid lineage-specific genes and duplicated genes are DE in two aphid biotypes with distinct host specificity. The identified DE salivary genes are strong candidate genes that might be involved in host plant adaptation in the *A. pisum* biotypes and deserve further functional characterization.

## DATA AVAILABILITY STATEMENT

The datasets generated and analysed for this study can be found in the NCBI using accession numbers: SRX3969578, SRX3969577, SRX3969576, SRX5933263, SRX5933264, SRX5933265, SRX3969582, SRX3969581, SRX3969580, SRX5933266, SRX5933267, SRX5933268, SRX5936232, SRX5936233, SRX5936234, SRX5936235, SRX5936236, SRX5936237, SRX5936238, SRX5936239, SRX5936240, SRX5936240, SRX5936248, SRX5936249, SRX5936246, SRX5936247, SRX5936252, SRX5936253, SRX5936250, SRX5936251, SRX5936244, SRX5936245, SRX5936228, SRX5936227, SRX5936226, SRX5936225, SRX5936224, SRX5936223, SRX5936222, SRX5936221, SRX5936230, SRX5936229, SRX5936242, SRX5936243, SRX5936231.

## AUTHOR CONTRIBUTIONS

EG, SM, J-CS, and AS designed the experiments. HB, FL, JJ, EG, SM, J-CS, and AS conducted the experiments, analyzed the data, and wrote the manuscript. J-CS and AS provided funding for the project.

## FUNDING

This work was funded by the Agence Nationale de la Recherche (ANR) Bugspit (ANR-13-JSV7-0012-01) to AS and by the ANR Speciaphid (ANR-11-BSV7-005-01) to J-CS.

## ACKNOWLEDGMENTS

We thank Gaëtan Denis, Jean-François Le Gallic, Frédérique Mahéo, and Sylvie Tanguy for technical support and Dr. Richard Harrington for critical reading of a version of the manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2019.01301/full#supplementary-material>

## REFERENCES

- Atamian, H. S., Chaudhary, R., Cin, V. D., Bao, E., Girke, T., and Kaloshian, I. (2012). In planta expression or delivery of potato aphid *Macrosiphum euphorbiae* effectors Me10 and Me23 enhances aphid fecundity. *MPMI* 26, 67–74. doi: 10.1094/MPMI-06-12-0144-FI
- Bendtsen, J. D., Jensen, L. J., Blom, N., von Heijne, G., and Brunak, S. (2004a). Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel.* 17, 349–356. doi: 10.1093/protein/gzh037
- Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004b). Improved prediction of signal peptides: signalP 3.0. *J. Mol. Biol.* 340, 783–795. doi: 10.1016/j.jmb.2004.05.028
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300. doi: 10.1016/j.neuroimage.2012.07.004
- Birnbaum, S. S. L., and Abbot, P. (2018). Insect adaptations toward plant toxins in milkweed–herbivores systems – a review. *Entomol. Exp. Appl.* 166, 357–366. doi: 10.1111/eea.12659
- Bos, J. I. B., Prince, D., Pitino, M., Maffei, M. E., Win, J., and Hogenhout, S. A. (2010). A functional genomics approach identifies candidate effectors from the aphid species *Myzus persicae* (Green Peach Aphid). *PLOS Genet.* 6, e1001216. doi: 10.1371/journal.pgen.1001216
- Boulain, H., Legeai, F., Guy, E., Morlière, S., Douglas, N. E., Oh, J., et al. (2018). Fast evolution and lineage-specific gene family expansions of aphid salivary effectors driven by interactions with host-plants. *Genome Biol. Evol.* 10, 1554–1572. doi: 10.1093/gbe/evy097
- Butlin, R. K., and Smadja, C. M. (2017). Coupling, reinforcement, and speciation. *Am. Nat.* 191, 155–172. doi: 10.1086/695136
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinf.* 10, 421. doi: 10.1186/1471-2105-10-421
- Carolan, J. C., Fitzroy, C. I. J., Ashton, P. D., Douglas, A. E., and Wilkinson, T. L. (2009). The secreted salivary proteome of the pea aphid *Acyrtosiphon pisum* characterised by mass spectrometry. *Proteomics* 9, 2457–2467. doi: 10.1002/pmic.200800692
- Chen, Y., Lun, A. T. L., and Smyth, G. K. (2016). From reads to genes to pathways: differential expression analysis of RNA-seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Research* 5, 1438. doi: 10.12688/f1000research.8987.2
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635
- Elzinga, D. A., De Vos, M., and Jander, G. (2014). Suppression of plant defenses by a *Myzus persicae* (Green Peach Aphid) salivary effector protein. *MPMI* 27, 747–756. doi: 10.1094/MPMI-01-14-0018-R
- Eyres, I., Jaquière, J., Sugio, A., Duvaux, L., Gharbi, K., Zhou, J.-J., et al. (2016). Differential gene expression according to race and host plant in the pea aphid. *Mol. Ecol.* 25, 4197–4215. doi: 10.1111/mec.13771
- Forister, M. L., Novotny, V., Panorska, A. K., Baje, L., Bassett, Y., Butterill, P. T., et al. (2015). The global distribution of diet breadth in insect herbivores. *Proc. Natl. Acad. Sci. U. S. A.* 112, 442–447. doi: 10.1073/pnas.1423042112
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883
- Guy, E., Boulain, H., Aigu, Y., Le Pennec, C., Chawki, K., Morlière, S., et al. (2016). Optimization of agroinfiltration in *Pisum sativum* provides a new tool for studying the salivary protein functions in the pea aphid complex. *Front. Plant Sci.* 7, 1171. doi: 10.3389/fpls.2016.01171
- Jaquière, J., Stoeckel, S., Nouhaud, P., Mieuze, L., Mahéo, F., Legeai, F., et al. (2012). Genome scans reveal candidate regions involved in the adaptation to host plant in the pea aphid complex. *Mol. Ecol.* 21, 5251–5264. doi: 10.1111/mec.12048
- Kettles, G. J., and Kaloshian, I. (2016). The potato aphid salivary effector Me47 is a Glutathione-S-Transferase involved in modifying plant responses to aphid infestation. *Front. Plant Sci.* 7, 1142. doi: 10.3389/fpls.2016.01142
- Kriventseva, E. V., Tegenfeldt, F., Petty, T. J., Waterhouse, R. M., Simão, F. A., Pozdnyakov, I. A., et al. (2015). OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying free software. *Nucleic Acids Res.* 43, D250–D256. doi: 10.1093/nar/gku1220
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315
- Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinf.* 12, 323. doi: 10.1186/1471-2105-12-323
- Li, Y., Park, H., Smith, T. E., and Moran, N. A. (2019). Gene family evolution in the pea aphid based on chromosome-level genome assembly. *Mol. Biol. Evol.* 36, 2143–2156. doi: 10.1093/molbev/msz138
- Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656
- Lu, H., Zhu, J., Yu, J., Chen, X., Kang, L., and Cui, F. (2019). A symbiotic virus facilitates aphid adaptation to host plants by suppressing jasmonic acid responses. *MPMI J.* doi: 10.1094/MPMI-01-19-0016-R
- Lun, A. T. L., Chen, Y., and Smyth, G. K. (2016). “It’s DE-licious: a recipe for differential expression analyses of RNA-seq experiments using quasi-likelihood methods in edgeR,” in *Statistical genomics: methods and protocols methods in molecular biology*. Eds. E. Mathé and S. Davis (New York, NY: Springer New York), 391–416. doi: 10.1007/978-1-4939-3578-9\_19
- Lund, S. P., Nettleton, D., McCarthy, D. J., and Smyth, G. K. (2012). Detecting differential expression in RNA-sequence data using quasi-likelihood with shrunken dispersion estimates. *Stat. Appl. Genet. Mo. B.* 11, doi: 10.1515/1544-6115.1826
- Mathers, T. C., Chen, Y., Kaithakottil, G., Legeai, F., Muggford, S. T., Baa-Puyoulet, P., et al. (2017). Rapid transcriptional plasticity of duplicated gene clusters enables a clonally reproducing aphid to colonise diverse plant species. *Genome Biol.* 18, 27. doi: 10.1186/s13059-016-1145-3
- Meziadi, C., Blanchet, S., Geffroy, V., and Pflieger, S. (2017). Virus-Induced Gene Silencing (VIGS) and foreign gene expression in *Pisum sativum* L. using the “One-Step” Bean pod mottle virus (BPMV) viral vector. *Methods Mol. Biol.* 1654, 311–319. doi: 10.1007/978-1-4939-7231-9\_23
- Meziadi, C., Blanchet, S., Richard, M., Pilet-Nayel, M.-L., Geffroy, V., and Pflieger, S. (2016). Bean pod mottle virus: a new powerful tool for functional genomics studies in *Pisum sativum*. *Plant Biotechnol. J.* 14, 1777–1787. doi: 10.1111/pbi.12537
- Moreno, A., Garzo, E., Fernandez-Mata, G., Kassem, M., Aranda, M. A., and Fereres, A. (2011). Aphids secrete watery saliva into plant tissues from the onset of stylet penetration. *Entomol. Exp. Appl.* 139, 145–153. doi: 10.1111/j.1570-7458.2011.01117.x
- Mutti, N. S., Louis, J., Pappan, L. K., Pappan, K., Begum, K., Chen, M.-S., et al. (2008). A protein from the salivary glands of the pea aphid, *Acyrtosiphon pisum*, is essential in feeding on a host plant. *Proc. Natl. Acad. Sci. U. S. A.* 105, 9965–9969. doi: 10.1073/pnas.0708958105
- Mutti, N. S., Park, Y., Reese, J. C., and Reeck, G. R. (2006). RNAi knockdown of a salivary transcript leading to lethality in the pea aphid, *Acyrtosiphon pisum*. *J. Insect Sci.* 6, 1–7. doi: 10.1673/031.006.3801
- Naessens, E., Dubreuil, G., Giordanengo, P., Baron, O. L., Minet-Kebdani, N., Keller, H., et al. (2015). A secreted MIF cytokine enables aphid feeding and represses plant immune responses. *Curr. Biol.* 25, 1898–1903. doi: 10.1016/j.cub.2015.05.047
- Nouhaud, P., Gautier, M., Gouin, A., Jaquière, J., Peccoud, J., Legeai, F., et al. (2018). Identifying genomic hotspots of differentiation and candidate genes involved in the adaptive divergence of pea aphid host races. *Mol. Ecol.* 27, 3287–3300. doi: 10.1111/mec.14799
- Peccoud, J., Mahéo, F., de la Huerta, M., Laurence, C., and Simon, J.-C. (2015). Genetic characterisation of new host-specialised biotypes and novel associations with bacterial symbionts in the pea aphid complex. *Insect. Conserv. Divers.* 8, 484–492. doi: 10.1111/icad.12131
- Peccoud, J., Ollivier, A., Plantegenest, M., and Simon, J.-C. (2009). A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. *Proc. Natl. Acad. Sci. U. S. A.* 106, 7495–7500. doi: 10.1073/pnas.0811171106
- Peccoud, J., Simon, J.-C., von Dohlen, C., Coeur d’acier, A., Plantegenest, M., Vanlerberghe-Masutti, F., et al. (2010). Evolutionary history of aphid-plant associations and their role in aphid diversification. *C. R. Biol.* 333, 474–487. doi: 10.1016/j.crv.2010.03.004
- Petersen, T. N., Brunak, S., Heijne, G. von, and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785. doi: 10.1038/nmeth.1701
- Pierleoni, A., Martelli, P. L., and Casadio, R. (2008). PredGPI: a GPI-anchor predictor. *BMC Bioinf.* 9, 392. doi: 10.1186/1471-2105-9-392

- Pitino, M., Coleman, A. D., Maffei, M. E., Ridout, C. J., and Hogenhout, S. A. (2011). Silencing of aphid genes by dsRNA feeding from plants. *PLoS One* 6, e25709. doi: 10.1371/journal.pone.0025709
- Pitino, M., and Hogenhout, S. A. (2013). Aphid protein effectors promote aphid colonization in a plant species-specific manner. *MPMI* 26, 130–139. doi: 10.1094/MPMI-07-12-0172-FI
- Quinlan, A. R. (2014). BEDTools: The Swiss-army tool for genome feature analysis. *Curr. Protoc. Bioinf.* 47, 11.12.1–34. doi: 10.1002/0471250953.bi1112s47
- R Development Core Team. (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47–e47. doi: 10.1093/nar/gkv007
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616
- Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11, R25. doi: 10.1186/gb-2010-11-3-r25
- Simon, J.-C., d'Alençon, E., Guy, E., Jacquin-Joly, E., Jaquiéry, J., Nouhaud, P., et al. (2015). Genomics of adaptation to host-plants in herbivorous insects. *Brief. Funct. Genomics* 14, 413–423. doi: 10.1093/bfpp/elv015
- The International Aphid Genomics Consortium. (2010). Genome sequence of the pea aphid *Acyrtosiphon pisum*. *PLOS Biol.* 8, e1000313. doi: 10.1371/journal.pbio.1000313
- Wang, W., Dai, H., Zhang, Y., Chandrasekar, R., Luo, L., Hiromasa, Y., et al. (2015). Armet is an effector protein mediating aphid–plant interactions. *FASEB J.* 29, 2032–2045. doi: 10.1096/fj.14-266023
- Wu, T. D., Reeder, J., Lawrence, M., Becker, G., and Brauer, M. J. (2016). GMAP and GSNAP for genomic sequence alignment: enhancements to speed, accuracy, and functionality. *Methods Mol. Biol.* 1418, 283–334. doi: 10.1007/978-1-4939-3578-9\_15

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Boulain, Legeai, Jaquiéry, Guy, Morlière, Simon and Sugio. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Transcription and Activity of Digestive Enzymes of *Nezara viridula* Maintained on Different Plant Diets

Pablo Emiliano Cantón and Bryony C. Bonning\*

Department of Entomology and Nematology, University of Florida, Gainesville, FL, United States

## OPEN ACCESS

### Edited by:

Isgouhi Kaloshian,  
University of California, Riverside,  
United States

### Reviewed by:

Natraj Krishnan,  
Mississippi State University,  
United States  
Keyan Zhu-Salzman,  
Texas A&M University, United States  
Livy Williams,  
Agricultural Research Service,  
United States Department  
of Agriculture, United States

### \*Correspondence:

Bryony C. Bonning  
bbonning@ufl.edu

### Specialty section:

This article was submitted to  
Invertebrate Physiology,  
a section of the journal  
Frontiers in Physiology

**Received:** 26 September 2019

**Accepted:** 09 December 2019

**Published:** 08 January 2020

### Citation:

Cantón PE and Bonning BC  
(2020) Transcription and Activity  
of Digestive Enzymes of *Nezara*  
*viridula* Maintained on Different Plant  
Diets. *Front. Physiol.* 10:1553.  
doi: 10.3389/fphys.2019.01553

*Nezara viridula* is a polyphagous stink bug that feeds on crops of economic importance such as corn, soybean and cotton. To increase understanding of the ability of this pest insect to feed on such diverse cropping systems, we analyzed the impact of an exclusive diet of corn or green bean on the enzymatic activity and transcriptomic profile of digestive enzymes. Growth rate and survival were reduced when insects were reared exclusively on green bean compared to corn. However, the overall protease and nuclease activity profiles were comparable between the two treatments. Distinct differences in inhibitor sensitivity and activity were seen in some cases, particularly for serine proteases in some regions of the midgut. The transcription profiles from *N. viridula* fed on corn versus green bean were distinct on principal component analysis of RNA-seq data. While specific transcripts differentially transcribed according to diet and across several tissues were identified, a large number of these transcripts remain unannotated. Further annotation for identification of these genes will be important for improved understanding of the remarkable polyphagy of *N. viridula*.

**Keywords:** diet, digestive enzymes, host plant, midgut, transcriptome, stink bug

## INTRODUCTION

Crop damage resulting from insect herbivory is a primary source of economic loss in agriculture. Species that feed on multiple host plants are of particular concern, as they can move between adjacent crops, feeding on plants that mature at different times during the year or persist in refuge plants or grasses. Species of the family Pentatomidae (Hemiptera), such as the southern green stink bug, *Nezara viridula*, possess all of these traits (Panizzi, 1997). Changes in agricultural practices, such as reduction in insecticide use or no tillage strategies, as well as changes in climate, have allowed *N. viridula* along with other species in this family to rise in prominence as crop pests (Panizzi, 2015).

*Nezara viridula* is a cosmopolitan polyphagous species that feeds on the seeds and fruits of more than 100 plant species in 30 different families (Todd, 1989). Some of these plants include crops of high economic importance such as soybean, cotton, and corn (Tillman, 2010). *N. viridula* feeds by a piercing-sucking mechanism in two phases: digestion initiates with the extra-oral secretion of saliva with digestive enzymes such as trypsin and chymotrypsin into plant tissues. Heteropteran species such as *N. viridula* use a non-reflux system and a remarkable maneuverability of their stylets to pierce and deliver extra-oral secretions, which provides a rate of recovery of liquefied tissue of >90% (Cohen, 1995). Following ingestion, this plant matter is completely degraded and absorbed in the midgut (Lomate and Bonning, 2016). *N. viridula* has four

anatomically and physiologically distinct midgut regions: M1, M2, M3, and M4 (Hirose et al., 2009). Most proteolytic activity occurs in M2 and M3 (Cantón and Bonning, 2019) mediated by the cysteine proteases Cathepsin B and L, while M4 functions to house endosymbiotic bacteria (Hosokawa et al., 2016). The proteases and nucleases present in saliva, salivary glands, and midgut tissues of *N. viridula* have been cataloged (Liu et al., 2018). The manner in which these myriad digestive enzymes are employed for digestion of plant material of varied composition is unclear.

The production of digestive protease inhibitors and a wealth of secondary compounds with toxic characteristics are among the main mechanisms by which plants defend against or deter herbivory. The regulation of insect response to plant defense mechanisms is of ongoing interest. In the case of protease inhibitors, insects can respond by general upregulation of digestive enzymes, production of specific enzymes that circumvent inhibition, or by detoxifying the toxic agents (Zhu-Salzman and Zeng, 2015). Some corn varieties produce cysteine proteases that damage the peritrophic membrane lining the insect gut (Pechan et al., 2000), but insects can upregulate inhibitors of these enzymes (Li et al., 2009). Changes in protease activities or gene transcription profiles have been noted in several insects including in response to phytotoxins (Halon et al., 2015), diet source (Coudron et al., 2007; Huang et al., 2017; Rivera-Vega et al., 2017), or adaptation to the presence of plant protease inhibitors (Lara et al., 2000; Oppert et al., 2005; Brioschi et al., 2007). Key proteases or defense mechanisms are potential targets for disruption for stink bug management through the application of protease inhibitors (Schlüter et al., 2010) or gene knockdown (Joga et al., 2016; Ghosh et al., 2017).

The goal of this study was to examine the effect of diet source on the proteases and nucleases of *N. viridula*. We used both transcriptomic and enzymatic assay approaches to generate data for gene transcription and biochemical profiles. Although enzyme activity profiles were similar, diet-dependent variation was sufficient to differentiate the transcriptomes derived from *N. viridula* maintained on corn versus green bean.

## MATERIALS AND METHODS

### Reagents

For nuclease substrates, calf thymus DNA was obtained from Sigma-Aldrich (St. Louis, MO, United States) and baker's yeast RNA was purchased from Fisher Scientific/Alfa Aesar (Haverhill, MA, United States). RNAlater stabilization solution was purchased from Invitrogen (Carlsbad, CA, United States). The protease substrates azocasein,  $N_{\alpha}$ -Benzoyl-D,L-arginine 4-nitroanilide hydrochloride (BAPNA), N-Succinyl-Ala-Ala-Pro-Phe p-nitroanilide (SAAPFpNA), L-Leucine p-nitroanilide (LpNA), pGlu-Pro-Leu p-nitroanilide (pGFLpNA) were obtained from Sigma-Aldrich. Z-Arg-Arg p-nitroanilide (zRRpNA) was acquired from Bachem (Bubendorf, Switzerland). The inhibitors Phenylmethylsulfonyl fluoride (PMSF),  $N_{\alpha}$ -Tosyl-L-lysine chloromethyl ketone hydrochloride (TLCK),  $N_{\alpha}$ -Tosyl-L-phenylalaninechloromethyl ketone (TPCK), E-64,

Ethylenedinitrilotetraacetic acid (EDTA) were also purchased from Sigma-Aldrich.

### Rearing and Dissection of *Nezara viridula*

The *N. viridula* colony was established August 12, 2014 with insects provided by Dr. Jeffrey Davis, Louisiana State University. The colony was reared on mixed diet at 28°C, 65% relative humidity, and a 16:8 hr light/dark photoperiod. The colony was supplemented once yearly with field caught *N. viridula* from Florida. For this study, *N. viridula* sub-colonies were maintained under the same conditions on exclusive diets of either corn cobs with kernels (*Zea mays*) or organic green bean (*Phaseolus vulgaris*) pods. Diet was changed twice weekly. For containers with green beans, pods were arranged in a conical pile and complemented with a moist cotton plug in a 1 oz plastic cup. Cages were inspected each morning for the presence of adults. Following the molt to adult, insects were moved to a new container and allowed to feed for 24 h on their respective fresh diet prior to dissection the following morning. Salivary glands and midgut M1, M2, and M3 were dissected from adults in 0.1 M PBS, pH 7.4. For each biological replicate, tissues from approximately 12 adults were dissected and pooled for enzymatic assays. Midgut sections were dissected from five adults and salivary glands from 12 adults for RNA extraction. Tissues were either flash frozen in liquid nitrogen and stored at -80°C for protein extraction or submerged in RNAlater solution and kept at -20°C until RNA extraction was performed. Three biological replicates were conducted for all experiments.

### Assessment of Insect Growth Rate

For each diet, forty 2nd instar nymphs were put on corn or green bean diet and reared as described above. At each diet change, survival was recorded and the length from the end of the abdomen to tip of the head was measured for all live individuals. Adults were weighed. A two-tailed *t*-test was performed to evaluate significance in differences between the slope and standard error of the regression curves for insect growth (length) on both diets. For survival, the three independent replicates were used to obtain Kaplan Meier curves, and a log rank test was performed between both diets, with a right tailed Chi test evaluation.

### Plant Protein Inhibitor Identification

The ENSEMBL Plant genome annotations of *Z. mays* (corn) and *P. vulgaris* (green bean) were queried through BioMart (Kinsella et al., 2011) for genes associated with the GO:0030414 term "peptidase inhibition activity." Gene IDs were recovered for each plant, and these IDs were used to retrieve their orthologs and PFAM domains from the same database. For each plant, each gene was classified according to the plants in which orthologs were found, if any, and the number of genes in each class counted. The number of PFAM domains associated with each gene in each plant was also counted and classified by the PFAM domain.

## Preparation of Protein Extracts From Tissues

For each of three biological replicates, homogenization of tissues was performed with a Polytron 2500E device (Kinematica, Luzern, Switzerland) in a 1.5 ml microcentrifuge tube on ice, using a 3:1 v/w ratio of PBS 0.1 M pH 7.4 and 10 000 rpm for 30 s. Debris was removed by two centrifugation steps at 10 000 g, 4°C for 10 min. Final protein concentration in the supernatant was determined by the Bradford method (Bio-Rad, Hercules, CA, United States) with BSA as a standard.

## General and Class-Specific Proteolytic Activity Assays

Proteolytic activity of midgut extracts was determined by degradation of azocasein as described previously (Lomate and Bonning, 2016) with optimization (Cantón and Bonning, 2019). In a reaction tube, 50 µg (M1) or 30 µg (M2 and M3) of tissue extract was incubated for 30 min at 37°C with or without the following inhibitors: 10 mM EDTA, 10 µM E-64, 100 µM TLCK, 100 µM TPCK, or 5 mM PMSF. The reaction was made up to a volume of 10 µl with 0.1 M acetate buffer, pH 5.0. After incubation, 200 µl of a 1% azocasein solution in 0.1 M acetate buffer pH 5.0 was added. The tubes were then incubated at 37°C for 2 hr (M2 and M3) or overnight (M1). To stop reactions, 300 µl of chilled 5% trichloroacetic acid was added to the tubes and centrifuged for 10 min at 10 000 g, 4°C. In a 96-well clear bottom plate, 150 µl of 1 M NaOH was added to neutralize 150 µl of supernatant. An iD3 SpectraMax plate reader (Molecular Devices, San Jose, CA, United States) was used to measure absorbance at 450 nm.

For class-specific protease activity, 1 mM solutions of synthetic substrates were obtained by solubilizing powders in DMSO and then slowly adding 0.1 M acetate buffer pH 5.0. The final concentration of DMSO was 10% in all cases, except for pGFLpNA with a concentration of 30%. Similar to the reactions described above, 50 µg (M1) or 30 µg (M2 and M3) of midgut extract was mixed in 20 µl of acetate buffer. The reactions were incubated for 30 min at 37°C, and afterward 100 µl of substrate solution was added and again incubated at 37°C. One hundred µl of 30% acetic acid was used to stop the reaction. Absorbance was measured at 410 nm for 200 µl from each reaction in 96-well plates.

Activity units in the enzyme assays were defined as the change of 0.1 of absorbance per minute per mg of protein. Statistical comparison of equivalent treatments between diets was performed with a *t*-test with a Bonferroni correction for multiple testing after determining normality of the means of biological replicates by a Shapiro–Wilks test.

## Nucleic Acid Degradation Assays

To test nuclease activity on different substrates, we prepared 0.1 mg/ml solutions of calf thymus DNA or baker's yeast RNA in buffer A (25 mM NaCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> in a 20 mM Tris–HCl pH 8.0 nuclease-free buffer). The TranscriptAid T7 High Yield Transcription Kit (ThermoFisher, Waltham, MA, United States) was used to prepare GFP dsRNA from a 502 bp

PCR template amplified from the pGlo plasmid [the primers used (**Supplementary Table S1**) include the T7 promoter sequence]. Dilutions of dsRNA were prepared in buffer A after purification with the PureLink RNA Mini Kit (Ambion, Foster City, CA, United States). All nucleic acid quantifications were performed by Nanodrop at 260 nm.

To monitor nuclease activity, we measured the absorbance due to the release of free nucleotides from nucleic acid substrates (Fraser, 1980). A 10 µl reaction mixture of buffer A and 10 µg of tissue extract was prepared, and then 200 µl of 0.1 mg/ml DNA or RNA in buffer A were added. Reactions were incubated for 30 min at 37°C and then stopped with 300 µl of chilled 10% trichloroacetic acid in nuclease-free water, with 20 µM sodium pyrophosphate. For dsRNA, the final reaction volume was 20 µl, with 2 µg of dsRNA and 10 µg of tissue extract in buffer A. The reactions were stopped with 30 µl of 10% trichloroacetic acid with 20 mM sodium pyrophosphate. Positive control reactions were prepared using 1 µl of 1 U/µl DNase I or 10 µg/µl RNaseA (ThermoFisher), according to substrate. Negative controls included no enzyme or extract. Once reactions were stopped, they were incubated for 1 hr on ice and then centrifuged for 10 min at 10,000 g and 4°C. Absorbance of the supernatant was measured at 260 nm by Nanodrop. Background absorbance of extracts was determined by following the steps above but using buffer A without substrate. We defined one unit of nuclease activity as an increase in absorbance of 0.01 per minute.

## Extraction, Purification, and Sequencing of mRNA

To purify total RNA, the PureLink RNA Mini kit (Ambion) was used. All RNAlater was removed from tissue samples, and then 600 µl of lysis buffer with β-mercaptoethanol was added. A 25G needle and syringe were used to homogenize tissues. Six hundred microliters of 70% ethanol in nuclease-free water was added and thoroughly mixed. Afterward, samples were processed according to the manufacturer's instructions for spin columns. RNA was eluted from columns with 30 µl of nuclease-free water and quantified by Nanodrop at 260 nm. DNase I (ThermoFisher) was added to samples and incubated at 37°C for 10 min. EDTA was added to 20 mM and incubated 10 min at 65°C. Reactions were cleaned by precipitating with 100 µl of isopropanol, 5 µl of sodium acetate 3 M pH 5.2, and 2.5 µl of RNA grade glycogen and incubated overnight at –20°C. Then, samples were centrifuged at 12 000 g, 4°C. The pellet was washed with 70% ethanol in nuclease-free water and centrifuged at 7,600 g for 5 min at 4°C. The final pellet was resuspended in nuclease-free water, with concentration determined by Nanodrop. Only samples of high integrity and purity, and with at least 2 µg of total RNA were selected for RNA-Seq. Quality was reassessed at the Sequencing Facility (**Supplementary Figure S1**). Twenty-four RNA samples representing three biological replicates of each condition were used for the preparation of mRNA libraries that were paired-end sequenced on an Illumina HiSeq 3000 platform with 150 cycles (Genewiz, Inc., South Plainfield, NJ, United States).

## Read Processing, Assembly, and Annotation of the *de novo* Transcriptome

Reads from sequencing were processed to remove bad tiles and adapters with FilterByTile (BBMap Suite (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, 2019) (parameters: -length 36, -q 5, -stringency 1, -e 0.1), respectively. Two complete sets of duplicates of the four tissues analyzed from corn and green bean diet samples were pooled to perform *de novo* transcriptome assembly with Trinity 2.8.3 [(Grabherr et al., 2011), -normalize\_max\_read\_cov 200 and -min\_kmer\_cov 2]. Transcript abundance for all tissues in triplicate was obtained with Salmon v 0.12 (Patro et al., 2017) with the -gcBiasflag, N50, ExN50, and transcripts per kilobase million (TPM) values for the transcripts in the assembly were obtained using scripts in the Trinity package and the Salmon abundance files. Completeness of transcript recovery was evaluated with BUSCO v3.01 (Waterhouse et al., 2017) with reference to the arthropod ortholog dataset. To annotate the transcriptome, we followed the online vignette for Trinotate v3.0.1 (Haas et al., 2013).

## Differential Expression Analysis for Assembled Transcripts and Functional Enrichment

The Bioconductor package TxImport 1.8.0 (Soneson et al., 2015) was used to import transcript abundance data from the quant.sf files of the three biological replicates of each diet using the lengthScaledTPM and dropInfReps = TRUE parameters. The txOut = TRUE argument was used to retain abundance data at the transcript level. DESeq2 v 1.20.0 (Love et al., 2014) was used to perform the statistical analysis of differential expression. Statistical testing incorporated the fold change shrinkage with the apeglm algorithm (Zhu et al., 2018) and a fold change threshold above 1 or below -1. Cutoff for significance was set at an  $s$ -value of 0.005. DESeq2 and ggplot2 were used to prepare PCA plots. The Trinotate annotation information was filtered using the IDs of differentially transcribed sequences for each comparison between corn and green bean tissues. Transcripts with a BLASTp, BLASTx, or PFAM match were counted. A custom R script was written to identify IDs of statistically significant transcripts that were common to lists from the comparisons of salivary glands, M1, M2, and M3. TPM from all samples was recovered for transcripts that had a significant fold change in two or more comparisons.

Enrichment of GO terms was determined for transcripts with significant differences using the TopGO package (Alexa and Rahnenfuhrer, 2018). For this, the Trinotate results for GO terms inferred from BLAST and PFAM hits were used to create a custom GO annotation reference with rows containing the transcript ID and all associated GO terms. TopGO was used to build a GO graph object with annotated lists of significantly different transcripts and the custom annotation file. A node size of five or higher was used as cutoff for terms included in statistical testing. Terms in the “Molecular Function” topology were evaluated with the “weight” algorithm (Alexa et al., 2006). A  $p$ -value of 0.05 established as cutoff for the Fisher exact test.

## RESULTS

### Impacts of Diet Type in Stink Bug Growth

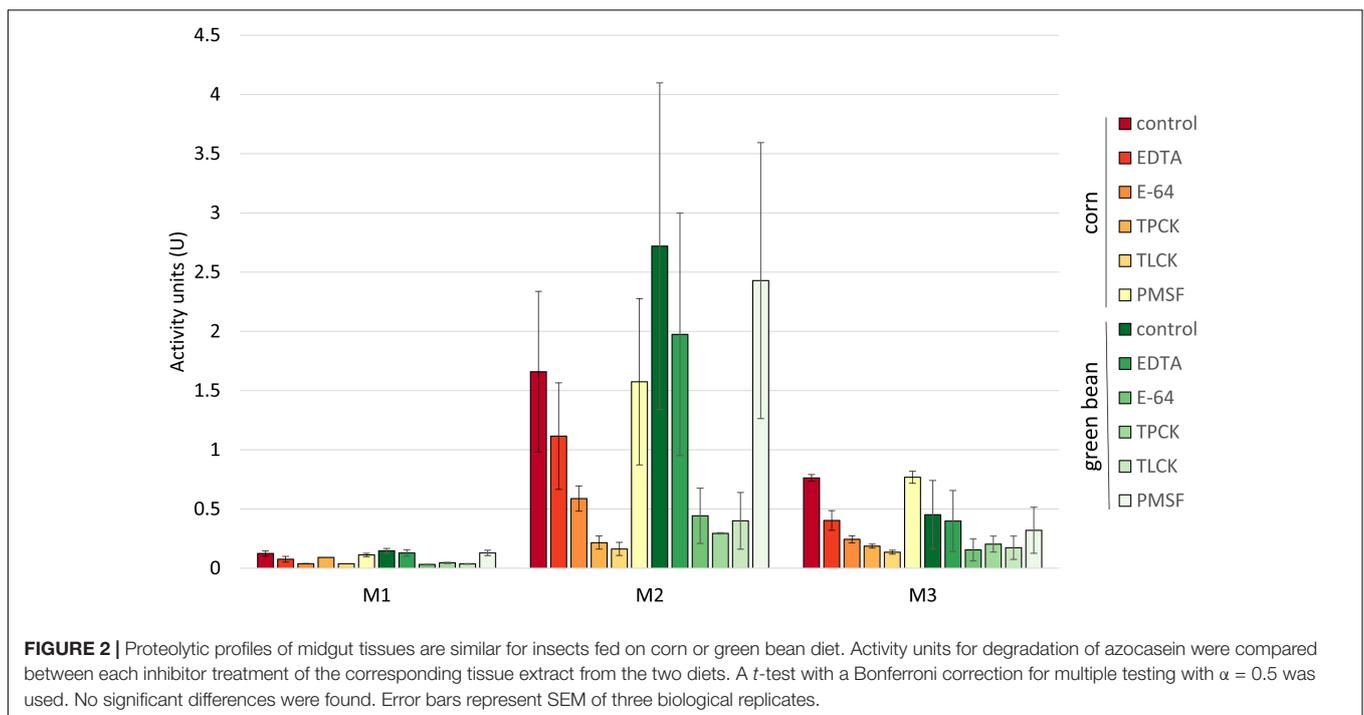
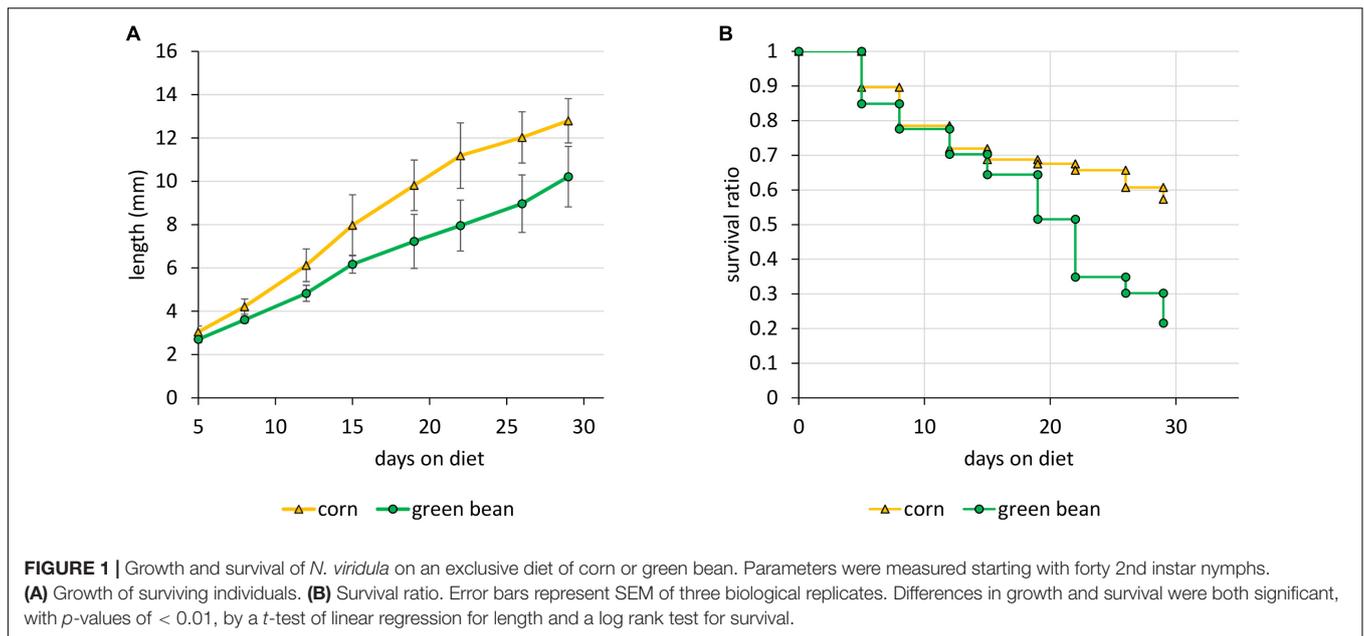
In order to explore the effects of specific diet types on the physiology of *N. viridula*, we first reared nymphs exclusively on either a graminoid (corn, *Z. mays*) or a legume (green bean, *P. vulgaris*). We monitored the growth and survival of the nymphs throughout their development. Under our rearing conditions, nymphs performed better on a diet of corncob with kernels than on green bean pods. Although both types of diet allowed some nymphs to molt into adults, green bean fed nymph growth lagged behind those fed on corn, and survival was higher throughout on corn (**Figure 1**). Differences in growth were significant ( $p$ -value < 0.0005), as were those in survival ( $p$ -value of < 0.009).

Plants produce protease inhibitors as a defense mechanism against herbivory (Zhu-Salzman and Zeng, 2015). We queried the publicly available annotation for corn and green bean to determine the differences in their repertoire of protease inhibitors. We identified 83 and 41 genes corresponding to the “peptidase inhibition activity” molecular function GO annotation for corn and green bean, respectively. However, only eight (9.6%, corn) or seven (17%, green bean) of the protease inhibitors genes were orthologs between the two plants (**Supplementary Figure S2A**). Additionally, we analyzed the PFAM domains for these genes where available. For both plants, the identified domains were mostly comprised by serine protease inhibitors, with around 20% of the genes corresponding to cysteine protease inhibitors (**Supplementary Figure S2B**). Among the serine protease inhibitor domains, in green bean the majority were identified as belonging to Kunitz STI type protein inhibitors, while in corn the largest proportion of PFAM domains belonged to the potato inhibitor family I.

### Comparison of Digestive Enzyme Activity Between Corn and Green Bean Diets

We assessed whether the differences in stink bug growth on the two diets would be reflected by changes in digestive enzyme activity. The digestive enzyme activity of protein extracts from M1, M2, and M3 tissues of *N. viridula* adults grown on green bean diet was compared to profiles previously determined for the same tissues from corn fed *N. viridula* (Cantón and Bonning, 2019). First, we tested the protein extracts from the M1, M2, and M3 midgut regions for proteolytic activity on an azocasein substrate. The overall activity profile and sensitivity to inhibitors was similar for both diets, with no significant differences between corresponding tissue assays (**Figure 2**). On both diets most proteolytic activity was detected in M2 extracts, followed by M3, with the lowest activity in M1. The inhibitors EDTA and PMSF had no effect on proteolytic activity in M2 for either diet.

We then used class-specific substrates to more precisely determine the type of proteases active in each of the extracts (**Figure 3**). As for the azocasein activity assays, no significant differences between the two diets were seen for most comparisons of the corresponding M1, M2, or M3 extracts. The most significant differences were observed for the degradation of the



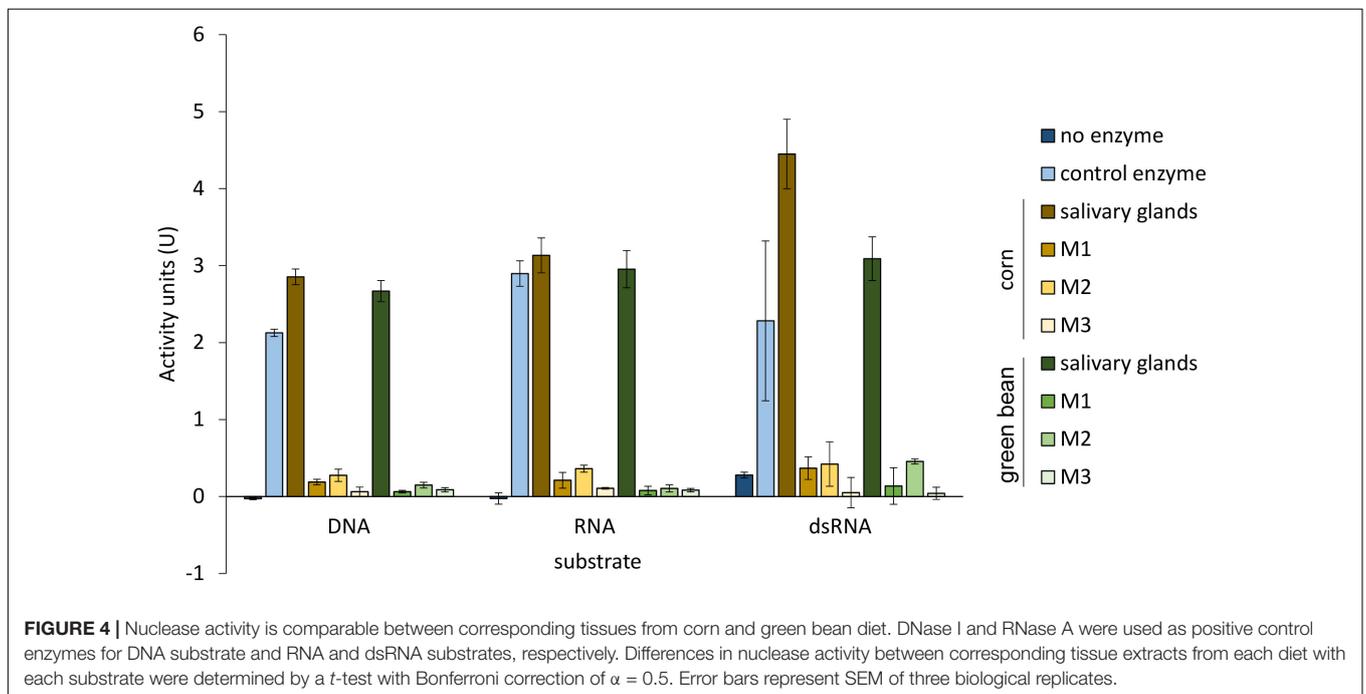
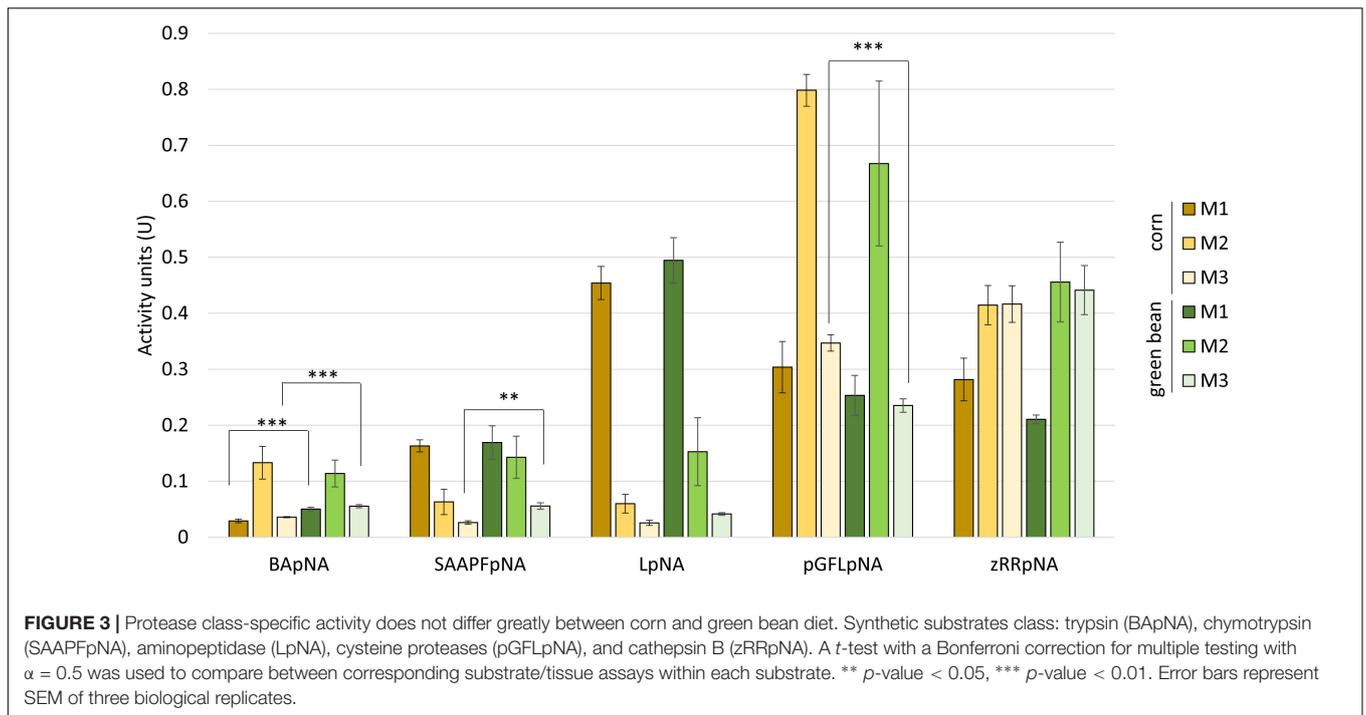
trypsin substrate BApNA, where the green bean diet extracts for M1 and M3 showed higher activity than the corn diet extracts (Figure 3). The lower degradation of the cysteine protease substrate pGFLpNA by extracts from M3 of green bean-fed compared to corn-fed stink bugs was also highly significant. The overall profile of substrate degradation was similar between the two diets for the remaining substrates.

The activity profiles for nucleases in the midgut and the salivary glands did not differ between the two diets for any of the tissue extracts analyzed (Figure 4). As reported previously

(Lomate and Bonning, 2016), salivary glands showed high nuclease and ribonuclease activity compared to the whole midgut. Data for the individually analyzed midgut regions were consistent with this result.

## Impact of Diet on Transcription

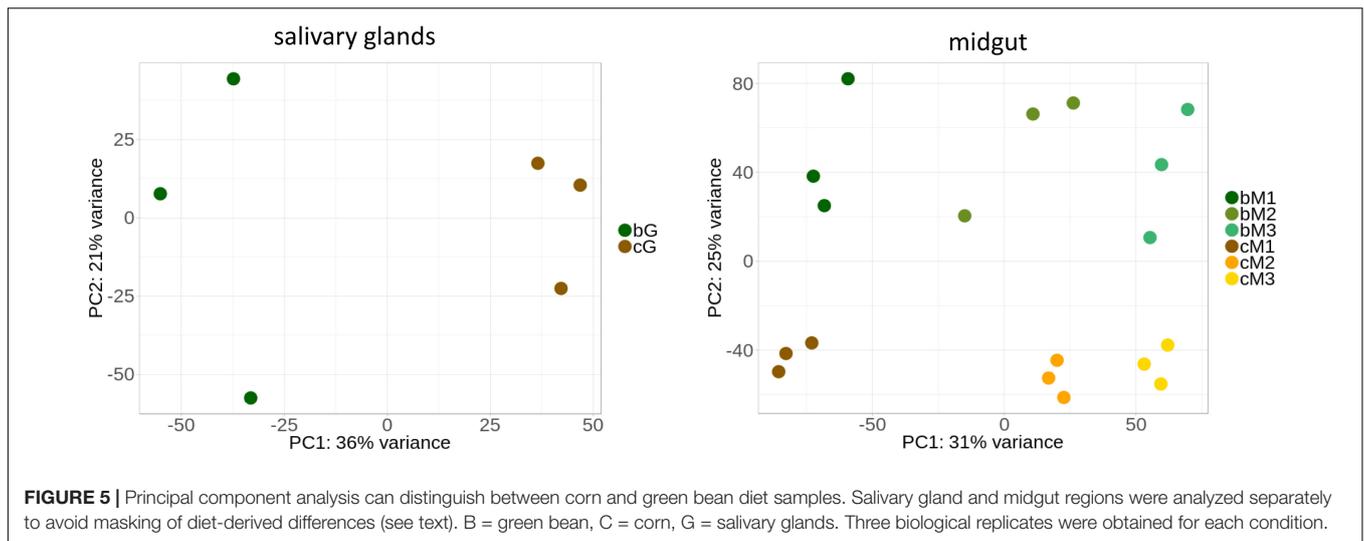
Although few significant differences were detected in enzyme activities between the two diets, we hypothesized that regulation at the transcriptional level could drive diet-related compensatory changes. We therefore compared the transcriptomes between



stink bugs fed on the two diets. We performed RNA-seq on the salivary glands, M1, M2, and M3 of adults grown on green bean diet and compared these results to our previous transcriptomic dataset of adult *N. viridula* fed on corn (Cantón and Bonning, 2019). We tested for differential expression using separate sets of samples for the salivary glands and the midgut regions. **Figure 5** shows the principal component analysis of these two datasets. For all tissues, the corn and green bean samples cluster

separately, indicating that differences in transcript abundance between diet types are distinct. The degree of separation of clusters for each midgut region for the two diets is comparable. The dispersion is notably higher for data from insects fed green bean compared to those fed corn, with the widest dispersion seen for salivary gland samples.

**Table 1** summarizes the number of transcripts from each of the comparisons between matching tissues in corn- and green



bean-fed insects, using the corresponding green bean-fed sample as reference. M2 had the most transcripts with a significant fold change with 244, while M3 had the least. Overall, pairwise comparisons showed that more transcripts were upregulated in corn than downregulated. We attempted GO enrichment analysis for the upregulated and downregulated transcripts, but a large number corresponded to sequences that could not be annotated through Trinotate. This led to unclear results with only a small number of transcripts per category (**Supplementary File S1**). In light of this, we identified transcripts that had a significant fold change in more than one pairwise comparison as potential diet specific transcripts regardless of annotation. By this method, we created sets of transcripts whose induction is more related to corn or green bean feeding (downregulation in corn being relative to upregulation in green bean). These results are summarized in **Table 2**. One transcript was significantly upregulated in all corn tissues, TRINITY\_DN1479\_c0\_g1, which encodes a serine protease. The unannotated transcript TRINITY\_DN1811\_c2\_g1\_i8 was downregulated in corn for three pairwise comparisons. In general, few of these transcripts with significant fold changes could be annotated through homology or PFAM domain identification.

Transcripts with significant fold-change in more than one comparison were separated into low, medium, and high TPM

transcripts (**Figures 6, 7 and Supplementary Figure S3**). TRINITY\_DN3872\_c0\_g1\_i6, TRINITY\_DN58\_c0\_g2\_i6, and TRINITY\_DN21766\_c0\_g2\_i2 all had significant upregulation in three comparisons, but their TPM was below 10 in all tissues (**Figure 6A**). The transcripts with medium TPM that were significantly upregulated in corn in three comparisons correspond to isoforms of TRINITY\_DN1406\_c0\_g1, with higher TPM in M3 tissue. Although these transcripts are unannotated they do contain a predicted secretion signal peptide. The unannotated transcript TRINITY\_DN2154\_c0\_g2\_i3 is significantly upregulated in corn in three comparisons, with very high TPM in salivary glands and very low TPM in the three midgut regions (although higher than the corresponding green bean-fed samples). A similar profile can be seen for the upregulated serine protease significant in four comparisons (**Figure 6B**). Transcript TRINITY\_DN870\_c0\_g1 has homology to Hrp65, a protein involved in RNA maturation. This transcript is significantly downregulated in corn compared to green bean in

**TABLE 1 |** Numbers of significantly upregulated or downregulated transcripts between tissues from insects fed on corn or green bean diets.

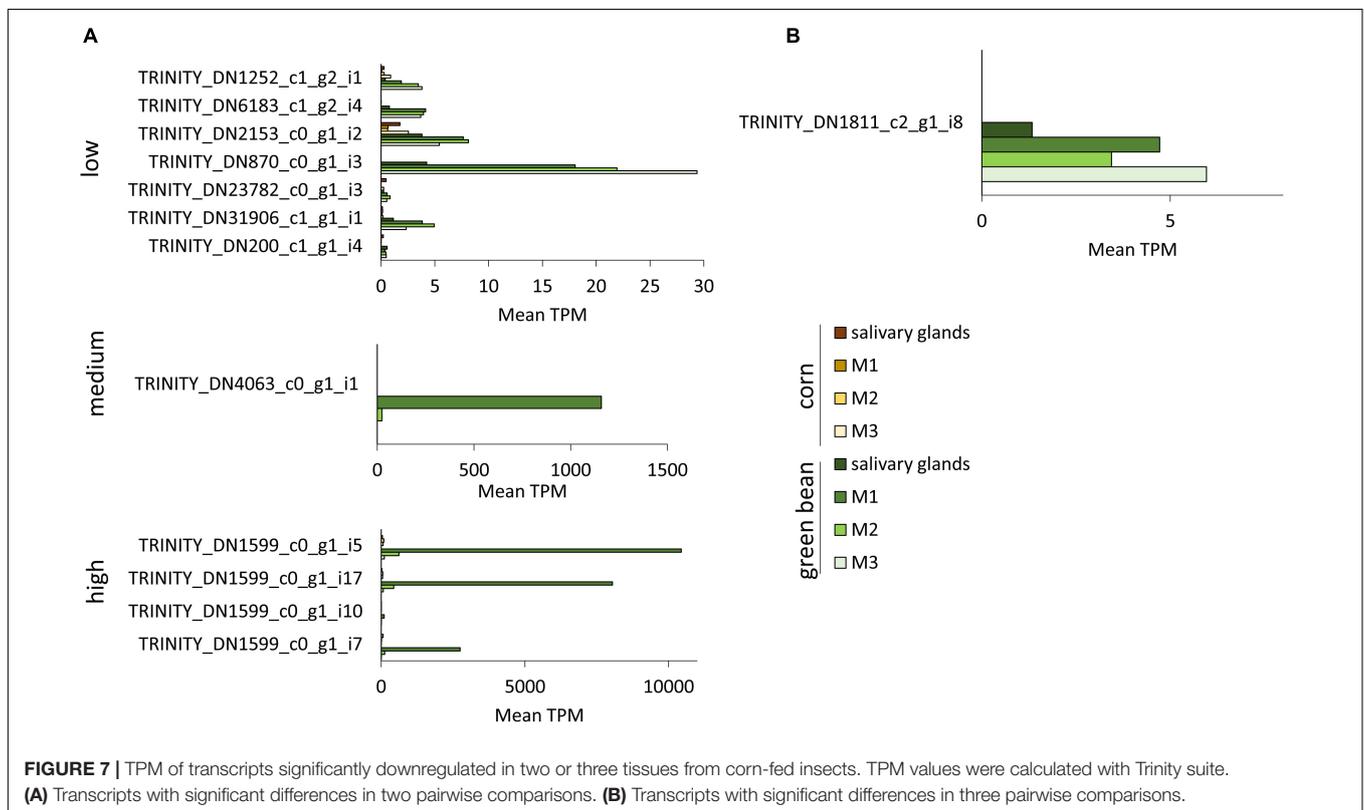
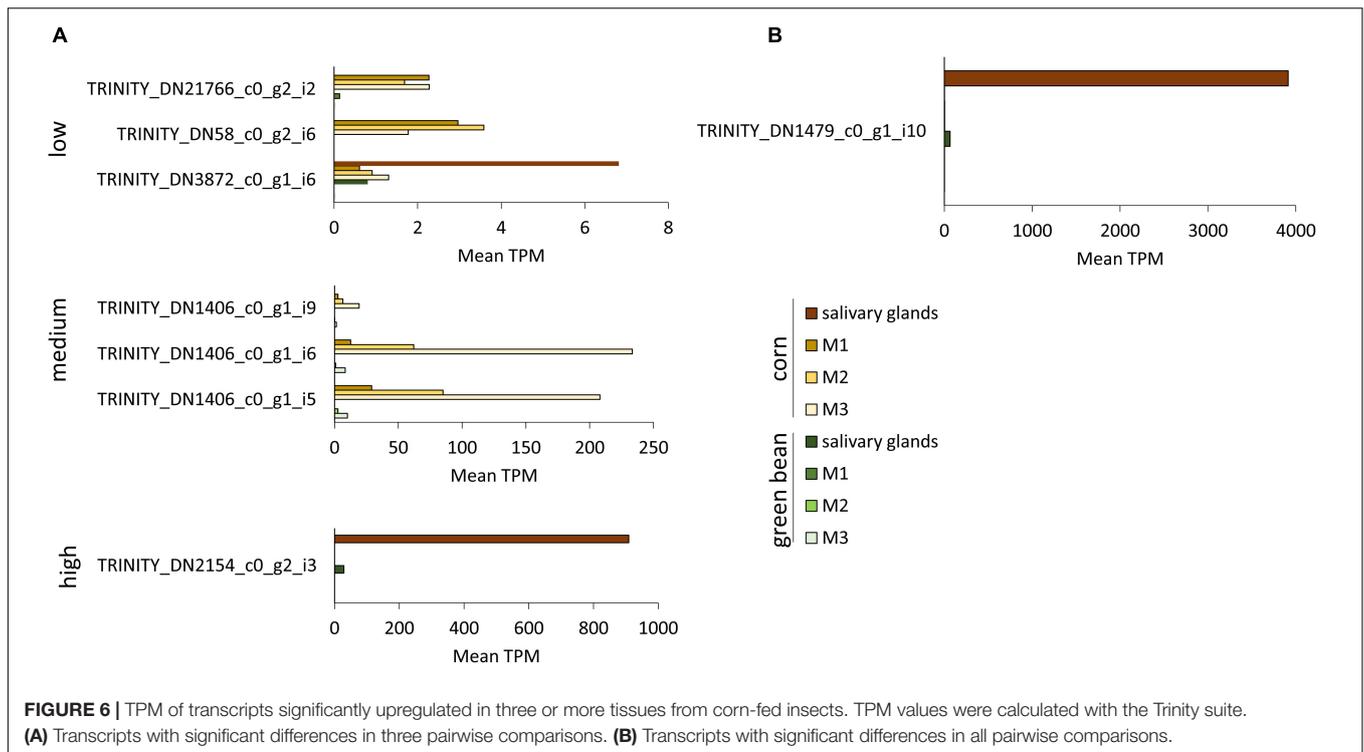
Tissue	Total	Upregulated	Downregulated	Analyzed
Salivary glands	123	110	13	186378
M1	124	100	24	281989
M2	244	156	88	281989
M3	35	32	3	281989

In each comparison, the corresponding tissue in the green bean diet samples was set as reference. The number of analyzed transcripts is those that passed the filtering step of DEseq2.

