UC Riverside
UC Riverside Electronic Theses and Dissertations

Title

Characterizing the Dual Transcriptomes of Woolly Apple Aphid, Eriosoma lanigerum (Hausmann), and its Host, Malus domestica (Borkh.), Across a Host Resistance Spectrum

Permalink
https://escholarship.org/uc/item/8d97j1f6

Author

Wemmer, Joshua Daniel

Publication Date

2019

Supplemental Material
https://escholarship.org/uc/item/8d97j1f6#supplemental

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Characterizing the Dual Transcriptomes of Woolly Apple Aphid, *Eriosoma lanigerum* (Hausmann), and its Host, *Malus domestica* (Borkh.), Across a Host Resistance Spectrum

A Thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Science

in

Entomology

by

Joshua D. Wemmer

December 2019

Thesis Committee:
Dr. Paul Nabity, Chairperson
Dr. Kerry Mauck
Dr. Allison Hansen

Copyright by
Joshua D. Wemmer
2019

The Thesis of Joshua D. Wemmer is approved:

____

____

____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

Dr. Paul Nabity, thank you for encouraging curiosity and providing honest guidance. I have taken your words of advice to heart and passed them on to others. You pushed me when I needed it, led thought-provoking discussions and challenged me to solve complex problems with creative solutions. With still so much to learn, I was fortunate to be entrusted to work autonomously on some great lab projects. Whether it was how to care for plants and insets, execute molecular lab techniques, analyze data or how to navigate graduate school, I can't thank you enough for all your help, big and small.

Dr. Chaoyang Zhao, thank you for the teaching me the nuances of working with nucleic acids on the lab bench and computationally. I will miss the laughs and stories we shared and will miss the time we spent getting to know one another.

Thank you to the Washington Tree Fruit Research Commission for funding this project.

Thank you to the faculty, graduate students, administrative staff and everyone else who has contributed to the inspiring and supportive culture at UCR Entomology.

Thank you to my friends and family, without you, none of this would be possible.

DEDICATION

To my loved ones, who have nurtured my passion for biology.

ABSTRACT OF THE THESIS

Characterizing the Dual Transcriptomes of Woolly Apple Aphid, *Eriosoma lanigerum* (Hausmann), and its Host, *Malus domestica* (Borkh.), Across a Host Resistance Spectrum

by

Joshua D. Wemmer

Master of Science, Graduate Program in Entomology
University of California, Riverside, December 2019
Dr. Paul Nabity, Chairperson

Aphids are economically important, phloem-feeding insects that can cause substantial feeding damage to crops. The mechanisms underlying plant immunity to aphids are poorly characterized, but recent advances have shown a paradigm exists in plant immune signaling during aphid and pathogen attack. To overcome plant defenses, aphids, like other plant parasites, secrete effectors to modulate host processes that can lead to suppressed immunity. Emerging models of plant immunity combined with powerful sequencing technologies allow in-depth characterization of plant immune responses to aphids. In this thesis, I focus on the transcriptomic response of plants and aphids in the early stages of infestation across a host resistance spectrum to study the genes and processes that underlie plant defense responses. Also, I identify and characterize aphid candidate effectors to reveal insect proteins important for host colonization, and further refine our hypotheses of effector-mediated susceptibility.

Through a systems biology approach, this thesis investigates the molecular underpinnings mediating the interaction between the domesticated apple, *Malus domestica,* and the woolly apple aphid, *Eriosoma lanigerum*. I found plant signal transduction, primary metabolism, photophosphorylation, cell wall modification and phytohormone mediated signaling were altered by woolly apple aphid feeding, and chloroplastic-ROS production and retrograde signaling may be linked to plant resistance phenotypes. Furthermore, host genes related to gall formation, such as the upregulation of xylogen-like arabinogalactan proteins and proteins involved in cell wall loosening, suggest xylem differentiation and cell expansion occur immediately upon colonization. Salivary gland effector characterization confirmed the presence of aphid effector orthologs such as C002, Shp and Armet. Furthermore, several enzymes possibly involved in reactive oxygen species scavenging and plant cell wall degradation were also identified. Overall, the research improves our understanding of apple transcriptional responses to aphid feeding, identifies woolly apple aphid candidate effectors that likely alter host processes, and provides new hypotheses for the mechanisms of gall induction by this iconic pest of apple.

TABLE OF CONTENTS

INTRODUCTION ... 1
Plant immunity to aphids: The pathogen paradigm ... 1
Mechanisms underlying gall induction ... 4
Natural and altered life histories of the invasive woolly apple aphid ... 5
Apple rootstock resistance to WAA ... 8
Characterizing apple WAA responses across a resistance spectrum ... 11
MATERIALS AND METHODS ... 12
Plant and Insect Care ... 12
Tissue-specific RNA extraction and sequencing ... 13
Evaluation of Rootstock Resistance ... 13
Statistical analyses of Aphid Survival ... 14
Apple and aphid responses to colonization RNA-Seq experimental design ... 15
RNA extraction, library preparation, and sequencing ... 15
Malus bioinformatics ... 16
Malus differential gene expression and gene ontology ... 18
WAA *de novo* transcriptome assembly, annotation, and transcript quantification ... 20
WAA effector prediction ... 22
Host associated WAA differential gene expression and gene ontology ... 23
RESULTS AND DISCUSSION ... 24
Aphid Performance on Host Genotypes ... 24
Apple transcriptome undergoes remodeling shortly after colonization ... 26
G.16 decreased auxin biosynthesis in response to WAA feeding ... 33
Apple genotypes vary in constitutive expression of defense-related genes ... 34
Aphid salivary glands encode plant manipulating effector genes ... 35
CONCLUSION ... 39
LITERATURE CITED ... 41
TABLES AND FIGURES ... 53

LIST OF TABLES & FIGURES

Table 1. RNAseq data for 40 apple libraries: read quality, GC content, duplication percentage, reads mapped and total genes present in metabolism bins for deduplicated and non-deduplicated datasets ……………………………………………………...…..53

Table 2. Aphid RNAseq data and de novo transcriptome assembly summary for 12 whole body, and 3 salivary gland libraries. Read quality, reads mapped and total genes present in several metabolism GO terms are included. Transcriptome assembly and clustering were performed with Trinity and CD-HIT …………………………………………54-55

Table 3. Woolly apple aphid transcripts orthologous to previously characterized aphid effectors. BLASTP against the NCBI non-redundant database was used to determine orthology of WAA transcripts to effectors in other aphid species. The effectors C002, Shp, and Armet are orthologous to three WAA candidate effectors …...………..…..56-57

Figure 1. Aphid survival analysis on three host genotypes. Aphids performed equally well on G.16 and G.87, and significantly better on G.935. Host plants were challenged with WAA and aphids were counted 2, 4 and 6 DPI. Differences in rootstock resistances were determined with a Kaplan-Meier log-rank survival analysis. No aphids were found feeding on G.202 for the duration of the experiment, and therefore was omitted from the survival analysis …………………………………………………………………...58

Figure 2. Venn diagram of up and downregulated genes after aphid feeding on three host genotypes and total up and downregulated genes across MapMan bins for all aphid-infested vs all control plants. Aphid feeding induced greater changes to host gene expression in G.16 vs G.87 and G.935. Aphid feeding also resulted in more downregulated genes (blue values) compared to upregulated genes (red values). An FDR threshold of <0.1 was used to determine significantly differentially expressed genes. In all aphid-infested vs control plants, 32428 total genes were tested for differential expression, and 675 were differentially expressed. Photosynthesis and cell wall organization bins were enriched as determined by Wilcoxon-rank sum tests ……..…59-60

Figure 3. Numbers of up and downregulated genes across MapMan bins for G.16 infested vs control plants. Cell wall organization, carbohydrate metabolism, RNA biosynthesis, and auxin synthesis bins were enriched as determined by Wilcoxon-rank sum tests. 32428 total genes were tested for differential expression, and 790 were differentially expressed at an FDR threshold of <0.1 …………………………………………………..…..61-62

Figure 4. Differentially expressed genes across uninfested genotypes. Control plants of each genotype were contrasted to obtain profiles of constitutively expressed genes at FDR <0.1. G.16 constitutive expression was highly dissimilar to G.935 and G.87, in terms of numbers of DE genes and enriched bins, which included cell wall organization, protein modification, and enzyme classification ……………………………………63-64

INTRODUCTION

Aphids are phloem-feeding insects that evolved in remarkable ways to suppress plant defenses and enhance the nutritional quality of host tissue (Elzinga et al. 2014; Sandström et al. 2000; Will et al. 2007). Aphids have specialized, syringe-like mouthparts termed stylets that facilitate their piercing-sucking mode of feeding. More than 4000 aphid species have been described, with about 10% of them classified as global crop pests because they cause substantial damage by depleting plant fluids, injecting compounds that induce abnormal plant phenotypes, such as leaf curls and galls, and transmitting disease-causing viruses and bacteria (Blackman & Eastop 2000). To survive and reproduce aphids must overcome significant hurdles shared by all herbivores: suppressing plant immune responses and overcoming nutritional deficiencies.

Plant immunity to aphids: The pathogen paradigm

The mechanisms underlying plant immunity against aphids are not well understood, in part because of the subtle nature by which aphids feed (Züst & Agrawal 2016). Plants perceive biotic stress through generic and specific molecular patterns, which result in the activation of immune responses that can be conceptualized in two types of immunity: pattern-triggered immunity (PTI) and effector-triggered immunity (ETI) (Jones & Dangl 2006). Together, PTI and ETI operate in a zig-zag model where the general mechanisms are well studied in plant-microbe interactions and that has guided

research aimed at understanding the specific molecular mechanisms of plant defense to insect herbivores. Plants perceive pathogens through pattern-recognition receptors (PRRs) that recognize conserved pathogen-derived molecules known as pathogen-associated molecular patterns (PAMPs) (Zipfel 2014). PTI is activated upon PAMP binding by PRRs, followed by phytohormone, reactive oxygen species (ROS) and ion-mediated signal transduction to induce defense-related gene expression (Dodds & Rathjen 2010; Jones & Dangl 2006). Additionally, plant-derived "danger signals" produced during physical damage by pathogens called damage-associated molecular patterns (DAMPs) also elicit DAMP-triggered immunity (Heil & Land 2014). In response to plant defenses, pathogens have evolved secretory molecules called effectors to attenuate immune responses and successfully colonize their hosts (Toruño et al. 2016). However, plant recognition of specific effectors through intracellular nucleotide-binding leucine-rich repeat proteins, also known as resistance or R proteins, activates ETI, which is a strong and quick immune response (Dodds & Rathjen 2010).