**TABLE 2 |** Number of unique transcripts with significant fold change between corn and green bean diet in more than one tissue.

Number of comparisons	Number of transcripts	Annotated	Unannotated
<b>Upregulated</b>			
1	309	153	156
2	32	14	18
3	7	1	6
4	1	1	0
<b>Downregulated</b>			
1	101	67	34
2	12	4	8
3	1	0	1
4	0	-	-

In each comparison, the corresponding tissue in the green bean diet samples was set as reference.



two comparisons, although the TPM is higher in all green bean tissues (**Figure 6A**). Of note is the very high TPM in green bean tissues of isoforms of TRINITY\_DN1599\_c0\_g1, particularly for

M1. This transcript is unannotated but has a secretion signal peptide (**Figure 7A**). Finally, TRINITY\_DN1811\_c2\_g1, which was significantly downregulated in corn for three comparisons,

has a low TPM in all green bean tissues, but was not transcribed in any of the corn samples (Figure 7B).

## DISCUSSION

We sought to determine if and how the diet could change the digestive physiology of *N. viridula* toward improved understanding of *N. viridula* polyphagy. No major differences were detected by enzyme assay of digestive tissues from insects maintained on corn or green bean diets in protease inhibitor sensitivity, protease class profiles, or nuclease activity between any of the tissues analyzed. Although no single gene or category of genes was wholly responsible for the clustering of samples by diet type observed in our PCA plots, these results indicate that there is a measurable, diet-specific transcriptomic change. Some specific transcripts show significant differences between the diets, but many could not be assigned an annotation. Further assessment is needed to assign these diet-related, differentially transcribed genes to specific functions.

Differential regulation of insect gene expression has been noted in response to host plant composition and defenses against herbivory. Changes in the amount and type of protease expressed have been observed for several insects (Jongsma et al., 1995; Moon et al., 2004; Brioschi et al., 2007; de Oliveira et al., 2013; Fescemyer et al., 2013; Rivera-Vega et al., 2017), typically with a shift to enzymes less sensitive to the inhibitory compounds present in the diet. Other diet-induced changes involve the active degradation of anti-nutritional factors that might otherwise limit nutrition (Girard et al., 1998). The response to host plant is not limited to digestive enzymes, but involves other metabolic processes including detoxification, stress response, and immunity pathways (Huang et al., 2017). However, diet-dependent changes in transcription in hemipteran species may not be as pronounced or even involve digestive enzymes. Feeding on diets with different phytotoxins only induced changes in a few detoxification enzymes in *Bemisia tabaci* (Halon et al., 2015). For the scale insect *Paratachardina pseudolobata*, feeding on different host plants induced transcriptional changes, yet only a fraction of the genes was related to detoxification, being otherwise enriched for primary metabolism functions (Christodoulides et al., 2017). In the pea aphid *Acyrtosiphon pisum*, changes in transcription were more pronounced for race effect than host plant effect, with only a few hundred genes having expression changes attributable to diet type (Eyes et al., 2016). More recently, in the closely related stink bug *Halyomorpha halys*, no significant changes were detected on different host plants for cytochrome P450 enzymes involved in detoxification (Mittapelly et al., 2019). Our results of a somewhat moderate response in transcriptional changes after feeding on different diets are therefore not entirely unexpected and are similar to those observed for other hemipterans, in contrast to clear responses reported for lepidopteran and coleopteran species. It has been proposed that the feeding mechanism (i.e., sucking vs. chewing) could be an important factor in determining the elicited plant responses to herbivory (Ali and Agrawal, 2012), and consequently in the compensatory mechanisms that the insects have evolved (Will et al., 2013).

One consideration in the evolution of a diet dependent response is that adaptation to host plant composition requires resources that would otherwise be allocated to growth and development. In the Colorado potato beetle, the presence of Cathepsin D inhibitor in plant diet was accompanied by initial effects on growth (Brunelle et al., 2004). Another coleopteran, the cowpea bruchid, also showed reduced early instar growth when fed on diet with the soyacystatin inhibitor, but resumed normal development at the 4th instar (Moon et al., 2004). Despite a preference for legumes, in our experiments the development of *N. viridula* fed on green beans was delayed. The total nymphal survival was slightly below that reported for this species when grown exclusively on green bean pods (Panizzi and Slansky, 1991). We cannot rule out that as our test population was reared under laboratory conditions for several generations it had lost tolerance for this food source. Additionally, regional variation in survivorship has been detected (Panizzi, 1997), and the origin of the initial colony (Louisiana) and supplementation with individuals from Florida may have contributed to this effect. While insects in this study were fed on corn and green beans which were detached from the plants, additional responses may occur when insects are fed on live plants. In this scenario, however, responses to differing diet may be confounded by insect responses to induced plant defenses.

For *N. viridula*, most proteolytic digestion will take place in M2 and M3. In M2, digestion is mediated by cysteine proteases, such as Cathepsin B and L (Cantón and Bonning, 2019). Analysis of inhibitor domains indicated that while both green bean and corn had a similar number of cystatins, green beans have more Kunitz-type inhibitors than corn. These inhibitors generally target serine proteases, although at least one cysteine protease inhibitor in plants has been reported with a Kunitz-type domain (Rustgi et al., 2018). If Kunitz-type inhibitors from green bean inhibit cysteine proteases, this could contribute to the inferior nutritional state of insects maintained on green bean relative to corn. Although no changes in inhibitor sensitivity were detected in adults, nor specific cysteine proteases upregulated in green beans, other compensatory functions that deal with what appears to be a more difficult source of nutrition may be identified from the unannotated transcripts with higher transcription in green beans (such as TRINITY\_DN1811\_c2\_g1\_i8). Although legume seeds can have high levels of trypsin inhibitors (El-Morsi, 2001), the low activity of this enzyme class in the midgut tissues (as observed in this work) is characteristic of the hemipteran lineage (Terra and Ferreira, 1994). These inhibitors are less likely to have an effect on *N. viridula* digestion.

The biological bases for how generalist insects safely ingest a wide variety of plant compounds is highly relevant in the context of plant-insect interactions for pests that feed on multiple crops of economic interest. The work presented here improves our understanding of the physiological responses of the digestive tract of *N. viridula* when feeding on legume and graminoid diets, and provides leads for future investigation of the role of the differentially regulated, unannotated transcripts in adaptation to host plant.

## DATA AVAILABILITY STATEMENT

Raw sequence reads from this project can be accessed at NCBI Short Read Archive SRP193118. Transcript assembly and abundance tables are available at NCBI Gene Expression Omnibus ID GSE130097.

## AUTHOR CONTRIBUTIONS

PC performed the experiments and bioinformatics analysis. PC and BB devised the experiments and wrote the manuscript.

## FUNDING

This work was supported by the National Science Foundation I/UCRC, the Center for Arthropod Management Technologies,

## REFERENCES

- Alexa, A., and Rahnenfuhrer, J. (2018). *topGO: Enrichment Analysis for Gene Ontology. R package version 2.34.0*.
- Alexa, A., Rahnenfuhrer, J., and Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating go graph structure. *Bioinformatics* 22, 1600–1607. doi: 10.1093/bioinformatics/btl140
- Ali, J. G., and Agrawal, A. A. (2012). Specialist versus generalist insect herbivores and plant defense. *Trends Plant Sci.* 17, 293–302. doi: 10.1016/j.tplants.2012.02.006
- Brioschi, D., Nadalini, L. D., Bengtson, M. H., Sogayar, M. C., Moura, D. S., and Silva-Filho, M. C. (2007). General up regulation of *Spodoptera frugiperda* trypsins and chymotrypsins allows its adaptation to soybean proteinase inhibitor. *Insect Biochem. Mol. Biol.* 37, 1283–1290. doi: 10.1016/j.ibmb.2007.07.016
- Brunelle, F., Cloutier, C., and Michaud, D. (2004). Colorado potato beetles compensate for tomato cathepsin D inhibitor expressed in transgenic potato. *Arch. Insect Biochem. Physiol.* 55, 103–113. doi: 10.1002/arch.10135
- Bushnell, B. (2015). *BBMap: a Fast, Accurate, Splice-Aware Aligner*. Available at: <https://sourceforge.net/projects/bbmap/> (accessed August 1, 2018).
- Cantón, P. E., and Bonning, B. C. (2019). Proteases and nucleases across midgut tissues of *Nezara viridula* (Hemiptera: Pentatomidae) display distinct activity profiles that are conserved through life stages. *J. Insect Physiol.* 119:103965. doi: 10.1016/j.jinsphys.2019.103965
- Christodoulides, N., Van Dam, A. R., Peterson, D. A., Frandsen, R. J. N., Mortensen, U. H., Petersen, B., et al. (2017). Gene expression plasticity across hosts of an invasive scale insect species. *PLoS One* 12:e0176956. doi: 10.1371/journal.pone.0176956
- Cohen, A. C. (1995). Extra-oral digestion in predaceous terrestrial arthropoda. *Ann. Rev. Entomol.* 40, 85–103. doi: 10.1146/annurev.en.40.010195.000505
- Coudron, T. A., Brandt, S. L., and Hunter, W. B. (2007). Molecular profiling of proteolytic and lectin transcripts in *Homalodisca vitripennis* (Hemiptera: Auchenorrhyncha: Cicadellidae) feeding on sunflower and cowpea. *Arch. Insect Biochem. Physiol.* 66, 76–88. doi: 10.1002/arch.20200
- de Oliveira, C. F. R., de Paula Souza, T., Parra, J. R. P., Marangoni, S., de Castro Silva-Filho, M., and Macedo, M. L. R. (2013). Insensitive trypsins are differentially transcribed during *Spodoptera frugiperda* adaptation against plant protease inhibitors. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 165, 19–25. doi: 10.1016/j.cbpb.2013.02.008
- El-Morsi, E. M. A. E. F. (2001). Legume seed protease inhibitors: their functions, actions and characteristics. *Egypt. J. Biol.* 3, 164–173.

under Grant Nos. IIP-1338775 and 1821914, and by industry partners.

## ACKNOWLEDGMENTS

The authors would like to thank Dr. Ke Wu, University of Florida, for providing pGlo primer sequences and plasmid used for production of GFP dsRNA, the University of Florida Research Computing facility for computational resources and assistance, and Drs. Jeffrey A. Davis, Louisiana State University and Amanda Hodges, University of Florida for provision of insects.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphys.2019.01553/full#supplementary-material>

- Eyres, I., Jaquiere, J., Sugio, A., Duvaux, L., Gharbi, K., Zhou, J. J., et al. (2016). Differential gene expression according to race and host plant in the pea aphid. *Mol. Ecol.* 25, 4197–4215. doi: 10.1111/mec.13771
- Fescemyer, H. W., Sandoya, G. V., Gill, T. A., Ozkan, S., Marden, J. H., and Luthe, D. S. (2013). Maize toxin degrades peritrophic matrix proteins and stimulates compensatory transcriptome responses in *Fall armyworm* midgut. *Insect Biochem. Mol. Biol.* 43, 280–291. doi: 10.1016/j.ibmb.2012.12.008
- Fraser, M. J. (1980). Purification and properties of *Neurospora crassa* endonuclease, an enzyme which can be converted to a single-strand specific endonuclease. *Methods Enzymol.* 65, 255–263. doi: 10.1016/s0076-6879(80)65035-6
- Ghosh, S. K., Hunter, W. B., Park, A. L., and Gundersen-Rindal, D. E. (2017). Double strand RNA delivery system for plant-sap-feeding insects. *PLoS One* 12:e0171861. doi: 10.1371/journal.pone.0171861
- Girard, C., Le Metayer, M., Bonade-Bottino, M., Pham-Delegue, M. H., and Jouanin, L. (1998). High level of resistance to proteinase inhibitors may be conferred by proteolytic cleavage in beetle larvae. *Insect Biochem. Mol. Biol.* 28, 229–237. doi: 10.1016/s0965-1748(98)00027-7
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi: 10.1038/nprot.2013.084
- Halon, E., Eakteman, G., Moshitzky, P., Elbaz, M., Alon, M., Pavlidi, N., et al. (2015). Only a minority of broad-range detoxification genes respond to a variety of phytotoxins in generalist *Bemisia tabaci* species. *Sci. Rep.* 5:17975. doi: 10.1038/srep17975
- Hirose, E., Panizzi, A., De Souza, J., Cattelan, A., and Aldrich, J. (2009). Bacteria in the gut of southern green stink gug (Heteroptera: Pentatomidae). *Ann. Entomol. Soc. Am.* 99, 91–95. doi: 10.1603/0013-8746(2006)099%5B0091:bitgos%5D2.0.co;2
- Hosokawa, T., Matsuura, Y., Kikuchi, Y., and Fukatsu, T. (2016). Recurrent evolution of gut symbiotic bacteria in pentatomid stinkbugs. *Zool. Lett.* 2:24. doi: 10.1186/s40851-016-0061-4
- Huang, X., Whitman, D. W., Ma, J., McNeill, M. R., and Zhang, Z. (2017). Diet alters performance and transcription patterns in *Oedaleus asiaticus* (Orthoptera: Acrididae) grasshoppers. *PLoS One* 12:e0186397. doi: 10.1371/journal.pone.0186397

- Joga, M. R., Zotti, M. J., Smagghe, G., and Christiaens, O. (2016). RNAi efficiency, systemic properties, and novel delivery methods for pest insect control: what we know so far. *Front. Physiol.* 7:553. doi: 10.3389/fphys.2016.00553
- Jongsma, M. A., Bakker, P. L., Peters, J., Bosch, D., and Stiekema, W. J. (1995). Adaptation of *Spodoptera exigua* larvae to plant proteinase inhibitors by induction of gut proteinase activity insensitive to inhibition. *Proc. Natl. Acad. Sci. U.S.A.* 92, 8041–8045. doi: 10.1073/pnas.92.17.8041
- Kinsella, R. J., Kahari, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., et al. (2011). Ensembl biomarts: a hub for data retrieval across taxonomic space. *Database* 2011:bar030. doi: 10.1093/database/bar030
- Krueger, F. (2019). *TrimGalore: A Wrapper Around Cutadapt, and FastQC to Consistently Apply Adapter, and Quality Trimming to FastQ Files, with Extra Functionality for. (RRBS)Data*. Available at: <https://github.com/FelixKrueger/TrimGalore> (accessed August 1, 2018).
- Lara, P., Ortego, F., Gonzalez-Hidalgo, E., Castanera, P., Carbonero, P., and Diaz, I. (2000). Adaptation of *Spodoptera exigua* (Lepidoptera: Noctuidae) to barley trypsin inhibitor BTI-CMe expressed in transgenic tobacco. *Transgenic Res.* 9, 169–178.
- Li, C., Song, X., Li, G., and Wang, P. (2009). Midgut cysteine protease-inhibiting activity in *Trichoplusia ni* protects the peritrophic membrane from degradation by plant cysteine proteases. *Insect Biochem. Mol. Biol.* 39, 726–734. doi: 10.1016/j.ibmb.2009.08.008
- Liu, S., Lomate, P. R., and 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. doi: 10.1016/j.ibmb.2018.10.003
- Lomate, P. R., and 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. doi: 10.1038/srep27587
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550.
- Mittapelly, P., Bansal, R., and Michel, A. (2019). Differential expression of cytochrome P450 CYP6 genes in the brown marmorated stink bug, *Halyomorpha halys* (Hemiptera: Pentatomidae). *J. Econ. Entomol.* 112, 1403–1410. doi: 10.1093/jee/toz007
- Moon, J., Salzman, R. A., Ahn, J. E., Koiwa, H., and Zhu-Salzman, K. (2004). Transcriptional regulation in cowpea bruchid guts during adaptation to a plant defence protease inhibitor. *Insect Mol. Biol.* 13, 283–291. doi: 10.1111/j.0962-1075.2004.00485.x
- Oppert, B., Morgan, T. D., Hartzer, K., and Kramer, K. J. (2005). Compensatory proteolytic responses to dietary proteinase inhibitors in the red flour beetle, *Tribolium castaneum* (Coleoptera: Tenebrionidae). *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* 140, 53–58. doi: 10.1016/j.cca.2005.01.006
- Panizzi, A. R. (1997). Wild hosts of pentatomids: ecological significance and role in their pest status on crops. *Annu. Rev. Entomol.* 42, 99–122. doi: 10.1146/annurev.ento.42.1.99
- Panizzi, A. R. (2015). Growing problems with stink bugs (Hemiptera: Heteroptera: Pentatomidae): species invasive to the U.S. and potential neotropical invaders. *Am. Entomol.* 61, 223–233. doi: 10.1093/ae/tmv068
- Panizzi, A. R., and Slansky, F. Jr. (1991). Suitability of selected legumes and the effect of nymphal and adult nutrition in the southern green stink bug (Hemiptera: Heteroptera: Pentatomidae). *J. Econ. Entomol.* 84, 103–113. doi: 10.1093/jee/84.1.103
- Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419. doi: 10.1038/nmeth.4197
- Pečan, T., Ye, L., Chang, Y., Mitra, A., Lin, L., Davis, F. M., et al. (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. doi: 10.1105/tpc.12.7.1031
- Rivera-Vega, L. J., Galbraith, D. A., Grozinger, C. M., and Felton, G. W. (2017). Host plant driven transcriptome plasticity in the salivary glands of the cabbage looper (*Trichoplusia ni*). *PLoS One* 12:e0182636. doi: 10.1371/journal.pone.0182636
- Rustgi, S., Boex-Fontvieille, E., Reinbothe, C., von Wettstein, D., and Reinbothe, S. (2018). The complex world of plant protease inhibitors: insights into a kunitz-type cysteine protease inhibitor of *Arabidopsis thaliana*. *Commun. Integr. Biol.* 11:e1368599. doi: 10.1080/19420889.2017.1368599
- Schlüter, U., Benchabane, M., Munger, A., Kiggundu, A., Vorster, J., Goulet, M.-C., et al. (2010). Recombinant protease inhibitors for herbivore pest control: a multitrophic perspective. *J. Exp. Bot.* 61, 4169–4183. doi: 10.1093/jxb/erq166
- Soneson, C., Love, M. I., and Robinson, M. D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* 4:1521. doi: 10.12688/f1000research.7563.2
- Terra, W. R., and Ferreira, C. (1994). Insect digestive enzymes - properties, compartmentalization and function. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 109, 1–62. doi: 10.1016/0305-0491(94)90141-4
- Tillman, P. G. (2010). Composition and abundance of stink bugs (Heteroptera: Pentatomidae) in corn. *Environ. Entomol.* 39, 1765–1774. doi: 10.1603/En09281
- Todd, J. W. (1989). Ecology and behavior of *Nezara viridula*. *Annu. Rev. Entomol.* 34, 273–292. doi: 10.1146/annurev.en.34.010189.001421
- Waterhouse, R. M., Seppey, M., Simao, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., et al. (2017). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* 35, 543–548. doi: 10.1093/molbev/msx319
- Will, T., Furch, A. C., and Zimmermann, M. R. (2013). How phloem-feeding insects face the challenge of phloem-located defenses. *Front. Plant Sci.* 4:336. doi: 10.3389/fpls.2013.00336
- Zhu, A., Ibrahim, J. G., and Love, M. I. (2018). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* 35, 2084–2092. doi: 10.1093/bioinformatics/bty895
- Zhu-Salzman, K., and Zeng, R. (2015). Insect response to plant defensive protease inhibitors. *Annu. Rev. Entomol.* 60, 233–252. doi: 10.1146/annurev-ento-010814-020816

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Cantón and Bonning. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Chemical, Physiological and Molecular Responses of Host Plants to Lepidopteran Egg-Laying

Cinzia Margherita Berte<sup>1</sup>, Luca Pietro Casacci<sup>2,3</sup>, Simona Bonelli<sup>2</sup>, Arianna Zampollo<sup>2</sup> and Francesca Barbero<sup>2\*</sup>

<sup>1</sup> Plant Physiology Unit, Department of Life Sciences and Systems Biology, Turin University, Turin, Italy, <sup>2</sup> Zoolab, Department of Life Sciences and Systems Biology, Turin University, Turin, Italy, <sup>3</sup> Museum and Institute of Zoology, Polish Academy of Sciences, Warsaw, Poland

## OPEN ACCESS

### Edited by:

Akiko Sugio,  
Environnement et Protection des  
Plantes, France

### Reviewed by:

Philippe Reymond,  
Université de Lausanne, Switzerland  
Nina E. Fatouros,  
Wageningen University and Research,  
Netherlands

### \*Correspondence:

Francesca Barbero  
francesca.barbero@unito.it

### Specialty section:

This article was submitted to  
Plant Microbe Interactions,  
a section of the journal  
Frontiers in Plant Science

**Received:** 25 September 2019

**Accepted:** 17 December 2019

**Published:** 30 January 2020

### Citation:

Berte<sup>1</sup> CM, Casacci LP, Bonelli S,  
Zampollo A and Barbero F (2020)  
Chemical, Physiological and Molecular  
Responses of Host Plants to  
Lepidopteran Egg-Laying.  
*Front. Plant Sci.* 10:1768.  
doi: 10.3389/fpls.2019.01768

Plant-lepidopteran interactions involve complex processes encompassing molecules and regulators to counteract defense responses they develop against each other. Lepidoptera identify plants for oviposition and exploit them as larval food sources to complete their development. In turn, plants adopt different strategies to overcome and limit herbivorous damages. The insect egg deposition on leaves can already induce a number of defense responses in several plant species. This minireview deals with the main features involved in the interaction between plants and lepidopteran egg-laying, focusing on responses from both insect and plant side. We discuss different aspects of direct and indirect plant responses triggered by lepidopteran oviposition. In particular, we focus our attention on the mechanisms underlying egg-induced plant defenses that can i) directly damage the eggs such as localized hypersensitive response (HR)-like necrosis, neoplasm formation, production of ovicidal compounds and ii) indirect defenses, such as production of oviposition-induced plant volatiles (OIPVs) used to attract natural enemies (parasitoids) able to kill the eggs or hatching larvae. We provide an overview of chemical, physiological, and molecular egg-mediated plant responses induced by both specialist and generalist lepidopteran species, also dealing with effectors, elicitors, and chemical signals involved in the process. Egg-associated microorganisms are also discussed, although little is known about this third partner participating in plant-lepidopteran interactions.

**Keywords:** butterflies, moths, egg-associated microorganisms, interactions, elicitors

## THE INSECT SIDE: HOW LEPIDOPTERA USE PLANT SIGNALS TO SELECT OVIPOSITION SITES

Lepidoptera mainly depend on plants to complete their development. The choices of gravid females for a suitable oviposition site will severely affect their offspring performances, thus impacting the whole population's survival (García-Barros and Fartmann, 2009). The allocation of eggs on specific larval host plants (LHPs) could be determined by a dynamic hierarchy of biotic and abiotic factors (Carrasco et al., 2015). Not only the plant species and its quality, but also the microclimatic conditions in the surroundings, the intra- or interspecific brood competition, and the occurrence of

symbionts or predators might regulate egg-laying behavior in Lepidoptera (Renwick and Chew, 1994; Ghidotti et al., 2018).

Females searching for an ideal LHP have to combine multifarious sensory information mainly made of chemical, visual, or tactile stimuli (Brévault and Quilici, 2010). Strategies and signals involved are extremely variable and can be summarized as follows: (i) blends of plant volatiles and (ii) visual cues enhance the flight towards the oviposition site and reveal where to land, (iii) substrate compounds are assessed using legs, ovipositor, or proboscis and function as proxies for quality and suitability of the plant site (Reisenman et al., 2010).

Although plants benefit from attracting pollinators, the majority of butterflies and moths should be considered foes as their larvae can be voracious herbivores. Thus, there is a trade-off between resources employed by plants to attract insects for their reproduction and those used to repel enemies. Wounds, bites, or the simple glueing of eggs are signs of current or future herbivore threat and can trigger striking chemical, physiological, and systemic reactions in plants (revised by Hilker and Fatouros, 2015; Schuman and Baldwin, 2016). If constitutive plant compounds usually act as attractants, blends of chemicals released as deterrents to eggs or herbivores may signal a resource already occupied. According to the lepidopteran species, the presence of conspecifics or heterospecifics could enhance (e.g., Anderson and Alborn, 1999) or deter (Sato et al., 1999; De Moraes et al., 2001) oviposition behavior.

Whatever the outcome (i.e., attraction or deterrence), the presence of prior egg deposition is detected by females not exclusively through sight or the perception of oviposition deterring pheromones, such as those released by *Pieris* spp. (Schoonhoven et al., 1990) or *Anthocharis cardamines* (Dempster, 1992), but also by discriminating oviposition-induced plant volatiles (OIPVs; see further section). For instance, by perceiving OIPVs released by *Brassica nigra*, *Pieris brassicae* selects egg-free plants as oviposition sites (Fatouros et al., 2012).

Beyond the ability of adult Lepidoptera to perceive and process plant cues, thus modifying their oviposition behavior, there is a deep gap in the knowledge of possible egg counteradaptations used to overcome the bulk of oviposition-induced plant defenses. More information is available on the diversity of plant responses elicited by egg-laying (Figure 1), which are reviewed hereafter by narrowing the discussion to the most recent literature.

## THE PLANT SIDE: LOCAL AND SYSTEMIC RESPONSES TO LEPIDOPTERAN EGG DEPOSITION

Insect oviposition on a host plant represents a particularly high risk for future herbivore attack and can enable plants to respond even before the actual damage occurs (Hilker and Fatouros, 2016).

### Egg-Induced Direct Plant Responses

Plant defense strategies can directly target insect eggs through desiccation, dropping, and crushing, eventually leading to egg

mortality (Hilker and Fatouros, 2015). Egg deposition of some herbivores can induce reactions in plants that resemble a hypersensitive-like response (HR). This mechanism usually activated by pathogens causes rapid cell death and results in the formation of necrotic plant tissue, leading to the isolation of the pathogens from healthy tissues (Lam et al., 2001). The formation of leaf necrosis in response to insect egg deposition leads to the detachment of eggs from leaves or to their desiccation. This process was observed for the first time in *B. nigra* in which a necrotic zone develops 24 hours after *Pieris rapae* oviposition; in 72 hours, the eggs dry out and often fall off (Shapiro and DeVay, 1987). HR-like necrosis following *P. brassicae* egg-laying was observed also in different plants belonging to the Brassicaceae family (Pashalidou et al., 2015; Griese et al., 2019). Probably a decrease of humidity due to cell apoptosis underneath the oviposition site can cause a release of water out of the eggs eventually leading to their shrinking (Fatouros et al., 2014; Griese et al., 2017).

Recently, Griese and colleagues (2017) demonstrated that the effectiveness of HR-like necrosis in *B. nigra* varies with plant genotype, plant individual, and the type of egg-laying behavior (singly or clustered). Egg bunching could be a strategy to overcome plant defenses by keeping eggs from dehydration. Thus, in *P. brassicae*, egg clusters are more effective to avoid egg-killing compared to the single egg deposition, while the plant genetic background defines the likelihood and severity of HR under natural conditions. The authors hypothesized that the formation of HR-like necrosis evolved as a defensive trait against lepidopteran specialists of brassicaceous plants (Griese et al., 2017). This hypothesis was tested by the same research group who showed that elicitation of HR-like necrosis is specific to the Pierinae subfamily, whose species are adapted to brassicaceous host plants. Non-brassicaceous feeding species were not shown to induce HR-like necrosis (Griese et al., 2019).

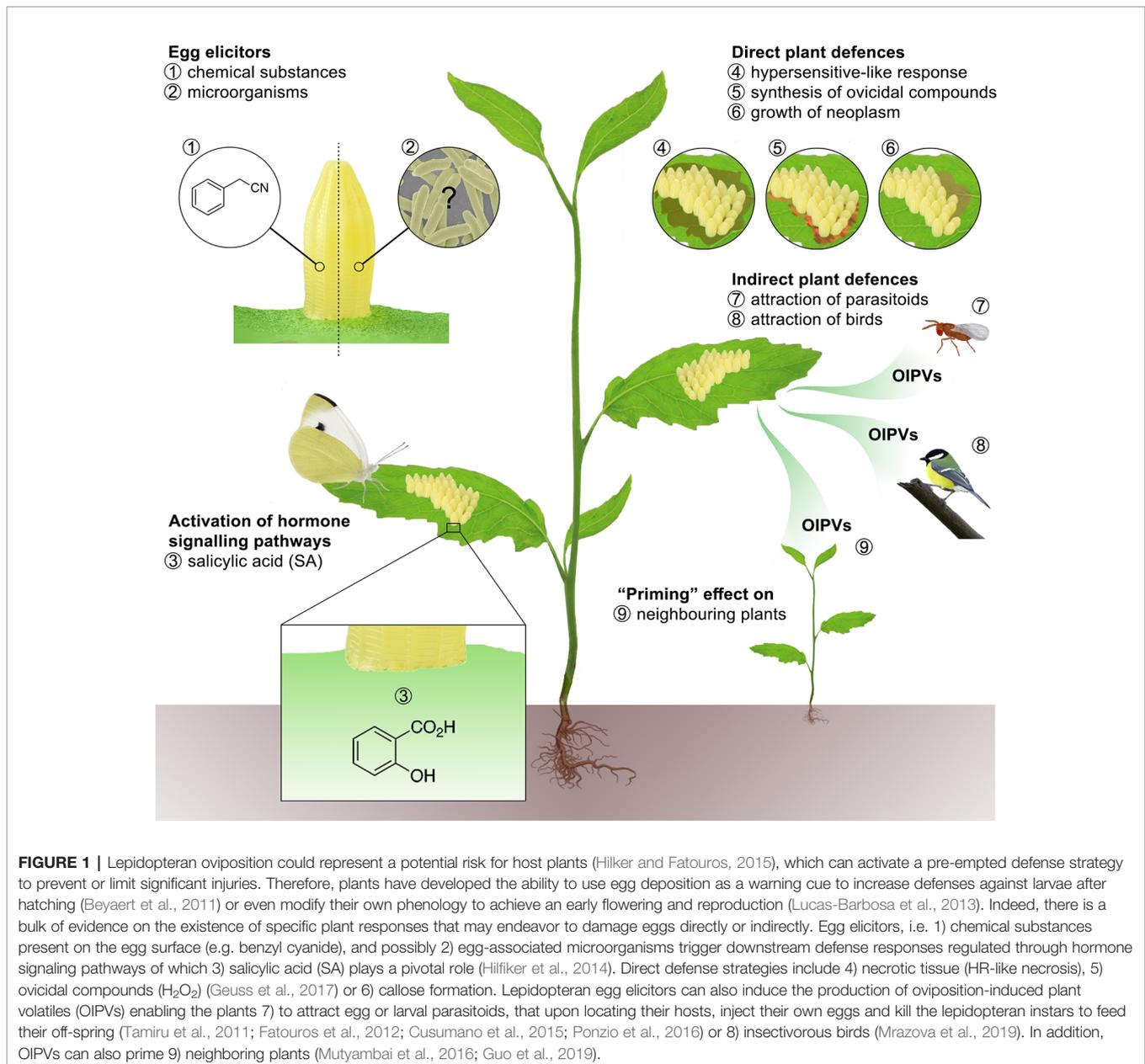
Localized cell death was also observed in *Arabidopsis thaliana* after *P. brassicae* egg-laying (Little et al., 2007; Gouhier-Darimont et al., 2019); however, the response in this plant species is less strong and specific compared to *Brassica* spp., being *A. thaliana* not a foodplant for these butterflies (Harvey et al., 2007).

FA second morphological plant response to insect eggs is neoplasm formation (Petzold-Maxwell et al., 2011; Geuss et al., 2017). This process consists of the growth of a new plant tissue (callus) below insect eggs, which may lead to egg detachment (Petzold-Maxwell et al., 2011). Neoplasm formation in combination with HR-like necrosis was shown to be an egg-killing response in several solanaceous species. Oviposition by a specialist moth *Heliothis subflexa* induced such responses in two ground-cherry species (*Physalis* spp.) (Petzold-Maxwell et al., 2011).

More recently, Geuss et al. (2017) demonstrated that *Solanum dulcamara* responds to *Spodoptera exigua* eggs with the formation of neoplasms and chlorotic tissue. The accumulation of high levels of ovicidal hydrogen peroxide at the oviposition site leads to egg-killing.

### Egg-Induced Indirect Plant Responses

FOviposition can induce changes in the leaf chemistry (Fatouros et al., 2008) or trigger the production of volatile organic compounds



(VOCs) called OIPVs (oviposition-induced plant volatiles) acting as synomones, i.e. indirectly harming eggs or imminent herbivores through the attraction of their natural enemies.

Alterations of the leaf chemistry composition that can be perceived by egg parasitoids after landing have been demonstrated in several crops and wild species following lepidopteran and hemipteran oviposition (Fatouros et al., 2005; Fatouros et al., 2008; Conti et al., 2010). For example, higher quantities of tetratriacontanoic acid and lower quantities of tetracosanoic acid (two important components of the epicuticular wax) were found in *A. thaliana* leaves after *P. brassicae* oviposition. These changes in molecule levels were

shown to be fundamental in retaining *Trichogramma* wasps to egg-infested leaves (Blenn et al., 2012).

Lepidopteran egg-laying does not cause obvious damages in plants (Tamiru et al., 2011; Fatouros et al., 2012), as it occurs in other herbivores, e.g. leafhoppers and beetles (Hilker et al., 2002). Therefore, in contrast to the significant or qualitative changes prompted by herbivory in the plant volatile blends, OIPVs involve primarily quantitative variations (Hilker and Fatouros, 2015), yet effective in attracting parasitoids of lepidopteran eggs and larvae and even insectivorous birds (Mäntylä et al., 2018). This has been demonstrated on egg-laden black mustard (*B. nigra*) and landrace maize varieties (*Zea mays*), which induce

emission of volatiles able to attract *Trichogramma* egg parasitoids (Tamiru et al., 2011; Fatouros et al., 2012; Cusumano et al., 2015; Ponzio et al., 2016).

While the ability of “warning” neighboring plants by means of volatile compounds released against herbivorous attacks is known to occur in various species (Heil and Ton, 2008), the existence of priming by OIPVs has been proven only recently. The study by Mutyambai and colleagues (2016) demonstrated that OIPVs released from the maize landrace ‘Nyamula’ are able to attract the parasitoid wasp (*Cotesia sesamiae*) of the stem borer, *Chilo partellus*. These OIPVs also trigger an indirect defense response in neighboring conspecific plants even when they are not directly exposed to eggs. Among the volatiles released from maize following *C. partellus* egg-laying or exposed to OIPVs, the authors detected a strong emission of (E)-4,8-dimethyl-1,3,7-nonatriene (DMNT), a key homoterpene known as a mediator of herbivore-parasitoid system, with other terpenoids (limonene and myrcene), phenylpropanoids (methyl salicylate) and decanal, compounds often involved in tritrophic interactions.

Egg deposition or treatment with elicitors did not show particular effects in commercial standard maize hybrids, indicating a possible loss of defense traits in plants subjected to artificial selection and breeding (Mutyambai et al., 2016; Tamiru et al., 2017) and, as in the case of HR-like necrosis in *B. nigra* (Griese et al., 2017), highlighting the role of plant genotype in defense mechanisms.

The role of OIPVs in inducing defenses in neighboring plants was not only demonstrated in maize, but also in two clones of *Populus* egg-laden by the moth pest, *Micromelalopha sieversi* (Guo et al., 2019). The authors observed that neighboring plants are able to activate defense responses triggered by the release of volatiles cues (3-carene and  $\beta$ -pinene) from oviposited plants, including the production of VOCs aimed to prevent egg-laying.

Eggs laid by herbivorous insects on a plant leaf indicate that larval feeding will soon occur. Recent studies have demonstrated that, in addition to the enhanced attraction of larval parasitoids (e.g., Pashalidou et al., 2015), “early herbivore alert” responses can also increase plant defense against future herbivory (revised by Hilker and Fatouros, 2015; Hilker and Fatouros, 2016). While a few studies indicate that insect egg deposition may suppress plant anti-herbivore defenses (Bruessow et al., 2010; Peñafior et al., 2011), additional studies comparing plant responses to egg-laying by several generalist and specialist insects are necessary to elucidate the mechanisms involved in this process.

## Defense Pathways and Gene Expression

It is well known that elicitors (see below), associated to egg deposition, trigger electrical signals and change  $Ca^{2+}$  homeostasis. This is subsequently followed by downstream defense responses regulated through hormone signaling pathways, whose jasmonic acid (JA) and salicylic acid (SA) are the major players involved (Reymond, 2013). Both the individual

hormones and their crosstalk play an essential role in fine-tuning defense responses to specific herbivores (Proietti et al., 2018).

The induction of the JA pathway by herbivore-associated elicitors has been extensively reported; however, there is no clear evidence that the JA-pathway is induced by insect egg deposition.

The response to oviposition by *P. brassicae* on *Arabidopsis* or *Brassica* spp., where eggs are laid on the leaf surface without any damage, appears mainly controlled by SA signaling pathway. In *Arabidopsis* plants, SA accumulated at high levels underneath *Pieris* eggs and several SA-responsive genes were upregulated by egg-laying also in systemic leaves (Hilfiker et al., 2014; Bonnet et al., 2017). These responses were absent in some *Arabidopsis* mutants lacking the SA-signaling pathway (Gouhier-Darimont et al., 2013). This defense mechanism is similar to the response triggered by pathogens (Gouhier-Darimont et al., 2013).

It is clear that lepidopteran oviposition induces different morphological, physiological, and chemical responses in plants that are strongly correlated to the variation in gene expression levels. The first study of *P. brassicae* egg-induced transcriptional changes performed with *Arabidopsis* whole-genome DNA microarrays showed the up-regulation of several defense-related genes, including some regulating cell death and innate immunity, and others involved in stress responses and in secondary metabolite biosynthesis (Little et al., 2007). More recently, a transcriptome comparison of *Arabidopsis* feeding-damaged leaves, with and without prior oviposition, revealed the up-regulation of *PR5*, a gene involved in SA-signaling, an increase in SA levels and flavanol accumulation in egg-laden but not yet damaged plants (Lortzing et al., 2019). Also Geuss et al. (2017) showed that feeding larvae of *S. exigua* induced an increase in *S. dulcamara* resistance, by changing its transcriptional and metabolic responses at both the local and systemic level. In particular, genes involved in phenylpropanoid metabolism were upregulated in previously oviposited plants, suggesting a crucial role of these molecules in oviposition-primed plant resistance.

Moreover, a study conducted on maize landrace Braz1006 demonstrated that both *C. partellus* egg deposition and a treatment with an elicitor that mimics herbivory can induce the up-regulation of the gene coding for the terpene synthase TPS23, which catalyzes the final step in the biosynthetic pathway of (E)-caryophyllene, an important signaling molecule involved in plant-herbivore interactions (Tamiru et al., 2017).

## Egg-Derived Elicitors

During oviposition, insects produce a vast range of substances from the ovary and accessory glands, which can act as **elicitors** of the above-mentioned plant defenses.

These secretions can provide eggs with protection against biotic and abiotic threats, facilitate their deposition (lubrication) or their substrate attachments. Beyond being found on the egg surface or at the plant-egg interface, bioactive compounds can also be found within the egg. Yet, the role of the inner compounds in eliciting plant responses seems unlikely due to the presence of physical barriers (e.g. eggshell, adhesive glue)

hindering the access to plant cell targets (Hilker and Fatouros, 2015). Bruessow et al. (2010) suggested that elicitors should be found within the eggs, in the embryo, as no reaction was observed when empty *P. brassicae* eggshells were applied at the leaf surface. However, the lack of any response could be due to external egg elicitor inactivation (instead of their absence) that occurs in the period between deposition and hatching event (Fatouros et al., 2015).

Experiments conducted with crushed egg extracts (EE) mimicked the response observed upon egg-laying in *A. thaliana* (Little et al., 2007). Using an *Arabidopsis* transgenic line containing the promoter of the egg-induced gene *PR1* coupled to the  $\beta$ -glucuronidase (GUS) reporter gene, Little et al. (2007) demonstrated that the application of soluble *P. brassicae* EE activates GUS and triggers plant responses. Similar results were obtained when EE from distantly related insects, either generalists or specialists, were applied to *A. thaliana* transgenic plants.

Although a very few compounds have been isolated, benzyl cyanide was identified as a molecule responsible for surface chemical changes induced by *P. brassicae* oviposition on *Brassica oleracea* var. *gemmifera*. The application of this male-derived anti-aphrodisiac mimicked the egg-induced arrestment of *Trichogramma brassicae* (egg parasitoids) in *B. oleracea* and *Arabidopsis* leaves (Fatouros et al., 2005; Blenn et al., 2012). Moreover, *P. rapae* females receive methyl salicylate and indole as anti-aphrodisiac compounds during mating. When applied onto the leaf, indole induced changes in the foliar chemistry that arrested *T. brassicae* wasps (Fatouros et al., 2009).

Besides the extensive research on plant-insect interactions and although it is generally assumed that plants detect elicitors through cell-surface receptors, to date, no such protein has been isolated and described. Following different attempts, in 2019, Gouhier-Darimont and co-workers identified an important component of *A. thaliana* perception system for insect eggs, LecRK-I.8, a L-type lectin receptor kinase. This protein seems to play a key role in early signal transduction steps by controlling several responses to *P. brassicae* egg-laying. The authors demonstrated that a lipidic fraction from *P. brassicae* eggs triggers localized cell death and that this response is significantly attenuated in *lecrk-I.8* mutant plants, suggesting that LecRK-I.8 is involved in the sensing of an egg-derived lipidic compound (Gouhier-Darimont et al., 2019).

## A THIRD PLAYER: EGG-ASSOCIATED MICROORGANISMS

Symbiotic bacteria play a pivotal role in the development and survival of their insect hosts, providing a full array of molecules for digestion, detoxification, and defense against pathogens (Douglas, 2015). There is still a scant knowledge on Lepidoptera-associated microbiomes, because the majority of studies is (i) merely descriptive, (ii) focused on single bacterial taxon, (iii) a few butterfly/moth species have been extensively surveyed, or (iv) only rarely endosymbionts have been

compared across different developmental instars (Di Salvo et al., 2019; Gao et al., 2019; Szenteczki et al., 2019). Nevertheless, an increasing number of experiments provide evidence for a crucial function of microbes in basic physiological processes of Lepidoptera (Paniagua Voirol et al., 2018), e.g. through the modulation of salivary elicitor biosynthesis (Wang et al., 2018).

Since data gathered until now suggest a remarkable diversity of (gut) microbiomes across diets and stages, it is questioned whether Lepidoptera harbor resident beneficial microbes or more likely acquire from food and/or environment a plastic microbial community, which favors them under changing conditions (Hammer et al., 2017). If confirmed, this scenario implies that eggs might not serve as the means for achieving the vertical transmission of core gut microbiomes, but only of other microbial symbionts. The inherited microbes could also be present on the egg surface and transferred by eggshell ingestion to newly hatched larvae (Duploux and Hornett, 2018), but their characterization and function are completely lacking.

The occurrence of egg-associated bacteria has been reported for a few species including *Manduca sexta*, *Rothschildia lebeau*, *Spodoptera littoralis*, and *Lymantria dispar* (Paniagua Voirol et al., 2018), but there are no insights about potential roles of egg-associated bacteria in eliciting plant responses.

## CONCLUSION

Egg-laying patterns are the outcomes of complex evolutionary dynamics shaped by physical, physiological, and ecological characteristics of the host plants. Although plant responses to both eggs and herbivores have been extensively explored (Hilker and Fatouros, 2015; Schuman and Baldwin, 2016), only a few studies have dealt with herbivore counteradaptations (Karban and Agrawal, 2002) and even less with egg defensive/offensive traits (Bruessow et al., 2010; Peñaflor et al., 2011). However, an increasing number of insights suggests that (i) the female ability to identify plants with inadequate plant defenses could be an evolutionarily advantageous strategy and (ii) the biochemical apparatus of plants could be subverted by egg compounds to inhibit or lower the LHP defenses against the incoming larval instars.

Unfortunately, the advance of this research is constrained by the lack of upstream knowledge about basic mechanisms fostering the specificity of plant responses. The latter are likely based on still undiscovered egg-associated compounds (elicitors) and their plant receptors, which therefore should be among the first issues to be tackled.

## AUTHOR CONTRIBUTIONS

CB and FB led the writing of the manuscript to which all authors contributed critically and gave final approval for publication.

## REFERENCES

- Anderson, P., and Alborn, H. (1999). Effects on oviposition behaviour and larval development of *Spodoptera littoralis* by herbivore-induced changes in cotton plants. *Entom. Exp. Appl.* 92 (1), 45–51. doi: 10.1046/j.1570-7458.1999.00523.x
- Beyaert, I., Köpke, D., Stiller, J., Hammerbacher, A., Yoneya, K., Schmidt, A., et al. (2011). Can insect egg deposition 'warn' a plant of future feeding damage by herbivorous larvae? *Proc. R. Soc. Lond. [Biol]* 279 (1726), 101–108. doi: 10.1098/rspb.2011.0468
- Blenn, B., Bandoly, M., Kuffner, A., Otte, T., Geiselhardt, S., Fatouros, N. E., et al. (2012). Insect egg deposition induces indirect defense and epicuticular wax changes in *Arabidopsis thaliana*. *J. Chem. Ecol.* 38 (7), 882–892. doi: 10.1007/s10886-012-0132-8
- Bonnet, C., Lassueur, S., Ponzio, C., Gols, R., Dicke, M., and Reymond, P. (2017). Combined biotic stresses trigger similar transcriptomic responses but contrasting resistance against a chewing herbivore in *Brassica nigra*. *BMC Plant Biol.* 17 (1), 127. doi: 10.1186/s12870-017-1074-7
- Brévault, T., and Quilici, S. (2010). Interaction between visual and olfactory cues during host finding in the tomato fruit fly *Neoceratitis cyanescens*. *J. Chem. Ecol.* 36 (3), 249–259. doi: 10.1007/s10886-010-9766-6
- Bruessow, F., Gouhier-Darimont, C., Buchala, A., Metraux, J. P., and Reymond, P. (2010). Insect eggs suppress plant defence against chewing herbivores. *Plant J.* 62 (5), 876–885. doi: 10.1111/j.1365-313X.2010.04200.x
- Carrasco, D., Larsson, M. C., and Anderson, P. (2015). Insect host plant selection in complex environments. *Curr. Opin. Insect Sci.* 8, 1–7. doi: 10.1016/j.cois.2015.01.014
- Conti, E., Salerno, G., Leombruni, B., Frati, F., and Bin, F. (2010). Short-range allelochemicals from a plant-herbivore association: a singular case of oviposition-induced synomone for an egg parasitoid. *J. Exp. Biol.* 213 (22), 3911–3919. doi: 10.1242/jeb.045922
- Cusumano, A., Weldegergis, B. T., Colazza, S., Dicke, M., and Fatouros, N. E. (2015). Attraction of egg-killing parasitoids toward induced plant volatiles in a multi-herbivore context. *Oecologia* 179 (1), 163–174. doi: 10.1007/s00442-015-3325-3
- De Moraes, C. M., Mescher, M. C., and Tumlinson, J. H. (2001). Caterpillar-induced nocturnal plant volatiles repel conspecific females. *Nature* 410 (6828), 577. doi: 10.1038/35069058
- Dempster, J. P. (1992). Evidence of an oviposition-detering pheromone in the orange-tip butterfly, *Anthocharis cardamines* (L.). *Ecol. Entomol.* 17 (1), 83–85. doi: 10.1111/j.1365-2311.1992.tb01043.x
- Di Salvo, M., Calcagnile, M., Talà, A., Tredici, S. M., Maffei, M. E., Schönrogge, K., et al. (2019). The microbiome of the *Maculinea-Myrmica* host-parasite interaction. *Sci. Rep.* 9 (1), 8048. doi: s41598-019-44514-7
- Douglas, A. E. (2015). Multiorganismal insects: diversity and function of resident microorganisms. *Annu. Rev. Entomol.* 60, 17–34. doi: 10.1146/annurev-ento-010814-020822
- Duploup, A., and Hornett, E. A. (2018). Uncovering the hidden players in Lepidoptera biology: the heritable microbial endosymbionts. *PeerJ.* 6, e4629. doi: 10.7717/peerj.4629
- Fatouros, N. E., Bukovinszkine'Kiss, G., Kalkers, L., Gamborena, R., Dicke, M., and Hilker, M. (2005). Oviposition-induced plant cues: do they arrest *Trichogramma* wasps during host location? *Entomol. Exp. Appl.* 115 (1), 207–215. doi: 10.1111/j.1570-7458.2005.00245.x
- Fatouros, N. E., Broekgaarden, C., Bukovinszkine'Kiss, G., van Loon, J. J., Mumm, R., Huigens, M. E., et al. (2008). Male-derived butterfly anti-aphrodisiac mediates induced indirect plant defense. *Proc. Natl. Acad. Sci. U. S. A.* 105 (29), 10033–10038. doi: 10.1073/pnas.0707809105
- Fatouros, N. E., Pashalidou, F. G., Aponte Cordero, W. V., van Loon, J. J., Mumm, R., Dicke, M., et al. (2009). Anti-aphrodisiac compounds of male butterflies increase the risk of egg parasitoid attack by inducing plant synomone production. *J. Chem. Ecol.* 35 (11), 1373–1381. doi: 10.1007/s10886-009-9714-5
- Fatouros, N. E., Lucas-Barbosa, D., Weldegergis, B. T., Pashalidou, F. G., van Loon, J. J., Dicke, M., et al. (2012). Plant volatiles induced by herbivore egg deposition affect insects of different trophic levels. *PLoS One* 7 (8), e43607. doi: 10.1371/journal.pone.0043607
- Fatouros, N. E., Pineda, A., Huigens, M. E., Broekgaarden, C., Shimwela, M. M., Candia, Figueroa, et al. (2014). Synergistic effects of direct and indirect defences on herbivore egg survival in a wild crucifer. *Proc. R. Soc. B Biol. Sci.* 281 (1789), 20141254. doi: 10.1098/rspb.2014.1254
- Fatouros, N. E., Paniagua Voirol, L. R., Drizou, F., Thi Doan, Q., Pineda, A., Frago, E., et al. (2015). Role of large cabbage white male-derived compounds in elicitation of direct and indirect egg-killing defenses in the black mustard plants. *Front. Plant Sci.* 6, 794. doi: 10.3389/fpls.2015.00794
- Gao, X., Li, W., Luo, J., Zhang, L., Ji, J., Zhu, X., et al. (2019). Biodiversity of the microbiota in *Spodoptera exigua* (Lepidoptera: Noctuidae). *J. Appl. Microbiol.* 126 (4), 1199–1208. doi: 10.1111/jam.14190
- García-Barros, E., and Fartmann, T. (2009). "Butterfly oviposition: sites, behaviour and modes," in *Ecology of Butterflies in Europe* (Cambridge: Cambridge University Press), 29–42.
- Geuss, D., Stelzer, S., Lortzing, T., and Steppuhn, A. (2017). *Solanum dulcamara*'s response to eggs of an insect herbivore comprises ovicidal hydrogen peroxide production. *Plant Cell Env.* 40 (11), 2663–2677. doi: 10.1111/pce.13015
- Ghidotti, S., Cerrato, C., Casacci, L. P., Barbero, F., Paveto, M., Pesce, M., et al. (2018). Scale-dependent resource use in the *Euphydryas aurinia* complex. *J. Insect Conserv.* 22 (3–4), 593–605. doi: 10.1007/s10841-018-0088-2
- Gouhier-Darimont, C., Schmiesing, A., Bonnet, C., Lassueur, S., and Reymond, P. (2013). Signalling of *Arabidopsis thaliana* response to *Pieris brassicae* eggs shares similarities with PAMP-triggered immunity. *J. Exp. Bot.* 64 (2), 665–674. doi: 10.1093/jxb/ers362
- Gouhier-Darimont, C., Stahl, E., Glauser, G., and Reymond, P. (2019). The *Arabidopsis* Lectin Receptor Kinase LecRK-I. 8 Is Involved in Insect Egg Perception. *Front. Plant Sci.* 10, 623. doi: 10.3389/fpls.2019.00623
- Griese, E., Dicke, M., Hilker, M., and Fatouros, N. E. (2017). Plant response to butterfly eggs: inducibility, severity and success of egg-killing leaf necrosis depends on plant genotype and egg clustering. *Sci. Rep.* 7 (1), 7316. doi: 10.1038/s41598-017-06704-z
- Griese, E., Pineda, A., Pashalidou, F. G., Iradi, E. P., Hiker, M., Dicke, M., et al. (2019). Plant responses to butterfly oviposition partly explain preference-performance relationships on different brassicaceous species. *BioRxiv* preprint July, 2019. doi: 10.1101/706044
- Guo, L., Liu, F., Zhang, S., Kong, X., and Zhang, Z. (2019). Egg deposition of *Micromelalopha sieversi* (Staudinger) on clones of *Populus* from Section *Aigeiros* induces resistance in neighboring plants. *Forests* 10 (2), 110. doi: 10.3390/f10020110
- Hammer, T. J., Janzen, D. H., Hallwachs, W., Jaffe, S. P., and Fierer, N. (2017). Caterpillars lack a resident gut microbiome. *Proc. Natl. Acad. Sci. U. S. A.* 114 (36), 9641–9646. doi: 10.1073/pnas.1707186114
- Harvey, J. A., Witjes, L. M., Benkirane, M., Duyts, H., and Wagenaar, R. (2007). Nutritional suitability and ecological relevance of *Arabidopsis thaliana* and *Brassica oleracea* as foodplants for the cabbage butterfly, *Pieris rapae*. *Plant Ecol.* 189 (1), 117–126. doi: 10.1007/s11258-006-9204-6
- Heil, M., and Ton, J. (2008). Long-distance signalling in plant defence. *Trends Plant Sci.* 13 (6), 264–272. doi: 10.1016/j.tplants.2008.03.005
- Hilfiker, O., Groux, R., Bruessow, F., Kiefer, K., Zeier, J., and Reymond, P. (2014). Insect eggs induce a systemic acquired resistance in *Arabidopsis*. *Plant J.* 80 (6), 1085–1094. doi: 10.1111/tpj.12707
- Hilker, M., and Fatouros, N. E. (2015). Plant responses to insect egg deposition. *Annu. Rev. Entomol.* 60, 493–515. doi: 10.1146/annurev-ento-010814-020620
- Hilker, M., and Fatouros, N. E. (2016). Resisting the onset of herbivore attack: plants perceive and respond to insect eggs. *Curr. Opin. Plant Biol.* 32, 9–16. doi: 10.1016/j.pbi.2016.05.003
- Hilker, M., Rohfritsch, O., and Meiners, T. (2002). "The plant's response towards insect egg deposition," in *Chemoeology of insect eggs and egg deposition*. Eds. Hilker, M. and T. Meiners (Oxford: Blackwell), 205–233.
- Karban, R., and Agrawal, A. A. (2002). Herbivore offense. *Annu. Rev. Ecol. Evol. Syst.* 33 (1), 641–664. doi: 10.1146/annurev.ecolsys.33.010802.150443
- Lam, E., Kato, N., and Lawton, M. (2001). Programmed cell death, mitochondria and the plant hypersensitive response. *Nature* 411 (6839), 848. doi: 10.1038/35081184
- Little, D., Gouhier-Darimont, C., Bruessow, F., and Reymond, P. (2007). Oviposition by pierid butterflies triggers defense responses in *Arabidopsis*. *Plant Physiol.* 143 (2), 784–800. doi: 10.1104/pp.106.090837