Aphids, like other plant parasites, interact with plants on a molecular level to suppress defenses and enhance nutrition. Aphids probe and feed, traversing their piercing stylet through the apoplast to the phloem, but also puncturing cells and interacting with cytoplasm and xylem for short durations (Tjallingii & Esch 1993). Plant defenses are triggered through the recognition of aphid-derived compounds, generally known as herbivore-associated molecular patterns (HAMPs). Aphids overcome the triggered defenses by deploying effectors present in the saliva that manipulate plant signaling

pathways, thus promoting susceptibility (Hogenhout & Bos 2011; Rodriguez & Bos 2013). Some plants may carry *R* (Resistance) genes whose products recognize specific effectors to activate ETI (Kaloshian et al. 2014; Stuart 2015), yet the function of numerous predicted effectors remains unknown. Bioinformatically driven examinations of whole body and tissue-specific transcriptomes, in addition to proteome screens of salivary secretions predict aphids retain large arsenals of secretory molecules with various enzymatic or protein binding properties akin to those validated as effectors in microbes (Bos et al. 2010; Carolan et al. 2011). Despite some similarities to effectors in other organisms, the degree of host specificity, manner of induced phenotype, and rapid evolution of gene families suggests aphids evolved lineage-specific effector arsenals (Boulain et al. 2018). Characterizing effector functions and their plant targets in insects remains a major challenge as many significant pest species represent non-model organisms where functional genetics techniques are lacking. Thus, comparative gene expression profiling of salivary tissues with whole bodies, and gene expression profiling of plants during the first stages of successful colonization are robust techniques to predict aphid effectors and link them to initial changes in plant signaling. So far, our current hypotheses of plant immunity to aphids are informed by the models developed in plant-pathology, but the specific mechanisms underlying plant perception of aphids, signal transduction and subsequent gene expression of defenses remain elusive.

Mechanisms underlying gall induction

Some aphids manipulate their plant hosts by inducing the formation of abnormal tissue called galls, which provide the aphid with a nutritionally stable and safe feeding environment typically devoid of plant defenses (Stone & Schönrogge 2003). The drastic changes to plant host physiology and anatomy illustrate the complexities of gall formation, but the mechanisms of how insects induce these extended phenotypes are unclear. Current hypotheses of gall induction and maintenance suggest either plant hormones initiate the process (Tooker et al 2014), or that effectors drive cell differentiation (Chen et al. 2008; Nabity 2016; Zhao et al. 2015). In addition to modifying the final plant phenotype, galling insects reprogram host transcriptomes to alter nutritional and defensive statuses in favor of the galler (Giron et al. 2016). A systems biology approach that involves transcriptomic profiling of both aphids and infested host tissue under varying conditions of susceptible and resistant host genotypes can link gene expression patterns in both insect and plant host, thereby revealing genes interacting during gall initiation. Here we characterize the dual transcriptomes of the tree-galling woolly apple aphid (WAA) and its host the domesticated apple during early stages of infestation across a resistance spectrum. We also perform salivary gland specific tissue extractions and RNAseq followed by an *in silico* prediction protocol to identify effector candidates in the secretome.

Natural and altered life histories of the invasive woolly apple aphid

The woolly apple aphid (WAA), *Eriosoma lanigerum* (Hausmann), has plagued apple orchards since domesticated apples, *Malus domestica*, were first cultivated in eastern North America by European settlers over 200 years ago (Baker 1915). The WAA soon invaded other apple growing regions around the world, most likely through the transportation of nursery rootstocks (Shoene & Underhill 1935). Now a globally invasive pest of apples, the WAA biology and ecology have been well studied by many entomologists to determine effective methods of control (Shoene & Underhill 1935).

The ancestral, primary host of WAA is the American elm, *Ulmus americana,* where both sexual and asexual/parthenogenic reproduction occur, as well as overwintering of the egg form (Patch 1916). Both the American elm and WAA are native to eastern North America. The emergence of the stem mothers, or fundatrices, from the egg form, occurs in the spring. This fundatrix is a specialized form that moves to the terminal buds of young elm branches where rosette-shaped galls form. Elm rosette galls, which are characterized by the bunched growth of leaves at the meristematic terminal buds, provide feeding sites for multiple generations of wingless parthenogenetic females. The emergence of the winged spring forms marks the start of migration by flight to summer hosts (Baker 1915). The genus, *Eriosoma*, contains 38 described species of woolly aphids and most of them host alternate between galls on *Ulmus* (elms) and summer, secondary hosts in a variety of genera from the families Rosaceae and Asteraceae (Sano & Akimoto 2011). WAA has been reported to feed on summer host

plants in the genera *Malus, Crataegus, Cotoneaster, Cydonia, Pyrus* and *Sorbus* (Blackman et al. 1994; Blackman & Eastop 2000). Upon moving to the summer hosts, wingless females will undergo multiple rounds of parthenogenetic reproduction. At the end of summer, winged Fall migrants migrate back to American elm to asexually produce wingless males and females that mate and generate the overwintering egg form. A single egg will nearly occupy the entire abdomen of a female, and after deposition in the crack of the elm bark, the female dies shortly thereafter (Baker 1915).

Invasive populations of WAA exist in apple-growing regions where the primary host, American elm, does not exist, revealing a shift in WAA life history. During the winter in temperate regions, wingless parthenogenetic females are able to escape cold conditions by feeding on roots below the frost line, where clonal reproduction can continue indefinitely (Theobald 1922). In WAA populations where the American elm is absent, the frequency and importance of sexual reproduction remain unclear (B. Lavandero et al. 2009; Sandanayaka & Bus 2005; Timm et al. 2005). Sexual morphs of WAA have been recorded in areas without the American elm in 14 countries across 4 continents (summarized in Table 1 by Sandanayaka and Bus, 2005). Experimentally mated WAA males and females have been observed depositing eggs on apple leaves in greenhouse and outdoor conditions, but the emergence of stem mothers from these eggs was not confirmed (Sandanayaka & Bus 2005). In the Western Cape Province of South Africa, where winters are mild, asexual forms of WAA predominate, and although winged forms are found in the Fall, they do not produce sexual males and females

(Pringle & Heunis 2001). Furthermore, low levels of genetic variation were estimated based on analyzing amplified fragment length polymorphisms of 192 individuals from four different apple-growing regions in the Western Cape (Timm et al. 2005). These results suggest high relatedness among WAA and a lack of sexual recombination in the Western Cape of South Africa. Relative to the cold winters in the native range of WAA, winters in the Western Cape are mild with higher temperatures and sunshine. For many aphid species changes in sexuality are triggered in response to environmental cues such as short-day lengths and low temperatures (Blackman & Eastop 2000; Ogawa & Miura 2014). Therefore, low sexual reproduction and genetic diversity of WAA in South Africa is possibly due to a lack of the required environmental cues to trigger polyphenic sexual development.

In contrast, the genetic diversity based on microsatellite markers of WAA populations in Chile is geographically structured, with high levels of heterozygosity and low linkage disequilibrium, which suggests sexual reproduction readily occurs (B. Lavandero et al. 2009a; B. Lavandero et al. 2009b; Lavandero et al. 2011). In Chile, sexuality change of WAA may be triggered by short-day lengths and low temperatures, but future studies will need to determine the environmental conditions that trigger sexuality in WAA as well as scan WAA populations in orchards for sexual morphs, eggs, and fundatrices. Sexual reproduction of WAA can have important effects on genomic diversity especially as it relates to agricultural control strategies for this pest (Domes et

al. 2007; Gilabert et al. 2009; Hoffmann et al. 2008; Miller et al. 2012; Valenzuela et al. 2010).

Apple rootstock resistance to WAA

In some regions, WAA is controlled with broad-spectrum pesticides, but these practices disrupt the natural enemy populations and lead to pest resurgence (Cohen et al. 1996; Penman & Chapman 1980; Rogers et al. 2011). Additionally, the detrimental effects of pesticides select for resistant aphid genotypes and increase management costs or prohibit product entry into organic markets through lingering residues and environmental contamination (Gill & Garg 2014). Integrated pest management strategies, which include biological control from natural enemies and planting resistant varieties of apple provide effective, economically and environmentally sustainable methods of WAA control (Blommers 1994; Bus et al. 2008; Bus et al. 2010; Nicholas et al. 2005; Wearing et al. 2010). While above-ground WAA colonies can be controlled by natural enemies and pesticides, root colonies exhibit low levels of predation and parasitism and are difficult to control with pesticides (Gontijo et al. 2012). Controlling root-feeding WAA with resistant rootstocks is a preferred tactic that has led to the discovery of novel sources of apple resistance and the development of commercial breeding programs (Bus et al. 2008; Bus et al. 2010; Crane et al. 1937; Cummins et al. 1981; Fazio et al. 2015).

Definitions of resistance in crop plants often depend on the specific plant-insect interactions studied, with resistance defined from the perspective of either the insect, the plant, or both (Painter 1951). From the plant perspective, resistance can be determined through experimental comparisons across host genotypes that display different phenotypes or alter insect performance relative to another cultivar (Painter 1951). In the early 1900s, severe outbreaks of WAA in apple-growing regions prompted the East Malling Research Station and the John Innes Horticultural Institution to collaboratively breed for rootstocks immune to WAA (Crane et al. 1937). A particular cultivar, 'Northern Spy', was a promising source of WAA immunity and was used in many experimental crosses (Crane et al. 1937). The breeding experiments resulted in the production of commercially available WAA resistant rootstock varieties termed the Malling-Merton rootstock series (Crane et al. 1937). The 'Northern Spy' derived WAA resistance was later identified to be controlled by a single dominant locus, which was named *Er (Knight et al. 1962). Er* was later renamed to *Er1* due to the identification of another WAA resistance locus, *Er2*, which is derived from *Malus* x robusta species. *Er1*-based resistance was an essential component of managing WAA until populations overcame *Er1*-containing genotypes and resurged in high numbers in some major apple-growing regions (Ateyyat & Al-Antary 2009; Giliomee et al. 1968; Rock & Zeiger 1974). The increasing importance of WAA as a pest has prompted rootstock breeders to search for new genes that can provide durable resistance to WAA.