- Lortzing, V., Oberländer, J., Lortzing, T., Tohge, T., Steppuhn, A., Kunze, R., et al. (2019). Insect egg deposition renders plant defence against hatching larvae more effective in a salicylic acid-dependent manner. *Plant Cell Environ.* 42 (3), 1019–1032. doi: 10.1111/pce.13447
- Lucas-Barbosa, D., van Loon, J. J., Gols, R., van Beek, T. A., and Dicke, M. (2013). Reproductive escape: annual plant responds to butterfly eggs by accelerating seed production. *Funct. Ecol.* 27 (1), 245–254. doi: 10.1111/1365-2435.12004
- Mäntylä, E., Kleier, S., Lindstedt, C., Kipper, S., and Hilker, M. (2018). Insectivorous birds are attracted by plant traits induced by insect egg deposition. *J. Chem. Ecol.* 44, 1127–1138. doi: 10.1007/s10886-018-1034-1
- Mrazova, A., Sam, K., and Amo, L. (2019). What do we know about birds' use of plant volatile cues in tritrophic interactions? *Curr. Opin. Insect Sci.* 32, 131–136. doi: 10.1016/j.cois.2019.02.004
- Mutyambai, M. D., Bruce, T. J. A., van den Berg, J., Midega, C. A. O., Pickett, J. A., and Khan, Z. R. (2016). An indirect defence trait mediated through egg-induced maize volatiles from neighbouring plants. *PLoS One* 11 (7), e0158744. doi: 10.1371/journal.pone.0158744
- Paniagua Voirol, L. R., Frago, E., Kaltenpoth, M., Hilker, M., and Fatouros, N. E. (2018). Bacterial symbionts in lepidoptera: their diversity, transmission, and impact on the host. *Front. Microbiol.* 9, 556. doi: 10.3389/fmicb.2018.00556
- Pashalidou, F. G., Fatouros, N. E., Van Loon, J. J., Dicke, M., and Gols, R. (2015). Plant-mediated effects of butterfly egg deposition on subsequent caterpillar and pupal development, across different species of wild Brassicaceae. *Ecol. Entomol.* 40 (4), 444–450. doi: 10.1111/een.12208
- Peñaflor, M. F. G. V., Erb, M., Robert, C. A. M., Miranda, L. A., Werneburg, A. G., Dossi, F. C. A., et al. (2011). Oviposition by a moth suppresses constitutive and herbivore-induced plant volatiles in maize. *Planta* 234 (1), 207–215. doi: 10.1007/s00425-011-1409-9
- Petzold-Maxwell, J., Wong, S., Arellano, C., and Gould, F. (2011). Host plant direct defence against eggs of its specialist herbivore, *Heliothis subflexa*. *Ecol. Entomol.* 36, 700–708. doi: 10.1111/j.1365-2311.2011.01315.x
- Ponzio, C., Cascone, P., Cusumano, A., Weldegergis, B. T., Fatouros, N. E., Guerrieri, E., et al. (2016). Volatile-mediated foraging behaviour of three parasitoid species under conditions of dual insect herbivore attack. *Anim. Behav.* 111, 197–206. doi: 10.1016/j.anbehav.2015.10.024
- Proietti, S., Caarls, L., Coolen, S., Van Pelt, J. A., Van Wees, S. C., and Pieterse, C. M. (2018). Genome-wide association study reveals novel players in defense hormone crosstalk in *Arabidopsis*. *Plant Cell Environ.* 41 (10), 2342–2356. doi: 10.1111/pce.13357
- Reisenman, C. E., Riffell, J. A., Bernays, E. A., and Hildebrand, J. G. (2010). Antagonistic effects of floral scent in an insect–plant interaction. *Proc. R. Soc. Lond. [Biol]* 277 (1692), 2371–2379. doi: 10.1098/rspb.2010.0163
- Renwick, J. A. A., and Chew, F. S. (1994). Oviposition behavior in Lepidoptera. *Annu. Rev. Entomol.* 39 (1), 377–400. doi: 10.1146/annurev.en.39.010194.002113
- Reymond, P. (2013). Perception, signaling and molecular basis of oviposition-mediated plant responses. *Planta* 238 (2), 247–258. doi: 10.1007/s00425-013-1908-y
- Sato, Y., Yano, S., Takabayashi, J., and Ohsaki, N. (1999). *Pieris rapae* (Lepidoptera: Pieridae) females avoid oviposition on *Rorippa indica* plants infested by conspecific larvae. *Appl. Entomol. Zool.* 34 (3), 333–337. doi: 10.1303/aez.34.333
- Schoonhoven, L. M., Beerling, E. A. M., Braaksma, R., and Van Vugt, Y. (1990). Does the imported cabbageworm, *Pieris rapae*, use an oviposition deterring pheromone? *J. Chem. Ecol.* 16 (5), 1649–1655. doi: 10.1007/BF01014097
- Schuman, M. C., and Baldwin, I. T. (2016). The layers of plant responses to insect herbivores. *Annu. Rev. Entomol.* 61, 373–394. doi: 10.1146/annurev-ento-010715-023851
- Shapiro, A. M., and DeVay, J. E. (1987). Hypersensitivity reaction of *Brassica nigra* L. (Cruciferae) kills eggs of *Pieris* butterflies (Lepidoptera: Pieridae). *Oecologia* 71 (4), 631–632. doi: 10.1007/BF00379310
- Szentezki, M. A., Pitteloud, C., Casacci, L. P., Kešnerová, L., Whitaker, M. R., Engel, P., et al. (2019). Bacterial communities within *Phengaris (Maculinea) alcon* caterpillars are shifted following transition from solitary living to social parasitism of *Myrmica* ant colonies. *Ecol. Evol.* 9 (8), 4452–4464. doi: 10.1002/ece3.5010
- Tamiru, A., Bruce, T. J. A., Woodcock, C. M., Caulifield, C. J., Midega, C. A. O., and Ogot, C. K. P. O. (2011). Maize landraces recruit egg and larval parasitoids in response to egg deposition by a herbivore. *Ecol. Lett.* 14 (11), 1075–1083. doi: 10.1111/j.1461-0248.2011.01674.x
- Tamiru, A., Bruce, T. J. A., Richter, A., Woodcock, C. M., Midega, C. A. O., Degenhardt, J., et al. (2017). A maize landrace that emits defense volatiles in response to herbivore eggs possesses a strongly inducible terpene synthase gene. *Ecol. Evol.* 7 (8), 2835–2845. doi: 10.1002/ece3.2893
- Wang, J., Yang, M., Song, Y., Acevedo, F. E., Hoover, K., Zeng, R., et al. (2018). Gut-associated bacteria of *Helicoverpa zea* indirectly trigger plant defenses in maize. *J. Chem. Ecol.* 44 (7–8), 690–699. doi: 10.1007/s10886-018-0976-7

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Berteza, Casacci, Bonelli, Zampollo and Barbero. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# AcDCXR Is a Cowpea Aphid Effector With Putative Roles in Altering Host Immunity and Physiology

Jacob R. MacWilliams<sup>1</sup>, Stephanie Dingwall<sup>2</sup>, Quentin Chesnais<sup>3</sup>, Akiko Sugio<sup>4</sup> and Isgouhi Kaloshian<sup>1,5,6\*</sup>

<sup>1</sup> Graduate Program in Biochemistry and Molecular Biology, University of California, Riverside, Riverside, CA, United States, <sup>2</sup> Department of Biochemistry, University of California, Riverside, Riverside, CA, United States, <sup>3</sup> Université de Strasbourg, INRAE, SVQV UMR-A1131, Colmar, France, <sup>4</sup> INRAE, UMR1349, Institute of Genetics, Environment and Plant Protection, Le Rheu, France, <sup>5</sup> Department of Nematology, University of California, Riverside, Riverside, CA, United States, <sup>6</sup> Institute for Integrative Genome Biology, University of California, Riverside, Riverside, CA, United States

## OPEN ACCESS

### Edited by:

Carolina Escobar,  
University of Castilla-La Mancha,  
Spain

### Reviewed by:

Sam T. Mugford,  
John Innes Centre, United Kingdom  
Eduard Venter,  
University of Johannesburg,  
South Africa

### \*Correspondence:

Isgouhi Kaloshian  
isgouhi.kaloshian@ucr.edu

### Specialty section:

This article was submitted to  
Plant Microbe Interactions,  
a section of the journal  
Frontiers in Plant Science

**Received:** 04 February 2020

**Accepted:** 21 April 2020

**Published:** 15 May 2020

### Citation:

MacWilliams JR, Dingwall S,  
Chesnais Q, Sugio A and Kaloshian I  
(2020) AcDCXR Is a Cowpea Aphid  
Effector With Putative Roles in Altering  
Host Immunity and Physiology.  
*Front. Plant Sci.* 11:605.  
doi: 10.3389/fpls.2020.00605

Cowpea, *Vigna unguiculata*, is a crop that is essential to semiarid areas of the world like Sub-Saharan Africa. Cowpea is highly susceptible to cowpea aphid, *Aphis craccivora*, infestation that can lead to major yield losses. Aphids feed on their host plant by inserting their hypodermal needlelike flexible stylets into the plant to reach the phloem sap. During feeding, aphids secrete saliva, containing effector proteins, into the plant to disrupt plant immune responses and alter the physiology of the plant to their own advantage. Liquid chromatography tandem mass spectrometry (LC-MS/MS) was used to identify the salivary proteome of the cowpea aphid. About 150 candidate proteins were identified including diacetyl/L-xylulose reductase (DCXR), a novel enzyme previously unidentified in aphid saliva. DCXR is a member of short-chain dehydrogenases/reductases with dual enzymatic functions in carbohydrate and dicarbonyl metabolism. To assess whether cowpea aphid DCXR (AcDCXR) has similar functions, recombinant AcDCXR was purified and assayed enzymatically. For carbohydrate metabolism, the oxidation of xylitol to xylulose was tested. The dicarbonyl reaction involved the reduction of methylglyoxal, an  $\alpha$ - $\beta$ -dicarbonyl ketoaldehyde, known as an abiotic and biotic stress response molecule causing cytotoxicity at high concentrations. To assess whether cowpea aphids induce methylglyoxal in plants, we measured methylglyoxal levels in both cowpea and pea (*Pisum sativum*) plants and found them elevated transiently after aphid infestation. Agrobacterium-mediated transient overexpression of AcDCXR in pea resulted in an increase of cowpea aphid fecundity. Taken together, our results indicate that AcDCXR is an effector with a putative ability to generate additional sources of energy to the aphid and to alter plant defense responses. In addition, this work identified methylglyoxal as a potential novel aphid defense metabolite adding to the known repertoire of plant defenses against aphid pests.

**Keywords:** cowpea aphid, salivary proteins, effector, diacetyl/L-xylulose reductase, DCXR, methylglyoxal, host defense

## INTRODUCTION

Cowpea (*Vigna unguiculata*) is one of the most important agronomic plant species grown in semiarid tropical regions of the world. Cowpea is well adapted to biotic and abiotic stresses and provides an excellent source of nutrition (Singh et al., 2002; Timko and Singh, 2008). However, a stress that is a limiting factor in cowpea production is infestation by the cowpea aphid, *Aphis craccivora* (Jackai and Daoust, 1986). Cowpea aphid infestation can cause devastating effects; it has been reported that young plants of highly susceptible cowpea cultivars were killed by an infestation of cowpea aphids initiated with fewer than ten aphids (Ofuya, 1995). Cowpea aphid feeding induced damage includes chlorosis, leaf curling, and stunted growth resulting in a decrease in yield (Blackman and Eastop, 2000; Kamphuis et al., 2012; Choudhary et al., 2017). In addition to cowpea aphid being a deadly pest, this aphid species is also known to vector over 50 plant viruses (Chan et al., 1991).

There are about 4500 species of aphids reported to date (Remaudiere and Remaudiere, 1997; Blackman and Eastop, 2000; Sorenson, 2009). Of these species, only 100 are considered to have an economic impact and 14 are considered to be serious pests, among which is the cowpea aphid (Sorenson, 2009). Aphids feed differently from chewing insects, which generate massive mechanical tissue damage. Aphids insert their specialized and flexible mouthparts, the stylets, through plant tissues to reach their source of food, the phloem sap, thus avoiding much of the mechanical tissue damage (Tjallingii and Esch, 1993; Tjallingii, 2006). *En route* to the phloem, aphids puncture cells and deposit saliva in the plant apoplast and the punctured cells to facilitate feeding and interfere with plant defenses (Miles, 1999; Will et al., 2007). Aphid feeding and colonization damage the plant, and aphids are categorized based on the type of damage they incur onto their hosts. Aphids that cause extensive direct damage are considered phytotoxic, whereas others that cause indirect damage – for example, by transmitting viruses – are considered non-phytotoxic (Nicholson et al., 2012). Phytotoxic aphids, such as the Russian wheat aphid (*Diuraphis noxia*) and greenbug (*Schizaphis graminum*), cause damage in low numbers and are believed to secrete salivary proteins into the plant that are responsible for the increased manifestation of the damage symptoms. In contrast, the non-phytotoxic aphids, like the pea aphid (*Acyrtosiphon pisum*) and potato aphid (*Macrosiphum euphorbiae*), do not cause damage at low numbers and secrete salivary proteins to enhance feeding and interfere with plant defenses (Nicholson et al., 2012; Nicholson and Puterka, 2014; Choudhary et al., 2015).

Aphid saliva has been shown to contain effector proteins that are necessary for successful aphid colonization (Mutti et al., 2006, 2008; Bos et al., 2010; Atamian et al., 2013; Pitino and Hogenhout, 2013; Elzinga et al., 2014; Naessens et al., 2015; Wang et al., 2015; Will and Vilcinskas, 2015; Guy et al., 2016; Kaloshian and Walling, 2016). To characterize aphid salivary protein content, the saliva of several aphid species has been investigated with liquid chromatography

tandem mass spectrometry (LC-MS/MS) (Harmel et al., 2008; Carolan et al., 2009; Cooper et al., 2010, 2011; Rao et al., 2013; Vandermoten et al., 2014; Chaudhary et al., 2015; Thorpe et al., 2016; Boulain et al., 2018; Loudit et al., 2018; Yang et al., 2018). These studies have identified numerous conserved salivary proteins common among the different aphid species as well as some that have only been identified in a single aphid species. The conserved proteins are presumed to be a core set of aphid effectors that are used by aphids to facilitate feeding or disrupt general plant defenses, while the unique proteins identified in only a single aphid species or biotype, act in a species-specific host-aphid interaction (Thorpe et al., 2016). This recent wealth of salivary protein identification stems from the release of additional aphid genomes and transcriptomes. Since the first aphid genome was released for the pea aphid, five additional aphid genomes are publicly available (International Aphid Genomics Consortium, 2010; Nicholson et al., 2015; Mathers et al., 2017; Wenger et al., 2017; Thorpe et al., 2018). Numerous aphid transcriptomes are also available including a transcriptome for the cowpea aphid (Agunbiade et al., 2013). Three main criteria have been used to identify putative aphid effectors: (1) expression of the candidate transcripts in aphid heads or salivary glands with prediction for secretion, (2) presence in saliva, and (3) sequence similarity to previously identified aphid effectors.

In general, microbial, nematode and pest effectors are diverse, lacking consensus sequences and features, making it difficult to predict effectors. This has led to reporting of mostly specific subclasses of effectors. For example, effectors from plant pathogenic fungi are small sized proteins with high cysteine content while those from *Phytophthora* contain a RXLR motif (Jiang et al., 2008; Stergiopoulos and de Wit, 2009; Petre and Kamoun, 2014; Sperschneider et al., 2015). To enhance plant fungal effector predictions, EffectorP was developed as a machine-learned predictor for fungal effectors that does not rely only on predetermined thresholds based on criteria including protein size and cysteine content (Sperschneider et al., 2016, 2018). It is therefore likely that the repertoire of aphid effectors can be enhanced with the development of machine learned effector identification programs.

Numerous studies have functionally characterized aphid effectors. These included overexpression of the candidate effector *in planta* or silencing it, through plant-mediated RNAi or injection with RNAi constructs, in the aphid and determining aphid performance on the plants. Of the effectors experimentally tested, about a dozen have shown altered aphid colonization phenotypes (Mutti et al., 2006, 2008; Bos et al., 2010; Atamian et al., 2013; Pitino and Hogenhout, 2013; Elzinga et al., 2014; Abdellatef et al., 2015; Naessens et al., 2015; Wang et al., 2015; Will and Vilcinskas, 2015; Guy et al., 2016; Kettles and Kaloshian, 2016). The altered survival/colonization phenotypes determined by some of these effectors act in species-specific and host-specific manner (Atamian et al., 2013; Pitino and Hogenhout, 2013; Elzinga et al., 2014; Rodriguez et al., 2017).

To date, the plant targets for only Mp1 and Me10 aphid effectors have been identified and the mechanism of effector function partially elucidated (Rodriguez et al., 2017; Chaudhary et al., 2019). The function of two additional aphid effectors MIF1 (Naessens et al., 2015) and Armet (Wang et al., 2015) have been predicted based on the function of homologous sequences from other organisms. Both MIF1 and Armet are highly conserved proteins in the animal kingdom. MIF1 encodes a macrophage migration inhibitory factor that is a cytokine deposited in aphid saliva during feeding (Calandra, 2003; Naessens et al., 2015). Armet in mammalian systems and in *Drosophila* has been reported in the cell as part of the unfolded protein response and extracellularly as a neurotrophic factor (Lindholm et al., 2007, 2008; Palgi et al., 2009, 2012). Both MIF1 and Armet are important for the pea aphid survival as knockdown of their expressions results in shortened lifespan (Naessens et al., 2015; Wang et al., 2015). The function of an additional effector, Me47 encoding a Glutathione S-transferase (GST), was shown based on its GST enzymatic activity and its ability to detoxify isothiocyanates that are implicated in herbivore defense (Kettles and Kaloshian, 2016).

Here we report the salivary proteome of a California population of the cowpea aphid using LC-MS/MS and publicly available aphid genomes and transcriptomes. We also characterize the function of a novel salivary protein, diacetyl/L-xylulose reductase (DCXR). DCXR is a member of short-chain dehydrogenases/reductases (Nakagawa et al., 2002). Mammalian orthologs of DCXR are involved in NADPH-dependent reduction of both carbohydrates and dicarbonyls (Nakagawa et al., 2002; Ishikura et al., 2003; Ebert et al., 2015). The reversible oxidative reduction of the carbohydrates xylitol and L-xylulose can lead to an additional energy source through the pentose phosphate pathway (Sochor et al., 1979; Nakagawa et al., 2002). The reduction of dicarbonyls detoxifies and prevents the formation of advanced glycation end-products (AGEs), also known as glycotoxins, associated with development of numerous degenerative human diseases (Chen et al., 2009; Gkogkolou and Bohm, 2012; Kizer et al., 2014). In plants, the build-up of dicarbonyls leads to oxidative stress and cell death resulting in stunted growth (Hoque et al., 2012; Ray et al., 2013; Sankaranarayanan et al., 2015; Li, 2016). One of these dicarbonyls, generated through multiple pathways in plants and animals, is methylglyoxal (Yadav et al., 2005a,b; Hoque et al., 2016; Mostofa et al., 2018). Depending on concentration, methylglyoxal can act as defense signaling molecule or as a cytotoxin during abiotic stress in plants (Li, 2016). Recently methylglyoxal has also been implicated in plant defense against biotic stresses (Melvin et al., 2017). Here we report the identification of DCXR in cowpea aphid saliva. We show that the recombinant cowpea aphid DCXR, AcDCXR, is able to catalyze the reversible xylitol to xylulose reaction as well as to utilize methylglyoxal as substrate. We also demonstrate that aphid feeding induced methylglyoxal accumulation and that expression of AcDCXR in *planta* enhanced aphid fecundity contributing to the success of the aphid as a pest.

## MATERIALS AND METHODS

### Plants and Growth Condition

Cowpea California blackeye cultivar 46 (CB46) and pea (*Pisum sativum*) cv ZP1130 were grown in UC Mix 3 soil<sup>1</sup> in 32 oz plastifoam cups in a pesticide free room at 22–24°C with a 16:8 light:dark photoperiod. Plants were fertilized weekly with MiracleGro (18-18-21; Stern's MiracleGro Products).

### Aphid Colony

A colony of cowpea aphids, collected from a field in Riverside, California, in summer of 2016, was reared on cowpea cv CB46. A second colony, taken from the cowpea plants, was reared on pea cv ZP1130 for 3 months before use. The colonies was maintained separately in insect cages in growth chambers at 26–30°C with a 16:8 light:dark photoperiod. The colony on cowpea was used for aphid saliva collection and the colony on pea was used for aphid bioassays.

### Saliva Collection

Cowpea aphid saliva was collected by feeding mixed developmental stages of the aphid on a water diet as previously described (Chaudhary et al., 2015). About 100–200 mixed stage aphids were loaded in a feeding chamber, consisting of a plastic cylinder with one end containing the diet inside a parafilm sachet, and the other end secured with a cheesecloth. Aphids were allowed to feed on the 200 µL of ultrapure autoclaved water for 16 h under yellow light. The components of the chamber were sterilized or treated with alcohol and all materials were handled in a laminar flow hood using aseptic conditions. After feeding the diet was collected aseptically using a pipet and stored at –80°C. A new cohort of aphids were used for each overnight collection and saliva was collected from an estimated 10,000 aphids over a three-month period.

### Saliva Preparation for MS/MS

Saliva was vacuum concentrated down to protein pellets and dissolved in 100 µL trypsin buffer (50 mM ammonium bicarbonate, pH 8.0, 10% v/v acetonitrile) containing 1 µg trypsin and treated overnight at 37°C. After trypsin digestion, the sample was centrifuged, the supernatant was collected, pelleted with a speedvac concentrator and suspended in 24 µL 0.1% formic acid for LC-MS/MS analysis.

### LC-MS/MS

A MudPIT approach was employed to analyze the trypsin-treated samples. A two-dimension nanoAcquity UPLC (Waters) and an Orbitrap Fusion MS (ThermoFisher Scientific) were configured together to perform online 2D-nano LC-MS/MS analysis. The 2D-nanoLC was operated with a 2D-dilution method that was configured with nanoAcquity UPLC. Two mobile phases for the first dimension LC fractionation were 20 mM ammonium formate (pH 10) and acetonitrile, respectively.

<sup>1</sup><https://agops.ucr.edu/soil-mixing>

Online fractionation was achieved by 5-min elution off a NanoEase trap column (PN# 186003682; Waters) using stepwise-increased concentration of acetonitrile. A total of five fractions were generated with 13, 18, 21.5, 27, and 50% of acetonitrile, respectively. A final flushing step used 80% acetonitrile to clean up the trap column. Each and every fraction was then analyzed online using a second dimension LC gradient.

For the second-dimension LC, a BEH130 C18 column (1.7  $\mu\text{m}$  particle, 75  $\mu\text{m}$  i.d., 20 cm long, PN# 186003544; Waters) was used for peptide separation. A Symmetry C18 (5  $\mu\text{m}$  particle, 180  $\mu\text{m}$  i.d., 20 mm long, PN# 186003514; Waters) served as a trap/guard column for desalting and pre-concentrating the peptides for each MudPIT fraction. The solvent components for peptide separation were as follows: mobile phase A was 0.1% formic acid in water, and mobile phase B was 0.1% formic acid in acetonitrile. The separation gradient was as follows: at 0 to 1 min, 3% B; at 2 min, 8% B; at 50 min, 45% B; at 52–55 min, 85% B; at 56–70 min, 3% B. The nano-flow rate was set at 0.3  $\mu\text{l}/\text{min}$  without flow-splitting.

Spectra were obtained using Orbitrap Fusion MS (Thermo Fisher Scientific). The Orbitrap Fusion MS was in positive ion mode with an ion transfer tube temperature of 275°C. The isolation window used was 2 Da. Three different types of dissociation were used: Collision Induced Dissociation (CID), High-energy Collision Induced Dissociation (HCD), and Electron Transfer Dissociation (ETD). The energy for each of these was 30%. Three scan ranges were used (300–1800, 300–2000 400–1400 Da) with 30 s dynamic exclusion.

## Proteome Data Analysis

The MS/MS spectra were filtered for high confident peptides with strict FDR (1%), with enhanced peptide and protein annotations using the software Proteome Discoverer v2.3 (Thermo Fisher). Spectra with peptide sequences less than 6 residues were removed. The search parameters allowed for 0.5 Da mass tolerance and 2 missed cleavage sites. The following modifications were included: modification of Met Oxidation  $\pm$  15.99492 D, Lys Acetyl  $\pm$  42.01057 D, Ser, Thr, Tyr Phospho  $\pm$  79.966333 D, *N*-Terminus Formyl  $\pm$  27.99492 Da, Pyro-Glu  $\pm$  17.02655 Da, *N*-Terminus Acetyl  $\pm$  42.01057 Da. The identified peptides were then searched against an aphid proteome database compiled from every aphid genome available on NCBI and AphidBase (Pea aphid, Russian wheat aphid, soybean aphid (*Aphis glycines*), bird cherry-oat aphid (*Rhopalosiphum padi*), green peach aphid (*Myzus persicae*), and black cherry aphid (*Myzus cerasi*) and other aphid proteins deposited in NCBI in 2017. These other proteins included six-frame translations of a cowpea aphid transcriptome and the transcriptome of the potato aphid). The 13,330 PSMs identified corresponded to 2,119 proteins and were further filtered to 721 protein group hits. Only high confidence (99%) were considered further filtering the protein groups to 521 protein groups. Spectra that came up when filtering out possible contaminants with a FASTA file containing common contaminants. To accept proteins, they needed to have at least 3 peptides in at least 2 of the 3 replicates (CID, HCD, ETD). The raw peptide spectra were deposited in the Mass Spectrometry

Interactive Virtual Environment (MassIVE) repository with the proteome ID: PXD017323.

## Annotation

The MS/MS identified proteins were annotated with BLASTP using OmicsBox (V 1.1.135 Hotfix) and the NCBI non-redundant protein database with the taxonomy filter for aphids, Aphidomorpha (3380) (*e*-value =  $1e^{-3}$ ) (Gotz et al., 2008). The proteins were then subjected to BLASTP to the pea aphid annotation v2.1b proteins on Aphidbase to identify the corresponding ACYPI homologs (BIPAA, 2017). Gene ontology (GO) was determined for molecular function, biological process, and cellular component using InterProScan (v5.36-75.0) (Gotz et al., 2008; Jones et al., 2014). The identified proteins were screened with SignalP (V3.0 and V5.0) and SecretomeP 1.0 using eukaryote and mammalian filters, respectively, and by TMHMM V2.0 (Krogh et al., 2001; Bendtsen et al., 2004a,b; Armenteros et al., 2019). The proteins were further analyzed using EffectorP 2.0 (Sperschneider et al., 2018).

## Clone Construction

RNA was extracted from 10 mixed developmental stage aphids using Trizol (Invitrogen), and cDNA was synthesized using SuperScript III reverse transcriptase (Invitrogen) according to manufactures instructions. Using *AcDCXR* (MN855408) gene-specific Gateway recombination primers (DCXRF-ACAAGTTTGTACAAAAAAGCAGGCTCCATGGAAGAATTC TTTGTGCGAAAAAAGTTTCAT, DCXRR- GGGGACCACT TTGTACAAGAAAGCTGGGTCCTGGCCAAAAATCCACCA TC), the DCXR coding region, excluding the secretion signal peptide, was amplified using Q5<sup>®</sup> High-Fidelity DNA Polymerase (New England Biolabs) with the following conditions: an initial 98°C for 30 s, 98°C for 7 s, 54°C for 20 s, 72°C for 30 s, for 30 cycles and a final cycle of 72°C for 3 min. DCXR was purified using GeneJET PCR Purification Kit (Thermo Scientific) and recombined into vector pDONR207 (Invitrogen) using BP Clonase (Invitrogen). Following Sanger sequencing pDONR207-DCXR was recombined into the expression vectors pDEST17 (Invitrogen; pDEST17-DCXR), pEAQ-HT-DEST1 (Sainsbury et al., 2009; pEAQ-HT-DEST1-AcDCXR), or pCAMBIA1300-GW-mScarlet (pCAMBIA1300-AcDCXR-mScarlet). pCAMBIA1300-GW-mScarlet was developed by modifying pCAMBIA1300 using parts from pGWB614 and p#128060 by restriction digestion and ligations. After transformation into *E. coli* strain DH5 $\alpha$  and the purified pDEST17-DCXR was transformed into *E. coli* strain ArcticExpress (Agilent) while pEAQ-HT-DEST1-AcDCXR and pCAMBIA1300-AcDCXR-mScarlet were transformed into *Agrobacterium tumefaciens* strains AGL01 and GV3101, respectively.

## Protein Purification

The pDEST17-AcDCXR was purified in a similar manner as previously described for the aphid effector Me47 (Kettles and Kaloshian, 2016). Briefly, pDEST17-AcDCXR (*N*-terminal 6xHis tag) in ArcticExpress was grown in LB media at 37°C to an OD<sub>600</sub> of 0.8 and the expression induced by adding of 0.5 mM IPTG followed by incubation at 10°C for 16 h. After centrifugation

(6,000 × *g* for 20 min) the cells were resuspended in chilled lysis buffer (300 mM NaCl, 50 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>, pH 7.0). The cells were lysed using sonication (4 × 15 s pulses), the soluble protein fraction was separated by centrifugation (10,000 × *g* for 45 min) and incubated with Ni-NTA agarose beads (Qiagen) for 1 h at 4°C with gentle agitation. The column was washed with the lysis buffer containing 40 mM imidazole to remove non-specifically bound proteins. After four washes, DCXR was eluted with three washes of lysis buffer containing 150, 200, and 200 mM of imidazole, respectively. The eluted fractions were concentrated with VivaSpin 500 Centrifugal Concentrator PES (Sartorius, United Kingdom) and monitored using Bradford assay with BSA as the standard. The recombinant DCXR was analyzed on a 12% SDS-PAGE using Coomassie Brilliant Blue G-250 staining.

### AcDCXR Enzyme Activity Assays

Oxidation of xylitol to xylulose by recombinant DCXR was measured through the reduction of NADP<sup>+</sup> to NADPH as previously described (Yang et al., 2017) with minor modification. A 0.5 mL reaction mixture containing 10 μg AcDCXR 100 mM glycine buffer, pH 9.5, 3 mM MgCl<sub>2</sub>, NADP<sup>+</sup>, and 200 mM xylitol were used in 1 mL cuvettes and a Beckman Coulter Du<sup>®</sup> 730 Life Sciences spectrophotometer. Reactions began after the addition of AcDCXR, and changes in absorbance at 340 nm were monitored. The reaction rates were calculated based on the NADP<sup>+</sup> concentrations.

Methylglyoxal reduction by recombinant DCXR was measured through the oxidation of NADPH to NADP<sup>+</sup> using 1 mL cuvettes as previously described (Misra et al., 1996) and the Beckman Coulter Spectrophotometer. The 0.5 mL reaction was composed of 10 μg DCXR, 100 μM sodium phosphate buffer (NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>, pH 6.5), 200 μM NADPH and methylglyoxal. The reaction was initiated with the addition of NADPH and monitored by the decrease in absorbance at 340 nm. The reaction rates were calculated based on the methylglyoxal concentrations.

### Transient Expression in Pea and Western Blot Analysis

*Agrobacterium tumefaciens* strain AGL01, carrying either pEAQ-HT-GFP or pEAQ-HT-DEST1-AcDCXR, were used in transient expression of pea, *Pisum sativum*, cv. ZP1130 as described previously (Guy et al., 2016). Bacterial cells, grown up overnight in YEP media, were harvested, washed three times in infiltration buffer (10 mM MgCl<sub>2</sub>, 10 mM MES pH 5.6, and 150 μM acetosyringone) and resuspended at a final OD<sub>600</sub> of 0.5. The youngest expanded leaf of a 2-week-old plant was infiltrated with a needleless syringe.

The duration of GFP expression in pEAQ-HT-GFP infiltrated leaves was monitored with Western blot analysis. Three 1 cm diameter leaf disks were cut from the same agroinfiltrated leaf using a cork borer after 2, 3, 5, 7, 8, 9, and 10 days post infiltration. Protein was extracted from the leaf disks by grinding in 200 μl lysis buffer (6 M Urea, 2 M Thiourea, 1% Protease inhibitor cocktail [Sigma P9599]). Samples were centrifuged at 14,000 × *g* for 5 min and the supernatant was

resuspended in equal volume 2x loading buffer (100 mM Tris-HCl pH 6.8, 100 mM DTT, 10% glycerol, 2% SDS, 0.01% bromophenol blue). About 25 μg of protein were loaded per sample on 12% SDS-PAGE and transferred to a nitrocellulose membrane. The membrane was probed with mouse anti-GFP antibody (Sigma) and secondary antibody, goat anti-mouse HRP-conjugated (Santa Cruz Biotechnology). Primary antibody was used at 1:2000 and secondary antibody was used at 1:2000 dilution. Pierce ECL Western Blotting Substrate (Thermo Scientific) was used to detect the signal with autoradiography film (Denville Scientific Inc.).

### *in planta* Localization of AcDCXR

*Agrobacterium tumefaciens* GV3101 carrying pCAMBIA1300-DCXR-mScarlet or pCAMBIA1300-GFP were grown and prepared as previously described for transient agroexpression. At an OD<sub>600</sub> = 0.5 each, the constructs were co-infiltrated in *Nicotiana benthamiana* leaves. Three days post infiltration, leaf epidermal cells were analyzed using a Leica SP5 confocal microscope. GFP and mScarlet were excited by 488 nm and 543 nm filters, respectively, and images were collected through band emission filters at 498–520 nm and 553–650 nm, respectively.

### Aphid Bioassays

A day after agroinfiltration, five adult cowpea aphids were caged onto the adaxial side of an agroinfiltrated leaf of 2-week-old pea plants. After 24 h (i.e., 2 days post infiltration; dpi), the adult aphids were removed, and 5 to 6 new-born nymphs were left on the leaf with both the adaxial and abaxial sides of the leaf accessible to the aphids. Eight days later (10 dpi), the surviving aphids were counted and transferred to a new infiltration site on a plant infiltrated 2 days earlier. The fecundity of these aphids was monitored two and five days later (i.e., when the aphids were 12 and 15 day-old). The nymphs were removed after each counting. This experiment was performed three times. Each experiment consisted of 13–15 plants per construct. All experiments were conducted at 22°C, 16:8 light:dark photoperiod.

### Determining Methylglyoxal Levels

Methylglyoxal levels were evaluated in 2-week-old cowpea and pea plants following the protocol by Borysiuk et al. (2018). Highly infested leaves were harvested at day 1, 2, and 3 after infestation. Briefly, samples were homogenized in 5% perchloric acid and centrifuged at 13,000 × *g* for 10 min at 4°C. The supernatant was decolorized with charcoal and neutralized with 1 M potassium carbonate. After centrifuging at 13,000 × *g* at 4°C the supernatant was used to estimate the methylglyoxal concentration in sodium dihydrogen phosphate buffer (pH 7.0). The absorbance was recorded after 10 min incubation with *N*-acetyl-L-cysteine to monitor the *N*-a-acetyl-S-(1-hydroxy-2-oxo-prop-1-yl) cysteine formation (Wild et al., 2012). Methylglyoxal concentration was determined using a standard curve of known methylglyoxal concentrations. The experiment with pea was performed once and with cowpea was performed twice.

## Statistical Analyses

We used generalized linear models (GLM) with a likelihood ratio and chi-square test to assess whether AcDCXR expression had an effect on aphid survival and fecundity. Data on aphid survival were analyzed with GLM following a binomial distribution and data on aphid fecundity were assumed to follow a Poisson distribution. The fit of all generalized linear models was checked by inspecting residuals and QQ plots. Methylglyoxal levels in plants were analyzed using a nested ANOVA (biological replicates treated as random factor) (package R: 'nlme'). When a significant effect was detected, a pairwise comparison using multiple comparisons of the means (package R: 'multcomp') (Tukey contrasts, *p*-values adjustment with 'fdr' method) at the 0.05 significance level was used to test for differences between days after infestation. Statistical analyses were performed using the R software (version 3.6.0) (R Core Team, 2019).

## RESULTS

### Aphid Salivary Proteome Analyses and Annotation

To identify the protein composition of the cowpea aphid saliva, aphid saliva was collected in parafilm feeding pouches containing water. The contents of the pouches were concentrated and subjected to proteome analyses. The peptides identified by LC-MS/MS were searched against a custom aphid protein database. The database was composed of proteomes based on all aphid genomes available in the summer of 2017, as well as cowpea aphid-specific transcriptome and a transcriptome from the potato aphid, both with six-frame translations. Around 175 candidate proteins were identified with at least three peptides from at least two replicates and having at least one unique peptide (**Supplementary Table 1**). The identified proteins were then annotated using BLASTP with OmicsBox (TaxID: Aphididae 27482). Among these annotated proteins, 18/175 (10.29%) were uncharacterized. In addition, functional redundancies were recorded among the proteins with annotations. To eliminate these redundancies, the proteins were subjected to BLASTP on AphidBase to identify their corresponding ACYPI homolog using the pea aphid protein database annotation v2.1b. Among these proteins, 47/175 (26.86%) shared one of 21 ACYPI top hits. Although these 47 proteins had at least one unique peptide, we grouped them as 21 proteins, resulting in a total of 149 salivary proteins (**Supplementary Table 1**).

Annotation of these proteins presented a wide range of functional attributes to the cowpea aphid salivary proteins. Among the 149 identified proteins, 33 proteins with similar functional annotations have been previously reported in the saliva of a cowpea aphid population from Gabon, Africa (Loudit et al., 2018). Among these 33 proteins are glucose dehydrogenases, carbonic anhydrases and a trehalase (**Supplementary Table 1**).

Of the 149 identified cowpea aphid proteins, gene ontology (GO) assigned 123 proteins with at least one GO term in the three most common ontological designations: molecular

function, biological process and cellular component. The three most abundant biological process designations were carbohydrate metabolic process (19%), translation (16%) and catabolic process (11%) (**Figure 1**). The three most abundant molecular function designations were oxidoreductase activity (20%), structural constituent of ribosome (16%) and ATP binding (13%) (**Figure 1**). As for the most abundant cellular component designations, they were for protein-containing complexes (33%) and cytosol (29%) (**Figure 1**).

### Effector Prediction

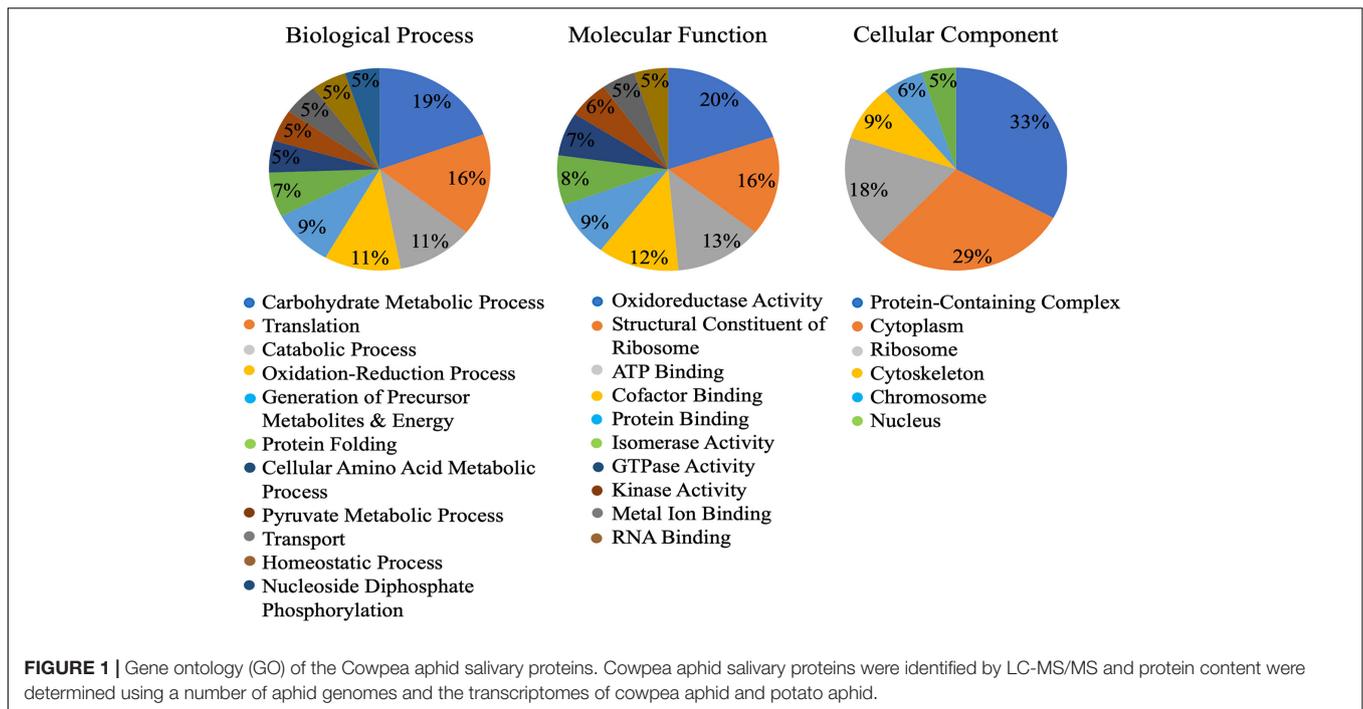
Since the cowpea aphid genome has not been sequenced, homologous proteins from the different aphid species or those based on cowpea aphid transcriptome, used in our custom database, were used for these analyses. Multiple bioinformatics tools were harnessed to screen the identified salivary proteins for putative effector function. First, the salivary proteins were evaluated for secretion using tools that predict classical and non-classical secretions, SignalP and SecretomeP (Bendtsen et al., 2004a,b; Armenteros et al., 2019), respectively. Using SignalP, a secretion signal was detected in 29 (19.46%) proteins, while SecretomeP predicted the secretion of an additional 23 (15.44%) of the 149 salivary proteins (**Table 1**). To eliminate proteins with transmembrane domains, presence of transmembrane helices was evaluated using TMHMM (Sonnhammer et al., 1998). Six of these predicted secreted proteins contained transmembrane helices.

A machine learning approach was recently used to develop novel prediction program for fungal effectors (Sperschneider et al., 2016, 2018). We wondered whether this tool, EffectorP, could be used to predict aphid effectors. To test this, we first subjected known aphid effectors for EffectorP analysis. We tested the C002 effector, identified first in pea aphid (Mutti et al., 2008), and Me10, identified in potato aphid (Atamian et al., 2013). Both C002 and Me10 were identified as effectors by EffectorP indicating that EffectorP can be utilized as a tool to screen for aphid effectors. Using EffectorP, 20/149 (13.4%) of the cowpea aphid salivary proteins were identified as putative effectors (**Table 1**). Only eight of the 20 were identified for secretion by SignalP or SecretomeP. Taken together 58 proteins were predicted for secretion or for effector function encoding a wide range of functions with eight being unknowns (**Figure 2** and **Table 1**).

### Selection and *in vitro* Characterization of AcDCXR

A set of criteria were applied to choose a putative effector protein identified by EffectorP for functional characterization. These included a previously unidentified effector predicted for secretion or with secretion signal peptide, a protein with predicted enzymatic activity, and high abundance in cowpea aphid saliva based on the SEQUEST score. Based on these criteria, DCXR was selected for further analysis.

Sequence prediction indicated that cowpea aphid DCXR (AcDCXR; GAJW01000401.1) consists of at least 263 amino acids, with the first 23 amino acids encoding a predicted



signal peptide, and a conserved enzymatic domain for short-chain dehydrogenases/reductases (**Supplementary Figure 1**). Using AcDCXR in BLASTP searches identified DCXR homologs in seven aphid species. Interestingly, only the DCXR from cotton melon aphid (*Aphis gossypii*; XP\_027848224.1) contains a secretion signal peptide (**Supplementary Figure 1**). Consistent with this information, DCXR has been reported previously from other aphid species but has not been previously identified in aphid saliva (Nguyen et al., 2008, 2009; Pinheiro et al., 2014).

Diacetyl/L-xylulose reductase is a multifunctional enzyme. Mammalian orthologs of DCXR have been shown to function in the glucuronic acid/uronate cycle, in a reversible reaction either oxidizing or reducing xylitol and xylulose, respectively (Sochor et al., 1979; Yang et al., 2017), as well as having  $\alpha$ - $\beta$ -dicarbonyl reductase activity to metabolize toxic carbonyls like methylglyoxal (Ebert et al., 2015). Direct comparison between AcDCXR and XP\_027848224.1 showed 100% identity at the amino acid level with perfect conservation of the enzyme active site (**Supplementary Figure 1**). To test whether AcDCXR has similar functions as the mammalian orthologs, we expressed recombinant AcDCXR and performed enzymatic assays.

Aphid diacetyl/L-xylulose reductase, amplified from cDNAs developed from the whole bodies of mixed stages of the aphid, was cloned into the pDEST17 expression vector and expressed in *E. coli* strain ArcticExpress. Purified AcDCXR (**Supplementary Figure 2**) was used in two distinct enzymatic assays to check its functionality. To verify whether AcDCXR is able to oxidize xylitol to xylulose, AcDCXR was assayed using xylitol as the substrate and NADP<sup>+</sup> as co-substrate. The reduction of NADP<sup>+</sup> to NADPH was spectroscopically monitored by the increase of absorbance at 340 nm. AcDCXR was able to oxidize xylitol to xylulose in a NADP<sup>+</sup> concentration-dependent

manner (**Figure 3A**). Analysis of the Lineweaver-Burke plot data determined the enzymatic constants to be:  $k_{cat} = 1.85 \text{ s}^{-1}$ , a  $K_m = 0.56 \text{ mM}$  and a  $V_{max} = 79.4 \text{ } \mu\text{M}/\text{min}$  (**Figure 3B**).

To determine whether AcDCXR was able to use methylglyoxal as a substrate, we tested the reduction of methylglyoxal by spectroscopically measuring the decrease in absorption of concomitant NADPH oxidation at 340 nm. We found that AcDCXR was able to reduce methylglyoxal in a concentration-dependent manner (**Figure 4A**). Analysis of the Lineweaver-Burke plot data determined the enzymatic constants to be:  $k_{cat} = 0.23 \text{ s}^{-1}$ , a  $K_m = 1.3 \text{ mM}$  and a  $V_{max} = 13.8 \text{ } \mu\text{M}/\text{min}$  (**Figure 4B**). The control reactions, in the presence of AcDCXR and absence of a substrate, showed neither oxidation nor reduction (**Figures 3A, 4A**). Similarly, the control reactions in the absence of the enzyme showed neither oxidation nor reduction, indicating the AcDCXR's presence was necessary to complete the reactions (**Figures 3A, 4A**). The kinetic constants in AcDCXR show that, *in vitro*, it was more efficient oxidizing xylitol with a  $k_{cat}/K_m$  of  $3.32 \text{ mM}^{-1} \text{ s}^{-1}$  compared to reducing methylglyoxal that had only a  $k_{cat}/K_m$  of  $0.174 \text{ mM}^{-1} \text{ s}^{-1}$ , nearly a 20-fold difference in activity.

### Functional Analysis of AcDCXR *in planta*

To functionally evaluate the role of AcDCXR on cowpea aphid colonization, AcDCXR was cloned into the binary vector pEAQ-DEST1 for Agrobacterium-mediated transient expression. Since Agrobacterium-mediated transient expression in cowpea has not yet been developed, pea plants were used for this experiment. Pea is a host for cowpea aphid and has been previously used successfully in transient expression experiments for evaluation of aphid effectors (Guy et al., 2016). Using the same cultivar of pea cv ZP1130, we first transiently expression GFP using

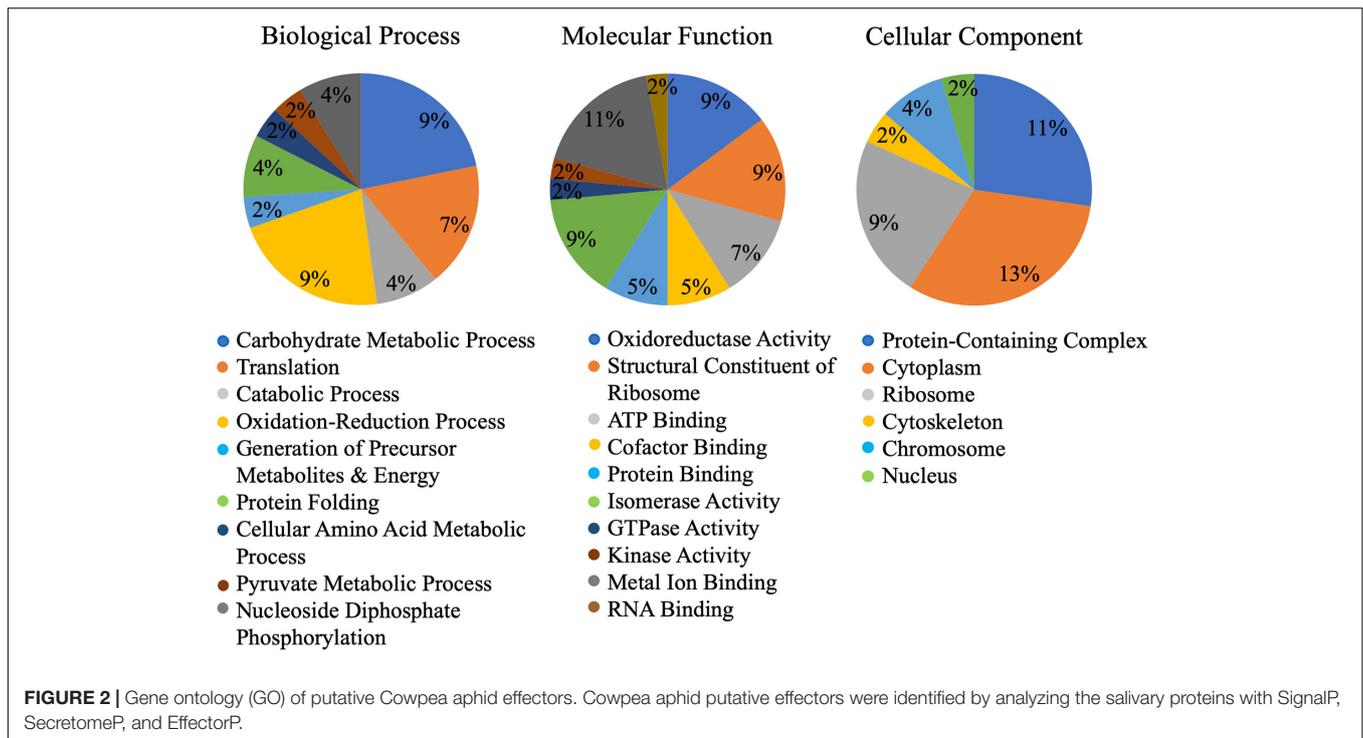
**TABLE 1** | Cowpea aphid salivary proteins identified for secretion or effector function using bioinformatic programs.

Accession	ACYPI	Description (BLASTP)	SignalP	SecretomeP	EffectorP
GAJW01000730.1_3	009769	glyceraldehyde-3-phosphate dehydrogenase [ <i>Aphis gossypii</i> ]			X
AG009127-PA	088094	histone H4-like, partial [ <i>Sipha flava</i> ]		X	X
NP_240299.1	085083	protein lin-28 homolog isoform X1 [ <i>Sipha flava</i> ]			X
AG010231-PA	23752	carbonic anhydrase 2-like [ <i>Aphis gossypii</i> ]	X		X
GAJW01000939.1_3	006727	triosephosphate isomerase [ <i>Aphis gossypii</i> ]			X
GAJW01000401.1_4	000057	putative diacetyl/L-xylulose reductase [ <i>Aphis citricidus</i> ]	X		X
Rpa07060.t1-protein	000028	nucleoside diphosphate kinase-like isoform X1 [ <i>Rhopalosiphum maidis</i> ]			X
Rpa13763.t1-protein	004622	probable pseudouridine-5'-phosphatase [ <i>Rhopalosiphum maidis</i> ]			X
GAJW01000399.1_1	001643	ribosomal protein S19e-like [ <i>Acyrtosiphon pisum</i> ]		X	X
NP_239859.1	000693	co-chaperonin GroES [ <i>Buchnera aphidicola</i> str. APS ( <i>Acyrtosiphon pisum</i> )]		X	X
GAJW01002612.1_4	002607	ras-related protein Rab-7a-like [ <i>Rhopalosiphum maidis</i> ]			X
Rpa11900.t1-protein	003541	peptidyl-prolyl cis-trans isomerase-like [ <i>Aphis gossypii</i> ]			X
GAJW01001625.1_4	008224	Uncharacterized protein LOC114123729 [ <i>Aphis gossypii</i> ] (Me10/Mp58)	X		X
GAJW01005083.1_5	006909	dihydropyridine reductase [ <i>Aphis gossypii</i> ]			X
GAJW01000315.1_5	000041	cytochrome c-like [ <i>Aphis gossypii</i> ]			X
GAJW01000773.1_5	000058	twinstar [ <i>Acyrtosiphon pisum</i> ]			X
GAJW01001573.1_3	007471	Superoxide dismutase [Cu-Zn] [ <i>Aphis gossypii</i> ]		X	X
AG013923-PA	006002	protein dj-1beta-like isoform X2 [ <i>Melanaphis sacchari</i> ]			X
AG015946-PA	26018	uncharacterized protein LOC114132136 [ <i>Aphis gossypii</i> ]		X	X
GAJW01000973.1_1	000031	ribosomal protein S12 [ <i>Acyrtosiphon pisum</i> ]			X
AG001995-PA	000422	apolipoproteins-like [ <i>Aphis gossypii</i> ]	X		
AG007466-PA	002298	soluble trehalase [ <i>Aphis glycines</i> ]	X		
AG008787-PA	000288	glucose dehydrogenase [FAD, quinone]-like [ <i>Aphis gossypii</i> ]	X		
AG009504-PA	005582	glucose dehydrogenase [FAD, quinone]-like [ <i>Melanaphis sacchari</i> ]	X		
Rpa18947.t1-protein	009964	endochitinase [ <i>Rhopalosiphum maidis</i> ]	X		
AG005369-PA	000986	uncharacterized protein LOC114133079 [ <i>Aphis gossypii</i> ]	X		
AG011655-PA	009915	endoplasmic reticulum chaperone [ <i>Aphis gossypii</i> ]	X		
AG005025-PA	002622	calreticulin [ <i>Aphis gossypii</i> ]	X		
GAJW01001996.1_3	009755	protein disulfide-isomerase [ <i>Melanaphis sacchari</i> ]	X		

(Continued)

TABLE 1 | Continued

Accession	ACYPI	Description (BLASTP)	SignalP	SecretomeP	EffectorP
Rpa07735.t1-protein	37407	uncharacterized protein LOC113549759 [ <i>Rhopalosiphum maidis</i> ]	X		
Rpa04366.t1-protein	005594	protein disulfide-isomerase A3-like [ <i>Rhopalosiphum maidis</i> ]	X		
AG011982-PA	004904	uncharacterized protein LOC114121223 [ <i>Aphis gossypii</i> ]	X		
AG008256-PA	004432	uncharacterized protein LOC114132390 [ <i>Aphis gossypii</i> ]	X		
Rpa01411.t1-protein	008926	protein disulfide-isomerase A6 homolog [ <i>Rhopalosiphum maidis</i> ]	X		
AG011980-PA	25151	uncharacterized protein LOC107882155 [ <i>Acyrtosiphon pisum</i> ]	X		
AG001228-PA	001479	phospholipase A1-like [ <i>Aphis gossypii</i> ]	X		
AG014046-PA	000975	calumenin-A-like [ <i>Aphis gossypii</i> ]	X		
AG001716-PA	007690	uncharacterized protein LOC100168851 precursor [ <i>Acyrtosiphon pisum</i> ]	X		
AG000496-PA	50398	RCC1 domain-containing protein 1 [ <i>Aphis gossypii</i> ]	X		
AG002007-PA	003669	puromycin-sensitive aminopeptidase isoform X2 [ <i>Aphis gossypii</i> ]	X		
AG012271-PA	000817	peroxidase-like [ <i>Aphis gossypii</i> ]	X		
GAJW01000521.1_1	008165	60S ribosomal protein L4 [ <i>Rhopalosiphum maidis</i> ]	X		X
GAJW01002804.1_6	000056	calmodulin [ <i>Bombyx mori</i> ]			X
MYZPE13164_0_v1.0_000072220.1	086010	histone H2B [ <i>Myzus persicae</i> ]			X
AG009875-PA	007671	histone H3-like [ <i>Diuraphis noxia</i> ]			X
AG015809-PA	003625	citrate synthase 1 [ <i>Aphis gossypii</i> ]			X
Rpa01379.t1-protein	007342	annexin B9-like isoform X2 [ <i>Rhopalosiphum maidis</i> ]			X
GAJW01000299.1_6	002483	eukaryotic initiation factor 4A [ <i>Acyrtosiphon pisum</i> ]			X
GAOM01000916.1_4	000087	bicaudal [ <i>Acyrtosiphon pisum</i> ]			X
BAH71147.1	010042	40S ribosomal protein S7-like [ <i>Acyrtosiphon pisum</i> ]			X
AG001482-PA	002577	adenylate kinase isoenzyme 1 [ <i>Aphis gossypii</i> ]			X
Mca25739.t1-protein	006705	ump-cmp kinase isoform X2 [ <i>Myzus persicae</i> ]			X
AG001456-PA	009018	inositol monophosphatase 1-like [ <i>Aphis gossypii</i> ]			X
AG004283-PA	21663	lipase member H-B-like isoform X1 [ <i>Aphis gossypii</i> ]			X
GAJW01001081.1_4	004620	complement component 1 Q subcomponent-binding protein, mitochondrial [ <i>Aphis gossypii</i> ]			X
AG012940-PA	006879	importin-5 [ <i>Aphis gossypii</i> ]			X
GAJW01001854.1_3	009959	protein lethal(2)essential for life [ <i>Aphis gossypii</i> ]			X
Mca14412.t1-protein	000100	40s ribosomal protein s4 [ <i>Diuraphis noxia</i> ]			X



*A. tumefaciens* strain AGL01. Monitoring GFP expression by western blot analysis, GFP was detected as early as 2 days after agroinfiltration and lasted at least for 10 days (**Supplementary Figure 3**). Based on the GFP expression in pea, a cowpea aphid bioassay was developed.