Currently, genetically based resistance to WAA is known to be associated with at least four quantitative trait loci (QTLs), *Er1, Er2, Er3,* and *Er4* (Bus et al. 2008; Bus et al. 2010). The *Er1* QTL is derived from 'Northern Spy' and is the first locus that was discovered to confer resistance to WAA (Crane et al. 1937). Aphid performance assays revealed contrasting results for resistance to 'Northern Spy' with reports of aphids being able to colonize plants and form galls, and other studies reporting very low survival with no gall formation (Sandanayaka et al. 2003). Compared to very susceptible apple cultivars, WAA on 'Northern Spy' ingested less phloem, which may suggest that resistance factors are active in the phloem or reduce access to the phloem (Sandanayaka et al. 2003).

The *Er2* locus confers cross-resistance to WAA and powdery mildew and originated from the crab apple cultivar 'Robusta 5' (King et al. 1991). With WAA resistance as a prominent feature, 'Robusta 5' has been used as a parent for many of the Geneva series of rootstocks including G.202, G.41, G.214 and G.210 (Fazio et al. 2015). Aphid performance on 'Robusta 5' resulted in very low survival, with significantly shorter phloem ingestion and a high percentage of aphids not reaching the phloem, suggesting a form of antixenotic resistance (Sandanayaka et al. 2003). A microarray analysis including a segregating population of 48 F1 generation apples between a cross of 'Ottawa 3' and 'Robusta 5' cultivars showed patterns of steady-state expression levels between trees resistant and susceptible to WAA. Only 7 genes were differentially expressed between resistant and susceptible trees, with 5 of the genes physically

clustering on chromosome 17 in the location of the previously mapped Er2 QTL (Jensen et al. 2014). Steady-state differential expression of the 5 genes clustered in the Er2 locus, is suggestive of a constitutive, antixenotic defense and consistent with the findings of Sandanayaka et al. (2003).

WAA performance assays have been used to assess rootstock resistance, while few studies have attempted to characterize host phenotypes associated with resistance loci or the molecular mechanisms underlying genetically based resistance for any of the 4 *Er* loci. Understanding the fundamental processes that underlie WAA colonization and host immune responses are crucial for predicting the efficacy and durability of resistance in the field.

Characterizing apple WAA responses across a resistance spectrum

A challenge for breeders in search of durable resistance to aphids is the lack of a fundamental understanding of the mechanisms underlying plant immunity to aphids. Characterizing gene expression during early stages of aphid colonization on resistant and susceptible host genotypes can shed light on important plant immune processes, such as aphid recognition, wound response, early signal transduction, changes to cell wall organization and changes to primary metabolism (Kuśnierczyk et al. 2008; Niu et al. 2018; Tetreault et al. 2019; Woldemariam et al. 2011). Transcriptional profiling of colonized plants coupled with the knowledge of the effectors deployed by the attacking aphid can provide insight into the cellular processes occurring at the plant-stylet

interface. Here I profiled the transcriptomes of resistant and susceptible apple rootstock genotypes during early WAA infestation. I also characterized the genes present in salivary glands and active during the initial phases of colonization.

MATERIALS AND METHODS

Plant and Insect Care

Apple rootstock varieties G.202, G.935, G.87, and G.16 were grown at the USDA Plant Genetic Resources Unit (PGRU) in Geneva, NY and shipped to Riverside, CA for aphid survival and RNA-seq experimentation. Both G.87 and G.935 are progeny of crosses between two highly heterozygous parents *Malus domestica* 'Ottawa 3' and *Malus robusta* 'Robust 5', and G.16 is the progeny of a cross between *Malus floribunda* and 'Ottawa 3' (Norelli et al. 2003). These crosses have been used extensively in the Apple Rootstock Breeding Program in Geneva, NY, resulting in phenotypically diverse populations that segregate for resistance to fire blight (*Erwinia amylovora*), powdery mildew (*Podosphaera leucotricha*) and WAA (Jensen et al. 2014; Robinson et al. 2002). Plants were grown under greenhouse conditions (28°C +/- 2°C) and watered with 20:20:20 NPK fertilizer regularly from June to October 2017. Populations of WAA were collected from apple trees in Quincy, Washington, USA on 'Fuji', and Placerville, California, USA on 'Mark' rootstocks. Washington (WA) and California (CA) populations were reared on *M. domestica* B.9 rootstocks and maintained in separate cages

within the greenhouse. Field collected aphids from Quincy, Washington, USA on ‘Fuji’ were established on M9 Nic□ 29 and used for tissue-specific RNA-seq and effector prediction.

Tissue-specific RNA extraction and sequencing

To isolate the transcripts most strongly expressed in salivary glands (SG), SG pairs from fourth instar larvae and wingless adult WAA were dissected under a Zeiss Stemi 508 stereoscope in phosphate-buffered saline (pH 8.0) and placed immediately into TRIzol solution for storage. For each RNAseq sample, dissected SG pairs were pooled from approximately 300 individuals. Samples of whole-body (WB) aphids were comprised of 10 fourth instar larvae and wingless adult WAA. Total RNA of all pools was extracted using a combination of a TRIzol RNA isolation protocol and an RNeasy Mini Kit (Qiagen, Hilden, Germany). Extracted RNA was assessed for quality and quantity using an Advanced Analytics Fragment Analyzer and RiboGreen quantification kit, respectively. 100bp paired-end libraries were built with the Illumina TruSeq RNA kit and assessed again for quality as above. Library preparation and RNA-Seq of the WB and SG samples was completed at the WSU Genomics Core on an Illumina HiSeq2500 in 2016.