Aphid bioassays were performed to evaluate the effect of AcDCXR overexpression in pea plants on cowpea aphid. Plants were agroinfiltrated with AcDCXR or GFP control constructs as described earlier for the western blot analysis. A day post infiltration (dpi), adult cowpea aphids, maintained on pea cv ZP1130, were placed on a leaf, at the site of the infiltration, in a clip cage. After 24 h (2 dpi), all adult aphids were removed and six newborn nymphs were left at the infiltration site. At ten dpi, similar number of aphids were counted on GFP and AcDCXR infiltrated leaves indicating no effect on nymph survival rate (GLM,  $\text{Chisq} = 0.034$ ,  $P = 0.854$ ) (**Figure 5A**). To evaluate the fecundity of these aphids, one aphid per cage was transferred to a freshly agroinfiltrated (2 dpi) plant, with the same construct, and aphid survival and fecundity was monitored 4 and 7 days later. Sixteen days after initiation of the aphid bioassay, no difference in adult survival was detected between aphids feeding on AcDCXR compared to those feeding on the GFP infiltrated leaves (GLM,  $\text{Chisq} = 0.367$ ,  $P = 0.544$ ) (**Figure 5B**). However, a significant difference (GLM,  $\text{Chisq} = 16.901$ ,  $P < 0.001$ ) in aphid fecundity was observed between the aphids feeding on AcDCXR compared to those feeding on the GFP control indicating a role for AcDCXR in cowpea aphid colonization (**Figure 5C**). To determine the subcellular localization of AcDCXR *in planta*, AcDCXR was cloned into the binary vector pCAMBIA-1300-mScarlet and used in Agrobacterium-mediated transient expression in *N. benthamiana*. pCAMBIA-1300-AcDCXR-mScarlet was

co-infiltrated with a GFP construct. As expected, GFP was detected throughout the cell including the nucleus, while AcDCXR-mScarlet was localized to the cytoplasm (**Figure 6**).

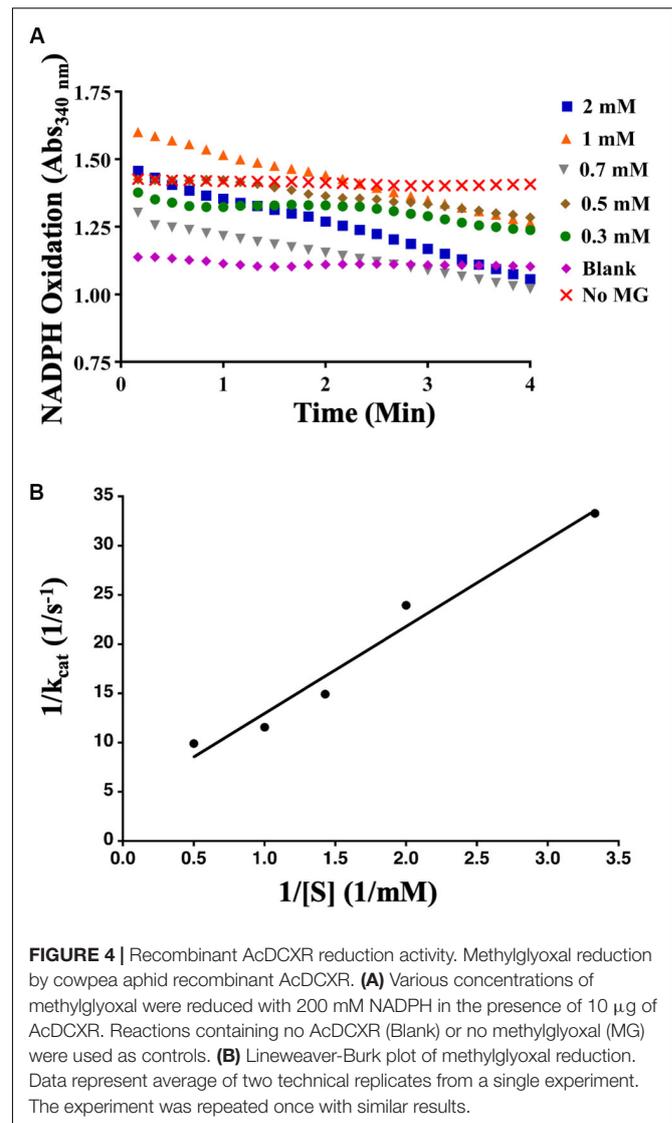
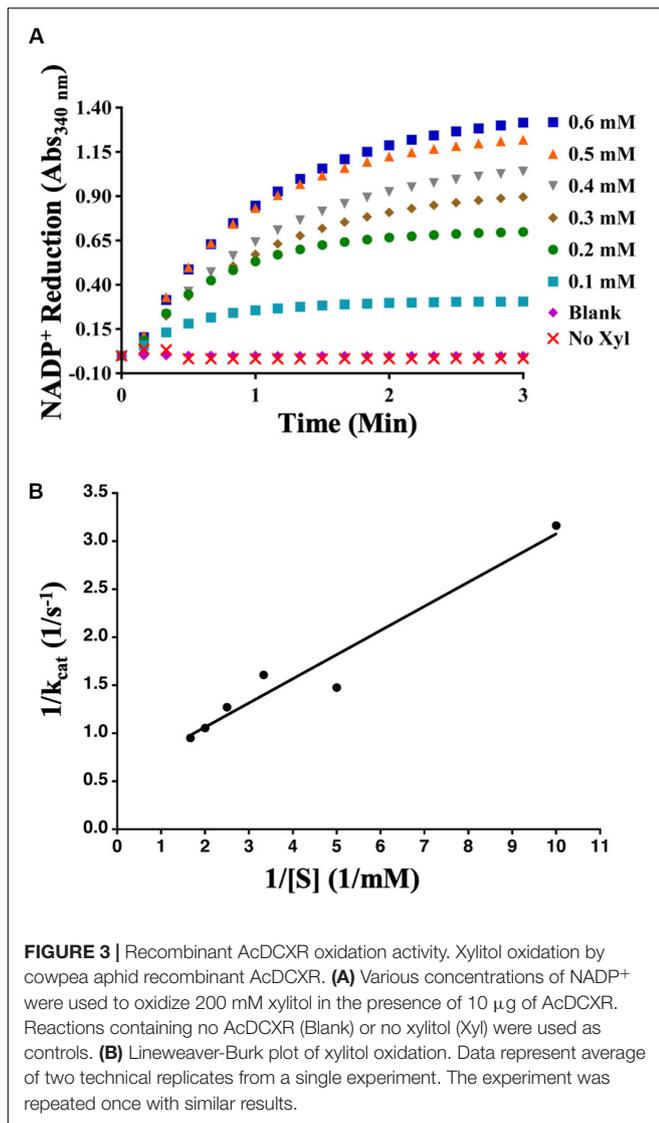
## Aphid Induce Methylglyoxal Accumulation

Methylglyoxal has been shown to accumulate in multiple plant species when exposed to abiotic stresses (Yadav et al., 2005a; Hossain et al., 2009; Mustafiz et al., 2014). Recently, it was also shown that methylglyoxal accumulates in plants exposed to biotic stresses (Melvin et al., 2017). To assess if methylglyoxal also accumulates by aphid infestation, methylglyoxal levels in cowpea and pea plants were monitored. A day after infestation of cowpea plants to cowpea aphids, a significantly higher (multiple comparisons,  $z = 2.812$ ,  $P = 0.015$ ) levels of methylglyoxal were detected in the infested leaves compared to the uninfested control leaves (**Figure 7A**). Methylglyoxal levels remained significantly higher (multiple comparisons,  $z = 3.832$ ,  $P < 0.001$ ) on day 2 but reduced to pre-infective levels on day 3 (multiple comparisons,  $z = 1.479$ ,  $P = 0.208$ ) (**Figure 7A**). A similar trend of methylglyoxal accumulation was detected in pea leaves exposed to cowpea aphids indicating that cowpea aphid feeding induces methylglyoxal levels irrespective of the host species (**Figure 7B**).

## DISCUSSION

### Cowpea Aphid Salivary Proteome

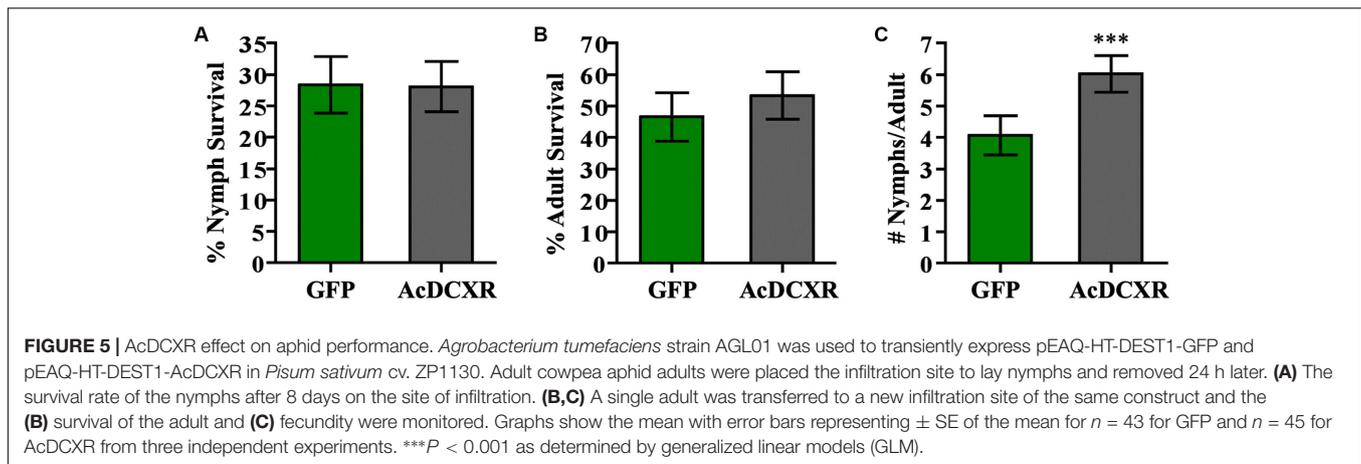
We carried out proteomics analysis to identify the salivary protein composition of a population of cowpea aphid from California. The identified proteins had a diverse range of



functions including some that are uncharacterized. We were conservative in assessing the salivary proteome and used strict cut-off measures to identify the proteins. Nevertheless, we identified 149 non-redundant proteins. Previously, the salivary proteome from an African cowpea aphid population was reported (Loudit et al., 2018). The majority of the proteins identified in our study were not reported from this African population suggesting that our approach allowed us to identify higher numbers of proteins. While the cowpea aphid saliva in this work was collected in water, the African cowpea aphid saliva was collected in a sucrose-based diet and required clean up steps before undergoing mass spectrometry and that could have contributed to the low number of proteins identified in the saliva. Interestingly, both studies did not identify a set of functionally characterized aphid effectors such as Armet, Me23, Ap25, Mp2, Mp55 (Atamian et al., 2013; Pitino and Hogenhout, 2013; Elzinga et al., 2014; Wang et al., 2015; Guy et al., 2016). While in our study we identified Me10/Mp58 and SHP, the structural sheath protein,

these two proteins were not identified in the African cowpea aphid saliva (Carolan et al., 2009; Chaudhary et al., 2015). The well characterized effector C002, was reported in the African population and not in this work (Mutti et al., 2006, 2008; Pitino and Hogenhout, 2013; Elzinga et al., 2014; Loudit et al., 2018). Although peptides for C002 and two additional effectors, Mp1 and MIF1, were detected in the saliva of the California cowpea aphids, this work, they did not fulfil the criteria used in our selection (Harmel et al., 2008; Naessens et al., 2015).

Unlike the salivary proteome of the African cowpea aphid, there were no proteins identified from secondary symbionts in the California cowpea aphid saliva (Loudit et al., 2018). The only bacterial proteins identified in the California cowpea aphid salivary proteome were from the primary endosymbiont *Buchnera aphidicola*, the chaperonin GroEL and GroES. GroEL has been previously identified in the saliva of several aphid species including the cowpea aphid (Chaudhary et al., 2014, 2015; Vandermoten et al., 2014; Loudit et al., 2018). GroEL is an



aphid-associated molecular pattern triggering immune responses in plants (Chaudhary et al., 2014).

Our work was limited by the absence of a cowpea aphid genome and a gland/head specific transcriptome that could have been used for the peptide searches. In addition, homologous sequences from different aphid species were used in the secretion prediction analyses including some originating from transcriptomes that could have been truncated. Therefore, the number of proteins predicted for secretion, 46 out of 149 (30.9%), based on the bioinformatic programs SignalP and SecretomeP, are likely an underestimate (**Table 1**). Previous work describing salivary proteome from aphids with genome sequences and gland/head specific RNAseq generated sequences, also identified a large number of proteins from aphid saliva, collected in sugar and amino acid-based diets, with no prediction for secretion (Thorpe et al., 2016; Boulain et al., 2018). Boulain et al. (2018) reported 37/51 (72.5%) of the pea aphid salivary proteins with a secretion prediction. Thorpe et al. (2016), studying three different aphid species, green peach aphid, black cherry aphid, and bird cherry-oat aphid, reported only 61/204 (30%) secretion prediction of the identified salivary proteins. Taken together, this information indicates that the current bioinformatic prediction programs are likely limited in their ability to identify aphid secreted proteins.

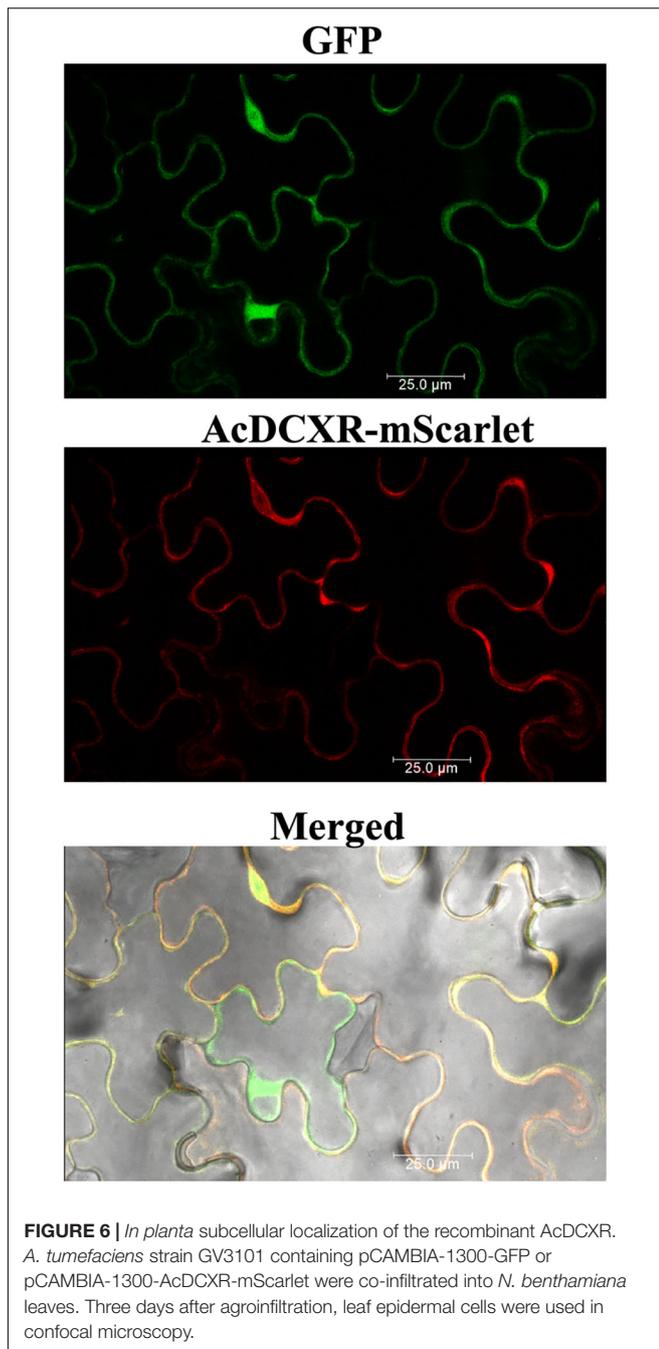
## Effector Prediction

Here we reported the use of a machine learning plant-pathogenic fungi effector prediction program, EffectorP, for prediction of aphid effectors (Sperschneider et al., 2016, 2018). We confirmed the use of EffectorP as a possible program for identifying aphid effector proteins by successfully subjecting the well-characterized aphid effectors C002 and Me10 to EffectorP analysis (Mutti et al., 2008; Atamian et al., 2013; Pitino and Hogenhout, 2013; Chaudhary et al., 2019). Interestingly, EffectorP predicted 20/149 of the cowpea aphid proteins as effectors. Among these 20 proteins, is the functionally characterized Me10 effector and three proteins which have been predicted for effector function (Atamian et al., 2013; Elzinga et al., 2014; Thorpe et al., 2016; Chaudhary et al., 2019). Orthologs of Me10 have been identified in multiple aphid species. Me10 has been detected in plant

tissues fed on by aphids and expression of Me10 in plants has been shown to enhance the performance of potato aphid on tomato and green peach aphid on *N. benthamiana* (Atamian et al., 2013; Chaudhary et al., 2015, 2019). In addition, Me10 was shown to interact with the tomato scaffold protein Fourteen-Three-Three isoform 7 (TFT7) and predicted to interfere with a mitogen-activated protein kinase defense signaling pathway (Chaudhary et al., 2019).

The remaining three previously predicted putative effectors are carbonic anhydrase, superoxide dismutase, and peptidyl-prolyl cis-trans isomerase (PPIase). The latter two proteins were identified in the proteomes of the pea aphid salivary glands (Carolan et al., 2011). While carbonic anhydrases have been identified in aphid saliva, superoxide dismutase and PPIase have not been previously reported in aphid saliva (Rao et al., 2013; Nicholson and Puterka, 2014; Chaudhary et al., 2015; Loudit et al., 2018). A carbonic anhydrase and a superoxide dismutase have been shown to be under positive selection further implicating these proteins as effectors (Thorpe et al., 2016). While clear roles for carbonic anhydrases and PPIases have not been characterized in plant immune responses, superoxide dismutases are attributed to detoxify reactive oxygen species (ROS), the well-known defense signaling molecule.

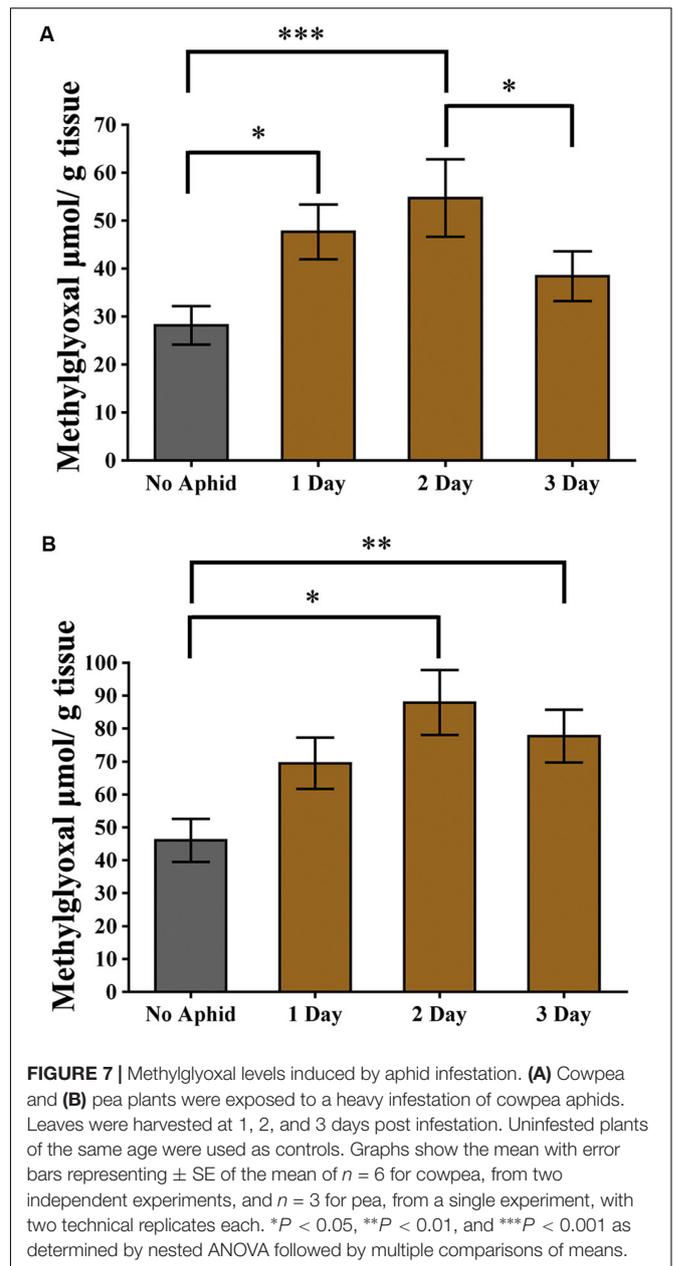
Among the EffectorP identified putative effector proteins, that had not been previously identified in aphid saliva or as a putative effector, is AcDCXR (**Table 1, Supplementary Table 1**). DCXR has been identified in the pea aphid salivary gland but has not been reported in the saliva of this aphid species (Carolan et al., 2011; Boulain et al., 2018). Interestingly, pea aphid homolog of AcDCXR as well as homologs from five additional aphid species with genome sequences, do not have a secretion signal peptide. The homolog from the cotton melon aphid does have a secretion signal suggesting that DCXR is one of the differential pest arsenals utilized by a subset of aphid species. An increase in DCXR accumulation was reported in a virulent biotype of greenbug infesting resistant wheat (Pinheiro et al., 2014). Additionally, enhanced accumulation of DCXR in response to heat/UV stress as well as predation by parasitoids in the potato aphid were reported from whole insects (Nguyen et al., 2008, 2009). Taken together, these information suggest that aphids may have evolved



different roles for DCXR to deal with stress conditions in the plant and within the aphid itself.

### Diacetyl/L-Xylulose Reductases

In mammals DCXRs are reported to be oxidoreductases for monosaccharides and dicarbonyls. Human DCXR was first discovered while investigating the disease pentosuria and found that an enzymatic defect in DCXR was the cause of the high excretion of L-xylulose. This led to the conclusion that L-xylulose is a possible substrate of DCXR



(Wang and Van Eys, 1970). DCXR has been shown also to catalyze reactions with other sugars. For example, xylitol is a sugar alcohol that is transported through the phloem as a carbon source (Lewis and Smith, 1967; Lemoine et al., 2013). Xylitol can be converted to xylulose and be used in the pentose phosphate pathway to generate glycolytic intermediates as a source of energy. Since the AcDCXR catalyzes the reversible reaction between xylulose and xylitol, the enzyme may provide the aphid an additional mode of generating energy.

Diacetyl/L-xylulose reductases also participates in the reductive metabolism of carbonyls. In this role, the enzyme is considered as a defense mechanism against harmful carbonyls (Nakagawa et al., 2002; Ebert et al., 2015; Yang et al., 2017). These

molecules lead to formation of AGEs by reacting with lysine, cysteine and arginine, thus inactivating proteins (Thornalley, 2006; Ahmed and Thornalley, 2007). One of these harmful carbonyls is methylglyoxal which is reactive  $\alpha$ - $\beta$ -dicarbonyl ketoaldehyde. Interestingly, methylglyoxal has been shown to accumulate in a number of plant species under various abiotic stresses (Yadav et al., 2005a; Hossain et al., 2009; Mustafiz et al., 2014; Rahman et al., 2015; Borysiuk et al., 2018). Recently, methylglyoxal has also been implicated in biotic stresses. Increases in methylglyoxal levels were detected in tobacco plants exposed to the bacterium *Pseudomonas syringae*, or the *Mungbean yellow mosaic virus*, or to the fungus *Alternaria alternata* (Melvin et al., 2017). In addition, exogenous application of methylglyoxal in wheat and rice plants upregulated antioxidant and defense-related genes indicating a role for methylglyoxal in plant defense (Kaur et al., 2015; Li et al., 2017). In this work we showed that aphid feeding also enhanced accumulation of methylglyoxal in cowpea and pea, suggesting methylglyoxal also functions in aphid defense. Since methylglyoxal levels in aphid infested leaves were mostly transient, this suggests that aphids are able to counteract methylglyoxal accumulation possibly through AcDCXR activity.

Transient expression of AcDCXR indicates that this enzyme is localized in the plant cell cytoplasm. Likewise, both AcDCXR substrates tested in this study, methylglyoxal and xylitol/xylulose, are also located in the cell cytoplasm. In plants, the pentose phosphate pathway where xylitol/xylulose are used, takes place in both the cytoplasm and plastids. Methylglyoxal is generated in multiple pathways in the cytoplasm and in various organelles (Phillips and Thornalley, 1993; Dennis and Blakeley, 2000; Kruger and von Schaewen, 2003).

The transient expression of AcDCXR increased the fecundity of the cowpea aphid most likely due to its effect on one or both of these two substrates; either by increasing the obtained nutrient content and/or through diminishing defense responses. This increase in fecundity was seen despite no differences in the survival of both adult and nymphal stages of the aphid. Transient or stable overexpression of a number of aphid effectors in various plant species including, Arabidopsis, tomato, pea and *N. benthamiana* also yielded increases in aphid fecundity but no effect on aphid survival suggesting that overexpression of multiple effectors may be needed to observe a pronounced change in aphid survival.

In this work, using a classical and a novel bioinformatics programs, SignalP and EffectorP, respectively, we identified a

novel aphid effector, AcDCXR. The functional annotation of DCXR and *in vitro* biochemical analysis of AcDCXR lead us to identify methylglyoxal as a potential novel metabolite involved in aphid defense. Therefore, identification of novel effectors may lead to the discovery of yet unknown defense pathways that may lead to novel approaches to engineer pest/pathogen resistance in crops.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in MassIVE, <https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp>, ID: PXD017323

## AUTHOR CONTRIBUTIONS

IK conceived the project. IK, SD, and AS designed the experiments. JM conducted the experiments. IK, JM, SD, and QC analyzed the data. IK and JM wrote the manuscript with help from all other authors.

## FUNDING

This research is partly funded by the USDA National Institute of Food and Agriculture Hatch project 1017522 to IK.

## ACKNOWLEDGMENTS

We thank Dr. Monika Ostaszewska-Bugajska (University of Warsaw) for sharing methylglyoxal evaluation protocol, Dr. Philip Roberts (UC Riverside) for the cowpea aphids, and Dr. Jiangman He (UC Riverside) for help with confocal microscopy and constructing the pCAMBIA-1300-GW-mScarlet vector.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.00605/full#supplementary-material>

## REFERENCES

- Abdellatef, E., Will, T., Koch, A., Imani, J., Vilcinskas, A., and Kogel, K. H. (2015). Silencing the expression of the salivary sheath protein causes transgenerational feeding suppression in the aphid *Sitobion avenae*. *Plant Biotechnol. J.* 13, 849–857. doi: 10.1111/pbi.12322
- Agunbiade, T. A., Sun, W., Coates, B. S., Djouaka, R., Tamo, M., Ba, M. N., et al. (2013). Development of reference transcriptomes for the major field insect pests of cowpea: a toolbox for insect pest management approaches in west Africa. *PLoS One* 8:e79929. doi: 10.1371/journal.pone.0079929
- Ahmed, N., and Thornalley, P. J. (2007). Advanced glycation endproducts: what is their relevance to diabetic complications? *Diabetes Obes. Metab.* 9, 233–245. doi: 10.1111/j.1463-1326.2006.00595.x
- Armenteros, J. J. A., Tsirigos, K. D., Sonderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. doi: 10.1038/s41587-019-0036-z
- Atamian, H. S., Chaudhary, R., Dal Cin, V., Bao, E., Girke, T., and Kaloshian, I. (2013). In planta expression or delivery of potato aphid *Macrosiphum euphorbiae* effectors Me10 and Me23 enhances aphid fecundity. *Mol. Plant Microbe Interact.* 26, 67–74. doi: 10.1094/Mpmi-06-12-0144-Fi

- Bendtsen, J. D., Jensen, L. J., Blom, N., von Heijne, G., and Brunak, S. (2004a). Feature based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel.* 17, 349–356. doi: 10.1093/protein/gzh037
- Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004b). Improved prediction of signal peptides: signalp 3.0. *J. Mol. Biol.* 340, 783–795. doi: 10.1016/j.jmb.2004.05.028
- BIPAA (2017). *AphidBase*. Available: <https://bipaa.genouest.org/is/aphidbase/> (accessed June 27, 2017).
- Blackman, R. L., and Eastop, V. F. (2000). *Aphids on the World's Crops*. New York, NY: John Wiley & Sons Inc.
- Borysiuk, K., Ostaszewska-Bugajska, M., Vaultier, M. N., Hasenfratz-Sauder, M. P., and Szal, B. (2018). Enhanced formation of methylglyoxal-derived advanced glycation end products in *Arabidopsis* under ammonium nutrition. *Front. Plant Sci.* 9:667. doi: 10.3389/fpls.2018.00667
- Bos, J. I. B., Prince, D., Pitino, M., Maffei, M. E., Win, J., and Hogenhout, S. A. (2010). A functional genomics approach identifies candidate effectors from the aphid species *Myzus persicae* (green peach aphid). *PLoS Genet.* 6:e1001216. doi: 10.1371/journal.pgen.1001216
- Boulain, H., Legeai, F., Guy, E., Morliere, S., Douglas, N. E., Oh, J., et al. (2018). Fast evolution and lineage-specific gene family expansions of aphid salivary effectors driven by interactions with host-plants. *Genome Biol. Evol.* 10, 1554–1572. doi: 10.1093/gbe/evy097
- Calandra, T. (2003). Macrophage migration inhibitory factor and host innate immune responses to microbes. *Scand. J. Infect. Dis.* 35, 573–576. doi: 10.1080/00365540310016277
- Carolan, J. C., Caragea, D., Reardon, K. T., Mutti, N. S., Dittmer, N., Pappan, K., et al. (2011). Predicted effector molecules in the salivary secretome of the pea aphid (*Acyrtosiphon pisum*): a dual transcriptomic/proteomic approach. *J. Proteome Res.* 10, 1505–1518. doi: 10.1021/pr100881q
- Carolan, J. C., Fitzroy, C. I. J., Ashton, P. D., Douglas, A. E., and Wilkinson, T. L. (2009). The secreted salivary proteome of the pea aphid *Acyrtosiphon pisum* characterised by mass spectrometry. *Proteomics* 9, 2457–2467. doi: 10.1002/pmic.200800692
- Chan, C. K., Forbes, A. R., and Raworth, D. A. (1991). *Aphid-Transmitted Viruses and their Vectors of the World*. Ottawa: Agriculture and Agri-Food Canada.
- Chaudhary, R., Atamian, H. S., Shen, Z., Briggs, S. P., and Kaloshian, I. (2014). GroEL from the endosymbiont *Buchnera aphidicola* betrays the aphid by triggering plant defense. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8919–8924. doi: 10.1073/pnas.1407687111
- Chaudhary, R., Atamian, H. S., Shen, Z., Briggs, S. P., and Kaloshian, I. (2015). Potato aphid salivary proteome: enhanced salivation using resorcinol and identification of aphid phosphoproteins. *J. Proteome Res.* 14, 1762–1778. doi: 10.1021/pr501128k
- Chaudhary, R., Peng, H. C., He, J. M., MacWilliams, J., Teixeira, M., Tsuchiya, T., et al. (2019). Aphid effector Me10 interacts with tomato TFT7, a 14-3-3 isoform involved in aphid resistance. *New Phytol.* 221, 1518–1528. doi: 10.1111/nph.15475
- Chen, L., Wang, T., Wang, X., Sun, B. B., Li, J. Q., Liu, D. S., et al. (2009). Blockade of advanced glycation end product formation attenuates bleomycin-induced pulmonary fibrosis in rats. *Respir. Res.* 10:55. doi: 10.1186/1465-9921-10-55
- Choudhary, A., Hussein, A., Samota, R. G., and Nehra, S. (2017). Effect of biotic and abiotic factors on the incidence of aphid, *Aphis craccivora* Koch on cowpea. *J. Pharmacogn. Phytochem.* 6, 1587–1590.
- Cooper, W. R., Dillwith, J. W., and Puterka, G. J. (2010). Salivary proteins of Russian wheat aphid (Hemiptera: Aphididae). *Environ. Entomol.* 39, 223–231. doi: 10.1603/En09079
- Cooper, W. R., Dillwith, J. W., and Puterka, G. J. (2011). Comparisons of salivary proteins from five aphid (Hemiptera: Aphididae) species. *Environ. Entomol.* 40, 151–156. doi: 10.1603/En10153
- Dennis, D. T., and Blakeley, S. D. (2000). “Carbohydrate metabolism,” in *Biochemistry & Molecular Biology of Plants*, eds B. B. Buchanan, W. Gruissem, and R. L. Jones (Waldorf, MD: American Society of Plant Physiologists), 630–675.
- Ebert, B., Kisiela, M., and Maser, E. (2015). Human DCXR - another ‘moonlighting protein’ involved in sugar metabolism, carbonyl detoxification, cell adhesion and male fertility? *Biol. Rev.* 90, 254–278. doi: 10.1111/brv.12108
- Elzinga, D. A., De Vos, M., and Jander, G. (2014). Suppression of plant defenses by a *Myzus persicae* (green peach aphid) salivary effector protein. *Mol. Plant Microbe Interact.* 27, 747–756. doi: 10.1094/Mpmi-01-14-0018-R
- Gkogkolou, P., and Bohm, M. (2012). Advanced glycation end products: key players in skin aging? *Dermatoendocrinol.* 4, 259–270. doi: 10.4161/derm.22028
- Gotz, S., Garcia-Gomez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., et al. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36, 3420–3435. doi: 10.1093/nar/gkn176
- Guy, E., Boulain, H., Aigu, Y., Le Pennec, C., Chawki, K., Morliere, S., et al. (2016). Optimization of agroinfiltration in *Pisum sativum* provides a new tool for studying the salivary protein functions in the pea aphid complex. *Front. Plant Sci.* 7:1171. doi: 10.3389/fpls.2016.01171
- Harmel, N., Letocart, E., Cherqui, A., Giordanengo, P., Mazzucchelli, G., Guillonneau, F., et al. (2008). Identification of aphid salivary proteins: a proteomic investigation of *Myzus persicae*. *Insect Mol. Biol.* 17, 165–174. doi: 10.1111/j.1365-2583.2008.00790.x
- Hoque, M. A., Uraji, M., Torii, A., Banu, M. N., Mori, I. C., Nakamura, Y., et al. (2012). Methylglyoxal inhibition of cytosolic ascorbate peroxidase from *Nicotiana tabacum*. *J. Biochem. Mol. Toxicol.* 26, 315–321. doi: 10.1002/jbt.21423
- Hoque, T. S., Hossain, M. A., Mostofa, M. G., Burritt, D. J., Fujita, M., and Tran, L. S. (2016). Methylglyoxal: an emerging signaling molecule in plant abiotic stress responses and tolerance. *Front. Plant Sci.* 7:1341. doi: 10.3389/fpls.2016.01341
- Hossain, M. A., Zakir, H. M., and Fujita, M. (2009). Stress-induced changes of methylglyoxal level and glyoxalase I activity in pumpkin seedlings and cDNA cloning of glyoxalase I gene. *J. Crop Sci.* 3:53.
- International Aphid Genomics Consortium (2010). Genome sequence of the pea aphid *Acyrtosiphon pisum*. *PLoS Biol.* 8:e1000313. doi: 10.1371/journal.pbio.1000313
- Ishikura, S., Usami, N., El-Kabbani, O., and Hara, A. (2003). Structural determinant for cold inactivation of rodent L-xylulose reductase. *Biochem. Biophys. Res. Comm.* 308, 68–72. doi: 10.1016/s0006-291x(03)01336-6
- Jackai, L., and Daoust, R. (1986). Insect pests of cowpeas. *Annu. Rev. Entomol.* 31, 95–119. doi: 10.1146/annurev.en.31.010186.000523
- Jiang, R. H., Tripathy, S., Govers, F., and Tyler, B. M. (2008). RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4874–4879. doi: 10.1073/pnas.0709303105
- Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031
- Kaloshian, I., and Walling, L. L. (2016). Hemipteran and dipteran pests: effectors and plant host immune regulators. *J. Integr. Plant Biol.* 58, 350–361. doi: 10.1111/jipb.12438
- Kamphuis, L. G., Gao, L., and Singh, K. B. (2012). Identification and characterization of resistance to cowpea aphid (*Aphis craccivora* Koch) in *Medicago truncatula*. *BMC Plant Biol.* 12:101. doi: 10.1186/1471-2229-12-101
- Kaur, C., Kushwaha, H. R., Mustafiz, A., Pareek, A., Sopory, S. K., and Singla-Pareek, S. L. (2015). Analysis of global gene expression profile of rice in response to methylglyoxal indicates its possible role as a stress signal molecule. *Front. Plant Sci.* 6:682. doi: 10.3389/fpls.2015.00682
- Kettles, G. J., and Kaloshian, I. (2016). The potato aphid salivary effector Me47 is a glutathione-S-transferase involved in modifying plant responses to aphid infestation. *Front. Plant Sci.* 7:1142. doi: 10.3389/fpls.2016.01142
- Kizer, J. R., Benkeser, D., Arnold, A. M., Ix, J. H., Mukamal, K. J., Djousse, L., et al. (2014). Advanced glycation/glycoxidation endproduct carboxymethyllysine and incidence of coronary heart disease and stroke in older adults. *Atherosclerosis* 235, 116–121. doi: 10.1016/j.atherosclerosis.2014.04.013
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315
- Kruger, N. J., and von Schaeuwen, A. (2003). The oxidative pentose phosphate pathway: structure and organisation. *Curr. Opin. Plant Biol.* 6, 236–246. doi: 10.1016/s1369-5266(03)00039-6

- Lemoine, R., La Camera, S., Atanassova, R., Dedaldechamp, F., Allario, T., Pourtau, N., et al. (2013). Source-to-sink transport of sugar and regulation by environmental factors. *Front. Plant Sci.* 4:272. doi: 10.3389/fpls.2013.00272
- Lewis, D. H., and Smith, D. C. (1967). Sugar alcohols (polyols) in fungi and green plants. *New Phytol.* 66, 185–204. doi: 10.1111/j.1469-8137.1967.tb05998.x
- Li, Z. G. (2016). Methylglyoxal and glyoxalase system in plants: old players, new concepts. *Bot. Rev.* 82, 183–203.
- Li, Z. G., Duan, X. Q., Min, X., and Zhou, Z. H. (2017). Methylglyoxal as a novel signal molecule induces the salt tolerance of wheat by regulating the glyoxalase system, the antioxidant system and osmolytes. *Protoplasma* 254, 1995–2006. doi: 10.1007/s00709-017-1094-z
- Lindholm, P., Peranen, J., Andressoo, J. O., Kalkkinen, N., Kokaia, Z., Lindvall, O., et al. (2008). MANF is widely expressed in mammalian tissues and differently regulated after ischemic and epileptic insults in rodent brain. *Mol. Cell. Neurosci.* 39, 356–371. doi: 10.1016/j.mcn.2008.07.016
- Lindholm, P., Voutilainen, M. H., Lauren, J., Peranen, J., Leppanen, V. M., Andressoo, J. O., et al. (2007). Novel neurotrophic factor CDNF protects and rescues midbrain dopamine neurons *in vivo*. *Nature* 448, 73–77. doi: 10.1038/nature05957
- Loudit, S. M. B., Bauwens, J., and Francis, F. (2018). Cowpea aphid-plant interactions: endosymbionts and related salivary protein patterns. *Entomol. Exp. Appl.* 166, 460–473. doi: 10.1111/eea.12687
- Mathers, T. C., Chen, Y., Kaithakottil, G., Legeai, F., Mugford, S. T., Baa-Puyoulet, P., et al. (2017). Rapid transcriptional plasticity of duplicated gene clusters enables a clonally reproducing aphid to colonise diverse plant species. *Genome Biol.* 18:27. doi: 10.1186/s13059-016-1145-3
- Melvin, P., Bankapalli, K., D'Silva, P., and Shivaprasad, P. V. (2017). Methylglyoxal detoxification by a DJ-1 family protein provides dual abiotic and biotic stress tolerance in transgenic plants. *Plant Mol. Biol.* 94, 381–397. doi: 10.1007/s11103-017-0613-9
- Miles, P. W. (1999). Aphid saliva. *Biol. Rev.* 74, 41–85. doi: 10.1017/S0006323198005271
- Misra, K., Banerjee, A. B., Ray, S., and Ray, M. (1996). Reduction of methylglyoxal in *Escherichia coli* K12 by an aldehyde reductase and alcohol dehydrogenase. *Mol. Cell. Biochem.* 156, 117–124. doi: 10.1007/bf00426333
- Mostofa, M. G., Ghosh, A., Li, Z. G., Siddiqui, M. N., Fujita, M., and Tran, L. P. (2018). Methylglyoxal - a signaling molecule in plant abiotic stress responses. *Free Radic. Biol. Med.* 122, 96–109. doi: 10.1016/j.freeradbiomed.2018.03.009
- Mustafiz, A., Ghosh, A., Tripathi, A. K., Kaur, C., Ganguly, A. K., Bhavesh, N. S., et al. (2014). A unique Ni<sup>2+</sup>-dependent and methylglyoxal-inducible rice glyoxalase I possesses a single active site and functions in abiotic stress response. *Plant J.* 78, 951–963. doi: 10.1111/tpj.12521
- Mutti, N. S., Louis, J., Pappan, L. K., Pappan, K., Begum, K., Chen, M. S., et al. (2008). A protein from the salivary glands of the pea aphid, *Acyrtosiphon pisum*, is essential in feeding on a host plant. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9965–9969. doi: 10.1073/pnas.0708958105
- Mutti, N. S., Park, Y., Reese, J. C., and Reeck, G. R. (2006). RNAi knockdown of a salivary transcript leading to lethality in the pea aphid, *Acyrtosiphon pisum*. *J. Insect Sci.* 6:38. doi: 10.1673/031.006.3801
- Naessens, E., Dubreuil, G., Giordanengo, P., Baron, O. L., Minet-Kebdani, N., Keller, H., et al. (2015). A secreted MIF cytokine enables aphid feeding and represses plant immune responses. *Curr. Biol.* 25, 1898–1903. doi: 10.1016/j.cub.2015.05.047
- Nakagawa, J., Ishikura, S., Asami, J., Isaji, T., Usami, N., Hara, A., et al. (2002). Molecular characterization of mammalian dicarbonyl/L-xylulose reductase and its localization in kidney. *J. Biol. Chem.* 277, 17883–17891. doi: 10.1074/jbc.M110703200
- Nguyen, T. T., Boudreault, S., Michaud, D., and Cloutier, C. (2008). Proteomes of the aphid *Macrosiphum euphorbiae* in its resistance and susceptibility responses to differently compatible parasitoids. *Insect Biochem. Mol. Biol.* 38, 730–739. doi: 10.1016/j.ibmb.2008.04.005
- Nguyen, T. T., Michaud, D., and Cloutier, C. (2009). A proteomic analysis of the aphid *Macrosiphum euphorbiae* under heat and radiation stress. *Insect Biochem. Mol. Biol.* 39, 20–30. doi: 10.1016/j.ibmb.2008.09.014
- Nicholson, S. J., Hartson, S. D., and Puterka, G. J. (2012). Proteomic analysis of secreted saliva from Russian Wheat Aphid (*Diuraphis noxia* Kurd.) biotypes that differ in virulence to wheat. *J. Proteomics* 75, 2252–2268. doi: 10.1016/j.jprot.2012.01.031
- Nicholson, S. J., Nickerson, M. L., Dean, M., Song, Y., Hoyt, P. R., Rhee, H., et al. (2015). The genome of *Diuraphis noxia*, a global aphid pest of small grains. *BMC Genomics* 16:429. doi: 10.1186/s12864-015-1525-1
- Nicholson, S. J., and Puterka, G. J. (2014). Variation in the salivary proteomes of differentially virulent greenbug (*Schizaphis graminum* Rondani) biotypes. *J. Proteomics* 105, 186–203. doi: 10.1016/j.jprot.2013.12.005
- Ofuya, T. (1995). Studies on the capability of *Cheilomenes lunata* (Fabricius) (Coleoptera: Coccinellidae) to prey on the cowpea aphid, *Aphis craccivora* Koch (Homoptera: Aphididae) in Nigeria. *Agric. Ecosyst. Environ.* 52, 35–38. doi: 10.1016/0167-8809(94)09006-S
- Palgi, M., Greco, D., Lindstrom, R., Auvinen, P., and Heino, T. I. (2012). Gene expression analysis of *Drosophila* *Manf* mutants reveals perturbations in membrane traffic and major metabolic changes. *BMC Genomics* 13:134. doi: 10.1186/1471-2164-13-134
- Palgi, M., Lindstrom, R., Peranen, J., Piepponen, T. P., Saarma, M., and Heino, T. I. (2009). Evidence that DmMANF is an invertebrate neurotrophic factor supporting dopaminergic neurons. *Proc. Natl. Acad. Sci. U.S.A.* 106, 2429–2434. doi: 10.1073/pnas.0810996106
- Petre, B., and Kamoun, S. (2014). How do filamentous pathogens deliver effector proteins into plant cells? *PLoS Biol.* 12:e1001801. doi: 10.1371/journal.pbio.1001801
- Phillips, S. A., and Thornalley, P. J. (1993). Formation of methylglyoxal and D-lactate in human red blood cells *in vitro*. *Biochem. Soc. Trans.* 21:163S. doi: 10.1042/bst021163s
- Pinheiro, P., Bereman, M. S., Burd, J., Pals, M., Armstrong, S., Howe, K. J., et al. (2014). Evidence of the biochemical basis of host virulence in the greenbug aphid, *Schizaphis graminum* (Homoptera: Aphididae). *J. Proteome Res.* 13, 2094–2108. doi: 10.1021/pr4012415
- Pitino, M., and Hogenhout, S. A. (2013). Aphid protein effectors promote aphid colonization in a plant species-specific manner. *Mol. Plant Microbe Interact.* 26, 130–139. doi: 10.1094/Mpmi-07-12-0172-Fi
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: <https://www.R-project.org/>
- Rahman, A., Mostofa, M. G., Alam, M. M., Nahar, K., Hasanuzzaman, M., and Fujita, M. (2015). Calcium mitigates arsenic toxicity in rice seedlings by reducing arsenic uptake and modulating the antioxidant defense and glyoxalase systems and stress markers. *Biomed Res. Int.* 2015:340812. doi: 10.1155/2015/340812
- Rao, S. A. K., Carolan, J. C., and Wilkinson, T. L. (2013). Proteomic profiling of cereal aphid saliva reveals both ubiquitous and adaptive secreted proteins. *PLoS ONE* 8:e57413. doi: 10.1371/journal.pone.0057413
- Ray, A., Ray, S., Mukhopadhyay, S., and Ray, M. (2013). Methylglyoxal with glycine or succinate enhances differentiation and shoot morphogenesis in *Nicotiana tabacum* callus. *Biol. Plant.* 57, 219–223.
- Remaudiere, G., and Remaudiere, M. (1997). *Catalogue des Aphididae du Monde (Homoptera Aphidoidea)*. Paris: Institut National de la Recherche Agronomique.
- Rodriguez, P. A., Escudero-Martinez, C., and Bos, J. I. B. (2017). An aphid effector targets trafficking protein VPS52 in a host-specific manner to promote virulence. *Plant Physiol.* 173, 1892–1903. doi: 10.1104/pp.16.01458
- Sainsbury, F., Thuenemann, E. C., and Lomonosoff, G. P. (2009). pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol. J.* 7, 682–693. doi: 10.1111/j.1467-7652.2009.00434.x
- Sankaranarayanan, S., Jamshed, M., and Samuel, M. A. (2015). Degradation of glyoxalase I in *Brassica napus* stigma leads to self-incompatibility response. *Nat. Plants* 1:15185. doi: 10.1038/nplants.2015.185
- Singh, B. B., Ehlers, J. D., Sharma, B., and Freire, F. R. (2002). "Recent progress in cowpea breeding," in *Proceedings of the World Cowpea Conference III*, eds C. A. Fatokun, B. B. Tarawali, B. B. Singh, P. M. Kormawa, and M. Tamo (Ibadan: IITA), 22–40.
- Sochor, M., Baquer, N. Z., and McLean, P. (1979). Glucose overutilization in diabetes: evidence from studies on the changes in hexokinase, the pentose phosphate pathway and glucuronate-xylulose pathway in rat kidney cortex in diabetes. *Biochem. Biophys. Res. Commun.* 86, 32–39. doi: 10.1016/0006-291x(79)90378-4

- Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998). "A hidden Markov model for predicting transmembrane helices in protein sequences," in *Proceedings of the 6th International Conference on Intelligent Systems for Molecular Biology*, Vol. 6, Montreal, 175–182.
- Sorenson, J. (2009). "Aphids," in *Encyclopedia of Insects*, 2nd Edn, eds V. Resh and R. Carde (San Diego, CA: Academic Press), 27–30.
- Sperschneider, J., Dodds, P., Gardiner, D., Singh, K. B., and Taylor, J. M. (2018). Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. *Mol. Plant Pathol.* 19, 2094–2110. doi: 10.1111/mpp.12682
- Sperschneider, J., Dodds, P. N., Gardiner, D. M., Manners, J. M., Singh, K. B., and Taylor, J. M. (2015). Advances and challenges in computational prediction of effectors from plant pathogenic fungi. *PLoS Pathog.* 11:e1004806. doi: 10.1371/journal.ppat.1004806
- Sperschneider, J., Gardiner, D. M., Dodds, P. N., Tini, F., Covarelli, L., Singh, K. B., et al. (2016). EffectorP: predicting fungal effector proteins from secretomes using machine learning. *New Phytol.* 210, 743–761. doi: 10.1111/nph.13794
- Stergiopoulos, I., and de Wit, P. J. (2009). Fungal effector proteins. *Annu. Rev. Phytopathol.* 47, 233–263. doi: 10.1146/annurev.phyto.112408.132637
- Thornalley, P. J. (2006). Advanced glycation end products in renal failure. *J. Ren. Nutr.* 16, 178–184. doi: 10.1053/j.jrn.2006.04.012
- Thorpe, P., Cock, P. J., and Bos, J. (2016). Comparative transcriptomics and proteomics of three different aphid species identifies core and diverse effector sets. *BMC Genomics* 17:172. doi: 10.1186/s12864-016-2496-6
- Thorpe, P., Escudero-Martinez, C. M., Cock, P. J. A., Eves-van den Akker, S., and Bos, J. I. B. (2018). Shared transcriptional control and disparate gain and loss of aphid parasitism genes. *Genome Biol. Evol.* 10, 2716–2733. doi: 10.1093/gbe/evy183
- Timko, M., and Singh, B. B. (2008). "Cowpea, a multifunctional legume," in *Genomics of Tropical Crop Plants*, eds P. H. Moore and R. Ming (New York, NY: Springer), 227–258.
- Tjallingii, W. F. (2006). Salivary secretions by aphids interacting with proteins of phloem wound responses. *J. Exp. Bot.* 57, 739–745. doi: 10.1093/jxb/erj088
- Tjallingii, W. F., and Esch, T. H. (1993). Fine-structure of aphid stylet routes in plant-tissues in correlation with EPG signals. *Physiol. Entomol.* 18, 317–328. doi: 10.1111/j.1365-3032.1993.tb00604.x
- Vandermoten, S., Harmel, N., Mazzucchelli, G., De Pauw, E., Haubruge, E., and Francis, F. (2014). Comparative analyses of salivary proteins from three aphid species. *Insect Mol. Biol.* 23, 67–77. doi: 10.1111/imb.12061
- Wang, W., Dai, H. E., Zhang, Y., Chandrasekar, R., Luo, L., Hiromasa, Y., et al. (2015). Armet is an effector protein mediating aphid-plant interactions. *FASEB J.* 29, 2032–2045. doi: 10.1096/fj.14-266023
- Wang, Y. M., and Van Eys, J. (1970). The enzymatic defect in essential pentosuria. *N. Engl. J. Med.* 282, 892–896. doi: 10.1056/NEJM197004162821604
- Wenger, J. A., Cassone, B. J., Legeai, F., Johnston, J. S., Bansal, R., Yates, A. D., et al. (2017). Whole genome sequence of the soybean aphid, *Aphis glycines*. *Insect Biochem. Mol. Biol.* doi: 10.1016/j.ibmb.2017.01.005 [Epub ahead of print].
- Wild, R., Ooi, L., Srikanth, V., and Munch, G. (2012). A quick, convenient and economical method for the reliable determination of methylglyoxal in millimolar concentrations: the N-acetyl-L-cysteine assay. *Anal. Bioanal. Chem.* 403, 2577–2581. doi: 10.1007/s00216-012-6086-4
- Will, T., Tjallingii, W. F., Thonnessen, A., and van Bel, A. J. E. (2007). Molecular sabotage of plant defense by aphid saliva. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10536–10541. doi: 10.1073/pnas.0703535104
- Will, T., and Vilcinskis, A. (2015). The structural sheath protein of aphids is required for phloem feeding. *Insect Biochem. Mol. Biol.* 57, 34–40. doi: 10.1016/j.ibmb.2014.12.005
- Yadav, S. K., Singla-Pareek, S. L., Ray, M., Reddy, M. K., and Sopory, S. K. (2005a). Methylglyoxal levels in plants under salinity stress are dependent on glyoxalase I and glutathione. *Biochem. Biophys. Res. Commun.* 337, 61–67. doi: 10.1016/j.bbrc.2005.08.263
- Yadav, S. K., Singla-Pareek, S. L., Reddy, M. K., and Sopory, S. K. (2005b). Transgenic tobacco plants overexpressing glyoxalase enzymes resist an increase in methylglyoxal and maintain higher reduced glutathione levels under salinity stress. *FEBS Lett.* 579, 6265–6271. doi: 10.1016/j.febslet.2005.10.006
- Yang, S., Jan, Y. H., Mishin, V., Heck, D. E., Laskin, D. L., and Laskin, J. D. (2017). Diacetyl/l-Xylulose reductase mediates chemical redox cycling in lung epithelial cells. *Chem. Res. Toxicol.* 30, 1406–1418. doi: 10.1021/acs.chemrestox.7b00052
- Yang, Z., Ma, L., Francis, F., Yang, Y., Chen, H., Wu, H., et al. (2018). Proteins identified from saliva and salivary glands of the Chinese gall aphid *Schlechtendalia chinensis*. *Proteomics* 18:e1700378. doi: 10.1002/pmic.201700378

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 MacWilliams, Dingwall, Chesnais, Sugio and Kaloshian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# BSA-Seq Discovery and Functional Analysis of Candidate Hessian Fly (*Mayetiola destructor*) Avirulence Genes

Lucio Navarro-Escalante<sup>1</sup>, Chaoyang Zhao<sup>2</sup>, Richard Shukle<sup>3†</sup> and Jeffrey Stuart<sup>4\*</sup>

<sup>1</sup> Department of Entomology, National Coffee Research Center, Manizales, Colombia, <sup>2</sup> Department of Botany and Plant Sciences, University of California, Riverside, Riverside, CA, United States, <sup>3</sup> USDA-ARS and Department of Entomology, Purdue University, West Lafayette, IN, United States, <sup>4</sup> Department of Entomology, Purdue University, West Lafayette, IN, United States

## OPEN ACCESS

### Edited by:

Isgouhi Kaloshian,  
University of California,  
Riverside, United States

### Reviewed by:

Georg Jander,  
Boyce Thompson Institute,  
United States  
Andrew Gloss,  
University of Chicago, United States

### \*Correspondence:

Jeffrey Stuart  
stuartj@purdue.edu

†Deceased (December 29, 2015)

### Specialty section:

This article was submitted to  
Plant Pathogen Interactions,  
a section of the journal  
Frontiers in Plant Science

Received: 21 October 2019

Accepted: 10 June 2020

Published: 25 June 2020

### Citation:

Navarro-Escalante L, Zhao C,  
Shukle R and Stuart J (2020) BSA-Seq  
Discovery and Functional Analysis of  
Candidate Hessian Fly (*Mayetiola  
destructor*) Avirulence Genes.  
Front. Plant Sci. 11:956.  
doi: 10.3389/fpls.2020.00956

The Hessian fly (HF, *Mayetiola destructor*) is a plant-galling parasite of wheat (*Triticum* spp.). Seven percent of its genome is composed of highly diversified signal-peptide-encoding genes that are transcribed in HF larval salivary glands. These observations suggest that they encode effector proteins that are injected into wheat cells to suppress basal wheat immunity and redirect wheat development towards gall formation. Genetic mapping has determined that mutations in four of these genes are associated with HF larval survival (virulence) on plants carrying four different resistance (*R*) genes. Here, this line of investigation was pursued further using bulked-segregant analysis combined with whole genome resequencing (BSA-seq). Virulence to wheat *R* genes *H6*, *Hdic*, and *H5* was examined. Mutations associated with *H6* virulence had been mapped previously. Therefore, we used *H6* to test the capacity of BSA-seq to map virulence using a field-derived HF population. This was the first time a non-structured HF population had been used to map HF virulence. *Hdic* virulence had not been mapped previously. Using a structured laboratory population, BSA-seq associated *Hdic* virulence with mutations in two candidate effector-encoding genes. Using a laboratory population, *H5* virulence was previously positioned in a region spanning the centromere of HF autosome 2. BSA-seq resolved *H5* virulence to a 1.3 Mb fragment on the same chromosome but failed to identify candidate mutations. Map-based candidate effectors were then delivered to *Nicotiana* plant cells via the type III secretion system of *Burkholderia glumae* bacteria. These experiments demonstrated that the genes associated with virulence to wheat *R* genes *H6* and *H13* are capable of suppressing plant immunity. Results are consistent with the hypothesis that effector proteins underlie the ability of HFs to survive on wheat.

**Keywords:** plant immunity, effector protein, resistance gene, plant gall, genome sequencing, wheat

## INTRODUCTION

The Hessian fly (HF, *Mayetiola destructor*) is an economically important, gall-forming, insect pest. It has a gene-for-gene relationship with its host plant, wheat (*Triticum* spp.). Recent investigations involving HF Resistance (*R*) genes *H13* and *H9* in wheat illustrate this relationship (Stuart, 2015): *H13* normally prevents HF larvae from galling wheat. These “*H13*-avirulent” larvae die as first instars on *H13*-plants. In contrast, “*H13*-virulent” HF larvae overcome this resistance; they both survive and gall *H13*-plants. This ability to survive and gall (*H13* virulence) is conditioned by recessive null mutations in a single HF gene, called *Avirulence* (*Avr*) gene *vH13* (Aggarwal et al., 2014). These *vH13* mutations are *H13*-specific. They do not, for example, allow plant galling and HF survival (virulence) on wheat plants carrying *R* gene *H9*. Instead, larvae that defeat *H9*-resistance are homozygous for recessive null mutations in a different *Avr* gene, *vH9* (Zhao et al., 2015). Wheat has at least 35 dominant, simply inherited, resistance (*R*) genes that prevent “avirulent” HF larval survival and plant galling (Miranda et al., 2010). The gene-for-gene hypothesis predicts that 35 different *Avr* genes correspond to each one of these *R* genes.

Similar gene-for-gene relationships exist between plants and plant pathogens (Harris et al., 2010). The study of these interactions has revealed two levels of defense in the plant immune system (Jones and Dangl, 2006). Basal plant immunity defends against non-adapted organisms. Highly adapted plant parasites use effector proteins to defeat this basal defense. To counter these parasites, plants have a second level of defense called Effector-Triggered Immunity (ETI). ETI uses *R*-gene-encoded proteins (*R* proteins) that recognize, either directly or indirectly, the presence of specific effectors. Upon effector detection, plant cells initiate a defense response that limits plant damage and infection. Natural selection then favors pathogens that have either masked or modified the effector beyond *R*-protein recognition or have lost the effector completely. This suggests that *Avr* genes are simply parasite genes that encode the effectors recognized by plant *R* proteins (Hogenhout et al., 2009).

Therefore, one hypothesis is that ETI underlies the HF-wheat gene-for-gene interaction. The corollary is that the HF uses effector proteins to defeat basal plant immunity. Additional evidence in favor of these hypotheses exist in both the plant and the insect. With respect to the plant, most *R* genes belong to gene families that encode proteins with nucleotide-binding (NB) and leucine-rich repeat (LRR) domains (Jones and Dangl, 2006). As natural selection has presumably favored their evolution in response to parasite adaptation, these are among the most prevalent and diverse genes in plant genomes. The genome of *Aegilops tauschii*, one diploid progenitor of hexaploid bread wheat (*T. aestivum*), contains over 1200 NB-LRR genes (Jia et al., 2013). Although the sequence of a HF *R* gene in wheat has yet to be published, mapping data indicates that they reside in clusters of NB-LRR genes (Gill et al., 1987; Raupp et al., 1993; Kong et al., 2005; Liu et al., 2005b; Liu et al., 2005c; Sardesai et al., 2005; Kong et al., 2008; Miranda et al., 2010).

With respect to the insect, hundreds of HF genes (seven percent of the HF genome) encode putative effectors. The majority of these are members of large, diverse gene families that were originally discovered as signal peptide-encoding transcripts in first-instar larval salivary glands (Chen et al., 2004; Chen et al., 2008; Chen et al., 2010). Some of these have been identified in HF-infested wheat tissue (Zhao et al., 2015; Wang et al., 2018). Like effector encoding genes in plant parasites (Hogenhout et al., 2009), these HF genes are experiencing diversifying selection (Chen et al., 2008; Zhao et al., 2015), presumably to remain adapted to wheat. Moreover, HF *Avr* gene mapping has shown a correspondence between HF *Avr* genes and putative effector-encoding genes (Stuart, 2015).

Here, we describe experiments that further tested this correspondence. Additional HF *Avr* gene mapping was performed using bulked-segregant analysis (Giovannoni et al., 1991; Michelmore et al., 1991) in combination with whole genome resequencing (BSA-seq). BSA uses pools of genomic DNA collected from individuals segregating for a trait of interest to identify polymorphic DNA markers linked to that trait. BSA-seq sequences pools of genomic DNA to identify linked single nucleotide polymorphisms (SNPs) and then directly positions those SNPs in the genome. BSA-seq has been successfully applied to gene mapping and identification in yeast (*Saccharomyces cerevisiae*) (Pomraning et al., 2011; Swinnen et al., 2012), zebrafish (*Danio rerio*) (Leshchiner et al., 2012), *Arabidopsis thaliana* (Austin et al., 2011; Schneeberger, 2014), rice (*Oryza sativa*) (Abe et al., 2012; Takagi et al., 2013); fruit fly (*Drosophila melanogaster*) (Bastide et al., 2013) and the malaria mosquito (*Anopheles gambiae*) (Redmond et al., 2015). It was also used to locate mutations in the brown planthopper (*Nilaparvata lugens*) *Avr* gene *vBph1* that defeat the *Bph1*-resistance in rice (*Oryza sativa*) (Kobayashi et al., 2014). Here, we were interested in three separate HF traits: virulence (defined as larval survival and plant-galling) to *H6*-, *Hdic*- and *H5*-resistant wheat seedlings. Virulence to *H6*, which had been mapped previously (Zhao et al., 2015), tested the accuracy of BSA-seq in the HF. We then mapped virulence to *Hdic* and *H5*.

To test putative HF effectors for plant immune suppression, we employed an assay that uses *Burkholderia glumae* bacteria, and the effector detector vector (pEDV) system to deliver HF candidate *Avr* proteins to *Nicotiana tabacum* and *N. benthamiana* via the bacterial type III secretion system (T3SS) (Sohn et al., 2007; Fabro et al., 2011). *B. glumae* is a bacterial rice pathogen that causes a rapid, localized cell death (a hypersensitive reaction, HR) in non-host *N. tabacum* and *N. benthamiana*. The pEDV system uses the type III secretion system (T3SS) to mediate foreign effector protein translocation into plant cells (Sohn et al., 2007; Sohn et al., 2014). Combined, *B. glumae* and pEDV effector protein translocation in *Nicotiana* plant cells enables the discovery of effectors with plant-immune suppression activity. The *B. glumae*/pEDV/*Nicotiana* system has been used to test effectors from the rice blast fungus *Magnaporthe oryzae* (Sharma et al., 2013), the false smut *Ustilaginoidea virens* (Zhang et al., 2014), the pathogenic fungus *Lasiodiplodia theobromae* (Yan

et al., 2018) and the root knot nematode *Meloidogyne incognita* (Shi et al., 2018a). Here, we examined proteins encoded by candidate HF *Avr* genes *vH13*, *vH6*, and *vHdic*.

## MATERIALS AND METHODS

### BSA-Seq Analysis

BSA-seq was used to perform genomic mapping of virulence to three HF *R* genes: *H6*, *Hdic*, and *H5*. To do this, DNA bulks were prepared using DNA isolated from individuals segregating for virulence. This required that populations segregating for virulence and avirulence to each *R* gene had to be identified and that individuals within these populations had to be separately genotyped. Virulence to each wheat *R* gene examined presented different challenges. The solution was to prepare separate HF populations for each *R* gene. A description of each population is presented below, followed by a description of whole-genome sequencing and data analysis. Later, we describe the preparation of PCR-based markers that were used to improve genetic resolution.

### Wheat *R* Genes and Insect Mapping Populations

Three different HF *R* genes (*H6*, *Hdic*, and *H5*), each in a different line of wheat, were examined in this investigation. Each wheat is maintained in the USDA-ARS Hessian fly laboratory at Purdue University. *H6* was discovered in bread wheat (*T. aestivum*) and is homozygous in the soft red winter wheat cultivar Caldwell (Patterson et al., 1982). Caldwell was developed, in part, to resist HF infestation in the Eastern United States. Caldwell seedlings were used to identify *H6*-virulent HF males as described below. *Hdic* was discovered in emmer wheat (*T. turgidum* ssp. *dicoccum*, PI 94641) and transferred to bread wheat, *T. aestivum* (Brown-Guedira et al., 2005). The homozygous *Hdic* hard winter wheat (KS99WGRC42) that was used to map *Hdic* in wheat (Liu et al., 2005a) was used to identify *Hdic*-virulent HF males in this investigation. *H5* was discovered in the Portuguese spring wheat cultivar Ribeiro (Shands and Cartwright, 1953). *H5* was backcrossed into the soft red winter wheat to produce the cultivar Abe (Patterson et al., 1975). Abe seedlings were used to identify *H5*-virulent males in a previous study (Behura et al., 2004). DNA extracted from some of those insects was used in the present investigation.

To map *H6* virulence, association mapping was performed using a non-structured Louisiana field population in which both virulence and avirulence to *H6*-wheat had been detected (Garcés-Carrera et al., 2014). Individual males in the Louisiana population were genotyped for *H6* virulence in separate testcrosses with homozygous *H6*-virulent virgin biotype-L females (Figure 1A). As described previously (Stuart et al., 1998), the ability of the offspring of each testcrossed male to gall and survive on *H6*-resistant (Caldwell) wheat seedlings determined male genotype. Genotyped males were collected, and their DNA was extracted using the DNeasy tissue kit (Qiagen, Chatsworth, CA).

To map *Hdic* virulence, we first isolated an *Hdic*-avirulent strain and an *Hdic*-virulent strain from an Israeli HF population.

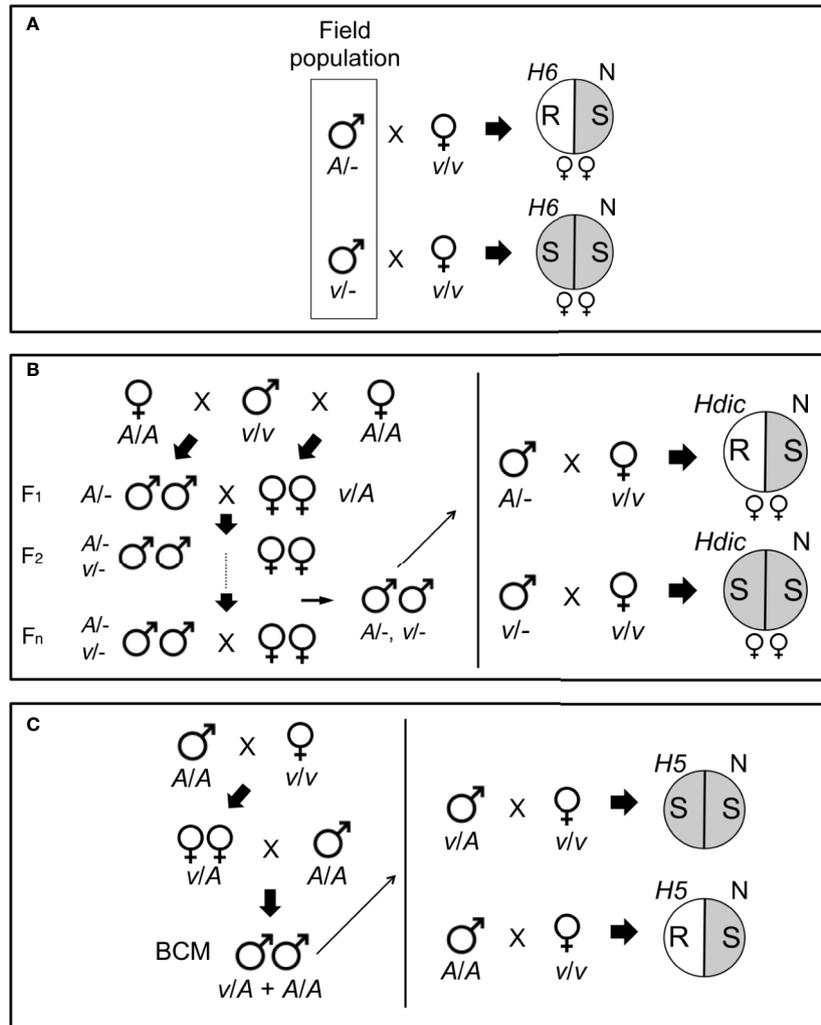
The *Hdic*-avirulent strain was selected using previously described methods (Zhao et al., 2015). Briefly, single mated females were caged and allowed to lay eggs on wheat seedlings in caged split-pots. One side of the pot contained susceptible Newton seedlings and the other side of the pot contained *Hdic*-resistant seedlings. Ten days after egg deposition, pots with *Hdic*-seedlings containing galled plants or living larvae were discarded. The larvae on susceptible (Newton) seedlings in the pots containing resistant *Hdic*-plants were allowed to develop. The emerging adult males and females were then intermated. This procedure was repeated for two additional generations, at which point no *Hdic* virulence was detected in the population. The *Hdic*-virulent strain was selected according to the method of Zantoko and Shukle (1997). For three generations, individual larvae were selected, one larva per plant, for the ability to survive, and gall *Hdic*-resistant seedlings. Surviving adults were collected and intermated to produce the *Hdic*-virulent strain.

The *Hdic* virulence mapping population was created by crossing a single virulent male with two avirulent sister females (one male-producing and one female-producing; Figure 1B). The resulting F1 male and female offspring were subsequently intercrossed to generate a *Hdic*-virulent advanced interbred population (vHdic-AIP). To maintain this population, individuals within vHdic-AIP were allowed to intermate and reproduce on *Hdic*-wheat. This process also served to disrupt linkage disequilibrium in the population. Individual F<sub>2</sub>, F<sub>6</sub>, and F<sub>10</sub> males were genotyped as hemizygous *Hdic*-virulent (v/-) or *Hdic*-avirulent (A/-) in testcrosses with individual, homozygous, *Hdic*-virulent (v/v), virgin females (Figure 1B). Genotyped males were used for genomic DNA extraction and samples were pooled as described below.

The *H5* virulence mapping population (vH5-BCM) was developed previously (Behura et al., 2004). Briefly, *H5*-avirulent males (Great Plains biotype; GP) and *H5*-virulent females (biotype L) were intermated and F<sub>1</sub> female offspring were backcrossed to a GP male to obtain vH5-BCM male offspring (Figure 1C). Since Hessian fly males transmit only their maternally derived chromosomes, the vH5-BCM males were testcrossed to homozygous, *H5*-virulent, biotype-L females, and their genotypes were determined by scoring the ability of their offspring to gall and survive on *H5*-resistant wheat (Abe) seedlings. Genotyped vH5-BCM males (n = 102) were collected for genomic DNA extraction. Behura et al. (2004) used DNA extracted from each of these males separately to map *H5* virulence to HF chromosome A2 (Table 1). DNA extracted from 48 of these males was used in the present investigation.

### Sample Pooling and Genome DNA Sequencing

DNA bulks were prepared by mixing approximately equal amounts of genomic DNA from each male used in the study. Paired-end (PE) sequencing libraries were prepared (100 bp PE reads, ~250bp insertion size) and genomic DNA sequencing (Illumina HiSeq2000) was performed by the Purdue Genomics Core Facility (Purdue University, West Lafayette, Indiana, USA; Table S1). The PE reads were later trimmed with Trimmomatic (v.0.3.2) (Bolger et al., 2014) to remove adapters (settings:



**FIGURE 1 |** Genotyping HF mapping populations. **(A)** Males collected from a Louisiana field population (boxed) were individually mated with single, homozygous *H6*-virulent (*v/v*), virgin females that produced offspring of only one sex. After mating, the females were placed separately in pots containing *H6*- and Newton (N) seedlings and allowed to oviposit on the plants. The eggs were allowed to hatch and the larvae were allowed to feed on the plants. Avirulent parental males (*A/-*) produced avirulent female (*v/A*) larvae incapable of galling *H6*-plants (R, plant resistance). Virulent parental males (*v/-*) produced virulent female (*v/v*) larvae capable of galling *H6*-plants (S, plant susceptibility). The sex of the offspring was determined by allowing larvae to develop into adults on the Newton seedlings in each pot. Matings that produced only male offspring were uninformative, because males carry only their mother’s X chromosome. **(B)** An advanced intercross population (AIP) segregating for *Hdic* virulence and avirulence was initiated with a cross between a single *Hdic*-virulent (*v/v*) male and two sister avirulent (*A/A*) females (left panel). One female produced only female offspring and the other produced only male offspring. Males and females in the *F*<sub>1</sub> and subsequent generations developed and were allowed to inter-mate and reproduce on susceptible plants. Males selected after the *F*<sub>1</sub> generation were genotyped individually as described in A where *Hdic* indicates *Hdic*-resistant plants (right panel). **(C)** An *H5* virulence mapping population was initiated from a single mating between a homozygous *H5*-virulent (*v/v*) female and a homozygous *H5*-avirulent (*A/A*) male (left panel). *F*<sub>1</sub> females developed on susceptible plants and were backcrossed to a single homozygous *H5*-avirulent (*A/A*) male. Backcross male offspring (BCM) were allowed to develop on susceptible plants and then selected for genotyping (right panel). Heterozygous (*v/A*) males mated to homozygous virulent (*v/v*) females produce two types of offspring: heterozygous (*v/A*) avirulent offspring and homozygous virulent (*v/v*) offspring capable of galling *H5*-seedlings (S). Homozygous (*A/A*) males produce only heterozygous (*v/A*) avirulent offspring incapable of galling *H5*-seedlings (R).

ILLUMINACLIP : TruSeq3-PE.fa:2:30:10:2:keepBothReads LEADING:3 TRAILING:3 MINLEN:50) and filtered for quality (Phred quality  $\geq$ Q20) with FASTX-toolkit (v.0.7) (settings: fastq\_quality\_filter -q 20 -p 80) ([http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html)).

### Read Mapping and SNP Analysis

The pre-processed and quality filtered Illumina PE reads from each bulked DNA sample were mapped to the HF reference genome (GenBank assembly accession number GCA\_000149185.1) using BWA v. 0.7.5a commands aln and sampe with default settings

**TABLE 1 |** HF *Avr* gene mapping approaches and results in this and previous investigations.

<b>R<sup>a</sup></b>	<b>Approach<sup>b</sup></b>	<b>Population (n)<sup>c</sup></b>	<b>Chrom.<sup>d</sup></b>	<b>Res<sup>e</sup></b>	<b>No.<sup>f</sup></b>	<b>Reference</b>
H3	BSA, AFLP	BC <sub>1</sub> (68)	A2	Un	0i	Behura et al., 2004
H5	BSA, AFLP	BC <sub>1</sub> (102)	A2	Un	0i	Behura et al., 2004
H5	BSA-seq	BC <sub>1</sub> (48)*	A2q	1.3	26i	This investigation
H6	BSA, M, B	F <sub>2</sub> -F <sub>8</sub> (335)	X2q, m	0.3	2i/1p	Zhao et al., 2015
H6	BSA-seq	NS (52)	X2q, m	6.1	1f	This investigation
H9	BSA, M, B	F <sub>2</sub> -F <sub>8</sub> (274) NS (92)	X1p, t	0.02	2i/1p	Zhao et al., 2015
H13	BSA, CW, M, B	F <sub>2</sub> (223) NS (79)	X2p, t	0.02	2i/1c/1f	Aggarwal et al., 2014
H24	BSA, M	F <sub>2</sub> (77) F <sub>4</sub> (66)	X1q, t	0.24	4i/1p	Zhao et al., 2016
Hdic	BSA-seq	F <sub>10</sub> (48)	X2q, c	2.1	Un	This investigation
Hdic	M	F <sub>10</sub> (48)**	X2q, c	1.1	8i/2p	This investigation

<sup>a</sup>Wheat *R* gene against which HF virulence was mapped. <sup>b</sup>Mapping approach, where BSA is bulked-segregant analysis, AFLP is amplified fragment length polymorphism, BSA-seq is BSA coupled with whole genome resequencing, M is microsatellite markers, B is bacterial artificial chromosome sequencing and CW is chromosome walking. <sup>c</sup>Mapping populations where BC<sub>1</sub> is a laboratory-produced backcross male population, F<sub>n</sub> is a laboratory recombinant inbred male population selected at the n generation and NS is one or more field-derived non-structured populations. \* is a subpopulation of that used by Behura et al. (2004). \*\* is the same population used in the Hdic BSA-seq experiment. <sup>d</sup>Chromosome position where A2 is autosome 2, X1 is X chromosome 1, X2 is X chromosome 2, q indicates long arm, p indicates short arm, m is the middle of the chromosome arm, t is telomeric, and c is centromeric. <sup>e</sup>Genomic resolution is the distance (Mb) between the closest flanking recombinant markers or the distance where the Fisher's exact test -Log<sub>10</sub>(p-value) is greater than 1.7 (H6 BSA-seq, **Figures 3**) or greater than 5.8 (H5 BSA-seq and Hdic BSA-seq, **Figures 3** and **4**). <sup>f</sup>Number of signal peptide-encoding genes identified (i) within the resolved region; the number of candidate *Avr* genes proposed (p) based on the presence and absence of transcription in avirulent and virulent larvae; the number of *Avr* genes confirmed (c) using RNA-interference gene silencing, and the number of genes showing effector activities in the functional assays performed in the present investigation (f). (Un, undetermined).

(Li and Durbin, 2010). SAMtools v.0.1.18 (Li et al., 2009) was used to remove ambiguously mapped and duplicated reads, keeping only those with a mapping quality higher than Q20 and proper mapped pairs. The SAMtools mpileup command was used to build a multiple-pileup file for SNP calling. SNPs around indels in the HF reference genome were filtered using the Perl scripts identify-indel-regions.pl (-indel-window = 5; window of 5bp in both directions) and filter-sync-by-gtf.pl. The final filtered mpileup file was synchronized using the java tool mpileup2sync.jar, filtering for base quality higher than Q20. SNP allele frequencies were estimated using the Perl script snp-frequency-diff.pl for bi-allelic SNPs using the following settings: -min-count = 4 (the minimum read count of the minor allele considering all bulks simultaneously); -min-coverage = 10 (the minimum read coverage per bulk used for SNP identification); and -max-coverage = 200 (the maximum read coverage per bulk used for SNP identification). These criteria were used in order to reduce the possibility of predicting false SNPs in genomic regions with poor sequencing coverage or repetitive DNA sequences. The statistical significance of allele frequency differences for each SNP position was determined with Fisher's exact test (FET) using Perl script fisher-test.pl. Fixation index (FST) values were determined for each SNP with Perl script fst-sliding.pl. The java tool mpileup2sync.jar and other Perl scripts used for SNP filtering and statistical analyses are included in the Popoolation2 tool (Kofler et al., 2011). The IGV genome viewer (Thorvaldsdóttir et al., 2013) was used to visualize the mapped reads as well as the FET and FST analyses. The average FET and FST values for 10-kb sliding-windows (5-kb steps) were plotted using R programming language [plot() function] as the cubic-smoothed line [smooth.spline() function] in order to reduce noise from sequencing variation across the HF genome. The Bonferroni correction method was used to establish the genome-wide statistical cutoff for FET analyses. Using an  $\alpha$  value of 0.05 and 31,600 10-kb genome windows across the 158-Mb HF genome established an FET significance cutoff value of 1.58e-6 (-Log<sub>10</sub>[FET] = 5.8).

## Data Availability

Whole-genome sequencing data for bulked samples are available at the NCBI Sequence Read Archive (SRA) under the NCBI Bioproject accession number PRJNA613640 (<https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA613640>).

## Genetic Mapping With PCR-Based Markers

The Hessian fly reference genome was used to map genes and identify PCR-based (microsatellite) markers and design PCR primers with the SSR Locator software (da Maia et al., 2008). The gene model identifiers are the names in the official HF gene set (OGS). These can be accessed at the USDA Arthropod i5k official workspace [https://i5k.nal.usda.gov/data/Arthropoda/maydes-\(Mayetiola\\_destructor\)/GCA\\_000149185.1/](https://i5k.nal.usda.gov/data/Arthropoda/maydes-(Mayetiola_destructor)/GCA_000149185.1/) and in the genome assembly curated at the National Center for Biotechnology Information (NCBI), GenBank assembly accession number GCA\_000149185.1. Molecular markers were used to genotype individuals taken from mapping populations and pooled DNA samples using standard PCR methods and the primers listed in **Table S2**.

## Fluorescent In Situ Hybridization (FISH)

The end-sequences of HF genomic bacterial artificial chromosomes (BACs) that had been mapped to HF polytene chromosomes (Aggarwal et al., 2009) were used as part of the HF genome sequencing project (Zhao et al., 2015). We used these data to identify HF BACs that reside within HF genome scaffolds A1Random.66 and X2.8. From among these BACs, we selected BAC HF07L11 as a probe for scaffold X2.8 and BAC Md23L24 as a probe for scaffold A1R.66. Using methodology described previously (Stuart et al., 2014), these BAC clones were fluorescently labeled, denatured, and allowed to hybridize to complementary bases on HF polytene chromosome preparations. Later, the chromosomes were stained with DAPI and the positions of BAC hybridizations were examined and photographed using fluorescence microscopy.

## Reverse Transcription PCR Analyses

To perform reverse transcription PCR (RT-PCR), total RNA was isolated from pools of 50 to 60 two-day-old first-instar larvae. RNA from avirulent and virulent strains were extracted separately using the RNeasy Mini Kit (Qiagen). Single-strand cDNA was reverse transcribed using the RNA of each individual pool separately and the SuperScript III First Strand kit (Invitrogen) according to the manufacturer's recommendations. Single-strand cDNA was then used in PCR experiments using gene-specific primers (Table S3). PCR was performed in 35 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 60 s, followed with a final step of 72°C for 5 min. The Hessian fly actin gene was included as a reference and internal control (Table S3). RT-PCR products were visualized on agarose gels. RT-PCR amplifications were performed with at least three biological replications.

## Phyre2 Structural Protein Modeling

Phyre2 (<http://www.sbg.bio.ic.ac.uk/~phyre2/>) is a free web-based service for prediction of the three-dimensional (3D) structure of a protein sequence using homology modeling against a database of Hidden Markov Model (HMM) profiles of known 3D protein domain structures from the Protein Data Bank (PDB, <http://www.wwpdb.org/>) (Kelley et al., 2015). Phyre2 was used to examine the predicted protein structures of the two best candidate *vHdic* genes in an attempt to identify similarities with the domains of other proteins.

## Candidate HF Effector Gene Cloning and Plasmid Construct Preparation

Total RNA was isolated from HF, first-instar, larval biotype GP using the RNeasy Mini Kit (Qiagen). RNA was subsequently reverse transcribed into first-strand cDNA using the SuperScript III First Strand Synthesis kit (Invitrogen). Double-stranded cDNA for effector genes was amplified from first-strand cDNA using gene-specific primers containing Gateway attB adapters (Table S4). These primers were designed to exclude the corresponding secretion signal peptides from each effector gene. Gene attB PCR products were recombined into pENTR/pDONR vectors using the Gateway BP reaction (Invitrogen) and chemically transformed into *Escherichia coli* OmniMAX2 cells (Invitrogen). Recombinant colonies were selected on 50 µg/ml kanamycin LB plates. Colonies carrying the recombinant plasmids were selected for plasmid isolation and DNA insert sequencing. Genes in pENTR/pDONR vectors were recombined by Gateway LR reactions (Invitrogen) into expression vector pEDV6 and transformed into *E. coli* OmniMAX2 cells. Recombinant colonies were selected on 10 µg/ml gentamicin LB plates and used for plasmid isolation. A pEDV-GFP construct was built by LR recombination between plasmids pENTR1AGFP-N2 (FR1) (Campeau et al., 2009) and pEDV6 (Sohn et al., 2014). Plasmid pEDV5 (Sohn et al., 2007) was used as an empty vector (EV) control. pENTR1A-GFP-N2 (FR1) was a gift from Eric Campeau (Addgene plasmid 19364). Plasmids pEDV5/6 were generously supplied by J. Jones (The Sainsbury Laboratory, Norwich U.K.).

## Mobilization of pEDV Constructs Into Bacteria

The pEDV constructs were transformed into *B. glumae* cells as follows: Electrocompetent *B. glumae* cells were prepared as previously described (Saitoh and Terauchi, 2013) with minor modifications. In brief, the *B. glumae* strain 336gr-1 was inoculated in 20 ml LB medium for 14–16 h at 28°C with shaking until OD600 = 0.8. The flask was then opened for 30 s under clean conditions and then maintained at 28°C with shaking for an additional 4 h. The cells were then pelleted twice at 4°C and 3000 rpm for 5 min and resuspended in cold 10% glycerol. The pellet was then dissolved in 600 µl of cold 10% glycerol and divided into 50-µl aliquots. These were stored at -80°C for transformations. Each plasmid construct (0.3 µg) was electroporated into *B. glumae* using a MicroPulse Electroporator (Bio-Rad). Transformant *B. glumae* strains were selected on gentamicin (25 µg/ml) LB agar. *B. glumae* strain 336gr-1 was a generous gift from J. H. Ham (Dept. Plant Pathology and Crop Physiology, Louisiana State University).

## Hypersensitivity Reaction (HR) Induction/Suppression Assays

For HR assays, *Bglu*-pEDV strains were plated on King's Broth (KB) agar 25 µg/ml gentamicin and incubated at 30°C for 14 to 16 h. *Bglu*-pEDV strains were dissolved in 0.9% NaCl solution at OD600 = 0.7. Bacteria suspensions were infiltrated with a 1-ml syringe without a needle into 4- to 5-week-old *Nicotiana tabacum* Burley 21 HA and *N. benthamiana* leaves. Infiltrated plants were maintained in a growth chamber at 24 ± 1°C with a 16:8 (light/dark) light cycle and 80 ± 5% relative humidity. HR was recorded after 24 h for *N. tabacum* and 48 h for *N. benthamiana*.

## Ion-leakage Assays

*Bglu*-pEDV cell suspensions were prepared and infiltrated in leaves of *N. tabacum* and *N. benthamiana* plants as described above. Leaf disks (150 mm diameter) were collected from the infiltrated areas using a cork borer 18-h post infiltration (hpi) for *N. tabacum* and 36 hpi for *N. benthamiana*. Leaf disks were floated on 15-ml nanopure water and incubated at 22°C with gentle shaking (100 rpm). Conductivity in the water was registered after 4 hours of incubation using a conductivity meter (Metler Toledo S30K) with a sensor probe (Conductivity Sensor LE703, Metler Toledo). Samples from three different plants were used as replicates for each treatment. Statistical analyses were performed using ANOVA and Tukey's HSD for significant differences ( $p < 0.05$ ) among treatments.

## RESULTS

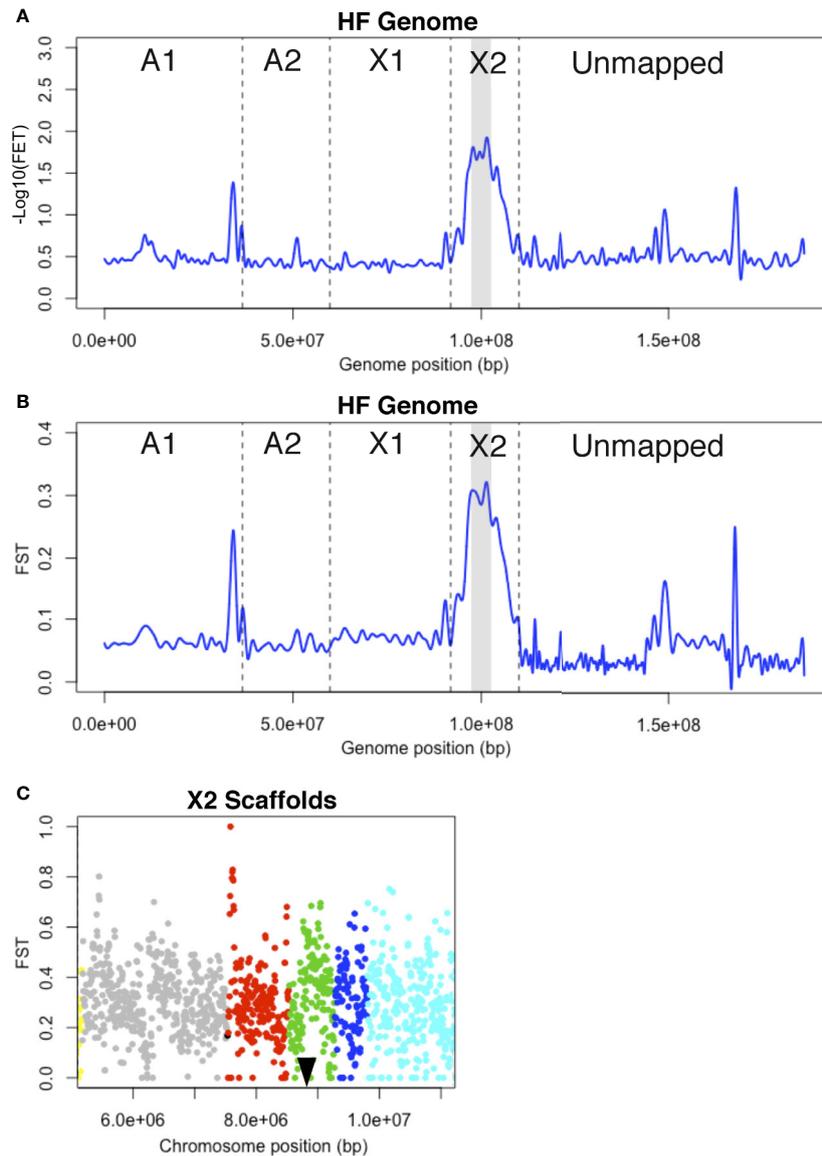
### BSA-Seq Confirms the Genomic Position of H6 Virulence

Zhao et al. (2015) mapped *H6* virulence on the long arm of HF chromosome X2 using four different structured mapping populations and PCR-based DNA markers (Table 1). This had been an arduous task. Therefore, we decided to test the capacity of

BSA-seq in the HF using *H6* virulence. *H6*-virulent bulk DNA was prepared from 23 males. *H6*-avirulent bulk was prepared from 19 males. Each male was collected from a field-derived population and genotyped in individual testcrosses (Figure 1A). These bulks were sequenced separately, resulting in 5.32 Gb of combined genomic data (Table S1). The BWA tool (Li and Durbin, 2010) aligned the reads from each bulk against the HF reference genome and then SAMtools (Li et al., 2009) used these mapping data to identify 1.5

million SNPs. Popoolation2 (Kofler et al., 2011) calculated SNP allele frequencies within each bulk and then determined the differences in allele frequency for each SNP. Using these data, Fisher's exact test (FET) was performed and average  $-\text{Log}_{10}(\text{FET})$  values within sliding 10-kb genome windows were plotted as a smoothed line across the HF genome (Figure 2A).

Unfortunately, the HF genome sequence is imperfectly assembled, and this was evident in the plotted data. Instead of



**FIGURE 2 |** *H6* virulence BSA-seq analysis. **(A)** Fisher's exact test (FET) for significance of SNP allele frequency differences between *H6*-avirulent and *H6*-virulent bulks plotted as a cubic-smoothed line. Average  $-\text{Log}_{10}(\text{FET})$  values for 10-kb sliding-windows (5-kb steps) were plotted across the HF genome. Vertical dashed lines separate scaffolds assigned to HF chromosomes (A1, A2, X1, and X2) and unassigned scaffolds (Unmapped). Gray shading indicates the genomic position of a 6.1-Mb chromosome-X2 region where  $-\text{Log}_{10}(\text{FET})$  is less than the 5.8 statistical cutoff, but greater than 1.7. **(B)** FST values for 10-kb sliding-windows (5-kb steps) plotted as a cubic-smoothed line across the HF genome. The 6.1-Mb region highlighted in A has the highest FST values (gray shading). **(C)** FST estimates within the 6.1-Mb chromosome-X2 region highlighted in B. Each dot represents the average FST values for 10-kb sliding-windows (5-kb steps). The scaffolds in this region are shown with dots of different colors; X2.8 is grey, X2.10 is red, X2.11 is green, X2.12 is dark blue and X2.13 is light blue. A black triangle indicates the position of the previously cloned candidate-*vH6* gene, Mdes009086-RA (Zhao et al., 2015).

a single peak that rose and fell over a single chromosomal position, several peaks were observed scattered along the genome map in both FET and FST plots (**Figure 2A, B**). These peaks indicate that linkage exists between *H6* virulence and the underlying genome scaffolds. Most of these scaffolds are located in the “unmapped” fraction of the genome map. Their peaks suggest that they should be assigned to the chromosome known to carry the *H6* virulence trait (Zhao et al., 2015), chromosome X2. Unmapped scaffold Un.18557, with an associated  $-\text{Log}_{10}[\text{FET}]$  of 1.3, was the clearest example. Another peak with the same elevation was associated with a single chromosome A1 scaffold (A1R.66). The same logic suggests that all or part of scaffold A1R.66 also belongs to chromosome X2. This was later confirmed when *Hdic* virulence was mapped (described below).

Importantly, the peak with the greatest elevation ( $-\text{Log}_{10}[\text{FET}] = 1.7$ ) identified a 6.1-Mb region on the long arm of chromosome X2 (**Figure 2A**). Although this value failed to meet the statistical cutoff, established using the Bonferroni correction method ( $-\text{Log}_{10}[\text{FET}] = 5.8$ ), the 6.1-Mb genome region under this peak includes the 300-kb genome window where *H6* virulence was previously mapped (Zhao et al., 2015). This region contains HF genome scaffolds X2.8, X2.10, X2.11, X2.12, and X2.13 (**Figure 2C**). Using Web Apollo (Kofler et al., 2011; Lee et al., 2013) and the HF genome reference sequence ([https://i5k.nal.usda.gov/Mayetiola\\_destructor](https://i5k.nal.usda.gov/Mayetiola_destructor)) we identified 945 gene models within this window. Gene model Mdes009086-RA, which was previously identified as the candidate gene conditioning *H6* virulence (*vH6*), resides within scaffold X2.11 (**Figure 2C**). Mdes009086-RA encodes a member of the putative HF effector protein family SSGP71. It is not transcribed in *H6*-virulent larvae (Zhao et al., 2015).

Although this investigation failed to resolve the position of *H6* virulence as well as the previous investigation (**Table 1**), the results demonstrated that non-structured field-derived populations can be used to identify SNPs linked to HF virulence and encouraged us to use BSA-seq to try to resolve the positions of virulence to other HF *R* genes. As discussed below, we believe that a larger mapping population would have improved *H6* virulence mapping resolution.

## BSA-Seq Resolves *vHdic* Position and Corrects the HF Genomic Map

To select a HF strain that was *Hdic*-virulent, the offspring of individual females taken from an Israeli field collection were examined for their ability to survive and stunt *Hdic*-wheat. During this process, we noted that all matings ( $n = 9$ ) between individual *Hdic*-virulent females and individual *Hdic*-avirulent males produced *Hdic*-virulent male offspring. This suggested that *Hdic* virulence is X-linked because male offspring do not inherit X chromosomes from their fathers; matings between virulent females (*v/v*) and avirulent males (*A/-*) produce only virulent male offspring (*v/-*).

To test this possibility, we looked for linkage between X-linked microsatellite markers using conventional BSA.  $F_2$  males collected from the *Hdic* virulence advanced intercross population (*vHdic*-AIP) were genotyped as *Hdic*-avirulent or *Hdic*-virulent individuals (**Figure 1B**). Separate avirulent and virulent DNA

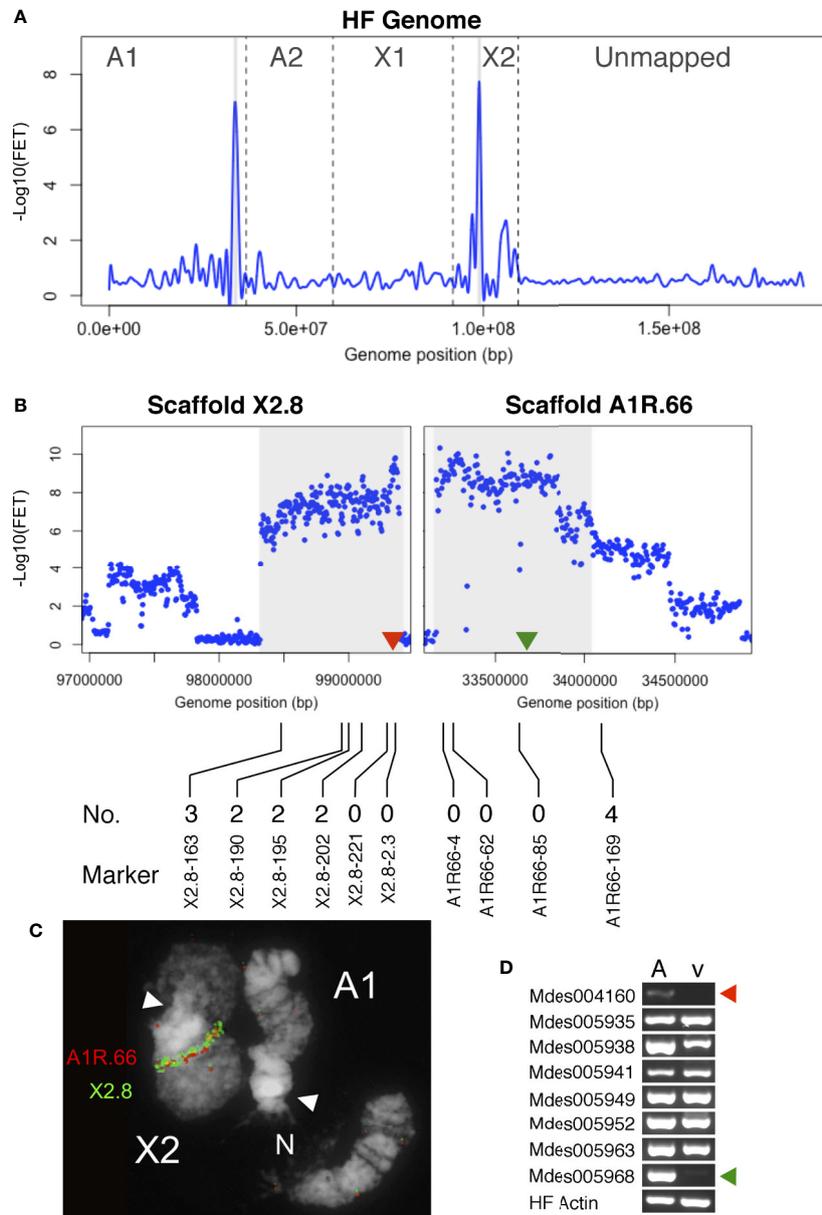
bulks were then prepared and used as template with primers designed for those microsatellites in separate PCR experiments. The two pools amplified alternative microsatellite alleles with six markers on scaffolds X2.7 and X2.8 (data not shown), indicating that those markers were linked to *Hdic* virulence. Genetic mapping using X2.7 and X2.8 markers and the DNA of 189 individual males collected from the  $F_2$ ,  $F_6$ , and  $F_{10}$  generations confirmed this linkage and suggested that *Hdic* virulence was present on either a 100-kb fragment on the proximal end of scaffold X2.8 or within the gap between X2.7 and X2.8.

For BSA-seq analysis, an *Hdic*-avirulent DNA bulk ( $n = 15$ ) and an *Hdic*-virulent DNA bulk ( $n = 33$ ) were prepared using DNA extracted from  $F_{10}$  males (**Figure 1B**). Each bulk was separately sequenced. This produced 12.2 Gb of combined genomic data containing 1.2 million SNPs (**Table S1**). The average FET values for SNPs across sliding 10-kb genome windows were plotted (**Figure 3A**). Two genomic regions were associated with peaks where  $-\text{Log}_{10}(\text{FET})$  values were greater than the statistical cutoff of 5.8. These appeared to be tightly linked to *Hdic* virulence. Consistent with the conventional BSA results, one region included a 1.1-Mb section of scaffold X2.8 (**Figure 3B**). Unexpectedly, the other region contained a 1.0-Mb section of scaffold A1R.66, the same A1 scaffold that showed X2 linkage in the *H6* virulence investigation described above.

We hypothesized that A1R.66 was contiguous with scaffold X2.8, and present in the gap between scaffolds X2.8 and X2.7 (**Figure S1**). That hypothesis was tested *via* genetic mapping using PCR-based A1R.66 markers (**Table S5**) and 48 individual *vHdic*-AIP  $F_{10}$  males. Consistent with the hypothesis, A1R.66 was linked to both X2 and *Hdic* virulence in those experiments (**Figure 3B**). It was tested again using fluorescence *in situ* hybridization (FISH) to the HF polytene chromosomes. An X2.8 BAC and an A1R.66 BAC used as probes hybridized together on the long arm of chromosome X2 (**Figure 3C**). Taken together, the genetic and FISH data determined that the most likely order of the scaffolds on the long arm of chromosome X2 is centromere-X2.7-A1R.66-X2.8. Therefore, disregarding the gap between scaffolds X2.8 and A1R.66, BSA-seq data positioned *Hdic* virulence within a contiguous 2.1-Mb region on the long arm of chromosome X2 (**Table 1, Figure 3B**). The PCR markers used in this investigation further resolved the *Hdic* virulence, to within a 1.1-Mb region flanked by the closest recombinant markers, X2.8-202 and A1R66-169 (**Figure 3B and Table 1**).

## Candidate *vHdic* Genes Identified

Using Web Apollo (Lee et al., 2013) and the HF genome reference sequence ([https://i5k.nal.usda.gov/Mayetiola\\_destructor](https://i5k.nal.usda.gov/Mayetiola_destructor)), 48 gene models on A1R.66 and six gene models on X2.8 were identified within the 1.1-Mb region identified above. The SignalP4.1 algorithm (Petersen et al., 2011) predicted that eight of these genes models encode proteins containing secretion signals (**Table S6**). Three of these had significant sequence similarities to other genes in insects (BLASTP,  $e \leq 3e-87$ ). Five others were predicted HF effector proteins. Four of these belong to gene family SSGP4, which consists of at least 64 predicted HF effector-encoding genes (Zhao et al., 2015). SSGP4-encoded proteins that have no



**FIGURE 3 |** *Hdcic* virulence BSA-seq analysis. **(A)** Fisher’s exact test (FET) analysis for significance of SNP allele frequency difference between *Hdcic*-avirulent and *Hdcic*-virulent bulks. The average  $-\text{Log}_{10}(\text{FET})$  values are plotted and the relative positions of scaffolds on chromosomes are presented as in **Figure 2A**. Genomic regions where  $-\text{Log}_{10}(\text{FET})$  values are statistically significant ( $> 5.8$ ) are highlighted in gray. These regions correspond to sequences within scaffolds A1R.66 and X2.8. **(B)** FET analysis of scaffolds X2.8 and A1R.66. Each dot represents average FET values for 10-kb sliding-windows (5-kb steps). Sequences where  $-\text{Log}_{10}(\text{FET})$  values are statistically significant ( $> 5.8$ ) are highlighted. Red and green triangles indicate the genomic positions of candidate genes Mdes004160 and Mdes005968, respectively. The positions of X2.8 and A1R.66 markers used to refine the genomic map are shown below the plots. The number of *Hdcic*-virulent recombinant individuals (out of 48 total) is shown directly above each marker’s designation. **(C)** Fluorescence *in situ* hybridization (FISH) of BAC-based probes to HF polytene chromosomes X2 and A1. Scaffold X2.8 BAC HF07L11 (green signal) and scaffold A1R.66 BAC Md23L24 (red signal) hybridized adjacent to each other on chromosome X2. White arrows indicate chromosome centromeres (N = nucleolus). **(D)** RT-PCR using total RNA isolated from *Hdcic*-avirulent (A) and -virulent (v) first instar HF larvae. Gene-specific primers for all eight putative signal-peptide encoding genes amplified *Hdcic*-avirulent cDNA. The Mdes004160 and Mdes005968 primers failed to amplify *Hdcic*-virulent cDNA. The HF actin gene was used as a positive control. Similar results were obtained in each of three independent biological replications.

sequence similarities to any other proteins thus far identified outside of the HF (BLASTP,  $e > 5.0$ ).

Using RT-PCR, we examined whether each of the eight signal-peptide-encoding genes is transcribed in first instar larvae. Transcripts of all eight genes were detected in *Hdic*-avirulent larvae (**Figure 3D**). However, transcripts of two SSGP4 genes, Mdes004160-RA and Mdes005968-RA, were not detected in *Hdic*-avirulent first-instar larvae. Because *Avr* gene loss-of-function often correlates with virulence, these observations make Mdes004160-RA and Mdes005968-RA good candidate *vHdic* genes. The predicted protein sequences of these genes are presented in **Figure S2**. Although their secretion-signal peptides are 75% identical, their mature proteins are only 32% identical.

Previous investigations have used Phyre2 (Kelley et al., 2015) to identify putative protein domains in predicted HF effectors (Zhao et al., 2015; Zhao et al., 2016). Therefore, we used Phyre2 to predict the 3D structures of the Mdes005968-RA- and Mdes004160-RA-encoded proteins in an attempt to identify similarities with known protein domain structures. This analysis failed to identify any protein with significant structural similarities with either predicted protein ( $e$ -value  $\geq 1$ ). Because it is located in the center of the *vHdic*-mapping window (**Figure 3D**), we selected gene Mdes004160 for the functional analysis described below.

## BSA-Seq Maps *H5* Virulence to HF Genomic Scaffold A2.7

Previously, *H5* virulence was mapped to a region spanning the centromere of chromosome A2. That region composed 30% of the chromosome's length (Behura et al., 2004), and in three independent experiments, displayed severe recombination suppression. Therefore, in the present investigation, we first examined microsatellite markers identified on scaffolds flanking the A2 centromere for linkage to *H5* virulence. To do this, we used the DNA extracted from 36  $F_2$  back-crossed males genotyped in the Behura et al. (2004) study (**Figure 1C**). The DNA of each male was examined separately for each marker. Consistent with the recombination suppression previously observed, each scaffold examined was linked to *H5* virulence (**Table S7**). However, recombination was completely lacking between *H5* virulence and markers on scaffolds A2.6 (458.5 kb) and A2.7 (1.3 Mb), suggesting that the resolution of the position of *H5* virulence might improve with a larger mapping population and additional markers.

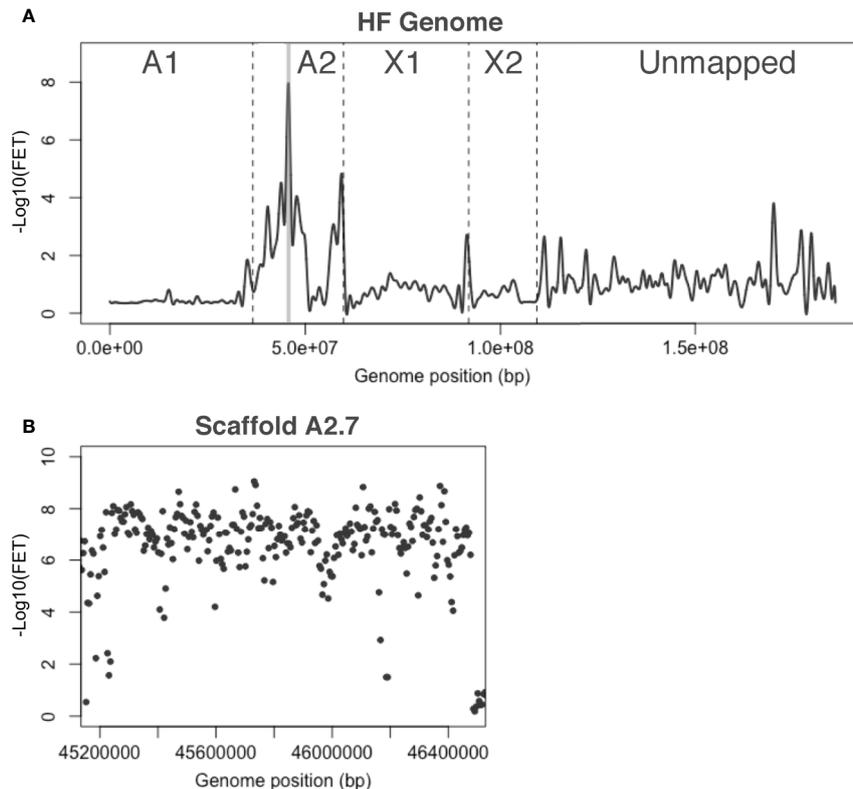
In an attempt to do this, BSA-seq was applied to DNA bulks from 24 additional *H5*-avirulent and 24 additional *H5*-virulent ( $n = 24$ ) males selected from the same  $F_2$  back-crossed mapping population (**Figure 1C**). Because *H5* virulence is autosomal, the *H5*-virulent bulk was developed from heterozygous ( $v/A$ ) males while the *H5*-avirulent bulk was developed from homozygous ( $A/A$ ) males. Therefore, SNPs in the *H5*-avirulent bulk were expected to be heterozygous in genomic regions unlinked to *H5* virulence and homozygous in genomic regions linked to *H5* virulence.

Whole-genome sequencing produced 13-Gb of high-quality reads, containing about 0.92 million SNPs (**Table S1**). Average FET values for SNPs were plotted as described above (**Figure 4A**). Recombination suppression was again observed across much of the A2 chromosome as most A2 scaffolds had  $-\text{Log}_{10}$  (FET) values of 2.0 and higher. In addition, one X1 scaffold and seven unmapped scaffolds had peaks with  $-\text{Log}_{10}$ (FET) scores of 2.0 or greater, suggesting that they should probably be assigned to chromosome A2. Nevertheless, *H5* virulence appeared to be most tightly linked to scaffold A2.7, with a statistically significant peak  $-\text{Log}_{10}$ (FET) value of 8.0 (**Figure 4A**). Plotting  $-\text{Log}_{10}$  (FET) values for 10-kb sliding-windows across the scaffold revealed that SNPs across the scaffold had similar FET values (**Figure 4B**). Consequently, we suspected that *H5* virulence is probably associated with an *Avr* gene within the 1.3-Mb sequence of scaffold A2.7 (**Table 1**).

Scaffold A2.7 contained 142 gene models. Thirty-seven of these encode predicted proteins carrying secretion signal peptides. BLASTP indicated that 26 of these are highly conserved and unlikely effector candidates. The remaining 11 genes have no similarities with genes in other insects. We discovered that one gene model, Mdes007142-RA, was composed of two putative effector-encoding genes (**Figure S3A**); one called SSGP47-1 (Chen et al., 2008) and another referred to here as Mdes007142 (b). We examined the transcription of all 12 of these genes using RT-PCR on first instar larvae in the *H5*-avirulent and *H5*-virulent strains. None of these were differentially transcribed between bulks (**Figure S3B**). Therefore, multiple alignment and manual comparisons were performed on DNA sequences of each bulk against each other and the reference sequence. These comparisons examined the sequences of the virulence- and avirulence-associated alleles of each gene for evidence of frameshift and nonsense mutations that might disrupt the proper translation of virulence-associated sequences. These comparisons failed to identify putative null mutations that might be associated with the virulence allele of a candidate *vH5* gene. Therefore, although BSA-seq was able to improve the resolution of *H5* virulence on chromosome A2, we were unable to associate *H5* virulence with a putative effector encoding gene. The possibility remains that a gene that has yet to be identified within the A2.7 sequence is responsible for *H5* virulence.

## Candidate-*vH6* and *vH13* Proteins Suppress Non-Host Plant Immunity

Based on the similarities that HF-wheat interaction has with pathogen-plant systems, we decided to explore the pEDV system for bacterial T3SS-dependent delivery of effector proteins into plant cells and test whether HF candidate effectors are capable of suppressing plant defense responses. As described above, candidate *vH6* (Mdes009086-RA) and one candidate *vHdic* (Mdes004160) were selected for this analysis. *vH13* was included as a third HF *Avr* gene (Aggarwal et al., 2014). Green fluorescent protein (GFP) was used as a negative control. Each gene was moved separately into the pEDV6 vector (**Figure 5A**) and each construct transformed into *B. glumae* for infiltration in



**FIGURE 4 |** *H5* virulence BSA-seq analysis. **(A)** Fisher's exact test (FET) analysis for significance of SNP allele frequency difference between *H5*-avirulent and *H5*-virulent bulks. The average  $-\text{Log}_{10}(\text{FET})$  values were plotted, and the relative positions of scaffolds on chromosomes are presented as in **Figure 2A**. A 1.3-Mb region corresponding to a single genome scaffold (A2.7) is highlighted in gray where the  $-\text{Log}_{10}(\text{FET})$  was statistically significant. **(B)** FET analysis of scaffold A2.7 (highlighted in panel **A**). Each dot represents the average  $-\text{Log}_{10}(\text{FET})$  value for 10-kb sliding-windows (5-kb steps).

non-host *N. tabacum* and *N. benthamiana* plants (Kang et al., 2008; Sharma et al., 2013). At 24 hpi in *N. tabacum* and 48 hpi in *N. benthamiana* (**Figures 5B, D**), the HR that *B. glumae* normally induces in *Nicotiana* was evident on leaf tissue infiltrated with *B. glumae* carrying an empty vector (EV), GFP and candidate *vHdic*. However, HR was reduced or absent on leaf tissue infiltrated with *B. glumae* harboring candidate *vH6* and *vH13*. These results were confirmed in a separate experiment in which ion leakage was used to assess plant cell integrity and membrane damage to the plant tissues (Rolny et al., 2011) (**Figures 5C, E**). Lack of HR suggests that, like many plant effectors, candidate-*vH6* and *vH13* proteins interfere with plant immunity.