Evaluation of Rootstock Resistance

To confirm rootstock resistance, aphid survival assays were conducted for the four Geneva rootstock varieties: G.202, G.16, G.87, and G.935. Survival assays were

conducted in a Conviron growth chamber (55-65% relative humidity, 28:25°C L:D, 15:9 L:D photoperiod, 400 µmol m^{-2} s^{-1}) where rootstocks were acclimated to growth chamber conditions for 4 days prior to aphid infestation. Fourteen individual plants of G.202, G.16, and G.87 and four individual plants of G.935 were used in the experiment. All 46 plants were randomly assigned positions within the growth chamber. To collect aphids for infestation, 10, 1st instar aphids were selected and placed into PCR tubes. A total of 920 aphids were infested, using 2 PCR tubes randomly assigned and securely fastened to each plant. In this setup, 20 aphids were introduced to each plant, and the vials were opened so aphids were able to freely migrate onto the plant. Aphid survival was quantified by counting the number of live aphids on each plant at 2, 4 and 6 days post infestation (DPI). Plants were thoroughly checked for WAA above the soil line, and 1 inch below the soil line.

Statistical analyses of Aphid Survival

The probability of aphid survival on a rootstock variety was calculated with Kaplan-Meier survival analysis using the 'survival' package in R version 3.3.1 (Therneau & Lumley 2014). To test for significant differences of aphid survival between the rootstock varieties, pairwise log-rank tests were conducted in the 'survminer' package (Kassambara et al. 2017). To decrease the likelihood of obtaining false-positive results, the calculated p-values were adjusted using a Bonferroni-Hochberg correction.

Apple and aphid responses to colonization RNA-Seq experimental design

M. domestica rootstock varieties with variable aphid performance (G.935, G.16 and G.87) were chosen for the transcriptomic analyses, whereas G202 was excluded due to zero established WAA feeding sites. Plants were transferred from the greenhouse and allowed to acclimate for four days with the same growth chamber conditions used in the aphid survival assay. Plants were randomly assigned treatments as uninfested (controls) or infested with a CA or WA strain of WAA. An equal number of mixed-age aphids were introduced to plants in PCR vials as described in the methods of the aphid survival assay. Each vial contained a set of either CA or WA strain of WAA comprised of 15 first instar, 3 fourth instar, and 3 wingless adults. Plant tissue and WAA were harvested from feeding sites 56-64h after aphid exposure. Harvested tissue was immediately flash-frozen in liquid nitrogen and stored in microcentrifuge tubes at -70 °C. Similar tissue types and amounts were harvested from control plants. Aphids were later separated from plant tissue and pooled in separate tubes for RNA extraction. Nine pooled samples of aphids each yielding approximately 20 aphids, three collected from each G.16 and G.87 and G.935, were prepared for RNAseq.

RNA extraction, library preparation, and sequencing

Insect and plant samples were ground using liquid nitrogen and total RNA was extracted using an RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) as per the manufacturer's instructions. The quality and quantity of RNA were confirmed with an

Agilent 2100 Bioanalyzer (Agilent, CA, USA) and Spectrophotometer (NanoDrop ND-8000, Thermo Scientific, MA). Only 40 apple samples from the original 51 were chosen for library construction and RNAseq based on quality assessments. Libraries were prepared, multiplexed and sequenced at UC Davis Genome Center in 2018 (Tables 1 & 2). 100ng of total RNA was used as input for the KAPA Hyper mRNA library prep kit. Strand-specific and barcode indexed RNA-seq libraries were generated from each after poly-A enrichment using the KAPA mRNA-seq Hyper kit (Kapa Biosystems, Cape Town, South Africa) following the instructions of the manufacturer. RNA was fragmented after poly-A enrichment at 94°C for 5 minutes. Libraries were amplified with 13 cycles of PCR. The fragment size distribution of the libraries was verified via micro-capillary gel electrophoresis on a Bioanalyzer 2100 (Agilent, Santa Clara, CA). The libraries were quantified by fluorometry on a Qubit fluorometer (LifeTechnologies, Carlsbad, CA) and pooled in equimolar ratios. The pool was quantified by qPCR with a Kapa Library Quant kit (Kapa Biosystems) and sequenced on three lanes of an Illumina HiSeq 4000 (Illumina, San Diego, CA) with paired-end 150bp reads. Forward and reverse reads of each sample were divided and sequenced across the three flow cell lanes to control for lane effects.

Malus bioinformatics

For all apple transcriptome analyses, we employed the *Malus* × *domestica* Golden Delicious doubled-haploid genome (GDDH13 version 1.1) available from

(https://iris.angers.inra.fr/gddh13/). We annotated the GDDH13 proteome with the Mercator4 pipeline, which assigns gene ontologies for use in the MapMan4 framework (Schwacke et al. 2019). Of the 45,116 proteins in the proteome 31,066 were annotated by Mercator4 with 20,864 being placed into functional categories called “bins”.