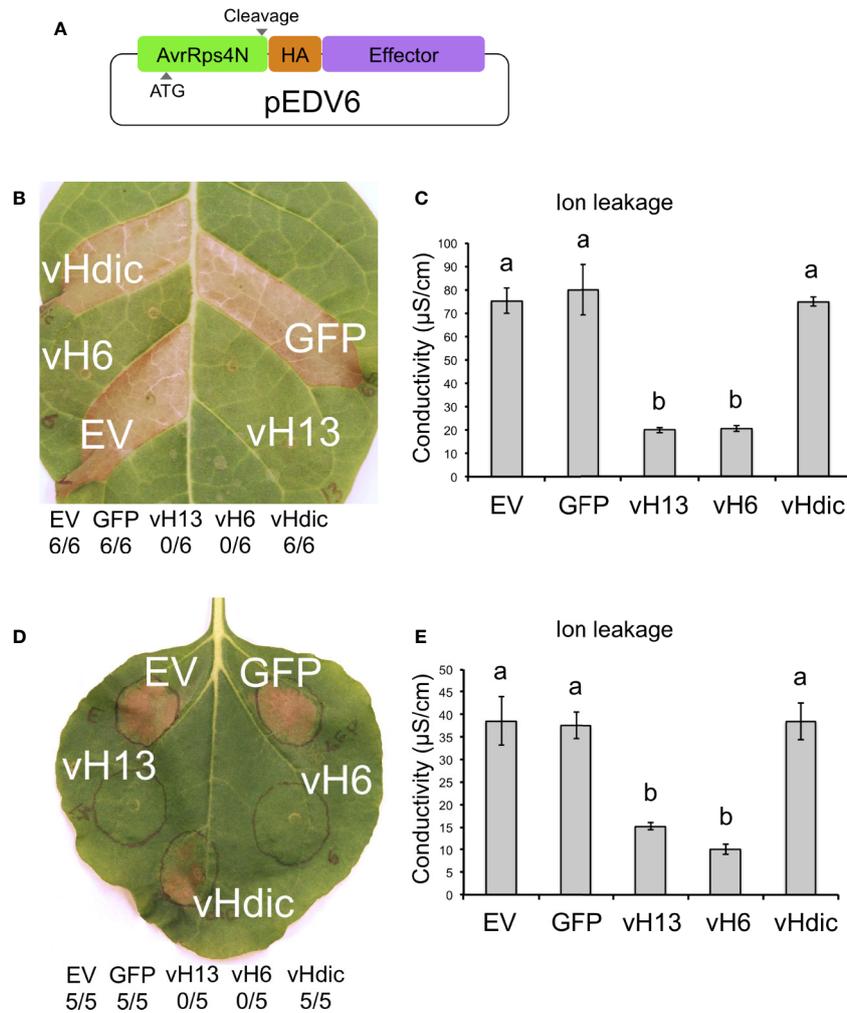
### Truncated *vH6* Failed to Suppress Bacterial-Induced Plant Cell Death

Candidate *vH6* is a member of a large family of HF effector-encoding genes (SSGP71). Like other SSGP71 proteins, candidate-*vH6* contains both F-box and leucine-rich-repeat (LRR) domains (**Figure 6A**). The F-box domain of candidate-*vH6* interacts with wheat Skp1-like proteins (Zhao et al., 2015), thereby mimicking a wheat E3 ligase. Therefore, we hypothesized that deleting the candidate-*vH6* F-box domain from the effector

would negatively impact its mode of action when delivered to *Nicotiana* cells by *B. glumae*. And as expected, unlike the complete candidate-*vH6* effector (Bglu-*vH6*), the truncated candidate (Bglu-*vH6*ΔFb) did not interfere with nonhost HR on *N. benthamiana* in either the leaf or ion leakage assays (**Figures 6B, C**).

## DISCUSSION

Virulence to seven different HF *R* genes in wheat has been positioned on HF chromosomes (**Table 1**). To our knowledge, outside of the HF, virulence to only one other plant *R* gene has been mapped, *Bph1* virulence in the brown planthopper (*Nilaparvata lugens*) (Kobayashi et al., 2014). In each case, BSA was used to identify linked DNA polymorphisms. Earlier HF investigations first sequenced BSA-discovered markers (*H5* and *H13*) and then, when it was feasible (*H13*), used the marker sequences to identify BACs that were later fully or partially sequenced themselves. The BAC sequence data was then used to develop new probes that permitted chromosome walking toward an *Avr* gene. PCR-based markers (microsatellites) identified in the sequence along the walk were then used to resolve the



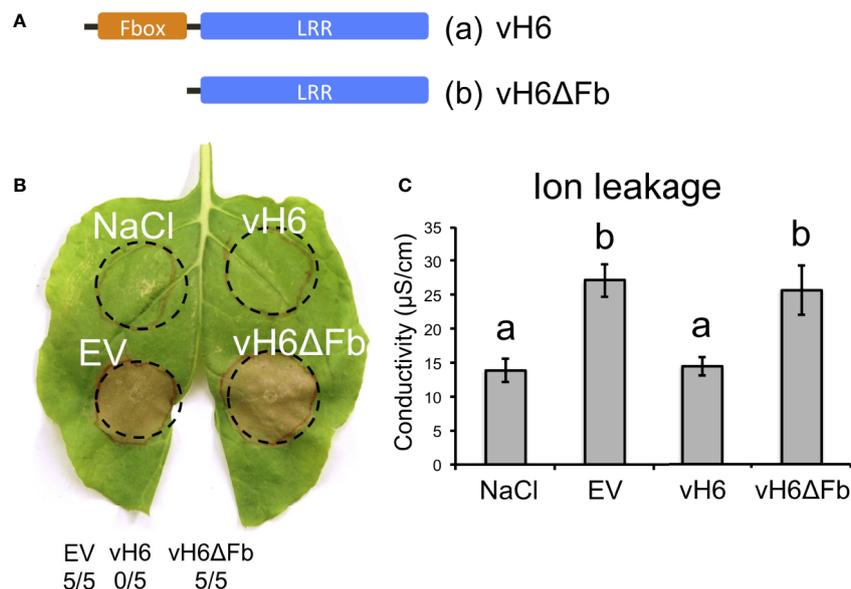
**FIGURE 5 |** Reactions of *Nicotiana tabacum* and *N. benthamiana* to *Burkholderia glumae* carrying T3SS-fusion constructs for candidate *Mayetiola destructor* effector proteins. **(A)** Schematic representation of the pEDV6 constructs. AvrRps4N: N-terminal region (T3SS-secretion signal) of the bacterial AvrRPS4 protein; HA: hemagglutinin peptide (HA) tag; Effector: candidate effector gene for testing. Initiation codon (ATG) and cleavage site are indicated. **(B)** *N. tabacum* reaction to the infiltration of *B. glumae* harboring the pEDV constructs (Bglu-pEDV) for the avirulence vH13 gene (vH13), avirulence vH6 gene (vH6), candidate avirulence vHdic gene (vHdic, Mdes004160), the green fluorescent protein gene (GFP) and empty vector control (EV). The number of times each Bglu-pEDV strain induced HR in 6 replications is shown. Pictures were taken 48 hpi for both plant species. HR was visible at 24 hpi in *N. tabacum*. **(C)** Ion leakage assays of infiltrated leaf areas in *N. tabacum*, measured at 18 hpi. **(D)** *N. benthamiana* reaction to the infiltration of Bglu-pEDV strains for vH13, vH6, candidate vHdic, GFP, and EV. The number of times each Bglu-pEDV strain induced HR in 5 replications is shown. **(E)** Ion leakage assays of infiltrated leaf areas in *N. benthamiana*, measured at 18 hpi. For panels **(C, E)**, each bar represents the average conductivity from 3 independent plants. Error bars represent the standard error. Statistical differences among the treatments were found with ANOVA (*N. tabacum*:  $F = 31.59$ ,  $p < 0.0001$ ; *N. benthamiana*:  $F = 17.76$ ,  $p = 0.0002$ ) and Tukey's test. Bars with different letters are significantly different ( $p < 0.05$ ).

position of virulence until the virulence-associated mutations themselves were discovered. The development of a fully sequenced HF genome greatly simplified this process as chromosome walking and BAC sequencing were no longer necessary for microsatellite discovery. BSA-seq further simplifies the process as the DNA polymorphisms linked to virulence are discovered and positioned simultaneously.

Nevertheless, we found that conventional microsatellite mapping provided better resolution for virulence to both *H6* and *Hdic* than the BSA-seq performed here. However, because BSA-seq performance depends on estimations of SNP allele

frequency within the bulked samples (Magwene et al., 2011; Rellstab et al., 2013), we believe that virulence resolution could be improved with DNA pools composed of greater numbers of individuals and better genome sequencing coverage, as larger sample sizes and higher-sequencing depth reduce the variability of SNP-allele frequency estimations and increase the chances of capturing recombination events.

In particular, the resolution of *H6* virulence was probably limited by the low genome-wide read coverage in relation to the bulk sizes (Table S1). Sequencing coverage is an important source of variation for allele frequency estimations because low



**FIGURE 6 |** *Burkholderia glumae* carrying truncated *vH6* elicits HR-related cell death in *Nicotiana benthamiana*. **(A)** Schematic representation for (a) complete *vH6* protein harboring the F-box domain (orange) and the leucine-rich repeats (LRR) domain (blue); and (b) truncated *vH6* lacking the Fbox domain (*vH6ΔFb*). **(B)** *B. glumae* pEDV constructs (Bglu-pEDV) for complete *vH6* (*vH6*), F-box-lacking *vH6* (*vH6ΔFb*) and the empty vector control (EV) were infiltrated in leaves of *N. benthamiana*. Infiltration buffer (NaCl) was also included as a control. HR elicitation was recorded at 48 hpi. *B. glumae* carrying an empty pEDV vector (EV) elicits cell death. *B. glumae* carrying pEDV-*vH6* fails to elicit cell death. Cell death is elicited when *B. glumae* carries truncated *vH6* (*vH6ΔFb*). The number of times each strain elicited cell death in 5 replications is shown. **(C)** Ion leakage assay for infiltrated leaf areas of *N. benthamiana*. Bars represent the average conductivity measured at 18 hpi in 3 independent plants. Error bars represent the standard error. Statistical differences among the treatments were found with ANOVA ( $F = 9.05$ ,  $p = 0.006$ ) and Tukey's test. Bars with different letters are significantly different ( $p < 0.05$ ).

coverage reduces the chance of capturing reads from each individual in the bulk. In general, genome sequencing coverage should be at least equal to the effective pool size (number of individuals multiplied by the ploidy level) in order to cover rare variants (Magwene et al., 2011; Rellstab et al., 2013). Because sequence coverage of the *H6*-avirulent and -virulent pools (14x and 14.6x) was less than each effective pool size (19 and 23), FET values were relatively low and linkage to *H6* virulence was relatively weak (FET  $p$ -value  $< 0.02$ ). In comparison, *Hdic* sequence coverage (16.5x and 49.2x) was greater than each effective pool size (15 and 33) and linkage was much stronger (FET  $p$ -value  $< 1e-6$ ). Although the *H5* effective pool sizes were higher (48 and 48, due to autosomal diploidy) the sequence coverage of each pool (36.9x and 33.1x) was still relatively high and linkage was also strong (FET  $p$ -value  $< 1e-8$ ).

Recombination rates also impact genetic resolution. *H13*, *H9*, and *H24* virulence are located near the telomeres of HF X chromosomes, where recombination frequency is extremely high. Mapping attempts resolved virulence to each of these *R* genes to a single candidate *Avr* gene. Attempts to map autosomal virulence, where recombination rates are much reduced, either disappointed (Behura et al., 2004), or failed (*H7H8*; Stuart, unpublished). Here, in comparison with the previous approach, BSA-seq efficiently improved the resolution of *H5* virulence on HF autosome A2 using the same mapping population used in a previous investigation (Behura et al., 2004). Recombination rates are typically low near the centromeres. Thus, we were impressed

with how BSA-seq was able to resolve the position of *Hdic* virulence near the X2 centromere.

The power of BSA-seq to map genes with imperfect genomic sequenced maps was evident. Each experiment identified scaffolds that were partially linked to the genes in question, particularly among the “unmapped” scaffolds. Using the more conventional mapping approach, linkage between A1R.66 sequence and *Hdic* virulence would have required a much-improved HF genome assembly. Using BSA-seq, it was possible to detect this linkage in spite of the imperfect assembly.

Mapping *Hdic* virulence strengthened the hypothesis that the HF uses effectors to defeat basal plant immunity. Virulence to five *R* genes in wheat (*H6*, *H9*, *H13*, *H24*, and *Hdic*) has now been associated with one or more candidate *Avr* genes (Table 1). Each of these genes belongs to the “predicted effector” genic fraction of the HF genome (Zhao et al., 2015). Candidate *vH6* and candidate *vH9* are members of the largest family of putative effector-encoding genes (SSGP71) (Zhao et al., 2015). These genes appear to encode E3-ligase mimicking effectors. Candidate *vH24* is a member of another small family that appears to encode secreted phosphatase 2C effectors (Zhao et al., 2016). *vH13* is a unique gene that encodes a highly variable protein. Both candidate *vHdic* genes are members of the putative effector encoding SSGP4 gene family.

The present investigation also provides direct evidence that two *Avr*-encoded proteins have effector functionality in susceptible plant-parasite interactions: candidate-*vH6* and *vH13* suppressed

the HR normally observed in *Nicotiana-B. glumae* interactions. Immune suppression is a well-established component of susceptible wheat-HF interaction. Infested susceptible wheat plants have lowered plant defense-related gene expression and reduced levels of defense-related phytohormones (Liu et al., 2007; Zhu et al., 2010). This inhibition is associated with the rapid development of the plant nutritive cells that are essential for HF larval survival (Harris et al., 2006; Subramanyam et al., 2018). The mechanisms underlying vH6 and vH13 immune suppression remain unknown.

Candidate-*vHdic* failed to suppress HR in the *Nicotiana-B. glumae* infiltration experiments. It is possible that the wrong candidate-*vHdic* was chosen for these experiments. It is also possible that this protein has lost its ability to suppress plant immunity. However, this observation does not eliminate the possibility that it is an effector protein. It simply may target other wheat physiological processes, or its target may not be intracellular. Moreover, Bacteria T3SS-based delivery, like any other heterologous method, has limitations. The machinery for protein synthesis in prokaryotes does not have the ability for post-translational modifications, which limits the analysis to eukaryotic effectors that do not require these modifications (Fabro et al., 2011).

The capacity of *B. glumae* to deliver eukaryotic T3SS-fusion effectors into plant cells expressed in pEDV system has been demonstrated previously and used to identify *M. oryzae* effectors with HR-suppressing effects in rice (Sharma et al., 2013). Suppression of *B. glumae*-induced HR in *N. benthamiana* has been used recently to identify several novel eukaryotic candidate effectors from the fungal pathogens *U. virens* and *L. theobromae* (Zhang et al., 2014), and the root knot nematode *M. incognita* (Shi et al., 2018a; Shi et al., 2018b). Here, we have added the HF to this list of eukaryotic plant pathogens and parasites. We anticipate that this system will be used to test hundreds of potential HF effectors for their effects on plant immunity.

## DATA AVAILABILITY STATEMENT

Whole-genome sequencing data for bulked samples are available at the NCBI Sequence Read Archive (SRA) under the NCBI

Bioproject accession number PRJNA613640 (<https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA613640>).

## AUTHOR CONTRIBUTIONS

LN-E performed BSA-seq and plant immunity experiments, designed experiments, analyzed data, and contributed to writing and editing the manuscript. CZ developed mapping populations, performed genetic mapping using microsatellite markers and analyzed data. RS analyzed data and made conceptual contributions. JS led the project, analyzed data and contributed to writing and editing the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

The authors gratefully acknowledge support for this work provided by USDA-NIFA AFRI Grants 2008-35302-18816 and 2010-03741, Hatch award IND011462 to JS, USDA-ARS Specific Agreement 58-3602-4-010, and a fellowship to JS from Fulbright-Colombia.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge assistance from Phillip San Miguel and the Purdue Genomics Core Facility and technical support from Sue Cambron. The authors thank the two reviewers whose comments and suggestions greatly improved this manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.00956/full#supplementary-material>

## REFERENCES

- Abe, A., Kosugi, S., Yoshida, K., Natsume, S., Takagi, H., Kanzaki, H., et al. (2012). Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat. Biotechnol.* 30, 174–178. doi: 10.1038/nbt.2095
- Aggarwal, R., Benatti, T. R., Gill, N., Zhao, C., Chen, M.-S., Fellers, J. P., et al. (2009). A BAC-based physical map of the Hessian fly genome anchored to polytene chromosomes. *BMC Genomics* 10, 293. doi: 10.1186/1471-2164-10-293
- Aggarwal, R., Subramanyam, S., Zhao, C., Chen, M.-S., Harris, M. O., and Stuart, J. (2014). Avirulence effector discovery in a plant galling and plant parasitic arthropod, the Hessian fly (*Mayetiola destructor*). *PLoS One* 9, e100958. doi: 10.1371/journal.pone.0100958
- Austin, R. S., Vidaurre, D., Stamatou, G., Breit, R., Provart, N. J., Bonetta, D., et al. (2011). Next-generation mapping of *Arabidopsis* genes. *Plant J.* 67, 715–725. doi: 10.1111/j.1365-3113X.2011.04619.x
- Bastide, H., Betancourt, A., Nolte, V., Tobler, R., Stöbe, P., Futschik, A., et al. (2013). A genome-wide, fine-scale map of natural pigmentation variation in *Drosophila melanogaster*. *PLoS Genet.* 9, e1003534. doi: 10.1371/journal.pgen.1003534
- Behura, S. K., Valicente, F. H., Rider, S. D. J. R., Shun-Chen, M., Jackson, S., and Stuart, J. J. (2004). A physically anchored genetic map and linkage to avirulence reveals recombination suppression over the proximal region of Hessian fly chromosome A2. *Genetics* 167, 343–355. doi: 10.1534/genetics.167.1.343
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Brown-Guedira, G. L., Hatchett, J. H., Liu, X. M., Fritz, A. K., Owuoche, J. O., Gill, B. S., et al. (2005). Registration of KS99WGRC42 Hessian fly resistant hard red winter wheat germplasm. *Crop Sci.* 45, 804–805. doi: 10.2135/cropsci2005.0804
- Campeau, E., Ruhl, V. E., Rodier, F., Smith, C. L., Rahmberg, B. L., Fuss, J. O., et al. (2009). A versatile viral system for expression and depletion of proteins in mammalian cells. *PLoS One* 4, e6529. doi: 10.1371/journal.pone.0006529
- Chen, M.-S., Fellers, J. P., Stuart, J. J., Reese, J. C., and Liu, X. (2004). A group of related cDNAs encoding secreted proteins from Hessian fly [*Mayetiola*

- destructor* (Say)] salivary glands. *Insect Mol. Biol.* 13, 101–108. doi: 10.1111/j.1365-2583.2004.00465.x
- Chen, M.-S., Zhao, H.-X., Zhu, Y. C., Scheffler, B., Liu, X., Liu, X., et al. (2008). Analysis of transcripts and proteins expressed in the salivary glands of Hessian fly (*Mayetiola destructor*) larvae. *J. Insect Physiol.* 54, 1–16. doi: 10.1016/j.jinsphys.2007.07.007
- Chen, M.-S., Liu, X., Yang, Z., Zhao, H., Shukle, R. H., Stuart, J. J., et al. (2010). Unusual conservation among genes encoding small secreted salivary gland proteins from a gall midge. *BMC Evol. Biol.* 10, 296. doi: 10.1186/1471-2148-10-296
- da Maia, L. C., Palmieri, D. A., de Souza, V. Q., Kopp, M. M., de Carvalho, F. I. F., and Costa de Oliveira, A. (2008). SSR Locator: tool for simple sequence repeat discovery integrated with primer design and PCR simulation. *Int. J. Plant Genomics* 2008, 412696. doi: 10.1155/2008/412696
- Fabro, G., Steinbrenner, J., Coates, M., Ishaque, N., Baxter, L., Studholme, D. J., et al. (2011). Multiple candidate effectors from the oomycete pathogen *Hyaloperonospora arabidopsidis* suppress host plant immunity. *PLoS Pathog.* 7, e1002348. doi: 10.1371/journal.ppat.1002348
- Garcés-Carrera, S., Knutson, A., Wang, H., Giles, K. L., Huang, F., Whitworth, R. J., et al. (2014). Virulence and biotype analyses of Hessian fly (Diptera: Cecidomyiidae) populations from Texas, Louisiana, and Oklahoma. *J. Econ. Entomol.* 107, 417–423. doi: 10.1603/EC13372
- Gill, B. S., Hatchett, J. H., and Raupp, W. J. (1987). Chromosomal mapping of Hessian fly-resistance gene *H13* in the D genome of wheat. *J. Hered.* 78, 97–100. doi: 10.1093/oxfordjournals.jhered.a110344
- Giovannoni, J. J., Wing, R. A., Ganai, M. W., and Tanksley, S. D. (1991). Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Res.* 19, 6553–6558. doi: 10.1093/nar/19.23.6553
- Harris, M. O., Freeman, T. P., Rohfritsch, O., Anderson, K. G., Payne, S. A., and Moore, J. A. (2006). Virulent Hessian fly (Diptera: Cecidomyiidae) larvae induce a nutritive tissue during compatible interactions with wheat. *Ann. Entomol. Soc. America* 99, 305–316. doi: 10.1603/0013-8746(2006)099[0305:vhfcd]2.0.co;2
- Harris, M. O., Freeman, T. P., Moore, J. A., Anderson, K. G., Payne, S. A., Anderson, K. M., et al. (2010). *H*-gene-mediated resistance to Hessian fly exhibits features of penetration resistance to fungi. *Phytopathology* 100, 279–289. doi: 10.1094/PHYTO-100-3-0279
- Hogenhout, S. A., Van der Hoorn, R. A. L., Terauchi, R., and Kamoun, S. (2009). Emerging concepts in effector biology of plant-associated organisms. *Mol. Plant Microbe Interact.* 22, 115–122. doi: 10.1094/MPMI-22-2-0115
- Jia, J., Zhao, S., Kong, X., Li, Y., Zhao, G., He, W., et al. (2013). *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature* 496, 91–95. doi: 10.1038/nature12028
- Jones, J. D. G., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286
- Kang, Y., Kim, J., Kim, S., Kim, H., Lim, J. Y., Kim, M., et al. (2008). Proteomic analysis of the proteins regulated by *HrpB* from the plant pathogenic bacterium *Burkholderia glumae*. *Proteomics* 8, 106–121. doi: 10.1002/pmic.200700244
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053
- Kobayashi, T., Yamamoto, K., Suetsugu, Y., Kuwazaki, S., Hattori, M., Jairin, J., et al. (2014). Genetic mapping of the rice resistance-breaking gene of the brown planthopper. *Nilaparvata lugens* Proc. R. Soc. B 281. doi: 10.1098/rspb.2014.0726
- Kofler, R., Pandey, R. V., and Schlotterer, C. (2011). PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* 27, 3435–3436. doi: 10.1093/bioinformatics/btr589
- Kong, L., Ohm, H. W., Cambron, S. E., and Williams, C. E. (2005). Molecular mapping determines that Hessian fly resistance gene *H9* is located on chromosome 1A of wheat. *Plant Breed.* 124, 525–531. doi: 10.1111/j.1439-0523.2005.01139.x
- Kong, L., Cambron, S. E., and Ohm, H. W. (2008). Hessian fly resistance genes *H16* and *H17* are mapped to a resistance gene cluster in the distal region of chromosome 1AS in wheat. *Mol. Breed.* 21, 183–194. doi: 10.1007/s11032-007-9119-5
- Lee, E., Helt, G. A., Reese, J. T., Munoz-Torres, M. C., Childers, C. P., Buels, R. M., et al. (2013). Web Apollo: a web-based genomic annotation editing platform. *Genome Biol.* 14, R93. doi: 10.1186/gb-2013-14-8-r93
- Leshchiner, I., Alexa, K., Kelsey, P., Adzhubei, I., Austin-Tse, C. A., Cooney, J. D., et al. (2012). Mutation mapping and identification by whole-genome sequencing. *Genome Res.* 22, 1541–1548. doi: 10.1101/gr.135541.111
- Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–595. doi: 10.1093/bioinformatics/btp698
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Liu, X. M., Brown-Guedira, G. L., Hatchett, J. H., Owuochi, J. O., and Chen, M.-S. (2005a). Genetic characterization and molecular mapping of a Hessian fly-resistance gene transferred from *T. turgidum* ssp. *dicoccum* to common wheat. *Theor. Appl. Genet.* 111, 1308–1315. doi: 10.1007/s00122-005-0059-3
- Liu, X. M., Fritz, A. K., Reese, J. C., Wilde, G. E., Gill, B. S., and Chen, M.-S. (2005b). *H9*, *H10*, and *H11* compose a cluster of Hessian fly-resistance genes in the distal gene-rich region of wheat chromosome 1AS. *Theor. Appl. Genet.* 110, 1473–1480. doi: 10.1007/s00122-005-1982-z
- Liu, X. M., Gill, B. S., and Chen, M.-S. (2005c). Hessian fly resistance gene *H13* is mapped to a distal cluster of resistance genes in chromosome 6DS of wheat. *Theor. Appl. Genet.* 111, 243–249. doi: 10.1007/s00122-005-2009-5
- Liu, X., Bai, J., Huang, L., Zhu, L., Liu, X., Weng, N., et al. (2007). Gene expression of different wheat genotypes during attack by virulent and avirulent Hessian fly (*Mayetiola destructor*) larvae. *J. Chem. Ecol.* 33, 2171–2194. doi: 10.1007/s10886-007-9382-2
- Magwene, P. M., Willis, J. H., and Kelly, J. K. (2011). The statistics of bulk segregant analysis using next generation sequencing. *PLoS Comput. Biol.* 7, e1002255. doi: 10.1371/journal.pcbi.1002255
- Michelmore, R. W., Paran, I., and Kesseli, R. V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci. U. S. A.* 88, 9828–9832. doi: 10.1073/pnas.88.21.9828
- Miranda, L. M., Bland, D. E., Cambron, S. E., Lyster, J. H., Johnson, J., Buntin, G. D., et al. (2010). Genetic mapping of an *Aegilops tauschii*-derived Hessian fly resistance gene in common wheat. *Crop Sci.* 50, 612. doi: 10.2135/cropsci2009.05.0278
- Patterson, F. L., Gallun, R. L., Roberts, J. J., Finney, R. E., and Shaner, G. E. (1975). Registration of Arthur 71 and Abe wheat 1 (Reg. Nos. 560 and 562). *Crop Sci.* 15, 736–736. doi: 10.2135/cropsci1975.0011183x001500050042x
- Patterson, F. L., Ohm, H. W., Shaner, G. E., Finney, R. E., Gallun, R. L., Roberts, J. J., et al. (1982). Registration of Caldwell wheat 1 (Reg. No. 660). *Crop Sci.* 22, 691–692. doi: 10.2135/cropsci1982.0011183X002200030081x
- Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi: 10.1038/nmeth.1701
- Pomraning, K. R., Smith, K. M., and Freitag, M. (2011). Bulk segregant analysis followed by high-throughput sequencing reveals the *Neurospora* cell cycle gene, *ndc-1*, to be allelic with the gene for ornithine decarboxylase, *spe-1*. *Eukaryot. Cell* 10, 724–733. doi: 10.1128/EC.00016-11
- Raupp, W. J., Amri, A., Hatchett, J. H., Gill, B. S., Wilson, D. L., and Cox, T. S. (1993). Chromosomal location of Hessian fly-resistance genes *H22*, *H23*, and *H24* derived from *Triticum tauschii* in the D genome of wheat. *J. Hered.* 84, 142–145. doi: 10.1093/oxfordjournals.jhered.a111300
- Redmond, S. N., Eiglmeier, K., Mitri, C., Markianos, K., Guelbeogo, W. M., Gneme, A., et al. (2015). Association mapping by pooled sequencing identifies TOLL 11 as a protective factor against *Plasmodium falciparum* in *Anopheles gambiae*. *BMC Genomics* 16, 779. doi: 10.1186/s12864-015-2009-z
- Relstab, C., Zoller, S., Tedder, A., Gugerli, F., and Fischer, M. C. (2013). Validation of SNP allele frequencies determined by pooled next-generation sequencing in natural populations of a non-model plant species. *PLoS One* 8, e80422. doi: 10.1371/journal.pone.0080422
- Rolny, N., Costa, L., Carrión, C., and Guaiamet, J. J. (2011). Is the electrolyte leakage assay an unequivocal test of membrane deterioration during leaf senescence? *Plant Physiol. Biochem.* 49, 1220–1227. doi: 10.1016/j.plaphy.2011.06.010

- Saitoh, H., and Terauchi, R. (2013). *Burkholderia glumae* competent cells preparation and transformation. *BIO-PROTOCOL* 3, e985. doi: 10.21769/bioprotoc.985
- Sardesai, N., Nemacheck, J. A., Subramanyam, S., and Williams, C. E. (2005). Identification and mapping of *H32*, a new wheat gene conferring resistance to Hessian fly. *Theor. Appl. Genet.* 111, 1167–1173. doi: 10.1007/s00122-005-0048-6
- Schneeberger, K. (2014). Using next-generation sequencing to isolate mutant genes from forward genetic screens. *Nat. Rev. Genet.* 15, 662–676. doi: 10.1038/nrg3745
- Shands, R. G., and Cartwright, W. B. (1953). A fifth gene conditioning Hessian fly response in common wheat 1. *Agron. J.* 45, 302–307. doi: 10.2134/agronj1953.00021962004500070007x
- Sharma, S., Sharma, S., Hirabuchi, A., Yoshida, K., Fujisaki, K., Ito, A., et al. (2013). Deployment of the *Burkholderia glumae* type III secretion system as an efficient tool for translocating pathogen effectors to monocot cells. *Plant J.* 74, 701–712. doi: 10.1111/tpj.12148
- Shi, Q., Mao, Z., Zhang, X., Ling, J., Lin, R., Zhang, X., et al. (2018a). The novel secreted *Meloidogyne incognita* effector MiISE6 targets the host nucleus and facilitates parasitism in *Arabidopsis*. *Front. Plant Sci.* 9, 252. doi: 10.3389/fpls.2018.00252
- Shi, Q., Mao, Z., Zhang, X., Zhang, X., Wang, Y., Ling, J., et al. (2018b). A *Meloidogyne incognita* effector MiISE5 suppresses programmed cell death to promote parasitism in host plant. *Sci. Rep.* 8, 7256. doi: 10.1038/s41598-018-24999-4
- Sohn, K. H., Lei, R., Nemri, A., and Jones, J. D. G. (2007). The downy mildew effector proteins ATR1 and ATR13 promote disease susceptibility in *Arabidopsis thaliana*. *Plant Cell* 19, 4077–4090. doi: 10.1105/tpc.107.054262
- Sohn, K. H., Segonzac, C., Rallapalli, G., Sarris, P. F., Woo, J. Y., Williams, S. J., et al. (2014). The nuclear immune receptor RPS4 is required for RRS1SLH1-dependent constitutive defense activation in *Arabidopsis thaliana*. *PLoS Genet.* 10, e1004655. doi: 10.1371/journal.pgen.1004655
- Stuart, J. J., Schulte, S. J., Hall, P. S., and Mayer, K. M. (1998). Genetic mapping of Hessian fly avirulence gene *vH6* using bulked segregant analysis. *Genome* 41, 702–708. doi: 10.1139/g98-071
- Stuart, J., Aggarwal, R., and Schemerhorn, B. (2014). Hessian Flies (Diptera). In *Protocols for Cytogenetic Mapping of Arthropod Genomes, 1st ed.* Ed. I. Sharakhov (Boca Raton: CRC Press), 63–78. doi: 10.1201/b17450-3
- Stuart, J. (2015). Insect effectors and gene-for-gene interactions with host plants. *Curr. Opin. Insect Sci.* 9, 56–61. doi: 10.1016/j.cois.2015.02.010
- Subramanyam, S., Shreve, J. T., Nemacheck, J. A., Johnson, A. J., Schemerhorn, B., Shukle, R. H., et al. (2018). Modulation of nonessential amino acid biosynthetic pathways in virulent Hessian fly larvae (*Mayetiola destructor*), feeding on susceptible host wheat (*Triticum aestivum*). *J. Insect Physiol.* 105, 54–63. doi: 10.1016/j.jinsphys.2018.01.001
- Swinnen, S., Schaerlaekens, K., Pais, T., Claesen, J., Hubmann, G., Yang, Y., et al. (2012). Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis. *Genome Res.* 22, 975–984. doi: 10.1101/gr.131698.111
- Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., et al. (2013). QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* 74, 174–183. doi: 10.1111/tpj.12105
- Thorvaldsdóttir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.* 14, 178–192. doi: 10.1093/bib/bbs017
- Wang, Z., Ge, J.-Q., Chen, H., Cheng, X., Yang, Y., Li, J., et al. (2018). An insect nucleoside diphosphate kinase (NDK) functions as an effector protein in wheat - Hessian fly interactions. *Insect Biochem. Mol. Biol.* 100, 30–38. doi: 10.1016/j.ibmb.2018.06.003
- Yan, J. Y., Zhao, W. S., Chen, Z., Xing, Q. K., Zhang, W., Chethana, K. W. T., et al. (2018). Comparative genome and transcriptome analyses reveal adaptations to opportunistic infections in woody plant degrading pathogens of Botryosphaeriaceae. *DNA Res.* 25, 87–102. doi: 10.1093/dnares/dsx040
- Zantoko, L., and Shukle, R. H. (1997). Genetics of virulence in the Hessian fly to resistance gene *H13* in wheat. *J. Hered.* 88, 120–123. doi: 10.1093/oxfordjournals.jhered.a023069
- Zhang, Y., Zhang, K., Fang, A., Han, Y., Yang, J., Xue, M., et al. (2014). Specific adaptation of *Ustilagoideae virens* in occupying host florets revealed by comparative and functional genomics. *Nat. Commun.* 5, 3849. doi: 10.1038/ncomms4849
- Zhao, C., Escalante, L. N., Chen, H., Benatti, T. R., Qu, J., Chellapilla, S., et al. (2015). A massive expansion of effector genes underlies gall-formation in the wheat pest *Mayetiola destructor*. *Curr. Biol.* 25, 613–620. doi: 10.1016/j.cub.2014.12.057
- Zhao, C., Shukle, R., Navarro-Escalante, L., Chen, M., Richards, S., and Stuart, J. J. (2016). Avirulence gene mapping in the Hessian fly (*Mayetiola destructor*) reveals a protein phosphatase 2C effector gene family. *J. Insect Physiol.* 84, 22–31. doi: 10.1016/j.jinsphys.2015.10.001
- Zhu, L., Liu, X., and Chen, M.-S. (2010). Differential accumulation of phytohormones in wheat seedlings attacked by avirulent and virulent Hessian fly (Diptera: Cecidomyiidae) larvae. *J. Econom. Entomol.* 103, 178–185. doi: 10.1603/ec09224

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a shared affiliation with one of the authors CZ at the time of the review.

Copyright © 2020 Navarro-Escalante, Zhao, Shukle and Stuart. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Juvenile Spider Mites Induce Salicylate Defenses, but Not Jasmonate Defenses, Unlike Adults

Jie Liu<sup>1,2</sup>, Saioa Legarrea<sup>1</sup>, Juan M. Alba<sup>1</sup>, Lin Dong<sup>1</sup>, Rachid Chafi<sup>1</sup>, Steph B. J. Menken<sup>1</sup> and Merijn R. Kant<sup>1\*</sup>

<sup>1</sup> Evolutionary and Population Biology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, Netherlands, <sup>2</sup> State Key Laboratory of Rice Biology & Ministry of Agriculture Key Lab of Molecular Biology of Crop Pathogens and Insects, Institute of Insect Sciences, Zhejiang University, Hangzhou, China

## OPEN ACCESS

### Edited by:

Brigitte Mauch-Mani,  
Université de Neuchâtel, Switzerland

### Reviewed by:

Isabel Diaz,  
Polytechnic University of Madrid,  
Spain  
Bettina Hause,  
Leibniz-Institut für Pflanzenbiochemie  
(IPB), Germany

### \*Correspondence:

Merijn R. Kant  
m.kant@uva.nl

### Specialty section:

This article was submitted to  
Plant Pathogen Interactions,  
a section of the journal  
Frontiers in Plant Science

**Received:** 06 March 2020

**Accepted:** 16 June 2020

**Published:** 10 July 2020

### Citation:

Liu J, Legarrea S, Alba JM, Dong L, Chafi R, Menken SBJ and Kant MR (2020) Juvenile Spider Mites Induce Salicylate Defenses, but Not Jasmonate Defenses, Unlike Adults. *Front. Plant Sci.* 11:980. doi: 10.3389/fpls.2020.00980

When plants detect herbivores they strengthen their defenses. As a consequence, some herbivores evolved the means to suppress these defenses. Research on induction and suppression of plant defenses usually makes use of particular life stages of herbivores. Yet many herbivorous arthropods go through development cycles in which their successive stages have different characteristics and lifestyles. Here we investigated the interaction between tomato defenses and different herbivore developmental stages using two herbivorous spider mites, i.e., *Tetranychus urticae* of which the adult females induce defenses and *T. evansi* of which the adult females suppress defenses in *Solanum lycopersicum* (tomato). First, we monitored egg-to-adult developmental time on tomato wild type (WT) and the mutant *defenseless-1* (*def-1*, unable to produce jasmonate-(JA)-defenses). Then we assessed expression of salivary effector genes (effector 28, 84, *SHOT2b*, and *SHOT3b*) in the consecutive spider mite life stages as well as adult males and females. Finally, we assessed the extent to which tomato plants upregulate JA- and salicylate-(SA)-defenses in response to the consecutive mite developmental stages and to the two sexes. The consecutive juvenile mite stages did not induce JA defenses and, accordingly, egg-to-adult development on WT and *def-1* did not differ for either mite species. Their eggs however appeared to suppress the SA-response. In contrast, all the consecutive feeding stages upregulated SA-defenses with the strongest induction by *T. urticae* larvae. Expression of effector genes was higher in the later developmental stages. Comparing expression in adult males and females revealed a striking pattern: while expression of effector 84 and *SHOT3b* was higher in *T. urticae* females than in males, this was the opposite for *T. evansi*. We also observed *T. urticae* females to upregulate tomato defenses, while *T. evansi* females did not. In addition, of both species also the males did not upregulate defenses. Hence, we argue that mite ontogenetic niche shifts and stage-specific composition of salivary secreted proteins probably together determine the course and efficiency of induced tomato defenses.

**Keywords:** life cycle, ontogenetic niche shift, plant defense, effector, suppression, induction, spider mite, tomato

## INTRODUCTION

Plants possess multilayered defenses against herbivores. These defenses may be constitutively present or be induced upon attack and serve to limit damage inflicted by the herbivore (Walling, 2000). Induced defenses include morphological reinforcements and accumulation of toxins and inhibitors of herbivore food digestion (Kessler and Baldwin, 2002). In addition, plants sometimes also establish so-called indirect defenses by attracting and/or arresting foraging predators or host seeking parasitoids, e.g., *via* the production of volatile attractants or the provision of shelter or alternative food (Sabelis, 1999; Sabelis et al., 2001). These defenses are regulated mainly by two central phytohormones: (a) jasmonic acid (JA) which orchestrates the defenses against herbivores (Howe and Jander, 2008) and necrotrophic pathogens (Glazebrook, 2005) and (b) salicylic acid (SA) which primarily organizes defenses against biotrophic pathogens and phloem-feeding herbivores (Kaloshian and Walling, 2005). The actions of these two central hormones are fine-tuned by a suite of ancillary hormones and their interplay is tightly linked to the local biotic and abiotic conditions, the plant's developmental stage and the particular tissues being attacked. SA- and JA-dependent responses were often—but not always—found to act antagonistically (Mur et al., 2006) and this was suggested to reflect an adaptive tailoring of distinct defenses against distinct attackers (Thaler et al., 2012). Feeding activities by several herbivores, e.g., aphids, whiteflies, and spider mites are known to induce both JA- and SA-dependent defense pathways (Moran and Thompson, 2001; Ament et al., 2004; Zhang et al., 2013; Cao et al., 2014). However, some herbivores can suppress the induction of plant defenses (Musser et al., 2002; Zarate et al., 2007; Kant et al., 2008; Kant et al., 2015). The generalist spider mite *Tetranychus urticae* Santpoort-2 has been shown to induce both JA- and SA-regulated defenses and produces a lower number of eggs on tomato WT plants than on JA-biosynthesis mutant *defenseless-1 (def-1)* (Li et al., 2002; Kant et al., 2008; Alba et al., 2015; Staudacher et al., 2017). In contrast, the spider mite *T. evansi* Viçosa-1 was found to suppress the induced defenses of tomato plants (Sarmiento et al., 2011b; Alba et al., 2015; Schimmel et al., 2018). However, suppression brings opportunities for non-suppressor mites to benefit from the lowered defenses when feeding on the same patch (Kant et al., 2008; Sarmiento et al., 2011a; Glas et al., 2014; Alba et al., 2015; Schimmel et al., 2017a; Schimmel et al., 2017b) giving rise to complex community interactions (Blaazer et al., 2018).

Spider mites (Acari: Tetranychidae) are stylet-feeding arthropods. Unfertilized females can produce male offspring through arrhenotokous parthenogenesis, but when fertilized their offspring is a mixture of both sexes (Wrensch, 1985; Carrière, 2003). They use their stylets to pierce plant cells, predominantly parenchyma, and to inject saliva in their host. Subsequently, they ingest and digest the cell contents (Tomczyk and Kropczynska, 1985; Bensoussan et al., 2018), which leads to visible chlorotic spots on the leaf surface of the plant (Kant et al., 2004; Bensoussan et al., 2016). The two-spotted spider mite, *T. urticae*, is highly polyphagous and can be found on numerous

host-plant species (Helle and Sabelis, 1985; Dermauw et al., 2012). Due to its high reproductive output (around 5–15 eggs per day, mostly depending on temperature, female age and host quality); its short generation cycle (around 14 d from egg to adult, mostly depending on temperature); its ability to rapidly adapt to novel hosts (>1,000 species recorded) and its ability to develop resistance to pesticides rapidly, this mite causes significant damage to crops worldwide (Fry, 1989; Agrawal, 2000; Van Leeuwen et al., 2010). On the contrary, *T. evansi* is more specialized and feeds predominantly on Solanaceae. It is widely present in South America and became invasive in Africa in the 1970s and, more recently, also in Europe (Ferragut et al., 2013). It is a threat to tomato cultivation as no biological control agents are available to control it (Sarmiento et al., 2011b; Navajas et al., 2013).

How plants perceive spider mites and mount specific defenses is still largely unclear. First, plants may respond to the mechanical stress due to spider mite feeding and the subsequent collapse of host cells (Bensoussan et al., 2016). Mechanical injury is well known for eliciting repair and defense responses (Mithöfer et al., 2005; Duran-Flores and Heil, 2016). Second, plants may respond to spider mite egg-deposition as has been demonstrated for the eggs of dipteran (Hilker et al., 2002; Bittner et al., 2017), lepidopteran (Fatouros et al., 2015), and coleopteran (Doss et al., 2000) insects, and was shown to sometimes benefit the insect (Hilker and Fatouros, 2015; Hilker and Fatouros, 2016). Third, plants may respond to spider mite secretions such as silk (Grbic et al., 2011; Doğan et al., 2017), feces (Santamaria et al., 2015), and especially the saliva they inject into host cells during feeding, reminiscent of herbivorous insects (Howe and Jander, 2008; Maffei et al., 2012). The saliva of *T. urticae* (Jonckheere et al., 2016) and *T. evansi* (Huang et al., 2019) contains roughly 100 proteins. A family of 13 secreted salivary *T. urticae* proteins, referred to as SHOT, was shown to exhibit strong host-dependent transcriptional plasticity (Jonckheere et al., 2018). Moreover, two additional secreted spider mite proteins, referred to as tetranins, were shown to upregulate plant defenses (Iida et al., 2019). In contrast, two salivary proteins, referred to as effector 28 and 84, were shown to suppress plant SA (Villaruel et al., 2016) and JA defenses (Schimmel et al., 2017b). How these proteins cause their effects on plants is still unknown but it has been suggested that plant receptor-like proteins may play a central role in the recognition of spider mite feeding (Zhurov et al., 2014; Santamaria et al., 2019).

The ontogenetic niche concept of Werner and Gilliam (1984) states that the use of resources of an organism depends on its developmental stage. It follows that if such resource is another organism, the ontogenetic niche shift of one may modulate the response of the other. For example, plants may respond differently to the consecutive life stages of a herbivore. The spider mite starts its life-cycle like an egg followed by four feeding stages: larva, protonymph (first nymphal stage), deutonymph (second nymphal stage), and finally the adult and these can be male or female. These stages obviously differ not only in size and morphology (Sabelis, 1985) but also in the

amount of food they need and the plant tissue or cell types they are able to utilize. In addition, the stylet of juvenile spider mites may be too shallow for reaching the palisade parenchyma (the cell type mites prefer to eat) especially when residing on the abaxial (lower) leaf surface (Bandurski et al., 1953; Bensoussan et al., 2018). Another clear difference is that adult females need to eat enough to produce eggs (roughly half of their body weight per day) while males (roughly eight times smaller than females) (Mitchell, 1973) and juveniles do not.

Spider mites are small ( $\leq 0.5$  mm) yet the adult females can be seen by the experienced naked eye; they are easy to distinguish from the other stages and are easier to handle than the smaller stages. In addition, the eggs laid by (young) females are considered a reliable proxy for host-plant quality; for mite population growth and for mite fitness. In standardized experiments on plant-mite interactions therefore (young) adult females are often used as representatives of the species as a whole (Li et al., 2002; Alba et al., 2015). Here we tested the robustness of this explicit assumption by monitoring the responses of the different spider mite developmental-stages to plant defenses as well as the cumulative responses of the plant to these consecutive developmental stages, similar to what will happen under natural conditions during the early stages of host colonization. We first followed the duration of the developmental stages of the two most common mite phenotypes on tomato: the first being maladapted to tomato and an inducer of tomato defenses (*T. urticae* Santpoort-2) and the second being adapted to tomato and a suppressor of tomato defenses (*T. evansi* Viçosa-1) (Alba et al., 2015) on WT tomato and on the mutant *def-1*. Subsequently, we submitted tomato plants to mite eggs and assessed the plant's cumulative defense response, in tandem with the mite's effector gene-expression, during the course of the mite's development into adulthood. Finally, we compared defenses induced by young adult males with those induced by young adult females and assessed effector-gene expression in the adult sexes.

## MATERIAL AND METHODS

### Plants and Mites

Seeds of tomato *Solanum lycopersicum* cv. Castlemart (WT) and jasmonate acid (JA) biosynthesis mutant *defenseless-1* (*def-1*, which is in the genetic background of cv. Castlemart) were germinated and grown in the greenhouse at 25°C, L16:D8 h, 50–60% relative humidity (RH). Three days before performing experiments, plants were transferred to a climate room (25°C, L16:D8 h, 60% RH, 300  $\mu\text{mol m}^{-2}\text{s}^{-1}$ ). The two-spotted spider mite *T. urticae* Santpoort-2 (for a detailed description of this strain see: Alba et al., 2015) was maintained on detached leaves of *Phaseolus vulgaris* cv. Speedy in a climate room (25°C, L16:D8 h, 60% RH, 300  $\mu\text{mol m}^{-2}\text{s}^{-1}$ ). The red spider mite *T. evansi* Viçosa-1 (for a detailed description of this strain see: Alba et al., 2015) was maintained on detached leaves of *S. lycopersicum* cv. Castlemart placed on wet cotton wool in a climate room (25°C, L16:D8 h, 60% RH, 300  $\mu\text{mol m}^{-2}\text{s}^{-1}$ ).

### Developmental Time, Survival, and Sex Ratio of *T. urticae* and *T. evansi* on WT and *def-1* Tomato

Developmental time, survival, and mite sex ratio were determined using single mites on leaf discs. Leaf discs (15 mm in diameter) were obtained from the leaflets of 28-d-old WT and *def-1* tomato plants using a metal hole puncher. The leaf disks were placed gently (with their adaxial side up) on a wet sponge covered with wet cotton wool in a plastic tray half filled with water. Leaf disks were infested with a single egg. These had been obtained by first habituating gravid females of *T. urticae* Santpoort-2, and *T. evansi* Viçosa-1, randomly taken from the mass rearings, on intact WT and *def-1* plants for 72 h. Then single habituated females were placed on a leaf disc and allowed to produce eggs for 12 h. From each leaf disk we removed all the females and removed all eggs except one. Subsequently we monitored each of these single eggs per leaf disk for egg hatching (egg survival) and survival and development of the feeding mite stages per disc were recorded twice per day at 8- and 16-h intervals until the mites reached adulthood or died. The developmental stage was determined by observing the shed skin of the previous life stage. We recorded the sex of the adults. For each of the four treatments we monitored 100 individual mites (i.e., 100 leaf discs). After 7 d, mites were transferred to a fresh leaf disc. This experiment was repeated three times independently in time, and the data were pooled for analysis. Developmental time was analyzed per life stage comparing WT and *def-1* data for the two mite species separately using the Student's t-test. The fraction of eggs that made it to adulthood on WT compared to *def-1* was determined after 384 h and the fraction of females relative to males among these adults on WT compared to *def-1* were analyzed separately for the two mite species after arcsine square root transformation using the Student's t-test in IBM SPSS Statistics 25.

### Collection of Mite and Tomato Material for Gene Expression Analysis

To obtain the material for simultaneous isolation of mite RNA and plant RNA we sampled leaflets of tomato plants infested with the consecutive mite life stages (Schimmel et al., 2017a). We monitored the effect of each developmental stage on tomato defense gene expression as a cumulative effect, i.e., we included the effect(s) of the previous stage(s) by infesting plants with eggs and sampling leaflets at the end of each of the consecutive developmental stages, i.e., at the end of the egg, larval, protonymph, and deutonymph stage and at the 2-d-old adult stage. One day before starting the experiment, we took random females from the mite rearing to put on new leaflets to collect their eggs. The next day we transferred 50 eggs to the second nonterminal leaflet of the third fully expanded leaf of 28-d-old WT plants using a soft brittle paintbrush. Control plants were touched 50 times in a similar manner with a clean brush. Lanolin was put around the petiole of the leaflets of control and infested plants to prevent mites from escaping during the course of the experiment. To determine the transition of one mite stage into

the other we used a parallel “experiment” on leaf disks. We prepared 60 leaf disks (with their adaxial side up) on a wet sponge covered with wet cotton wool in a plastic tray half filled with water and placed on each disc one egg. The disks were observed twice per day and the first mite on these disks that entered the next developmental stage—as shown by their shed skin—determined the moment we sampled the intact plants. Doing so we reasoned that we would sample the intact plants at the end of each mite developmental phase under the assumption that mite development on disks and intact leaflets is similar. The disks were observed until the mites had reached adulthood. Per developmental stage we sampled the infested leaflet of five plants (five distinct biological samples), and in parallel we sampled an uninfested leaflet of five control plants (five distinct biological samples) for each stage. This experiment was repeated four times independently in time. For sampling leaves induced by one of the two adult sexes we used a different protocol based on Alba et al. (2015). Briefly, eggs were allowed to hatch on intact plants and we waited until adults were 16 d old after oviposition. We then placed 15 adult mites, either males or females, to the second nonterminal leaflet of the third fully expanded leaf of 28-d-old WT plants using a soft brittle paintbrush and sampled these after 2 d of infestation. Per adult sex we sampled the infested leaflet of five plants (five distinct biological samples), and in parallel we sampled one uninfested leaflet from five control plants (five distinct biological samples) per stage. This experiment was repeated two times independently in time. All samples were flash frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until we extracted mRNA.

## Expression Analysis of Mite and Tomato Genes

RNA isolation, cDNA synthesis, and assessed transcript accumulation by means of RT-qPCR were performed as described in Kant et al. (2004) and Alba et al. (2015) using the protocol of Verdonk et al. (2003). For *PI-IIc* (SGN Solyc03g020050.2), *PR-1a* (SGN Solyc09g007010.1), *Actin* (SGN Solyc03g078400.2), *RP49* (GenBank XM\_015934205.2), 84 of *T. evansi* (GenBank KT182961), and *T. urticae* (GenBank XM\_015936396.2) and for effector 28 of *T. urticae* (GenBank XM\_025162299.1) we used the same primers as in Schimmel et al. (2017a). For *SHOT2b* of *T. urticae* (GenBank XM\_015940069.1) we used the following primers: Fw GATCTTCGCCGAAAACAAT and Rev TCATCTTCATGAACATTAGATTGA. For *SHOT3b* of *T. urticae* (GenBank XM\_015931098.1) we used the following primers: Fw TCGCCTCAACTGGAGCTT and Rev AGCAAGAGATGAACCGATTG. For *SHOT3b* of *T. evansi* (GenBank MH979735.1) we used the following primers: Fw GAAAATGGAGTCGCAACTGTC and Rev. ACCGAAAGTTGATAGGACACC. Quantitative PCR reactions were performed on each sample twice (two technical replicates per sample). The expression value per sample was calculated as the average of the two technical replicates. Expression was normalized using the tomato housekeeping gene *Actin* for all qPCRs because the expression of the mite housekeeping gene *RP49* varied too much during spider mite development (see *Results*). Expression was also

corrected for mite survival. For the figures, the normalized transcript abundances were scaled by dividing all values including standard errors by the lowest average value (setting the latter to 1 in a data neutral manner). Data were analyzed by means of a generalized linear model, assuming gamma distribution and a log link function. The independent time points at which experiments were repeated were used as random factor in the analysis. Means of each group were compared by LSD *post hoc* test in IBM SPSS Statistics 25.

## RESULTS

### Marginal Effects of JA-Defenses on Developmental Times of Consecutive Spider Mite Life-Stages

To assess the extent to which stage-specific developmental times of inducer and suppressor mites were affected by JA-dependent defense, we monitored the duration of the larval and the two nymphal stages of *T. urticae* and *T. evansi* males and females on leaf disks of Castlemart tomato plants (WT) and on disks of the JA-biosynthesis mutant *def-1*. The overall developmental time from egg to adult WT and *def-1* did not differ significantly for *T. urticae* Santpoort-2 (**Table 1**;  $t = 0.31$ ,  $P = 0.76$ ) or for *T. evansi* (**Table 1**;  $t = 0.882$ ,  $P = 0.40$ ) and such differences were also not seen when analyzing males and females separately (*T. urticae* Santpoort-2 females:  $t = -0.582$ ,  $P = 0.56$ ; males  $t = 1.157$ ,  $P = 0.25$ ; for *T. evansi* Viçosa-1 females:  $t = 0.562$ ,  $P = 0.58$ ; males:  $t = 0.686$ ,  $P = 0.50$ ). We did also not observe clear differences across mite species at the level of developmental stages. *T. urticae* Santpoort-2 did not exhibit significantly different developmental times for any of the stages or of the sexes on either WT or *def-1* (larva female:  $t = -1.269$ ,  $P = 0.21$ ; larva male:  $t = 0.577$ ,  $P = 0.57$ ; protonymph female:  $t = -1.179$ ,  $P = 0.24$ ; protonymph male:  $t = 0.811$ ,  $P = 0.42$ ; deutonymph female:  $t = -1.074$ ,  $P = 0.29$ ). For *T. evansi* the female protonymph stage lasted longer on WT ( $t = -2.216$ ,  $P = 0.03$ ). Interestingly, the developmental times of all nymphal stages of *T. evansi* males were significantly shorter on WT (**Table 1**; protonymph:  $t = 3.118$ ,  $P = 0.003$ ; deutonymph:  $t = 0.2873$ ,  $P = 0.006$ ). Also egg-to-adult survival was similar across the treatments ( $F = 1.950$ ,  $P = 0.159$ ). Finally, the sex ratio did not significantly differ across the treatments (**Table 1**;  $F_{5,12} = 0.43$ ,  $P = 0.819$ ).

### Feeding Juvenile Spider Mites Induce SA-, but No JA-, Responses in Tomato

To assess whether tomato plants respond differently to different spider mite life stages we infested tomato leaflets with 50 spider mite eggs and monitored the expression of tomato genes *PI-IIc* and *PR-1a* during the course of the development of the mites from egg to adult (**Figure 1**). Overall *PI-IIc* expression was significantly affected by mite infestation (Wald  $\chi^2 = 54.216$ ;  $P < 0.001$ ). The manually deposited egg batches of either *T. urticae* Santpoort-2 or *T. evansi* Viçosa-1 did not significantly

**TABLE 1** | Cumulative duration of the developmental stages, the egg-to-adult survival and the sex ratio of *Tetranychus urticae* Santpoort-2 and *T. evansi* Viçosa-1 on WT and *def-1* tomato plants.