Reads were adapter-trimmed using BBDuk v.0.35 (https://sourceforge.net/projects/bbmap/). Read quality was assessed with FastQC software version 0.11.3. (Andrews 2010). All FastQC reports were aggregated for visualization using MultiQC (https://multiqc.info/). Reads were then aligned to the GDDH13 genome using the RNA-seq aligner STAR v2.5.3a (Dobin et al. 2013). The average percentage of mapped reads for each sample was 94.3%, and the average percentage of uniquely mapped reads was 55.6%. RNAseq is susceptible to generating highly duplicated reads, which are artifacts in libraries and can bias differential gene expression analysis towards false positives (Klepikova et al. 2017). Duplicated reads may be removed *in silico*, however, choosing to do so may bias differential gene expression towards more false negatives (Klepikova et al. 2017). Because our RNAseq libraries had high duplication levels (mean=83.8%), we tested the effect deduplication had on detecting differentially expressed (DE) genes between treatments. Therefore, two sets of BAM files were created, one set with duplications retained and another set with duplicated reads removed using the MarkDuplicates tool from Picard 2.6.0 (https://broadinstitute.github.io/picard/). Mapped reads were quantified at the gene level with HTSeq□count software (Anders et al. 2015) and the GDDH13 v1.1 general feature

format file (.gff3) from (https://iris.angers.inra.fr/gddh13). A summary of library characteristics for each sample is available in Table 1.

Malus differential gene expression and gene ontology

To test for DE genes between plant contrasts of interest, we employed the R Bioconductor packages *edgeR* for normalization and *limma-voom* for variance-stabilized transformation, linear modeling, and an empirical Bayes moderated statistics computation. Non-expressed and lowly expressed genes were filtered from the count matrix by retaining genes with more than 40 counts across all samples. Filtered gene counts were normalized for library size with the Trimmed Means of M-values (TMM) method implemented in *edgeR*.

To assess the similarity of expression profiles among samples, we performed a multi-dimensional scaling (MDS) analysis, hierarchical clustering and for each sample, we quantified the number of genes assigned to the bins 'carbohydrate metabolism', 'lipid metabolism' and 'amino acid metabolism'. The assessment revealed several samples that deviated from normal expression patterns. To pinpoint the cause of high variation for these samples, we examined their initial RNA concentrations and Bioanalyzer results and found one sample (51c) as an outlier with low RNA quantity and quality due to degradation. Low starting sample quality is known to decrease library complexity and

bias measurements of differential gene expression (Gallego Romero et al. 2014); to avoid bias we removed 51c from the downstream analyses.

DE genes were detected using the voom method from the *limma* Bioconductor package in R software v.3.6.0 (Law et al. 2014). The voom method estimates mean-variance relationships and calculates precision weights for each individual observation. DE genes were then identified between contrasts of interest by performing linear modeling and empirical Bayes analyses in limma. Genes with a Benjamini-Hochberg adjusted *p*-value of less than 0.1 were considered significantly differentially expressed.

We assessed how high sequence duplication affects differential expression patterns by applying the above differential expression analysis to both deduplicated and non-deduplicated datasets. Deduplicating sequences increases the detection of DE genes across all contrasts. For all infested plants compared to all controls, we detected 675 DE genes in the deduplicated dataset, 461 DE genes in the duplicated dataset and 414 common to both. The genotype-specific contrasts of infested vs control plants also resulted in more DE genes in the deduplicated dataset. Therefore, all subsequent apple analyses were carried out with the deduplicated dataset.

We characterized the global responses induced by aphid feeding across infested vs control plants for all samples, for each genotype, and we also contrasted control plants of each genotype to characterize differences in their constitutive expression. Differential

expression results were interpreted by visualizing biological pathways and processes using MapMan (Thimm et al. 2004; Usadel et al. 2005). To highlight bins containing DE genes with significantly higher absolute log-fold changes, within the MapMan software, Wilcoxon rank-sum *p*-values were calculated for each bin by comparing the log-fold changes of the occupying DE genes compared to all other DE genes. We also examined the subcellular localization of DE genes with TargetP v.1.1 software, which predicts the presence of N terminal presequences: chloroplast transit peptide, mitochondrial transit peptide, and signal peptide.

WAA *de novo* transcriptome assembly, annotation, and transcript quantification

Raw reads were adapter trimmed assessed for read quality as described above (Andrews 2010). Two libraries, EL87_3 and EL16_3, were not used for *de novo* transcriptome assembly and downstream analyses due to high GC content and sequence duplication levels, which corresponded with abnormal expression profiles. Therefore, RNAseq data from 10 WB libraries and 3 SG libraries were *de novo* assembled using Trinity software, which predicts alternatively spliced products from the same gene, and therefore produces contigs labeled at the putative 'isoform' and 'gene' level. (Grabherr et al. 2011). To reduce transcript redundancy, CD-HIT software (Li & Godzik 2006) was used to cluster highly similar transcripts at a 95% identity threshold. The assembly produced 318,749 contigs which were reduced to 268,921 after clustering with CD-HIT. The remaining assessments and analyses were performed with the clustered assembly.