Treatment	Sex	Larva (hrs.)	Protonymph (hrs.)	Deutonymph (hrs.)	Adult (hrs.)	Fraction eggs reaching adulthood	Fraction adult females
<i>T. urticae</i> Santpoort-2							
<i>def-1</i>	♀+♂	147.1 ± 2.4 a	214.3 ± 2.2 a	259.5 ± 2.0 a	318.9 ± 2.7 a	0.5 ± 0.3 a	0.6 ± 0.02 a
WT	♀+♂	149.6 ± 2.6 a	216.1 ± 3.1 a	260.9 ± 2.9 a	317.5 ± 3.4 a	0.5 ± 0.2 a	0.6 ± 0.02 a
<i>def-1</i>	♀	144.9 ± 3.5 a	212.0 ± 3.3 a	258.0 ± 2.7 a	317.8 ± 3.9 a		
WT	♀	152.3 ± 4.7 a	219.2 ± 5.7 a	264.3 ± 5.2 a	321.9 ± 6.0 a		
<i>def-1</i>	♂	150.0 ± 3.1 a	217.2 ± 2.9 a	261.4 ± 2.8 a	320.2 ± 3.6 a		
WT	♂	147.6 ± 2.7 a	213.7 ± 3.2 a	258.3 ± 3.3 a	314.1 ± 3.8 a		
<i>T. evansi</i> Viçosa-1							
<i>def-1</i>	♀+♂	153.0 ± 1.9 a	213.9 ± 2.2 a	256.3 ± 2.6 a	315.5 ± 2.6 a	0.6 ± 0.2 a	0.7 ± 0.11 a
WT	♀+♂	152.4 ± 1.2 a	213.2 ± 1.5 a	251.4 ± 1.8 a	312.6 ± 2.4 a	0.7 ± 0.2 a	0.7 ± 0.08 a
<i>def-1</i>	♀	149.4 ± 2.1 a	209.0 ± 2.1 a	252.4 ± 2.8 a	317.4 ± 3.1 a		
WT	♀	152.2 ± 1.5 a	215.0 ± 1.7 b	253.0 ± 1.9 a	315.0 ± 2.8 a		
<i>def-1</i>	♂	162.5 ± 3.7 a	227.1 ± 5.1 a	266.8 ± 5.4 a	311.4 ± 4.6 a		
WT	♂	152.8 ± 1.8 a	208.8 ± 3.0 b	247.5 ± 4.0 b	306.3 ± 4.7 a		

The columns "Larva", "Protonymph", "Deutonymph" and "Adult" indicate the average duration in hours (hrs) it took to reach these stages from the start of the experiment. This experiment was conducted three times independently, each time starting with 100 eggs, each on a single leaf disc, per mite species per plant genotype. "Fraction eggs reaching adulthood" was calculated as the fraction of living adults after 384 h relative to the number of eggs that had been submitted to the test. "Fraction female" refers to the sex ratio expressed as the fraction of adult females. Statistics were applied to *def-1* and WT data pairs in each column using Student's *t*-test at  $\alpha = 0.05$  and data pairs marked with the same letter are not significantly different.

affect the expression of *PI-IIc* (Figure 1A). Subsequently, the larvae of *T. evansi* Viçosa-1, but not those of *T. urticae* Santpoort-2, downregulated *PI-IIc* expression. However, expression of *PI-IIc* remained near control levels during all subsequent developmental stages of both mite species until the adult stage was reached. As adults, only *T. urticae* Santpoort-2, but not *T. evansi* Viçosa-1, upregulated *PI-IIc*. In contrast to *PI-IIc*, the manually deposited eggs of *T. urticae* Santpoort-2 as well as *T. evansi* Viçosa-1 downregulated expression of *PR-1a* relative to control plants (Figure 1B). However, all feeding stages of both species upregulated *PR-1a* expression but *T. urticae* Santpoort-2 stronger than *T. evansi* Viçosa-1 (Wald  $\chi^2 = 47.292$ ;  $P < 0.001$ ).

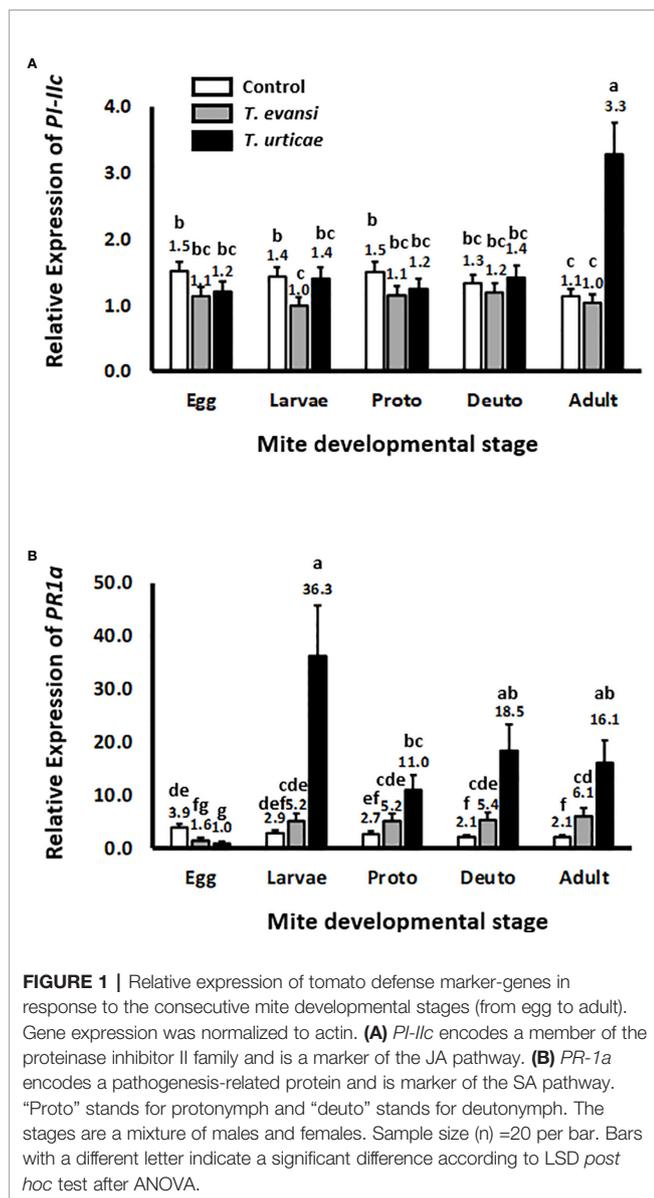
### Expression of Housekeeping Gene *RP49* Is Variable Across Spider Mite Developmental Stages

Expression of *T. evansi* Viçosa-1 or *T. urticae* Santpoort-2 genes by means of RT-qPCR is often normalized using housekeeping gene *RP49* (Morales et al., 2016; Villarroel et al., 2016; Schimmel et al., 2017a; Suzuki et al., 2017; Jonckheere et al., 2018; Yoon et al., 2018). However, Yang et al. (2015) warned that expression of *RP49* and other housekeeping genes may not be suitable for normalizing gene expression levels across developmental stages. Indeed, the levels of *RP49* expression we observed differed greatly between life stages. For *T. urticae* Santpoort-2 the average Ct (cycle threshold) of *RP49* in eggs was 30; in larvae and protonymphs 28 and in the other stages 27 (so a eight-fold difference between eggs and adults). Similarly, for *T. evansi* Viçosa-1 the average Ct of *RP49* in eggs was 29; in larvae and protonymphs 27 and in the other stages it was 25 (so a 16-fold difference between eggs and adults). Hence *RP49* was unsuitable to correct for sample-to-sample variation—i.e., variation in reverse transcription and PCR efficiency—in cDNA samples obtained from different developmental stages. However, mite RNA and tomato RNA had been collected together as total RNA (Schimmel et al., 2017a) and hence we could use tomato actin to correct for technical variation between samples. This illustrates

an advantage of collecting plant and mite RNA together although it will come at the expense of mite genes with low absolute expression levels.

### Effector 84 and *SHOT3b* Genes Are Expressed Higher in Nymphs and Adults Than in Eggs and Larvae

To assess whether spider mite effector-gene expression is plastic across their life stages, we infested tomato leaflets with 50 spider mite eggs and monitored the expression of salivary effector 84 (Figure 2A) and *SHOT3b* (Figure 2C) during the course of the development of the mites from egg to adult. The expression of effector 84 per *T. evansi* Viçosa-1 individual changed during development (Wald  $\chi^2 = 39.872$ ;  $P < 0.001$ ): it increased from egg to larva and from larva to protonymph but remained stable for the later life stages (Figure 2A). The expression of *SHOT3b* in *T. evansi* Viçosa-1 also changed during development (Wald  $\chi^2 = 18.672$ ;  $P = 0.001$ ) yet was not significantly different between egg, larva and deutonymph (Figure 2C). The pattern of expression of effector 84 in *T. urticae* Santpoort-2 individuals was similar to that of *T. evansi* Viçosa-1 individuals albeit at 10–30 fold lower levels (Figure 2A). Also the expression pattern of *SHOT3b* in *T. urticae* Santpoort-2 individuals was similar to that of *T. evansi* Viçosa-1 but here only the expression in eggs was significantly lower than in the feeding stages and the expression was 3–10 fold lower than in *T. evansi* Viçosa-1 except for the expression in eggs that was almost 50-fold lower (Figure 2C). We also assessed expression of effector *SHOT2b* but the expression of this gene cannot be detected in *T. urticae* Santpoort-2 mites feeding from tomato and is not present in the genome of *T. evansi* Viçosa-1 (Jonckheere et al., 2018). Therefore, we detected expression only in the isolated females of *T. urticae* Santpoort-2 (i.e., using the same cDNA as for Figures 2A, C) since these had been obtained from bean. We also assessed expression of effector 28 (Villarroel et al., 2016; Schimmel et al., 2017a). Expression of effector 28 in *T. urticae* Santpoort-2 paralleled the expression of its effector 84.



It was only detected for *T. urticae* Santpoort-2 and expression was similar across the developmental stages except that the expression in protonymphs relative to eggs was significantly eight-fold higher (Supplemental Figure S1). Finally, we did not include *SHOT2b* in a figure because expression was only detected in *T. urticae* Santpoort-2 females but the standard error is  $\pm 0.39$  when the average expression is set to 1.

### Effector Genes Are Expressed Higher by *T. evansi* Males Than Females but the Opposite Applies to *T. urticae*

Effector *84* was expressed four-fold higher in *T. evansi* Viçosa-1 males compared to females whereas for *T. urticae* females this gene was expressed almost 40-fold higher than in males (Figure 2B). This species-specific pattern was similar for *SHOT3b* since this gene was expressed almost four-fold higher in *T. evansi*

Viçosa-1 males than in females whereas in *T. urticae* Santpoort-2 females expression was 25-fold higher than in males (Figure 2D).

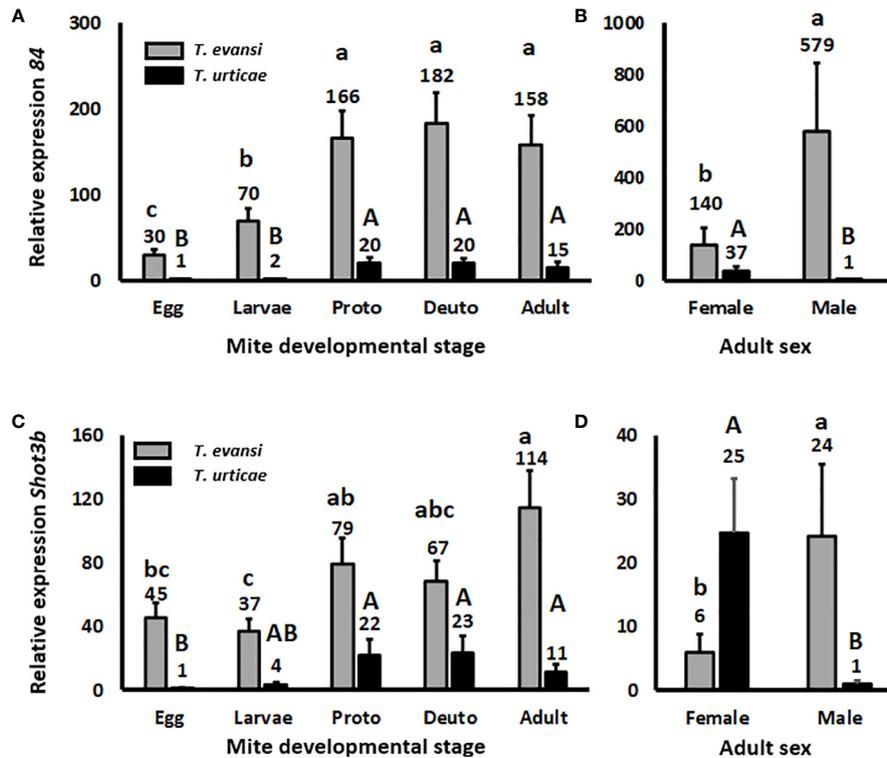
### Spider Mite Males Do Not Induce Defenses

To assess whether tomato plants respond differently to spider mite males or females we infested tomato plants with 15 individuals of the same sex and monitored the expression of tomato genes *PI-IIc* and *PR-1a* after 2 d (Figure 3). The expression of *PI-IIc* in plants infested with either *T. evansi* Viçosa-1 males or females did not exceed control levels. Also *T. urticae* Santpoort-2 males did not upregulate *PI-IIc* while females significantly upregulated its expression four-fold (Figure 3A). The expression of *PR-1a* was not upregulated by *T. evansi* Viçosa-1 females or *T. urticae* Santpoort-2 males. *T. evansi* Viçosa-1 females downregulated *PR-1a* expression while *T. urticae* Santpoort-2 females upregulated it 50-fold relative to the control plants (Figure 3B).

### DISCUSSION

Here we demonstrated that inducible JA defenses do not significantly alter developmental time or survival of *T. urticae* Santpoort-2 and *T. evansi* Viçosa-1 males and females and do not affect the spider mite sex ratio. In addition, we showed that only *T. urticae* Santpoort-2 adult females upregulate the expression of tomato JA-marker gene *PI-IIc*, while *T. evansi* Viçosa-1 larvae downregulate the expression of this gene. Eggs of both species suppressed the expression of the tomato SA-marker gene *PR-1a* but this gene was upregulated by the cumulative action of all subsequent feeding stages, especially by *T. urticae* Santpoort-2 larvae and adult females. Expression of mite effector gene *84* was lower in eggs and larvae than in the later stages of both species and a similar pattern we observed for *SHOT3b* although differences were not always significant. In addition, in *T. evansi* Viçosa-1, expression of the effector genes was higher in males than females but for *T. urticae* Santpoort-2 this was the other way around. Furthermore, we observed that only the females of *T. urticae* Santpoort-2 induce *PI-IIc* and *PR-1a* while *T. evansi* Viçosa-1 females suppress *PR-1a* expression below housekeeping levels after 2 d of infestation. Finally, feeding by spider mite males did not alter expression of *PI-IIc* and *PR-1a*.

Since developmental time to maturity has been considered a key life-history trait for evolutionary adaptation via natural selection (Cole, 1954), we tested if JA-defenses affect overall developmental time of spider mites. We also analyzed this for males and females separately since males are known to develop faster than females and eat less (Sabelis, 1985; Rajakumar et al., 2005). We found that that inducible JA defenses do not significantly alter developmental time, survival of either *T. urticae* Santpoort-2 or *T. evansi* Viçosa-1 males and females and mite sex ratio. In contrast to this observation, it was shown previously that the reproductive performance of adult *T. urticae*

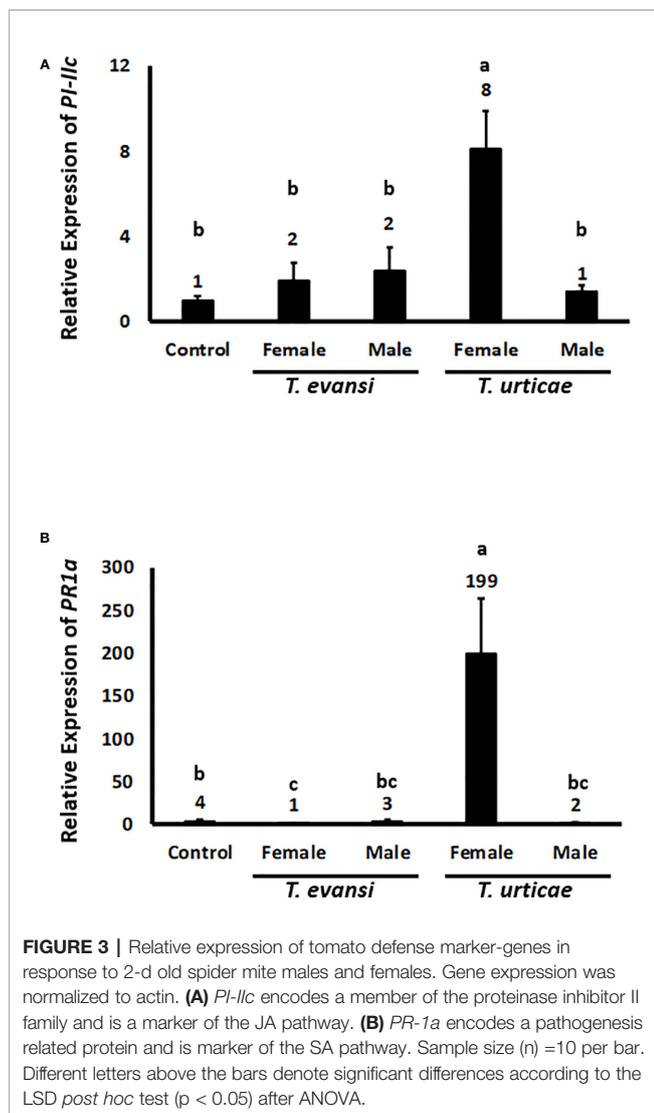


**FIGURE 2** | Relative expression of the mite effector gene *84* and *SHOT3b* in the consecutive spider mite developmental stages. Gene expression was normalized to actin. **(A)** Effector *84* expression in the developmental stages of *T. urticae* and *T. evansi*. **(B)** Effector *84* expression in *T. urticae* and *T. evansi* females and males. **(C)** *SHOT3b* expression in the developmental stages of *T. urticae* and *T. evansi*. **(D)** *SHOT3b* expression in *T. urticae* and *T. evansi* females and males. "Proto" stands for protonymph and "deuto" stands for deutonymph. The developmental stages in **(A, C)** are a mixture of males and females derived from 50 eggs and corrected for survival. The sample size (n) = 20 per bar in **(A, C)**. **(B, D)** were conducted with 15 individuals per treatment. The sample size (n) = 10 per bar in **(B, D)**. We divided the values in **(A, B)** by the lowest average to make relative expression comparable across the two panels. The same we did for **(C, D)**. Different letters above the bars denote significant differences according to the LSD *post hoc* test ( $p < 0.05$ ) after analysis by Generalized Linear Model performed per species independently.

Santpoort-2 is affected negatively by tomato JA defenses (Kant et al., 2008; Alba et al., 2015). Moreover, while the performance of tomato-adapted mites was not affected by tomato JA-defenses (Kant et al., 2008), these defenses were shown to decrease the hatching rate of their eggs (Ament et al., 2004). Finally, suppression of JA-defenses was shown to maximize fecundity of *T. evansi* Viçosa-1 (Sarmiento et al., 2011b; Alba et al., 2015; Ataide et al., 2016; Schimmel et al., 2017a; Schimmel et al., 2017b). Together this indicates that JA-defenses in general have detrimental effects on adult spider mites like *T. urticae* Santpoort-2 or *T. evansi* Viçosa-1. The observation that JA defenses do not significantly alter developmental time is in line with the observation that juvenile spider mites do not induce JA-defenses. This suggests that developmental times on WT plants do not differ from those on *def-1* because juveniles do not induce this defense in WT plants. However, the juvenile feeding stages do induce cumulative SA-defenses while adult mites were found to be significantly affected by this type of defense, although the effect sizes were small (Villarroel et al., 2016). Hence, possibly spider mite developmental time may change on the tomato SA-mutant *nahG* (Glas et al., 2014). Our main conclusion is that JA-defenses seem to be much more relevant for the interaction

between tomato plants and adult mites than between the plant and juveniles.

We can only speculate why JA-defenses are not induced by juveniles but we suggest it may relate to the kinds of cells/tissues the juvenile stages feed from in combination with the amount of feeding and their nutrient requirements. For example, also the juveniles of the generalist grasshopper *Schistocerca emarginata* were shown to have a much more narrow diet breadth than the adults (Sword and Dopman, 1999) while female grasshoppers were shown to often gain more weight than males (Unsicker et al., 2008) and have higher need for nitrogen for producing eggs (Chapman and Joern, 1990). Such differences may also apply to spider mites: protonymphs (3.7  $\mu\text{g}$ ) are three times heavier than larvae while in turn female deutonymphs are three times heavier than protonymphs (Sabelis, 1981). In addition, females are six times heavier than males (24 vs. 4  $\mu\text{g}$ ) and produce, depending on host quality, 5–15 eggs (1.2  $\mu\text{g}$  each) per day while their estimated food conversion efficiency is around 20% (Sabelis, 1981). Clearly this indicates that females have to take up and convert much more food than males or juveniles and will be therefore probably be responsible for most of the feeding damage on the plant. Apart from nutritional needs, also mite physiology,



especially stylet length, may affect the type and magnitude of the defenses juvenile mites induce. The spider mite's feeding parts include the pedipalps and the two cheliceral stylets. The cheliceral stylets can join to form a needle-like structure used for piercing plant cells and for transferring saliva while the pedipalps contain claws for rupturing plant cell walls as well as silk glands for producing web (Ragusa and Tsolakakis, 2000). The average stylet length of female *T. urticae* can vary from 103  $\mu\text{m}$  (larvae) to 157  $\mu\text{m}$  (adult females) (Park and Lee, 2002) and it was estimated they can reach between 70–120  $\mu\text{m}$  deep into a plant leaf (Tomczyk and Kropczynska, 1985). A tomato leaflet in turn has a thickness ranging from 150 to 250  $\mu\text{m}$  depending on water status and temperature (Sekhar and Sawhney, 1990; Lechowski et al., 2006; Sánchez-Rocha et al., 2008). The palisade parenchyma, the cell type mites prefer to eat, of leaflets of 170  $\mu\text{m}$  thick was found to be about 20  $\mu\text{m}$  under the adaxial (upper) surface but nearly 100  $\mu\text{m}$  away from the abaxial (lower) leaf surface (Bandurski et al., 1953; Bensoussan

et al., 2018). Since spider mites often reside on the lower leaf surface, probably to be shielded from harsh weather conditions and natural enemies, it can be difficult especially for the smaller stages to reach the palisade parenchyma. Accordingly, while chlorophyll is usually clearly visible in adults (*T. urticae* is rather transparent) it is often not in young juveniles or males. Hence larvae may feed from epidermal cells and mesophyll more than adult females do, and therefore elicit different responses, reminiscent of small mites like *Aculops lycopersici* that are also restricted to epidermal cell layers (Glas et al., 2014). Although not much is known about the abilities of different plant cell types to display JA- or SA-responses there are indications that such differences exist (Ohashi and Matsuoka, 1987; Huang et al., 1991; Uzunova and Popova, 2000) Together this indicates that ontogenetic niche shifts, e.g., characterized by a change in tissue or cell type usage by different herbivore developmental stages, may also shift the plant's defense response.

We observed that manually deposited spider mite eggs suppressed the expression of *PR-1a* while this gene was upregulated by all subsequent feeding stages. For a variety of insect species, it was shown that their eggs can induce (Hilker and Fatouros, 2015; Hilker and Fatouros, 2016) or suppress (Bruessow et al., 2010) plant defenses (Reymond, 2013). We deposited newly produced eggs manually on the leaf surface and this may differ from natural egg deposition by female mites. At higher population densities spider mites tend to deposit most of their eggs (around 0.001  $\text{mm}^3$  in size) in the web, probably to regulate egg humidity (Gerson, 1985), thereby not touching the leaf surface. When mites do deposit eggs onto the leaf surface (especially when mite densities are not so high yet) they occasionally cover these eggs with silk threads, composed of fibroin with a high serine content (Grbic et al., 2011), but there is no evidence for eggs being glued onto the leaf surface like some insects do (Voigt, 2016). Hence, the manual egg deposition we did may actually mimic natural deposition during the early stages of host plant colonization reasonably well. The egg itself has a wax layer on the outside, possibly surrounding a cement layer of oil and protein, while the embryo respire through the water resistant egg shell *via* air ducts and cone-shaped perforation organs—that are formed during embryo development—and that pierce through the shell and may conduct a lytic or plasticizing substances (Crooker, 1985). It is therefore well conceivable that substances produced during embryonic development are released on the outside of the egg; come into contact with the plant and cause physiological changes like the ones we observed. The biological significance of the *PR-1a* downregulation in response to spider mite eggs could maybe be determined using tomato SA-mutant *nahG* (Glas et al., 2014) but remains elusive at this stage. Finally, it would be interesting to assess if the mite's endosymbiont status (Staudacher et al., 2017) of the eggs and the consecutive juvenile and adult stages change in titer and differentially affect plant defense gene expression.

We monitored the expression of four effector genes: *SHOT2b* and *SHOT3b* (Jonckheere et al., 2018) and effector 28 and 84 (Jonckheere et al., 2016; Villarreal et al., 2016). Effector *SHOT2b* is unique for *T. urticae* and only expressed in mites after eating from

certain fabacean hosts like bean (*P. vulgaris*). The host-dependent regulation of *SHOT2* genes is asymmetric, i.e., it is upregulated rapidly (hours) in mites transferred to the fabacean host but downregulated slowly (possibly only in the next generation) after transfer to a non-fabacean host (Jonckheere et al., 2018). In our experiments only the separate males and females (Figures 2B, D) had been obtained from bean and, accordingly, we detected *SHOT2b* expression only in these (female) mites. Hence *SHOT2b* may play a role in the *T. urticae*-tomato interaction during the early phase of the colonization (i.e., by the first generation of mites) but not likely during later generations. However, the regulation of effector *SHOT3b* is opposite to that of *SHOT2b* and is expressed higher in mites on tomato compared to mites on beans (Jonckheere et al., 2018). In our experiments on tomato, expression of *SHOT3b* was lower in eggs and larvae than in the later stages of both mite species, similar to the expression pattern of effector 28 in *T. urticae* and of 84 in both species. Unlike in earlier studies (Villarroel et al., 2016; Schimmel et al., 2017a; Schimmel et al., 2017b) we did not detect expression of effector 28 in any of the stages of *T. evansi* Viçosa-1. Possibly this was due to the fact that we collected mite and tomato RNA together as total RNA thereby diluting *T. evansi* Viçosa-1 28 mRNA too much. The expression patterns of *SHOT3b* and effector 84 reinforce the notion that these proteins are produced and secreted primarily by the feeding stages (Jonckheere et al., 2016; Villarroel et al., 2016). Jonckheere et al. (2018) suggested the family of *SHOT3* genes to facilitate host-compatibility in a more generic manner than the *SHOT1* and *SHOT2* families. However, in contrast to *T. urticae*, expression of the *SHOT3b* and 84 genes in *T. evansi* Viçosa-1 was always higher in males than females. *T. evansi* is a gregarious species while *T. urticae* is not and possibly the *T. evansi* males play a role in creating a suitable feeding site for their kin. However, looking at the *PI-IIc* and *PR-1a* expression data also the females alone are capable of suppressing defenses (Figure 3) while in mixtures of males and females we observed slight yet significant *PR-1a* upregulation (Figure 1). Given the fact that the expression of spider mite genes associated with host defenses appeared to be rather plastic (Dermauw et al., 2012; Schimmel et al., 2017a; Jonckheere et al., 2018) it would be interesting to see how expression of effector (and detoxification) genes of *T. evansi* males is affected by the presence of related and unrelated *T. evansi* females (that both suppress defenses) as well as by the presence of defense-inducing competitors like *T. urticae* females (Schimmel et al., 2017a; Schimmel et al., 2017b). This could reveal if *T. evansi* males are capable of adjusting their magnitude of defense suppression depending on kinship with surrounding mites.

We observed that only the adult females of *T. urticae* Santpoort-2 induce expression of *PI-IIc* and *PR-1a* while adult males do not and while *T. evansi* Viçosa-1 females downregulate *PR-1a* expression below housekeeping levels after 2 d of infestation (Figure 3). Juveniles, on the other hand, upregulate *PR-1a* expression (Figure 1). These results suggest that adult males and juveniles, both being much smaller than adult females, may utilize their host plant differently than adult females. These observations also bring depth to data published previously on the timing of defense induction by adult female mites spanning a period of more than 4 d since in those

samples eggs will have hatched into larvae. These larvae may account for some of the late SA responses that were observed in such time courses (e.g., Alba et al., 2015). As noted earlier, *T. evansi* males and females separately did not upregulate *PR-1a* expression (Figure 3) while the mixed adults (Figure 1) did. There are two differences between these experiments that might explain this. The first is that the total number of individuals in the life-stages experiment was about three times higher than in the male/female trial. The second is that in the life-stages experiment induction of defenses by adults was preceded by the induction of defenses by all the juvenile stages (like in nature) but in the male/female trial it was not. Both factors can have contributed to the moderate upregulation of *PR-1a* observed in the life-stages experiment. Taken together, we provided evidence that mite ontogenetic niche shifts and stage-specific composition of their saliva together may determine the course and efficiency of induced tomato defenses.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: DOI 10.6084/m9.figshare.12630299.

## AUTHOR CONTRIBUTIONS

MK originally formulated the idea. JL, SL, and MK conceived and designed the experiments. JL, LD, JA, RC and SL performed the experiments. JL, JA, and SL analyzed the data. JL, SL, SM, and MK wrote the manuscript.

## FUNDING

JL was supported by the Chinese Scholarship Council (CSC). SL was supported by the Netherlands Organization for Scientific Research (STW-GAP/13550). RC and JA were supported by the Netherlands Organization for Scientific Research (STW-VIDI/13492). MK was supported under the European Union's Horizon 2020 research and innovation program (773 902-SuperPests and C-IPM/ALW.FACCE.6).

## ACKNOWLEDGMENTS

We would like to thank Ludek Tikovsky, Harold Lemereis and Thijs Hendrix for taking care of the plants and Alessandra Scala and Livia Ataide for helping with counting mite eggs. We thank Greg Howe (Department of Energy-Plant Research Laboratory, East Lansing, Michigan) for providing us with *def-1*.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.00980/full#supplementary-material>

## REFERENCES

- Agrawal, A. A. (2000). Host-range evolution: Adaptation and trade-offs in fitness of mites on alternative hosts. *Ecology* 81, 500–508. doi: 10.1890/0012-9658(2000)081[0500:HREAAAT]2.0.CO;2
- Alba, J. M., Schimmel, B. C., Glas, J. J., Ataide, L. M., Pappas, M. L., Villarroel, C. A., et al. (2015). Spider mites suppress tomato defenses downstream of jasmonate and salicylate independently of hormonal crosstalk. *New Phytol.* 205, 828–840. doi: 10.1111/nph.13075
- Ament, K., Kant, M. R., Sabelis, M. W., Haring, M. A., and Schuurink, R. C. (2004). Jasmonic acid is a key regulator of spider mite-induced volatile terpenoid and methyl salicylate emission in tomato. *Plant Physiol.* 135 (4), 2025–2037. doi: 10.1104/pp.104.048694
- Ataide, L. M., Pappas, M. L., Schimmel, B. C., Lopez-Orenes, A., Alba, J. M., Duarte, M. V., et al. (2016). Induced plant-defenses suppress herbivore reproduction but also constrain predation of their offspring. *Plant Sci.* 252, 300–310. doi: 10.1016/j.plantsci.2016.08.004
- Bandurski, R. S., Scott, F. M., Pflug, M., and Went, F. W. (1953). The effect of temperature on the color and anatomy of tomato leaves. *A J. Bot.* 40, 41–46. doi: 10.1002/j.1537-2197.1953.tb06447.x
- Bensoussan, N., Santamaria, M. E., Zhurov, V., Diaz, I., Grbic, M., and Grbic, V. (2016). Plant-herbivore interaction: dissection of the cellular pattern of *Tetranychus urticae* feeding on the host plant. *Front. Plant Sci.* 7, 1105. doi: 10.3389/fpls.2016.01105
- Bensoussan, N., Zhurov, V., Yamakawa, S., O'Neil, C. H., Suzuki, T., Grbić, M., et al. (2018). The Digestive System of the Two-Spotted Spider Mite, *Tetranychus urticae* Koch, in the Context of the Mite-Plant Interaction. *Front. Plant Sci.* 9, 1206. doi: 10.3389/fpls.2018.01206
- Bittner, N., Trauer-Kizilelma, U., and Hilker, M. (2017). Early plant defence against insect attack: involvement of reactive oxygen species in plant responses to insect egg deposition. *Planta* 245, 993–1007. doi: 10.1007/s00425-017-2654-3
- Blaazer, J. H., Villacis-Perez, E. A., Chafi, R., Van Leeuwen, T., Kant, M. R., and Schimmel, B. C. J. (2018). Why Do Herbivorous Mites Suppress Plant Defenses? *Front. Plant Sci.* 9, 1057. doi: 10.3389/fpls.2018.01057
- Bruessow, F., Gouhier-Darimont, C., Buchala, A., Metraux, J. P., and Reymond, P. (2010). Insect eggs suppress plant defence against chewing herbivores. *Plant J.* 62 (5), 876–885. doi: 10.1111/j.1365-313X.2010.04200.x
- Cao, H. H., Wang, S. H., and Liu, T. X. (2014). Jasmonate- and salicylate-induced defenses in wheat affect host preference and probing behavior but not performance of the grain aphid, *Sitobion avenae*. *Insect Sci.* 21, 47–55. doi: 10.1111/1744-7917.12023
- Carrière, Y. (2003). Haplodiploidy, sex, and the evolution of pesticide resistance. *J. Econ. Entomol.* 96 (6), 1626–1640. doi: 10.1603/0022-0493-96.6.1626
- Chapman, R. F., and Joern, A. (Eds.) (1990). *Biology of grasshoppers* (New York: Wiley).
- Cole, L. C. (1954). The population consequences of life history phenomena. *Q. Rev. Biol.* 29, 103–137. doi: 10.1086/400074
- Crooker, A. (1985). “Embryonic and juvenile development,” in *Spider Mites: Their Biology, Natural Enemies and Control*. Eds. W. Helle and M. W. Sabelis (Amsterdam, Netherlands: Elsevier Science Ltd.).
- Dermauw, W., Wybouw, N., Rombauts, S., Menten, B., Vontas, J., Grbic, M., et al. (2012). A link between host plant adaptation and pesticide resistance in the polyphagous spider mite *Tetranychus urticae*. *Proc. Natl. Acad. Sci. U.S.A.* 110 (2), E113–E122. doi: 10.1073/pnas.1213214110
- Doğan, S., Çağlar, M., Çağlar, B., Çırak, Ç., and Zeytun, E. (2017). Structural characterization of the silk of two-spotted spider mite *Tetranychus urticae* Koch (Acari: Tetranychidae). *Syst. Appl. Acarol* 22, 572–583. doi: 10.11158/saa.22.4.11
- Doss, R. P., Oliver, J. E., Proebsting, W. M., Potter, S. W., Kuy, S., Clement, S. L., et al. (2000). Bruchins: insect-derived plant regulators that stimulate neoplasm formation. *Proc. Natl. Acad. Sci. U.S.A.* 97 (11), 6218–6223. doi: 10.1073/pnas.110054697
- Duran-Flores, D., and Heil, M. (2016). Sources of specificity in plant damaged-self recognition. *Curr. Opin. Plant Biol.* 32, 77–87. doi: 10.1016/j.pbi.2016.06.019
- Fatouros, N. E., Voirol, L. R. P., Drizou, F., Doan, Q. T., Pineda, A., Frago, E., et al. (2015). Role of Large Cabbage White butterfly male-derived compounds in elicitation of direct and indirect egg-killing defenses in the black mustard. *Front. Plant Sci.* 6, 794. doi: 10.3389/fpls.2015.00794
- Ferragut, F., Garzon-Luque, E., and Pekas, A. (2013). The invasive spider mite *Tetranychus evansi* (Acari: Tetranychidae) alters community composition and host-plant use of native relatives. *Exp. Appl. Acarol* 60, 321–341. doi: 10.1007/s10493-012-9645-7
- Fry, J. D. (1989). Evolutionary adaptation to host plants in a laboratory population of the phytophagous mite *Tetranychus urticae* Koch. *Oecologia* 81, 559–565. doi: 10.1007/BF00378969
- Gerson, U. (1985). “Webbing,” in *Spider Mites: Their Biology, Natural Enemies and Control*, vol. World Crop Pests, vol. 1B. Eds. W. Helle and M. W. Sabelis (Amsterdam, Netherlands: Elsevier Science Ltd.).
- Glas, J. J., Alba, J. M., Simoni, S., Villarroel, C. A., Stoops, M., Schimmel, B. C., et al. (2014). Defense suppression benefits herbivores that have a monopoly on their feeding site but can backfire within natural communities. *BMC Biol.* 12, 98. doi: 10.1186/s12915-014-0098-9
- Glazebrook, J. (2005). Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annu. Rev. Phytopathol.* 43, 205–227. doi: 10.1146/annurev.phyto.43.040204.135923
- Grbic, M., Van Leeuwen, T., Clark, R. M., Rombauts, S., Rouze, P., Grbic, V., et al. (2011). The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. *Nature* 479, 487–492. doi: 10.1038/nature10640
- W. Helle and M. W. Sabelis (Eds.) (1985). *Spider mites: their biology, natural enemies and control* (Amsterdam: Elsevier).
- Hilker, M., and Fatouros, N. E. (2015). Plant responses to insect egg deposition. *Annu. Rev. Entomol.* 60, 493–515. doi: 10.1146/annurev-ento-010814-020620
- Hilker, M., and Fatouros, N. E. (2016). Resisting the onset of herbivore attack: plants perceive and respond to insect eggs. *Curr. Opin. Plant Biol.* 32, 9–16. doi: 10.1016/j.pbi.2016.05.003
- Hilker, M., Kobs, C., Varama, M., and Schrank, K. (2002). Insect egg deposition induces *Pinus sylvestris* to attract egg parasitoids. *J. Exp. Biol.* 205, 455–461.
- Howe, G. A., and Jander, G. (2008). Plant immunity to insect herbivores. *Annu. Rev. Plant Biol.* 59, 41–66. doi: 10.1146/annurev.arplant.59.032607.092825
- Huang, J.-F., Bantroch, D. J., Greenwood, J. S., and Staswick, P. E. (1991). Methyl Jasmonate Treatment Eliminates Cell-Specific Expression of Vegetative Storage Protein Genes in Soybean Leaves. *Plant Physiol.* 97, 1512–1520. doi: 10.1104/pp.97.4.1512
- Huang, H. J., Cui, J. R., Chen, L., Zhu, Y. X., and Hong, X. Y. (2019). Identification of Saliva Proteins of the Spider Mite *Tetranychus evansi* by Transcriptome and LC-MS/MS Analyses. *Proteomics* 19, e1800302. doi: 10.1002/pmic.201800302
- Iida, Y., Desaki, Y., Hata, K., Uemura, T., Yasuno, A., Islam, M., et al. (2019). Tetransins: new putative spider mite elicitors of host plant defense. *New Phytol.* 224, 875–885. doi: 10.1111/nph.15813
- Jonckheere, W., Dermauw, W., Zhurov, V., Wybouw, N., Van den Bulcke, J., Villarroel, C. A., et al. (2016). The salivary protein repertoire of the polyphagous spider mite *Tetranychus urticae*: a quest for effectors. *Mol. Cell. Proteomics* 15, 3594–3613. doi: 10.1074/mcp.M116.058081
- Jonckheere, W., Dermauw, W., Khalighi, M., Pavlidi, N., Reubens, W., Baggerman, G., et al. (2018). A gene family coding for salivary proteins (SHOT) of the polyphagous spider mite *Tetranychus urticae* exhibits fast host-dependent transcriptional plasticity. *Mol. Plant Microbe Interact.* 31, 112–124. doi: 10.1094/MPMI-06-17-0139-R
- Kaloshian, I., and Walling, L. L. (2005). Hemipterans as plant pathogens. *Annu. Rev. Phytopathol.* 43, 491–521. doi: 10.1146/annurev.phyto.43.040204.135944
- Kant, M. R., Ament, K., Sabelis, M. W., Haring, M. A., and Schuurink, R. C. (2004). Differential timing of spider mite-induced direct and indirect defenses in tomato plants. *Plant Physiol.* 135, 483–495. doi: 10.1104/pp.103.038315
- Kant, M. R., Sabelis, M. W., Haring, M. A., and Schuurink, R. C. (2008). Intraspecific variation in a generalist herbivore accounts for differential induction and impact of host plant defenses. *Proc. Biol. Sci.* 275, 443–452. doi: 10.1098/rspb.2007.1277
- Kant, M. R., Jonckheere, W., Knecht, B., Lemos, F., Liu, J., Schimmel, B. C. J., et al. (2015). Mechanisms and ecological consequences of plant defence induction and suppression in herbivore communities. *Ann. Bot.* 115, 1015–1051. doi: 10.1093/aob/mcv054
- Kessler, A., and Baldwin, I. T. (2002). Plant responses to insect herbivory: the emerging molecular analysis. *Annu. Rev. Plant Biol.* 53 (1), 299–328. doi: 10.1146/annurev.arplant.53.100301.135207
- Lechowski, Z., Mej, E., and Jan Bialczyk, J. (2006). Accumulation of biomass and some macroelements in tomato plants grown in media with (+)-usnic acid. *Environ. Exp. Bot.* 56, 239–244. doi: 10.1016/j.envexpbot.2005.03.001

- Li, C., Williams, M. M., Loh, Y. T., Lee, G. II, and Howe, G. A. (2002). Resistance of cultivated tomato to cell content-feeding herbivores is regulated by the octadecanoid-signaling pathway. *Plant Physiol.* 130, 494–503. doi: 10.1104/pp.005314
- Maffei, M. E., Arimura, G. II, and Mithöfer, A. (2012). Natural elicitors, effectors and modulators of plant responses. *Nat. Prod. Rep.* 29 (11), 1288–1303. doi: 10.1039/c2np20053h
- Mitchell, R. (1973). Growth and Population Dynamics of a Spider Mite (*Tetranychus urticae* K., Acarina: Tetranychidae). *Ecology* 54, 1349–1355. doi: 10.2307/1934198
- Mithöfer, A., Wanner, G., and Boland, W. (2005). Effects of feeding *Spodoptera littoralis* on lima bean leaves. II. Continuous mechanical wounding resembling insect feeding is sufficient to elicit herbivory-related volatile emission. *Plant Physiol.* 137 (3), 1160–1168. doi: 10.1104/pp.104.054460
- Morales, M. A., Mendoza, B. M., Lavine, L. C., Lavine, M. D., Walsh, D. B., and Zhu, F. (2016). Selection of Reference Genes for Expression Studies of Xenobiotic Adaptation in *Tetranychus urticae*. *Int. J. Biol. Sci.* 12, 1129–1139. doi: 10.7150/ijbs.16319
- Moran, P. J., and Thompson, G. A. (2001). Molecular responses to aphid feeding in *Arabidopsis* in relation to plant defense pathways. *Plant Physiol.* 125, 1074–1085. doi: 10.1104/pp.125.2.1074
- Mur, L. A., Kenton, P., Atzorn, R., Miersch, O., and Wasternack, C. (2006). The outcomes of concentration-specific interactions between salicylate and jasmonate signaling include synergy, antagonism, and oxidative stress leading to cell death. *Plant Physiol.* 140 (1), 249–262. doi: 10.1104/pp.105.072348
- Musser, R. O., Hum-Musser, S. M., Eichenseer, H., Peiffer, M., Ervin, G., Murphy, J. B., et al. (2002). Herbivory: caterpillar saliva beats plant defences. *Nature* 416, 599–600. doi: 10.1038/416599a
- Navajas, M., De Moraes, G. J., Auger, P., and Migeon, A. (2013). Review of the invasion of *Tetranychus evansi*: biology, colonization pathways, potential expansion and prospects for biological control. *Exp. Appl. Acarol.* 59, 43–65. doi: 10.1007/s10493-012-9590-5
- Ohashi, Y., and Matsuoka, M. (1987). Localization of Pathogenesis-Related Proteins in the Epidermis and Intercellular Spaces of Tobacco Leaves after Their Induction by Potassium Salicylate or Tobacco Mosaic Virus Infection. *Plant Cell Physiol.* 28, 1227–1235. doi: 10.1093/oxfordjournals.pcp.a077410
- Park, Y.-L., and Lee, J.-H. (2002). Leaf Cell and Tissue Damage of Cucumber Caused by Twospotted Spider Mite (Acari: Tetranychidae). *J. Econ. Entomol.* 95, 952–957. doi: 10.1093/jee/95.5.952
- Ragusa, S., and Tsolakis, H. (2000). Notes on the adaptation of some phytophagous and predacious mites to various ecological parameters in the Mediterranean countries. *Web Ecol.* 1, 35–47. doi: 10.5194/we-1-35-2000
- Rajakumar, E., Hugar, P. S., and Patil, B. V. (2005). Biology of red spider mite, *Tetranychus urticae* Koch (Acari: Tetranychidae) on jasmine. *Karnataka J. Agric. Sci.* 18, 147D 149.
- Reymond, P. (2013). Perception, signaling and molecular basis of oviposition-mediated plant responses. *Planta* 238 (2), 247–258. doi: 10.1007/s00425-013-1908-y
- Sánchez-Rocha, S., Vargas-Luna, M., Gutiérrez-Juárez, G., Huerta-Franco, R., and Olalde-Portugal, V. (2008). Benefits of the Mycorrhizal Fungi in Tomato Leaves Measured by Open Photoacoustic Cell Technique: Interpretation of the Diffusion Parameters. *Int. J. Thermophys* 29, 2206–2214. doi: 10.1007/s10765-008-0411-3
- Sabelis, M. W., Janssen, A., and Kant, M. R. (2001). The enemy of my enemy is my ally. *Science* 291 (5511), 2104–2105. doi: 10.1126/science.1059939
- Sabelis, M. W. (1981). *Biological control of two-spotted spider mites using phytoseiid predators. Part I. Agricultural Research Reports* (Wageningen: Pudoc), 910.
- Sabelis, M. W. (1985). “Reproductive strategies,” in *Spider Mites: their Biology, Natural Enemies and Control, vol 1A*. Eds. W. Helle and M. W. Sabelis (Amsterdam: Elsevier), 265–276.
- Sabelis, M. W. (1999). “Behavioral responses of predatory and herbivorous arthropods to induced plant volatiles: from evolutionary ecology to agricultural applications,” in *Induced Plant Defenses against Pathogens and Herbivores*. Eds. A. A. Agrawal, S. Tuzun and E. Bent (St. Paul, Minnesota: The American Phytopathological Society (APS) Press), pp 269–pp 297.
- Santamaria, M. E., González-Cabrera, J., Martínez, M., Grbic, V., Castañera, P., Diaz, L., et al. (2015). Digestive proteases in bodies and faeces of the two-spotted spider mite, *Tetranychus urticae*. *J. Insect Physiol.* 78, 69–77. doi: 10.1016/j.jinsphys.2015.05.002
- Santamaria, M. E., Martínez, M., Arnaiz, A., Rioja, C., Burow, M., Grbic, V., et al. (2019). An Arabidopsis TIR-Lectin Two-Domain Protein Confers Defense Properties against *Tetranychus urticae*. *Plant Physiol.* 179, 1298–1314. doi: 10.1104/pp.18.00951
- Sarmento, R. A., Lemos, F., Dias, C. R., Kikuchi, W. T., Rodrigues, J. C., Pallini, A., et al. (2011a). A herbivorous mite down-regulates plant defence and produces web to exclude competitors. *PLoS One* 6, e23757. doi: 10.1371/journal.pone.0023757
- Sarmento, R. A., Lemos, F., Bleeker, P. M., Schuurink, R. C., Pallini, A., Oliveira, M. G. A., et al. (2011b). A herbivore that manipulates plant defence. *Ecol. Lett.* 14, 229–236. doi: 10.1111/j.1461-0248.2010.01575.x
- Schimmel, B. C., Ataíde, L., Chafí, R., Villarreal, C. A., Alba, J. M., Schuurink, R. C., et al. (2017a). Overcompensation of herbivore reproduction through hyper-suppression of plant defenses in response to competition. *New Phytol.* 214 (4), 1688–1701. doi: 10.1111/nph.14543
- Schimmel, B. C. J., Ataíde, L. M. S., and Kant, M. R. (2017b). Spatiotemporal heterogeneity of tomato induced defense responses affects spider mite performance and behavior. *Plant Signal Behav.* 12 (10), e1370526. doi: 10.1080/15592324.2017.1370526
- Schimmel, B. C. J., Alba, J. M., Wybouw, N., Glas, J. J., Meijer, T. T., Schuurink, R. C., et al. (2018). Distinct Signatures of Host Defense Suppression by Plant-Feeding Mites. *Int. J. Mol. Sci.* 20, 19(10)pii: E3265. doi: 10.3390/ijms19103265
- Sekhar, K. N. C., and Sawhney, V. K. (1990). Leaf development in the normal and solanifolia mutant of tomato (*Lycopersicon esculentum*). *Am. J. Bot.* 77, 46–53. doi: 10.1002/j.1537-2197.1990.tb13526.x
- Staudacher, H., Schimmel, B. C. J., Lamers, M., Wybouw, N., Groot, A. T., and Kant, M. R. (2017). Independent effects of a herbivore’s bacterial symbionts on its performance and induced plant defences. *Int. J. Mol. Sci.* 18, 182. doi: 10.3390/ijms18010182
- Suzuki, T., Nunes, M. A., España, M. U., Namin, H. H., Jin, P., Bensoussan, N., et al. (2017). RNAi-based reverse genetics in the chelicerate model *Tetranychus urticae*: A comparative analysis of five methods for gene silencing. *PLoS One* 12, e0180654. doi: 10.1371/journal.pone.0180654
- Sword, G. A., and Dopman, E. B. (1999). Developmental Specialization and Geographic Structure of Host Plant Use in a Polyphagous Grasshopper, *Schistocerca emarginata* (=lineata) (Orthoptera: Acrididae). *Oecologia* 120, 437–445. doi: 10.1007/s004420050876
- Thaler, J. S., Humphrey, P. T., and Whiteman, N. K. (2012). Evolution of jasmonate and salicylate signal crosstalk. *Trends Plant Sci.* 17 (5), 260–270. doi: 10.1016/j.tplants.2012.02.010
- Tomczyk, A., and Kroczyńska, D. (1985). “Effects on the host plant,” in *Spider Mites: their Biology, Natural Enemies and Control, vol 1A*. Eds. W. Helle and M. W. Sabelis (Amsterdam: Elsevier), 317–327.
- Unsicker, S. B., Oswald, A., Köhler, G., and Weisser, W. W. (2008). Complementarity effects through dietary mixing enhance the performance of a generalist insect herbivore. *Oecologia* 156 (2), 313–324. doi: 10.1007/s00442-008-0973-6
- Uzunova, A. N., and Popova, L. P. (2000). Effect of salicylic acid on leaf anatomy and chloroplast ultrastructure of barley plants. *Photosynthetica* 38, 243–250. doi: 10.1023/A:1007226116925
- Van Leeuwen, T., Vontas, J., Tsagkarakou, A., Dermauw, W., and Tirry, L. (2010). Acaricide resistance mechanisms in the two-spotted spider mite *Tetranychus urticae* and other important Acari: a review. *Insect Biochem. Mol. Biol.* 40 (8), 563–572. doi: 10.1016/j.ibmb.2010.05.008
- Verdonk, J. C., De Vos, C. R., Verhoeven, H. A., Haring, M. A., van Tunen, A. J., and Schuurink, R. C. (2003). Regulation of floral scent production in petunia revealed by targeted metabolomics. *Phytochemistry* 62, 997–1008. doi: 10.1016/S0031-9422(02)00707-0
- Villarreal, C. A., Jonckheere, W., Alba, J. M., Glas, J. J., Dermauw, W., Haring, M. A., et al. (2016). Salivary proteins of spider mites suppress defenses in *Nicotiana benthamiana* and promote mite reproduction. *Plant J.* 86 (2), 119–131. doi: 10.1111/tpj.13152

- Voigt, D. (2016). In situ visualization of spider mite-plant interfaces. *J. Acarol Soc Jap* 25, 119–132. doi: 10.2300/acari.25.Suppl\_119
- Walling, L. L. (2000). The myriad plant responses to herbivores. *J. Plant Growth Regul.* 19 (2), 195–216. doi: 10.1007/s003440000026
- Werner, E. E., and Gilliam, J. F. (1984). The ontogenetic niche and species interactions in size-structured populations. *Annu. Rev. Ecol. Syst.* 15 (1), 393–425. doi: 10.1146/annurev.es.15.110184.002141
- Wrensch, D. L. (1985). “Reproductive parameters,” in *Spider Mites: their Biology, Natural Enemies and Control, vol 1A*. Eds. W. Helle and M. W. Sabelis (Amsterdam: Elsevier), 165–170.
- Yang, C., Pan, H., Liu, Y., and Zhou, X. (2015). Stably Expressed Housekeeping Genes across Developmental Stages in the Two-Spotted Spider Mite, *Tetranychus urticae*. *PLoS One* 10, e0120833. doi: 10.1371/journal.pone.0120833
- Yoon, J.-S., Sahoo, D. K., Maiti, I. B., and Palli, S. R. (2018). Identification of target genes for RNAi-mediated control of the Twospotted Spider Mite. *Sci. Rep.* 8, 14687. doi: 10.1038/s41598-018-32742-2
- Zarate, S.II, Kempema, L. A., and Walling, L. L. (2007). Silverleaf whitefly induces salicylic acid defenses and suppresses effectual jasmonic acid defenses. *Plant Physiol.* 143 (2), 866–875. doi: 10.1104/pp.106.090035
- Zhang, P. J., Xu, C. X., Zhang, J. M., Lu, Y. B., Wei, J. N., Liu, Y. Q., et al. (2013). Phloem-feeding whiteflies can fool their host plants, but not their parasitoids. *Funct. Ecol.* 27, 1304–1312. doi: 10.1111/1365-2435.12132
- Zhurov, V., Navarro, M., Bruinsma, K. A., Arbona, V., Santamaria, M. E., Cazaux, M., et al. (2014). Reciprocal Responses in the Interaction between Arabidopsis and the Cell-Content-Feeding Chelicerate Herbivore Spider Mite. *Plant Physiol.* 164, 384–399. doi: 10.1104/pp.113.231555

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Liu, Legarrea, Alba, Dong, Chafi, Menken and Kant. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Big Genes, Small Effectors: Pea Aphid Cassette Effector Families Composed From Miniature Exons

Matthew Dommel<sup>1†</sup>, Jonghee Oh<sup>2†</sup>, Jose Carlos Hugueta-Tapia<sup>1</sup>, Endrick Guy<sup>3</sup>, Hélène Boulain<sup>3</sup>, Akiko Sugio<sup>3</sup>, Marimuthu Murugan<sup>4</sup>, Fabrice Legeai<sup>3</sup>, Michelle Heck<sup>5</sup>, C. Michael Smith<sup>4</sup> and Frank F. White<sup>1\*</sup>

<sup>1</sup> Department of Plant Pathology, University of Florida, Gainesville, FL, United States, <sup>2</sup> Department of Plant Pathology, Kansas State University, Manhattan, KS, United States, <sup>3</sup> INRAE, UMR Institute of Genetics, Environment and Plant Protection, Le Rheu, France, <sup>4</sup> Department of Entomology, Kansas State University, Manhattan, KS, United States, <sup>5</sup> USDA-ARS, Cornell University, Ithaca, NY, United States

## OPEN ACCESS

### Edited by:

Zuhua He,  
Chinese Academy of Sciences, China

### Reviewed by:

Saskia A. Hogenhout,  
John Innes Centre, United Kingdom  
Chengshu Wang,  
Shanghai Institutes for Biological  
Sciences (CAS), China

### \*Correspondence:

Frank F. White  
fwhite@ufl.edu

<sup>†</sup>These authors share first authorship

### Specialty section:

This article was submitted to  
Plant Pathogen Interactions,  
a section of the journal  
Frontiers in Plant Science

Received: 22 February 2020

Accepted: 27 July 2020

Published: 02 September 2020

### Citation:

Dommel M, Oh J, Hugueta-Tapia JC,  
Guy E, Boulain H, Sugio A,  
Murugan M, Legeai F, Heck M,  
Smith CM and White FF (2020) Big  
Genes, Small Effectors: Pea Aphid  
Cassette Effector Families Composed  
From Miniature Exons.  
Front. Plant Sci. 11:1230.  
doi: 10.3389/fpls.2020.01230

Aphids secrete proteins from their stylets that evidence indicates function similar to pathogen effectors for virulence. Here, we describe two small candidate effector gene families of the pea aphid, *Acyrtosiphon pisum*, that share highly conserved secretory signal peptide coding regions and divergent non-secretory coding sequences derived from miniature exons. The KQY candidate effector family contains eleven members with additional isoforms, generated by alternative splicing. Pairwise comparisons indicate possible four unique KQY families based on coding regions without the secretory signal region. KQY1a, a representative of the family, is encoded by a 968 bp mRNA and a gene that spans 45.7 kbp of the genome. The locus consists of 37 exons, 33 of which are 15 bp or smaller. Additional KQY members, as well as members of the KHI family, share similar features. Differential expression analyses indicate that the genes are expressed preferentially in salivary glands. Proteomic analysis on salivary glands and saliva revealed 11 KQY members in salivary proteins, and KQY1a was detected in an artificial diet solution after aphid feeding. A single KQY locus and two KHI loci were identified in *Myzus persicae*, the peach aphid. Of the genes that can be anchored to chromosomes, loci are mostly scattered throughout the genome, except a two-gene region (KQY4/KQY6). We propose that the KQY family expanded in *A. pisum* through combinatorial assemblies of a common secretory signal cassette and novel coding regions, followed by classical gene duplication and divergence.

**Keywords:** pea aphid, *Acyrtosiphon pisum*, salivary gland, secretion protein, effector protein, gene family, proteomic analysis

## INTRODUCTION

Aphids are important pests of plants that can cause economical damage through loss of crop yield and dissemination of plant viruses through their feeding habits (Miles, 1999). There are many different species of aphid that have been found to cause crop damage, including the pea aphid, *Acyrtosiphon pisum*. The various species of aphid all display a diversity of host ranges, extending

from narrow to broad (Jaouannet et al., 2014). An aphid with a narrow host range consumes either one individual species or closely related plants within a single family. An aphid with a broad host range can feed on many different plant species spanning different taxonomic families. During this interaction, aphids extract phloem sap from the leaves and stems of the host plant through stylets, which are inserted into phloem cells. Plants possess both a constitutive and inducible immune response that fights insect consumption (Cook et al., 2015). Once fed upon, a plant can mount a defensive response to thwart parasitic processes. Aphid interactions with non-host plants are hypothesized to fail, in part, due to an immune reaction, while a successful aphid feeding involves suppressing the plant immune response (Jaouannet et al., 2014).

During feeding, aphids secrete saliva, which contains numerous proteins, enzymes, and other compounds, that assist stylet insertion, nutrient extraction, and host tissue interactions (Miles, 1999; Tjallingii, 2006; Will et al., 2007). Upon probing of a potential feeding plant, aphids secrete gelling saliva that acts to surround and protect the stylet. After puncturing the plant, the aphids secrete a watery saliva to thwart plant defenses (Miles, 1999). Components of the salivary proteins are hypothesized to play a role in facilitating the interaction with the host, in analogy to effectors of plant pathogenic bacteria and fungi. In contrast to pathogen effectors, functional evidence for effector action is limited in aphids. Nonetheless, variations in candidate effectors of aphids are hypothesized to contribute to the adaptation of aphid populations to specific host species (biotypes) and genotypes. Ectopic expression and silencing of some candidate aphid effectors have been shown to affect aphid fecundity and growth on host plants (Mutti et al., 2008; Pitino and Hogenhout, 2013).

Effector proteins are often relatively small proteins with no clear function based on relatedness to other proteins and are secreted into the host cell or extracellular milieu. A prominent example of a pea aphid effector is the protein C002. Identified initially from an EST library from the salivary glands, C002 is secreted into the target plant and hypothesized to assist in feeding (Mutti et al., 2008). Reduced expression through inhibitory RNA (RNAi) of the C002 transcript resulted in reduced feeding time of the aphids and, ultimately, premature death. Since the discovery of C002, C002 homologs and additional candidate aphid effectors have been identified (Elzinga and Jander, 2013; Rodriguez and Bos, 2013; Chaudhary et al., 2015; Thorpe et al., 2016; Boulain et al., 2018). One effector of *M. persicae*, Mp10, has been immunologically localized to the cytoplasm and chloroplasts of plant cells (Mugford et al., 2016).

The pea aphid is a model aphid species that exhibits a narrow host range feeding on legumes exclusively. The pea aphid genome and multiple other aphid genomes are available for analysis and comparisons (Richards et al., 2010; Burger and Botha, 2017; Wenger et al., 2017; Chen et al., 2019; Li et al., 2019; Quan et al., 2019). Additional genomic resources and salivary gland expressed sequence tag (EST) libraries of *A. pisum*, and other phytophagous aphids, provide numerous effector

candidates (International Aphid Genomics Consortium, 2010; Legeai et al., 2010; Shigenobu et al., 2010). Additionally, mass spectrometry proteomic analysis has been used to identify these proteins from salivary glands tissue and saliva secreted into artificial diets (Carolan et al., 2009; Cooper et al., 2010; Carolan et al., 2011; Rao et al., 2013; Chaudhary et al., 2015; Boulain et al., 2018). Despite this progress, much remains to be known about the effectors of aphid salivary proteins in aphid-host plant interactions (Mutti et al., 2006; Carolan et al., 2011; Rao et al., 2013; Boulain et al., 2018). Here, we report the identification of two candidate effector gene families of *A. pisum* and *M. persicae*.

## RESULTS

### Identification of Cassette Gene Families in Pea Aphid

Previously sequenced salivary gland cDNA sequences for *A. pisum* were retrieved from NCBI, and dataset was analyzed for sequences encoding predicted secreted peptides. Multiple transcripts were identified that encoded relatively short (100–450 aa) proteins and, upon alignment, could be divided into two families based on predicted amino acid sequence similarities (**Supplemental Figures 1 and 2**). Each family, named KQY and KHI, was composed of multiple genes and, in some cases, two to four isoforms, which were produced by alternative splicing (**Table 1**). At least one member of the loci, with the exception of KQY2, were found previously to be up-regulated in salivary glands (**Table 1**, Boulain et al., 2018). Three related sequences were also identified in the peach aphid (*Myzus persicae*) genome (**Table 1**). The notable feature of the predicted proteins is the conserved signal peptide region, ranging in size from 19 to 28 amino acids, combined with C-terminal divergent sequences (**Figures 1A, B**). The families were, hereafter, referred to as candidate cassette effectors, and the two families were named KQY and KHI after conserved amino acid sequences in the N-terminal region of all or most members (**Figures 1A, B**). The KQY family is comprised of eleven genes and seventeen different isoforms due to splicing variants (**Table 1**). One member was identified in *M. persicae*. The KHI family is composed of six genes and 10 isoforms. Two members were identified in *M. persicae*. The proteins range from 9.2 to 24.4 kDa.

A maximum likelihood phylogeny was produced, using the N-terminal nucleotides coding sequences for signal peptide region that is unique for each gene (**Figures 1C, D**). Two distinct groups of KQY genes cluster together through high bootstrap values; KQY1, KQY4, KQY6, and KQY8, KQY11. KQY1, KQY4, and KQY6 possess a related bootstrap value of 87, though the KQY4 and KQY6 are more distantly related within this group, only containing a bootstrap value of 36. KQY8 and KQY11 are highly similar, which is related in their bootstrap value of 99. Beyond the secretory signal peptide coding region, pairwise BLAST analysis of the KQY coding sequences indicates four possible gene families (KQY1, 4, 6; KQY2, 5, 9, 10; KQY3, 8, 11, Mp; KQY 7) at the probability level of  $1 \times 10^{-5}$  (**Supplemental**

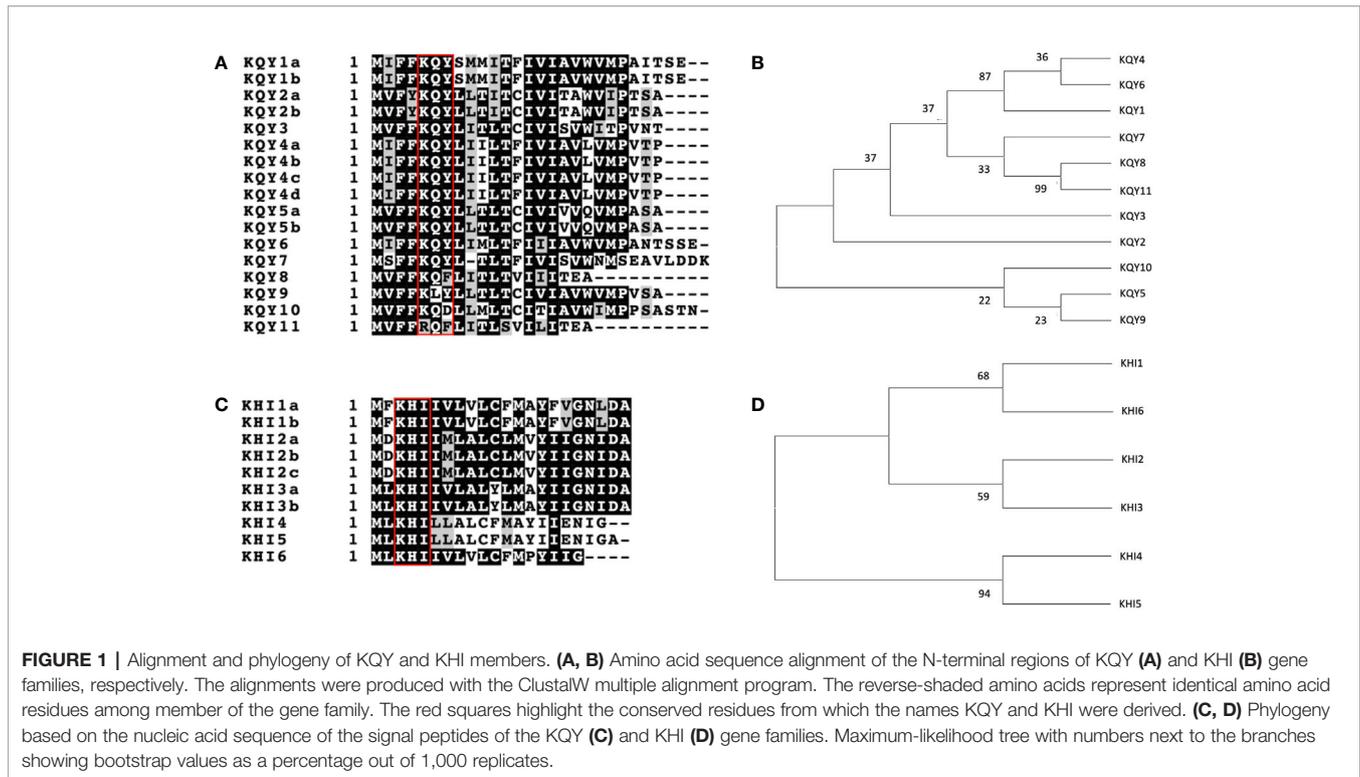
**TABLE 1** | Members of the KQY and KHI families.

Gene Family	Gene/Isoform <sup>a</sup>	Gene Locus	Transcript ID	SG Up <sup>b</sup>	Protein ID	Signal Peptide Predication(TargetP-2.0) <sup>c</sup>
<b>KQY</b>	KQY1a	LOC100158789 ACYPI000223	NM_001162442	+	NP_001155914	MIFF <b>KQY</b> SMMITFIVIAVWVMPAITSE (0.8578)
	KQY1b	LOC100158789 ACYPI000223	AK339882		BAH70584	MIFF <b>KQY</b> SMMITFIVIAVWVMPAITSE (0.855)
	KQY2a	LOC100302371	AK342599		BAH72568	MV <b>FYKQY</b> LLTITCIVITAWWIPTSA (0.9902)
	KQY2b	LOC100302371	AK342927		BAH72760	MV <b>FYKQY</b> LLTITCIVITAWWIPTSA (0.9902)
	KQY3	LOC100301916 ACYPI073633	AK341661	+	NP_001156348 BAH72003	MV <b>FFKQY</b> LITLTCIVISWITPVNT (0.9924)
	KQY4a	LOC100302370	AK342948	+	NP_001153885 BAH72770	MIFF <b>KQY</b> LILTFIVIAVLVMPVTP (0.9362)
	KQY4b	LOC100302370	AK342242		BAH72349	MIFF <b>KQY</b> LILTFIVIAVLVMPVTP (0.9324)
	KQY4c	LOC100302370	AK342690		BAH72627	MIFF <b>KQY</b> LILTFIVIAVLVMPVTP (0.9297)
	KQY4d	LOC100302370	AK342678		BAH72619	MIFF <b>KQY</b> LILTFIVIAVLVMPVTP (0.9537)
	KQY5a	LOC100302375	AK342406	+	NP_001156434 BAH72443	MV <b>FFKQY</b> LLTLTCIVIVQVMPASA (0.9963)
	KQY5b	LOC100302376	AK342378		NP_001156434 BAH72425	MV <b>FFKQY</b> LLTLTCIVIVQVMPASA (0.9857)
	KQY6	LOC100302481 (NV12)	AK340126	+	NP_001156817 BAH70788	MIFF <b>KQY</b> LIMLTFIIIAVWVMPANT (.9693)
	KQY7	LOC100302485 (NV22)	AK340563	+	NP_001156835 BAH71149	MS <b>FFKQY</b> LTLTFIVISVWNMSEA (0.9649)
	KQY8	LOC100302439 ACYPI24906	AK342473	+	NP_001156591 BAH72486	MV <b>FFKQF</b> LITLTVIITEA (0.9676)
	KQY9	LOC100302403	AK342683	+	NP_001156509 BAH72621	MV <b>FFKLY</b> LLTLTCIVIAVWVMPVSA (0.9981)
	KQY10	LOC100302381	AK342808	+	NP_001156441 BAH72693	MFNVLILSLISYTFEPSYTLKFKM <b>VFFKQD</b> LLML TCITIAVWIMPPSASTN (0.8081)
	KQY11	LOC100302480	AK340121	+	NP_001156816 BAH70784	MV <b>FFRQF</b> LITLTVIITEA (0.9787)
KQY <sup>Mp</sup>	LOC111039170 ( <i>Myzus persicae</i> )	XM_022322515		XP_022178207	MH <b>FFKHLY</b> LIVLTYIVISFWFMPASL (0.9345)	
<b>KHI</b>	KHI1a	LOC100159750 ACYPI001099	AK339863	+	NP_001155863 BAH70570	MF <b>KHII</b> VLVLCFMAYFVGNLDA (0.998)
	KHI1b	LOC100159750 ACYPI001099	AK339862		BAH70569	MF <b>KHII</b> VLVLCFMAYFVGNLDA (0.9983)
	KHI2a	LOC100302383	AK341162	+	NP_001156448 BAH71618	MD <b>KHII</b> MLALCLMVYIIGNIDA (0.9936)
	KHI2b	LOC100302383	AK341161		BAH71617	MD <b>KHII</b> MLALCLMVYIIGNIDA (0.9937)
	KHI2c	LOC100302383	AK342769		BAH72672	MD <b>KHII</b> MLALCLMVYIIGNIDA (0.9952)
	KHI3a	LOC100166702 ACYPI007553	AK340197	+	NP_001156548 BAH70850	ML <b>KHII</b> VLALYLMAYIIGNIDA (0.9965)
	KHI3b	LOC100166702 ACYPI007553	AK342603		NP_001155718 BAH72572	ML <b>KHII</b> VLALYLMAYIIGNIDA (0.9947)
	KHI4	LOC100570519 ACYPI46154	AK341077	+	NP_001280397 BAH71554	ML <b>KHILL</b> ALCFMAYIENIG (0.9586)
	KHI5	LOC100534636	AK341390	+	NP_001191953 BAH71796	ML <b>KHILL</b> ALCFMAYIENIGA (0.9976)
	KHI6	LOC100571631	AK340760	+	NP_001233103 BAH71306	ML <b>KHII</b> VLVLCFMPYIIG (0.9985)
	KHI <sup>Mp</sup> 1a	LOC111029516	XM_022308527	nd	XP_022164219	MV <b>RHII</b> MLAICIMFYIIGNAMALTPAERKA
	KHI <sup>Mp</sup> 1b	LOC111029516	XM_022308528	nd	XP_022164220	MV <b>RHII</b> MLAICIMFYIIGNAMALTPAERKA
	KHI <sup>Mp</sup> 1c	LOC111029516	XM_022308529	nd	XP_022164221	MV <b>RHII</b> MLAICIMFYIIGNAMALTPAERKA
	KHI <sup>Mp</sup> 2a	LOC111029518	XM_022308532	nd	XP_022164224	MSTMV <b>KHIN</b> MLALFIMFYIIGNAMALTPAERKA
KHI <sup>Mp</sup> 2b	LOC111029518	XM_022308533	nd	XP_022164225	MSTMV <b>KHIN</b> MLALFIMFYIIGNAMALTPAERKA	
KHI <sup>Mp</sup> 2c	LOC111029518	XM_022308534	nd	XP_022164226	MSTMV <b>KHIN</b> MLALFIMFYIIGNAMALTPAERKA	
KHI <sup>Mp</sup> 2d	LOC111029518	XM_022308536	nd	XP_022164228	MSTMV <b>KHIN</b> MLALFIMFYIIGNAMALTPAERKA	
<b>C002</b>	C002 <sup>Ap</sup>	LOC100167863 ACYPI008617	XM_001948323	+	XP_001948358	MGSYKLYVAVMAIAIAVQVEVRC (0.9704)

<sup>a</sup>Mp, *Myzus persicae*, Ap, *Acyrtosiphon pisum*.

<sup>b</sup>+, Identified as up-regulated in salivary glands in comparison to alimentary tract by Boulain et al. (2018). Up-regulation of locus is indicated, and no differential expression of isoforms is implied. nd, not detected, no salivary gland ESTs from *M. persicae* were identified by BLAST.

<sup>c</sup>KQY10 and KHI<sup>Mp</sup> isoforms predicted N-terminal peptide, which may be misannotated.



**FIGURE 1 |** Alignment and phylogeny of KQY and KHI members. **(A, B)** Amino acid sequence alignment of the N-terminal regions of KQY **(A)** and KHI **(B)** gene families, respectively. The alignments were produced with the ClustalW multiple alignment program. The reverse-shaded amino acids represent identical amino acid residues among member of the gene family. The red squares highlight the conserved residues from which the names KQY and KHI were derived. **(C, D)** Phylogeny based on the nucleic acid sequence of the signal peptides of the KQY **(C)** and KHI **(D)** gene families. Maximum-likelihood tree with numbers next to the branches showing bootstrap values as a percentage out of 1,000 replicates.

**Figure 3).** KQY7 shares no sequence relatedness beyond the secretory signal region at the DNA or protein level. At the same time, all members share some sequence identity in 3' region of the transcripts at the nucleotide level, with the exception of KQY7 and KQYmp (**Supplemental Figure 4**).

Similarly, within the KHI gene family, only two KHI members cluster close together according to bootstrap values, KHI4 and KHI5. KHI4 and KHI5 possess a bootstrap value of 94, indicated high homology. The remaining the KHI gene signal peptides are loosely related with KHI1, KHI6, and KHI2, KHI3 clustering with bootstrap values of 68 and 59, respectively.

KQY10 is annotated with twenty-nine additional N-terminal amino acid residues in comparison to the other family members (**Table 1**). Both the sequence, as annotated, or a shortened version are predicted to contain a signal peptide. Similarly, the KHI<sup>MP2</sup> locus, including all isoforms, are annotated with three additional amino acid residues at the N-terminus (**Table 1**).

## Identification of Candidate Cassette Effectors by Proteomic Analysis of Salivary Gland Proteins

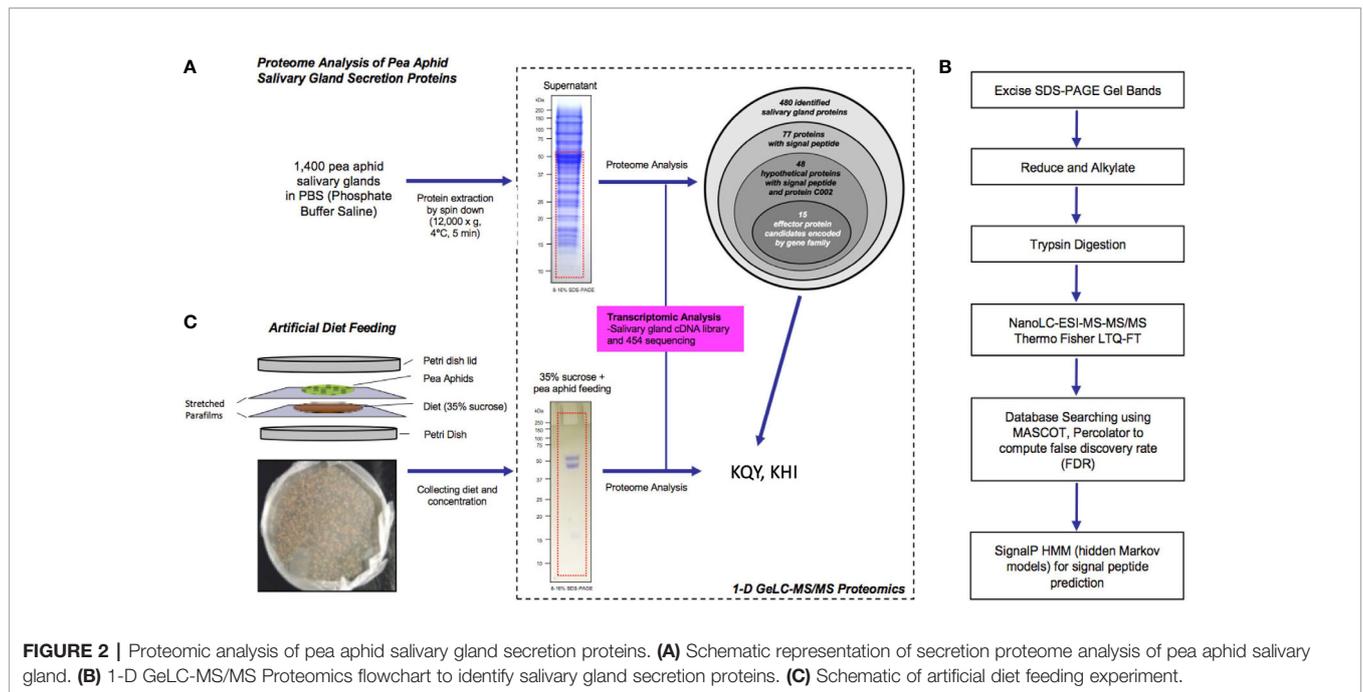
A proteomic analysis was conducted to determine whether member of the two families were present in salivary glands and secreted in salivary fluids (**Figure 2A**). Proteins were extracted from the salivary gland tissues of *A. pisum* and separated on a SDS-PAGE (1DE) gel (**Figure 2B**). Proteins in 10–60 kDa range were then subjected to 1-D GeLC-MS/MS. Of the 480 proteins, 77 proteins with predicted secretion signals were present

(**Supplemental Table 1**). Notably, 16 of the candidate secreted gland proteins were members of the KQY and KHI families (**Table 2; Supplemental Table 1**). Ten KQY proteins were found, namely KQY1a, KQY2a, KQY2b, KQY3, KQY4a, KQY4c, KQY5b, KQY9, KQY10, and KQY11. KQY2a and b, KQY4a and c, and KHI2a and b were the isoforms found concurrently. Five of the KHI gene family corresponding proteins of the six KHI genes were identified, and six out of the ten protein isoforms were found (**Table 2, Supplemental Table 1**). In addition to other candidate effectors, the analysis identified the conserved aphid effector C002 (**Table 2**).

Proteins were collected from artificial diet media after feeding by *A. pisum* to determine if any of the family members could be detected in extracellular fluids using an artificial diet (**Figure 2C**). Total protein was analyzed through 1-D GeLC-MS/MS (**Figure 2B**). A total of nine aphid proteins were identified, including KQY protein, KQY1a (**Table 3**). Additional proteins included amino peptidase and angiotensin converting enzymes and have been previously observed (Boulain et al., 2018).

## Large Gene Structure and Genomic Location of Candidate Cassette Effector Genes

The pea aphid genome consists of four different chromosomes. The gene for KQY1a protein (gi|241896885) is anchored on the A1 chromosome in the *A. pisum* strain AL4f genome (GCF\_005508785.1) (**Table 4**). Each of the pea aphid KQY genes have been found placed within the genome except for



**FIGURE 2 |** Proteomic analysis of pea aphid salivary gland secretion proteins. **(A)** Schematic representation of secretion proteome analysis of pea aphid salivary gland. **(B)** 1-D GeLC-MS/MS Proteomics flowchart to identify salivary gland secretion proteins. **(C)** Schematic of artificial diet feeding experiment.

**TABLE 2 |** Cassette effectors from proteomic analysis *A. pisum* salivary glands.

Gene Family	Protein Name	NCBI Accession Number	# of Peptides (Unique)	% Coverage
<b>KQY</b>	KQY1a	NP_001155914	17(14)	66
	ACYPI000223			
	LOC100158789			
	KQY2a	BAH72568	9(8)	47
	LOC100302371			
	KQY2b	NP_001156348	8(6)	36
	LOC100302371			
	KQY3	NP_001153885	2(1)	13
	LOC100301916			
	KQY4a	NP_001156343	10(9)	43
	LOC100302370			
	KQY4c	BAH72627	10(9)	45
	LOC100302370			
	KQY5b	NP_001156435	8(8)	45
	LOC100302376			
	KQY9	NP_001156509	3(3)	32
	LOC100302403			
	KQY9	NP_001156509	3(3)	24
	LOC100302403			
	KQY10	NP_001156441	1(1)	6
	LOC100302381			
KQY11	NP_001156816	3(3)	24	
LOC100302480				
<b>KHI</b>	KHI1a	NP_001155863	4(3)	33
	ACYPI001099			
	LOC100159750			
	KHI2a	BAH71618	2(2)	14
	LOC100302383			
	KHI2b	NP_001156448	2(2)	14
	LOC100302383			
	KHI3a	NP_001156548	12(9)	56
	LOC100166702			
	KHI5	NP_001191953	6(6)	38
	LOC100534636			
KHI6	NP_001233103	4(4)	44	
LOC100571631				
<b>C002</b>	C002	XM_001948323	5(4)	33

**TABLE 3** | *A. pisum* salivary gland proteins detected in synthetic diet using 1-D gel LC-MS/MS.

Protein Name	NCBI Accession	Top Ion E-value	# of Peptides (Unique)	% Coverage	Signal Peptide Probability (D-score)	Mw (kDa)	Predicted Protein Name/Function
KQY1a	NP_001155863	2.10E-04	1(1)	8	0.645	23.1	KQY gene family
Aminopeptidase N-like	gij193636568	2.20E-09	19(15)	51	0.822	70.1	M1 zinc dependent metalloprotease
Angiotensin converting enzyme-like	gij193669489	5.20E-07	7(6)	42.9	0.703	74.2	M2 zinc dependent angiotensin converting enzyme (peptidase)
Angiotensin converting enzyme-like	gij209571509	1.30E-04	4(4)	9.6	0.994	74.5	M2 zinc dependent angiotensin converting enzyme (peptidase)
Aminopeptidase N-like	gij193702193	4.00E-06	18(9)	29.2	0.265	90	M1 zinc dependent metalloprotease-
Leucyl-cystinyl aminopeptidase-like	gij193643503	3.30E-07	9(6)	31.5	0.974	105	M1 zinc dependent metalloprotease
Aminopeptidase N-like isoform 1	gij193634323	9.00E-10	16(13)	59.8	0.991	105.2	M1 zinc dependent metalloprotease
Aminopeptidase N-like	gij193657504	8.30E-06	5(3)	7.8	0.753	105.4	M1 zinc dependent metalloprotease
Aminopeptidase N-like	gij193575561	2.80E-10	30(21)	31.6	0.981	105.5	M1 zinc dependent metalloprotease

KQY11. KQY genes can be found placed on chromosome A1, A2, and X but not A3 (Table 4, Figure 3). Two of the KHI genes, KHI2 and KHI6, were unable to be placed within the pea aphid genome, and the remaining KHI genes were also found on chromosome A1, A2, and X, but not A3 (Table 4, Figure 3).

*KQY1a* covers approximately 45.7 kbp, and the transcript is comprised of 37 relatively small exons (6–416 bp) (Figure 4). This structure of a large gene coding for a small protein using many miniature exons is also observed with other KQY gene family members (Table 4). KHI members are also generated from relatively large genes. The *KHI6* transcript (gij239789352) is 943 bps long and comes from 10 exons in a gene that is 21.789 kbps long (Figure 4). The gene sizes of the mentioned gene families range from 12 kbp to 87 kbp. The first reported pea aphid effector/secretion protein, C002, shown here for contrast, has relatively small gene size (~6 kbp) with only two exons (Figure 4). No significant similar/conserved protein motifs and domains were found. The protein function of the gene family is unknown (hypothetical protein). A separate predicted locus (LOC100569066) can be found within an intron of *KQY2*. The gene product is highly conserved RAD50-interacting protein 1 (XP\_016658051).

## DISCUSSION

Here, we add to the characterization of candidate effectors of *A. pisum*, and, by sequence relatedness, possibly, *M. persicae* with the description of two families of genes, which by several criteria, appear to be variable secreted salivary gland proteins (Carolan et al., 2009; Carolan et al., 2011; Rao et al., 2013; Boulain et al., 2018). Twenty-seven protein candidates based on representative cDNAs could be assigned to either the KQY or KHI families, and most of the cDNA were represented in salivary gland RNAseq libraries. All of the loci, with exception of *KQY2*, were previously shown to have at least one isoform up-regulated in salivary gland in relation to alimentary tract expression, and all are predicted to encode secreted small molecular weight proteins (~12–28 kDa).

Furthermore, peptides from a majority of the loci were detected in protein extractions of washed salivary glands, and one was detected in artificial feeding media. In a previous analysis, unidentified isoforms of three cassette effectors were detected in an artificial diet, including *KQY2*, which lacked clear evidence for salivary gland expression (Boulain et al., 2018).

The members of the two families were named cassette effectors due to the conserved N-terminal region, which harbors the signal secretion motif, and the divergent coding sequences distal to the secretory signal region. The model implies that novel coding sequences could be swapped on to the signal cassette, generating novel secreted proteins, which, in turn, can then facilitate the adaptation process of the aphid to new hosts or host varieties. The KQY genes can be grouped into three gene subfamilies that have, at least in part, expanded by gene duplication and divergence. *KQY7* constitutes a single gene subfamily. Nonetheless, members of different families share sequence similarities beyond the coding regions indicating possible mosaic gene structure. The presence of a single KQY candidate from the related but distant green peach aphid (*M. persicae*) may be the result of amplification of a single gene during adaptation of pea aphids to various leguminous hosts. Whether cassette swapping was involved in adaptation to a new host cannot be definitively stated. Analysis of various biotypes of *A. pisum* may reveal subspecies cassette gene content. Cassette effectors analogous to KQY and KHI have been previously identified in the Hessian fly genome, where the SSSGP-1 family share a similar structure (Chen et al., 2010), and domain swapping with secretory domains has been proposed, to name a few, to drive complexity in scorpion venom, in the evolution of plastid nuclear encoded proteins, and new virulence in nematodes (Tonkin et al., 2008; Vanholme et al., 2009; Wang et al., 2016). Exon shuffling has long been proposed, in itself, as one benefit of eukaryotic gene structure (Koonin et al., 2013; Smithers et al., 2019). The KQY and KHI genes are represented by varying numbers of mRNAs isoforms. However, definitive conclusions with regards to the levels of individual isoforms or loci remain unclear.

Some of the candidate cassette effector genes are quite large. *KQY1a*, as an example, is produced from a 986 base mRNA, which,

**TABLE 4** | Genome locations of KQY and KHI families.

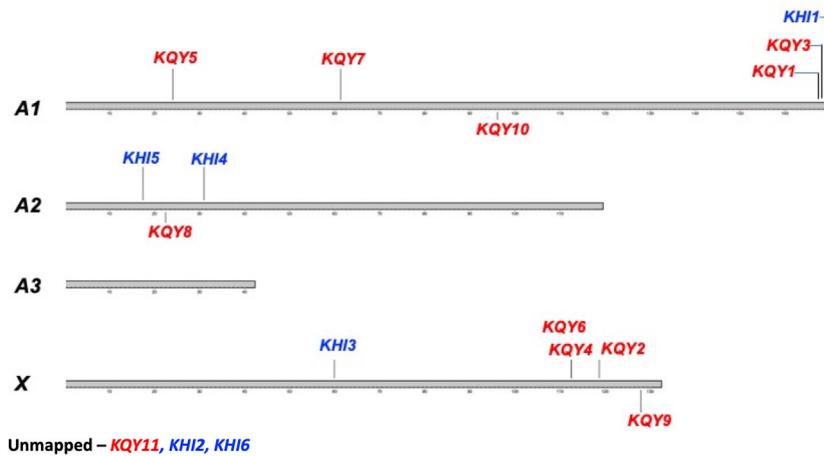
Gene Family	Gene/Isoform Designation	Gene Locus	Chromosome	Location	Size (kb)exons (range bp)/introns (range bp)
<b>KQY</b>	KQY1a	LOC100158789	Chr A1 NC_042494.1	167,823,999–167,869,655	45.7 37(6-416)/36(180-5168)
	KQY1b	LOC100158789			
	KQY2a	LOC100302371	Chr X NC_042493.1	118,679,504–118,712,594	33.1
	KQY2b	LOC100302371			31.2 >13(5-420)/>12(555-9442)
	KQY3	LOC100301916	Chr A1 NC_042494.1	Complement 168,127,940–168,141,540	13,6 15(9-582)/14(281-5949)
	KQY4a	LOC100302370	Chr X NC_042493.1	Complement 112,529,404–112,604,186	74.8 32(5-389)/>31(203-10515)
	KQY4b	LOC100302370			
	KQY4c	LOC100302370			
	KQY4d	LOC100302370			
	KQY5a	LOC100302375			87.8 >27(6-399)/>26(377-43413)
	KQY5b	LOC100302376	Chr A1 NC_042494.1	24,100,856–24,188,627	87.8
	KQY6	LOC100302481 (NV12)	Chr X NC_042493.1	112,481,406–112,505,954	24.5
	KQY7	LOC100302485 (NV22)	Chr A1 NC_042494.1	Complement 61,298,211–61,309,480	11.3
	KQY8	LOC100302439	Chr A2 NC_042495.1	Complement 22,473,653–22,510,347 Length: 36,695 nt	36.7
	KQY9	LOC100302403	Chr X NC_042493.1	Complement 127,899,608–127,978,192	78.6
	KQY10	LOC100302381	Chr A1 NC_042494.1	96,124,545–96,133,787	9.2
<b>KHI</b>	KQY11	LOC100302480	NW_021761267.1 unplaced		
	KHI1a	LOC100159750	Chr A1 NC_042494.1	complement 170,686,182–170,699,132	13.0 12.8, 11(27-564)/10(170-2537)
	KHI1b	LOC100159750			
	KHI2a	LOC100302383	NW_021771857.1 unplaced		
	KHI2b	LOC100302383			
	KHI2c	LOC100302383			
	KHI3a	LOC100166702	Chr X NC_042493.1	59,926,126–59,940,736	14.6 13(6-355)/12(173-2296)
	KHI3b	LOC100166702			14.6 14(6-355)/13(95-2296)
	KHI4	LOC100570519 (ACYPI46154)	Chr A2 NC_042495.1	complement 31,024,034–31,037,869	13.9
	KHI5	LOC100534636	Chr A2 NC_042495.1	complement 17,449,710–17,453,573	18.9 >13(16-254)/>12(67-6096)
KHI6	LOC100571631	NW_021770650.1 unplaced			

in turn, is spliced from 46 kb of DNA, containing 37 exons and 36 introns. The gene sizes are not the largest, but, given the protein product, they are remarkable. The human gene for type III collagen, for example, is 44 kb and has 52 exons. However, the mRNA is 5460 bases, encoding a protein of 1446 amino acid residues in length, compared to the 986 mRNA and 204 aa products. *KQY4* and *KQY5* may be nearly twice the size of *KQY1a*. Further conclusions regarding *KQY4* and *KQY5* and some other gene of the candidate cassette effectors await improved genome sequencing and assembly. General conclusions regarding the arrangement of the genes may change due to future assembly improvements. The gene that can be mapped are scattered throughout the genome and, at present, only one pair are present in tandem (*KQY4* and *KQY6*), despite the general view that highly evolving loci occur in multigenic loci. The contribution of cassette family genes to aphid adaptation awaits attempts to alter the expression of individual genes.

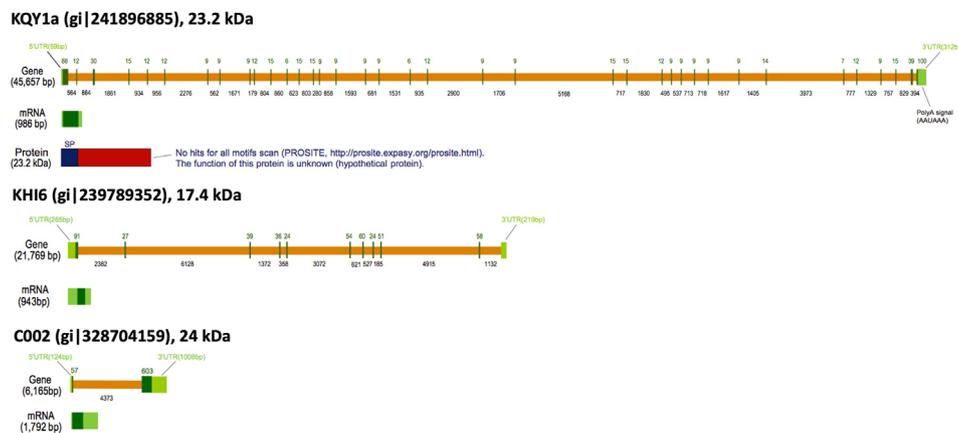
## MATERIALS AND METHODS

### Pea Aphids, Salivary Glands, Proteins Collection

Pea aphid (*A. pisum*) clone LSR1 was maintained on *Vicia faba* at 20°C. Salivary glands of feeding adult aphid on the host plants were dissected following a protocol of the previous study (Mutti et al., 2006). For salivary gland protein extraction, the dissected salivary glands of *A. pisum* were stored in PBS solution with protease inhibitor cocktail (Roche) and centrifuged at 12,000 × g for 15 min at 4°C without tissue homogenization to avoid cellular proteins. After centrifugation and collecting supernatant, salivary gland proteins of the supernatant were precipitated with 20% TCA (v/v) and incubated at -20°C, overnight. The protein pellet was collected by centrifugation (1,500 × g for 10 min, 4°C) and then washed with 100% acetone 3 times and



**FIGURE 3** | Placement of the *KQY* and *KHI* genes on the pea aphid chromosomes. The pea aphid chromosomes A1, A2, A3, and X. *KQY* genes are colored in red and the *KHI* genes are colored in blue. *KQY11*, *KHI2*, and *KHI6* are unmapped.



**FIGURE 4** | Gene structures of *KQY1a* and *KHI6*. *C002* is shown for comparison. Diagrams are based on the graphic sequence display generated from the NCBI genome sequence viewer. Orange bars = introns; Green bars = exons.

allowed the protein pellet to air dry. The protein pellet was dissolved in SDS-PAGE sample buffer [0.25 M Tris-HCl (pH6.8), 50% glycerol, 5% SDS, and 5%  $\beta$ -mercaptoethanol] for protein separation by 1-D SDS-PAGE for proteome analysis.

### Saliva Collection From Artificial Diet

Synthetic diet preparation and saliva collection were conducted under aseptic conditions (Will et al., 2007). Pea aphid saliva collection plates were prepared by stretching sterilized parafilm over the bottom of the 100 by 15 mm plastic petri dishes. Parafilm sheet surface sterilized and exposed to UV light for 30 min and the parafilms were stretched to 50% of the original size. Five milligrams chemically defined synthetic diet (35% sucrose solution) was placed on the stretched parafilm and cover with the other sterilized stretched parafilm (Figure 1A). Fifteen aphid saliva collection

plates (approximately 1,600 pea aphid on each plate) were prepared for the secreted saliva collection from the synthetic diet. The diet from a 24 h collection period was pooled to give a volume approximately 75 ml, followed by concentration using a Vivaspin concentrator (GE Healthcare) with 3,000 molecular weight cut-off PES membrane at 4°C. The concentrated proteins were separated by 1-D SDS-PAGE and visualized with Coomassie blue R-250.

### In Gel Sample Preparation for Mass Spectrometry

For salivary gland proteome analysis, we have identified salivary gland proteins using by 1-D GeLC-MS/MS proteome approach. Proteins from salivary gland tissues and artificial diet were separated on 8%–16% Tris-HCl precast gel (Bio-Rad) in a Mini-Protean Electrophoresis Unit (Bio-Rad) and stained with

Coomassie blue R-250 (**Figure 1A**). The stained protein bands of interest were excised using sterile surgical blades and the gel slices (no larger than  $2 \times 5$  mm) were transferred to individual 1.5 ml microcentrifuge tubes with 10  $\mu$ l HPLC grade water to prevent dehydration and prepared In-gel digestion. Proteins in the gel slices were reduced with 10 mM DTT in 200 mM ammonium bicarbonate at 60°C for 15 min, and then subjected to amidation in 20 mM iodoacetamide in 200 mM ammonium bicarbonate at room temperature in the dark for 30 min. The gel pieces were washed with 200 mM ammonium bicarbonate/50% acetonitrile (v/v) before addition of 250  $\mu$ l of acetonitrile and incubation at room temperature for 15 min. The remaining solvent was removed, and the gel slices were completely dried using SpeedVac system (Thermo Fisher Scientific). The proteins in the gel slices were digested with 5 ng/ml sequencing grade modified porcine trypsin (Promega) in 200 mM ammonium bicarbonate/10% acetonitrile (v/v) at 55°C for 2 h. Trypsin was inactivated by adding 0.1% trifluoroacetic acid after protein digestion and the supernatant was transferred into 0.5 ml microcentrifuge tube for mass spectrometric analysis.

### Capillary Liquid Chromatography-Mass Spectrometry Analysis for Protein Identification

Samples were analyzed by LC-MS/MS using a NanoAcquity chromatographic system (Waters Corp., Milford, MA) coupled to an LTQ-FT mass spectrometer (ThermoFinnigan, Bremen, Germany). Peptides were separated on a reverse-phase C<sub>18</sub> column, 5 cm, 500  $\mu$ m I.D. (CVC Microtech). A gradient was developed from 1% to 40% B (99.9% acetonitrile, 0.1% formic acid) in 50 min, ramped to 95% B in 4 min and held at 95% B for 5 min at a flow rate of 20  $\mu$ l/min with solvents, A (99.9% H<sub>2</sub>O, 0.1% formic acid) and B. NanoAcquity UPLC Console (Waters Corp., Version 1.3) was used to execute the injections and gradients. The ESI source was operated with spray voltage of 2.8 kV, a tube lens offset of 160 V and a capillary temperature of 200°C. All other source parameters were optimized for maximum sensitivity of the YGGFL peptide MH<sup>+</sup> ion at m/z 556.27. The instrument was calibrated using an automatic routine based on a standard calibration solution containing caffeine, peptide MRFA, and Ultramark 1621 (Sigma). Data-dependent acquisition method for the mass spectrometer (configured version LTQ-FT 2.2) was set up using Xcalibur software (ThermoElectron Corp., Version 2.0). Full MS survey scans were acquired at a resolution of 50,000 with an Automatic Gain Control (AGC) target of  $5 \times 10^5$ . Five most abundant ions were fragmented in the linear ion trap by collision-induced dissociation with AGC target of  $2 \times 10^3$  or maximum ion time of 300 ms. The ion selection threshold was 500 counts. The LTQ-FT scan sequence was adapted from the reference (Olsen and Mann, 2004).

### Database Searches

MS/MS spectra were analyzed using Mascot (Matrix Science, London, UK; Version 2.3). Mascot was set up to search the SwissProt database and our pea aphid salivary gland transcriptome data of *A. pisum* assuming the trypsin digestion. Search was performed with a fragment ion mass tolerance of 0.20 Da and a parent ion tolerance of 20 PPM. Iodoacetamide derivative of cysteine

was specified as a fixed modification. Oxidation of methionine was specified as a variable modification. Scaffold software (Version 3.6, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identification from the MS/MS data was performed using the MASCOT to correlate the data against NCBI non-redundant database and our salivary gland transcriptome data of *A. pisum*. To improve peptide identification accuracy, the results of protein identification were validated by multiple search engines (Mascot, Sequest and X! Tandem) using Scaffold software. Peptide identifications were accepted if they could be established at greater than 50.0% probability as specified by the Peptide Prophet algorithm (Keller et al., 2002). Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al., 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

### Protein Sequence and Domain/Motif Analysis

The amino acid sequence of the proteins of the gene family was analyzed with ClustalW alignment program for the gene family protein grouping (<https://www.genome.jp/tools-bin/clustalw>) using the slow parameters of a 10.0 gap open penalty and a 0.1 gap extension penalty with the BLOSUM (for protein) weight matrix. The amino acid alignment was produced by T-Coffee using default parameters ([ebi.ac.uk/Tools/msa/tcoffee/](http://ebi.ac.uk/Tools/msa/tcoffee/)) and illustrated using BoxShade ([embnet.vital-it.ch/software/BOX\\_form.html](http://embnet.vital-it.ch/software/BOX_form.html)). The MS-identified protein sequences were analyzed with the ScanProsite and SMART program at the ExPaSy (<http://expasy.org/>), and EMBL (<http://smart.embl-heidelberg.de/>) for the domain/motif analysis to predict protein functions. Signal peptide of the all MS-identified proteins was predicted by using SignalP 4.1 server (<http://www.cbs.dtu.dk/services/SignalP/>) with a eukaryote D-cutoff value of 0.6. The pea aphid genome map was produced using karyoploteR ([bioconductor.org/packages/release/bioc/html/karyoploteR.html](http://bioconductor.org/packages/release/bioc/html/karyoploteR.html)) (Gel and Serra, 2017). Transcript similarity analysis was done using BLASTN comparing two or more sequences ([https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE\\_TYPE=BlastSearch](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch)). The KQY3 transcript without the secretory peptide and polyA regions was analyzed using BLASTN against single members of the KQY gene families also without their signal peptide and polyA nucleotides.

### DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in **Supplementary Table 1**, further inquiries can be directed to the corresponding author/s.

### AUTHOR CONTRIBUTIONS

MD and JO are co-first authors. All authors contributed to the article and approved the submitted version.

## ACKNOWLEDGMENTS

The authors wish to thank Nadya Galeva at the Mass Spectrometry & Analytical Proteomics Laboratory, The University of Kansas for advice with mass spectrometry analysis. FW and JO wish to thank the Kansas State University Arthropod Genomics Center of Excellence for funds to conduct this project.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.01230/full#supplementary-material>

## REFERENCES

- Boulain, H., Legeai, F., Guy, E., Morlière, S., Douglas, N. E., Oh, J., et al. (2018). Fast evolution and lineage-specific gene family expansions of aphid salivary effectors driven by interactions with host-plants. *Genome Biol. Evol.* 10, 1554–1572. doi: 10.1093/gbe/evy097
- Burger, N. F. V., and Botha, A. M. (2017). Genome of Russian wheat aphid an economically important cereal aphid. *Stand. Genom. Sci.* 28, 90. doi: 10.1186/s40793-017-0307-6
- Carolan, J. C., Fitzroy, C. II, Ashton, P. D., Douglas, A. E., and Wilkinson, T. L. (2009). The secreted salivary proteome of the pea aphid *Acyrtosiphon pisum* characterised by mass spectrometry. *Proteomics* 9, 2457–2467. doi: 10.1002/pmic.200800692
- Carolan, J. C., Caragea, D., Reardon, K. T., Mutti, N. S., Dittmer, N., Pappan, K., et al. (2011). Predicted effector molecules in the salivary secretome of the pea aphid (*Acyrtosiphon pisum*): A dual transcriptomic/proteomic approach. *J. Proteome Res.* 10, 1505–1518. doi: 10.1021/pr100881q
- Chaudhary, R., Atamian, H., Shen, Z., Briggs, S., and Kaloshian, I. (2015). Potato aphid salivary proteome: Enhance salivation using resorcinol and identification of aphid phosphoproteins. *J. Proteome Res.* 14, 1792–1778. doi: 10.1021/pr501128k
- Chen, M. S., Liu, X., Yang, Z., Zhao, H., Shukle, R. H., Stuart, J. J., et al. (2010). Unusual conservation among genes encoding small secreted salivary gland proteins from a gall midge. *BMC Evol. Biol.* 28, 296. doi: 10.1186/1471-2148-10-296
- Chen, W., Shakir, S., Bigham, M., Richter, A., Fei, Z., and Jander, G. (2019). Genome sequence of the corn leaf aphid (*Rhopalosiphum maidis* Fitch). *Gigascience* 8, 1–12. doi: 10.1093/gigascience/giz033. pii: giz033.
- Cook, D. E., Mesarich, C. H., and Thomma, B. P. (2015). Understanding plant immunity as a surveillance system to detect invasion. *Annu. Rev. Phytopathol.* 53, 541–563. doi: 10.1146/annurev-phyto-080614-120114
- Cooper, W. R., Dillwith, J. W., and Puterka, G. J. (2010). Salivary proteins of Russian wheat aphid (Hemiptera: Aphididae). *Environ. Entomol.* 39, 223–231. doi: 10.1603/EN09079
- Elzinga, D. A., and Jander, G. (2013). The role of protein effectors in plant-aphid interactions. *Curr. Opin. Plant Biol.* 16, 451–456. doi: 10.1016/j.pbi.2013.06.018
- Gel, B., and Serra, E. (2017). KaryoploteR: An R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* 33, 3088–3090. doi: 10.1093/bioinformatics/btx346
- International Aphid Genomics Consortium (2010). Genome sequence of the pea aphid *Acyrtosiphon pisum*. *PLoS Biol.* 23, e1000313. doi: 10.1371/journal.pbio.1000313
- Jaouannet, M., Rodriguez, P. A., Thorpe, P., Lenoir, C. J. G., MacLeod, R., Escudero-Martinez, C., et al. (2014). Plant immunity in plant-aphid interactions. *Front. Plant Sci.* 5, 663. doi: 10.3389/fpls.2014.00663
- Keller, A., Nesvizhskii, A. II, Kolker, E., and Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* 74, 5383–5392. doi: 10.1021/ac025747h
- Koonin, E. V., Csuros, M., and Rogozin, I. B. (2013). Whence genes in pieces: reconstruction of the exon-intron gene structures of the last eukaryotic

**SUPPLEMENTAL FIGURE 1 |** Full amino acid sequence alignment of the KQY salivary gland secretion protein candidates. The alignment was produced with the ClustalW multiple alignment program.

**SUPPLEMENTAL FIGURE 2 |** Full amino acid sequence alignment of the KHI salivary gland secretion protein candidates. The alignment was produced with the ClustalW multiple alignment program.

**SUPPLEMENTAL FIGURE 3 |** Pairwise BLASTP analysis of KQY family. Number indicates probability of match by chance (expect value). Cells of the same color indicate member of possible gene family at probability below 1e-05. Only one isoform was used for each gene.

**SUPPLEMENTAL FIGURE 4 |** Alignment of KQY3 transcript with other members of the KQY family by BLAST. Colored boxes indicate alignment scores above 40. Red ticks indicate relative location of the stop codon for each gene.

- common ancestor and other ancestral eukaryotes. *Wiley Interdiscip. Rev. RNA* 4, 93–105. doi: 10.1002/wrna.1143
- Legeai, F., Shigenobu, S., Gauthier, J. P., Colbourne, J., Rispe, C., Collin, O., et al. (2010). AphidBase: a centralized bioinformatic resource for annotation of the pea aphid genome. *Insect Mol. Biol.* 19 Suppl 2, 5–12. doi: 10.1111/j.1365-2583.2009.00930.x
- Li, Y., Park, H., Smith, T. E., and Moran, N. A. (2019). Gene family evolution in the pea aphid based on chromosome-level genome assembly. *Mol. Biol. Evol.* 36, 2143–2156. doi: 10.1093/molbev/msz138
- Miles, P. W. (1999). Aphid saliva. *Biol. Rev.* 74, 41–85. doi: 10.1111/j.1469-185X.1999.tb00181.x
- Mugford, S. T., Barclay, E., Drurey, C., Findlay, K. C., and Hogenhout, S. A. (2016). An immune-suppressive aphid saliva protein is delivered into the cytosol of the plant mesophyll cells during feeding. *Mol. Plant Microbe* 29, 854–861. doi: 10.1094/MPMI-08-16-0168-R
- Mutti, N. S., Park, Y., Reese, J. C., and Reeck, G. R. (2006). RNAi knockdown of a salivary transcript leading to lethality in the pea aphid, *Acyrtosiphon pisum*. *J. Insect Sci.* 6, 1–7. doi: 10.1673/031.006.3801
- Mutti, N. S., Louis, J., Pappan, L. K., Pappan, K., Begum, K., Chen, M.-S., et al. (2008). A protein from the salivary glands of the pea aphid, *Acyrtosiphon pisum*, is essential in feeding on a host plant. *Proc. Natl. Acad. Sci.* 105, 9965–9969. doi: 10.1073/pnas.0708958105
- Nesvizhskii, A. II, Keller, A., Kolker, E., and Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* 75, 4646–4658. doi: 10.1021/ac0341261
- Olsen, J. V., and Mann, M. (2004). Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. *Proc. Natl. Acad. Sci.* 101, 13417–13422. doi: 10.1073/PNAS.0405549101
- Pitino, M., and Hogenhout, S. A. (2013). Aphid protein effectors promote aphid colonization in a plant species-specific manner. *Mol. Plant Microbe* 26, 130–139. doi: 10.1094/mpmi
- Quan, Q., Hu, X., Pan, B., Zeng, B., Wu, N., Fang, G., et al. (2019). Draft genome of the cotton aphid *Aphis gossypii*. *Insect Biochem. Mol. Biol.* 105, 25–32. doi: 10.1016/j.ibmb.2018.12.007
- Rao, S. A. K., Carolan, J. C., and Wilkinson, T. L. (2013). Proteomic profiling of cereal aphid saliva reveals both ubiquitous and adaptive secreted proteins. *PLoS One* 8, 57413. doi: 10.1371/journal.pone.0057413
- Richards, S., Gibbs, R. A., Gerardo, N. M., Moran, N., Nakabachi, A., Stern, D., et al. (2010). Genome sequence of the pea aphid *Acyrtosiphon pisum*. *PLoS Biol.* 8, e1000313. doi: 10.1371/journal.pbio.1000313
- Rodriguez, P. A., and Bos, J. I. (2013). Toward understanding the role of aphid effectors in plant infestation. *Mol. Plant Microbe Interact.* 26, 25–30. doi: 10.1094/MPMI-05-12-0119-FI
- Shigenobu, S., Richards, S., Cree, A. G., Morioka, M., Fukatsu, T., Kudo, T., et al. (2010). A full-length cDNA resource for the pea aphid, *Acyrtosiphon pisum*. *Insect Mol. Biol.* 19, 23–31. doi: 10.1111/j.1365-2583.2009.00946.x
- Smithers, B., Oates, M., and Gough, J. (2019). ‘Why genes in pieces?’-revisited. *Nucleic Acids Res.* 47, 4970–4973. doi: 10.1093/nar/gkz284
- Thorpe, P., Cock, P., and Bos, J. (2016). Comparative transcriptomics and proteomics of three different aphid species identifies core and diverse effector sets. *BMC Genomics* 12, 172. doi: 10.1186/s12864-016-2496-6

- Tjallingii, W. F. (2006). Salivary secretions by aphids interacting with proteins of phloem wound responses. *J. Exp. Bot.* 57, 739–745. doi: 10.1093/jxb/erj088
- Tonkin, C. J., Foth, B. J., Ralph, S. A., Struck, N., Cowman, A. F., and McFadden, G.II (2008). Evolution of malaria parasite plastid targeting sequences. *Proc. Natl. Acad. Sci. U. S. A.* 105, 4781–4785. doi: 10.1073/pnas.0707827105
- Vanholme, B., Kast, P., Haegeman, A., Jacob, J., Grunewald, W., and Gheysen, G. (2009). Structural and functional investigation of a secreted chorismate mutase from the plant-parasitic nematode *Heterodera schachtii* in the context of related enzymes from diverse origins. *Mol. Plant Pathol.* 10, 189–200. doi: 10.1111/j.1364-3703.2008.00521.x
- Wang, X., Gao, B., and Zhu, S. (2016). Exon shuffling and origin of scorpion venom biodiversity. *Toxins (Basel)* 9, E10. doi: 10.3390/toxins9010010
- Wenger, J. A., Cassone, B. J., Legeai, F., Johnston, J. S., Bansal, R., Yates, A. D., et al. (2017). Whole genome sequence of the soybean aphid, *Aphis glycines*. *Insect Biochem. Mol. Biol.* 102917, 1–10. doi: 10.1016/j.ibmb.2017.01.005. pii: S0965-1748(17)30005-X.
- Will, T., Tjallingii, W. F., Thö, A., and Van Bel, A. J. E. (2007). Molecular sabotage of plant defense by aphid saliva. *Proc. Natl. Acad. Sci.* 104, 10536–10541. doi: 10.1073/pnas.0703535104

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer SH declared a past co-authorship with one of the authors AS to the handling editor.

Copyright © 2020 Dommel, Oh, Huguet-Tapia, Guy, Boulain, Sugio, Murugan, Legeai, Heck, Smith and White. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Advantages of publishing in Frontiers



## OPEN ACCESS

Articles are free to read for greatest visibility and readership



## FAST PUBLICATION

Around 90 days from submission to decision



## HIGH QUALITY PEER-REVIEW

Rigorous, collaborative, and constructive peer-review



## TRANSPARENT PEER-REVIEW

Editors and reviewers acknowledged by name on published articles

## Frontiers

Avenue du Tribunal-Fédéral 34  
1005 Lausanne | Switzerland

Visit us: [www.frontiersin.org](http://www.frontiersin.org)

Contact us: [info@frontiersin.org](mailto:info@frontiersin.org) | +41 21 510 17 00



## REPRODUCIBILITY OF RESEARCH

Support open data and methods to enhance research reproducibility



## DIGITAL PUBLISHING

Articles designed for optimal readership across devices



## FOLLOW US

[@frontiersin](https://www.instagram.com/frontiersin)



## IMPACT METRICS

Advanced article metrics track visibility across digital media



## EXTENSIVE PROMOTION

Marketing and promotion of impactful research



## LOOP RESEARCH NETWORK

Our network increases your article's readership