The average contig length was 780 bases, while the GC content was 30.81%, which is similar to previously published aphid transcriptomes (Thorpe et al. 2016) (Table 2). The completeness of the transcriptome assembly was assessed with BUSCO (Benchmarking Universal Single-Copy Orthologs) v.3 (Simão et al. 2015) to obtain the number of conserved single-copy, insect-specific orthologs represented in the assembly. Out of 1658 conserved orthologs in nearly all insects, 96.4% of those orthologs were found in our assembly, indicating a high completeness. Fasta files with complete protein sequences with a minimum length of 70 amino acids and corresponding nucleotide sequences were produced with Transdecoder v.5.0.2, which predicted 69,099 total proteins, of which 37,159 were complete, meaning they contained a start and stop codon. We annotated the transcriptome with Gene Ontology (GO), Kyoto Encyclopedia of Genomes and Genes (KEGG), and Clusters of Orthologous Genes (COG) term assignments using Trinotate v.3.1.1 (https://trinotate.github.io/) and the following sequence analyses: functions were assigned with BLASTP and BLASTX v.2.2.3 searches against a Swissprot database with e-value cutoffs of 0.001 (Camacho et al. 2009), and searches against a PFAM database using HMMER v.3.0 (http://hmmer.org/). Predictions of transmembrane domains and signal peptides were conducted with TMHMM v.2.0c and SignalP v.4.1c, respectively. The results were parsed for storage in an SQLite database, and a tab-delimited annotation report file was generated with Trinotate. We quasi-mapped and quantified transcript abundance for each sample with Salmon v.0.14.1 (Patro et al. 2017), a quasi-mapping based index was built with the transcriptome assembly using a k value of 31. Salmon

quantification of transcript abundance resulted in an average of 93.6% mapping rate for SG and WB libraries.

WAA effector prediction

To identify putative effectors in WAA, we performed an *in silico* prediction by first identifying differentially upregulated transcripts in SG libraries, then passing these transcripts through a bioinformatic pipeline to identify potential proteins with secretory motifs and unique to herbivore lineages based on sequence homology. For DE analysis, a Salmon-quantified count matrix of 3 SG and 3 WB pooled libraries was filtered for lowly expressed genes by including a 1 count per million (CPM) cut-off in at least two samples, leaving a total of 26,740 transcripts tested for DE. The counts were *edgeR* TMM normalized, converted to log2 CPM and DE genes were determined with the limma-trend pipeline. For effector prediction, we considered transcripts upregulated in SGs with an adjusted *p*-value of less than 0.05 and a log2 fold change of greater than 1.5, which resulted in 5,377 'isoform' level upregulated transcripts. Of the SG-upregulated transcripts, we obtained 2,311 protein sequences with both start and stop codons predicted from the Transdecoder-produced peptide fasta file. To determine if the upregulated genes were candidate effectors, secretory proteins were identified based on having a signal peptide predicted by SignalP, no transmembrane domains predicted by TMHMM, and extracellular localization predicted by TargetP v.1.1. The remaining candidate effectors (n=390) were then blasted (BLASTP e-value cutoff=0.1) against an

in-house generated non-herbivorous insect database spanning across insect taxa. We assigned putative functions to candidate effectors using homology to proteins found in the NCBI non-redundant database (BLASTP e-value cutoff=0.1), which provided the top annotated hits to similar aphid proteins.

The secretome was predicted for the WAA transcriptome as well as all expressed SG transcripts that met a CPM>1 threshold in at least 2 libraries and contained a complete protein-coding sequence. Secretory pathway prediction and herbivore specificity were assessed with the same tools and parameters previously stated for the SG-upregulated transcripts.

Host associated WAA differential gene expression and gene ontology

RNAseq libraries generated from WAA collected on the genotypes G.87, G.16 and G.935 were tested for DE with the Salmon-quantified count matrix which included 2 sample replicates of G.16 and G.87 and 3 replicates of G.935. The count matrix was filtered for lowly expressed genes by including a 1 CPM cut-off in at least two samples, leaving a total of 29,599 transcripts tested for DE. The counts were *edgeR* TMM normalized, converted to log2 CPM values and the limma-trend pipeline was applied. Transcripts with an adjusted *p*-value of less than 0.1 were considered significantly differentially expressed between treatments.

RESULTS AND DISCUSSION

Aphid Performance on Host Genotypes

Aphid survival varied depending on the plant genotype and its genetic background ($\chi^2_{(2)} = 23.8$, $P < 0.0001$). On G.935, aphid numbers decreased between 0 and 2 DPI, but then remained constant at 71% until 6 DPI (Fig. 1), indicating tolerance to or a lack of inducible defenses. Aphids performed similarly on G.16 and G.87, where aphid survival declined rapidly to ~53% between 0 to 2 DPI and then declined steadily to ~40% by 6 DPI (Fig. 1). A log-rank Kaplan-Meier survival analysis revealed a significant difference between genotypes (Fig. 1), and pairwise log-rank tests between genotypes showed statistically significant differences between G.935 vs G.16 and G.935 vs G.87 ($P < 0.0001$). For all four genotypes, there were no visible signs of a hypersensitive response (i.e. necrosis), which is a common defense response against insects with established feeding sites, such as galling insects (Fernandes & Negreiros 2001). Furthermore, no aphid mortality was observed after thoroughly scanning infested plants. Therefore, we hypothesize that antixenotic factors played a role in early defense responses for these genotypes. Previous characterizations of WAA performance on apple genotypes derived from 'Robusta 5' and *M. floribunda* genetic backgrounds have shown similar WAA settling and feeding behaviors indicative of antixenotic characteristics (Sandanayaka et al. 2003; Sandanayaka et al. 2005).

Although G.87 and G.935 are both progenies from crosses between ‘Ottawa 3’ and ‘Robusta 5’, and thus share a similar genetic background compared to G.16 (‘Ottawa 3’ x *M. floribunda*), the difference in their resistance to WAA highlights the phenotypic diversity produced from crossbreeding highly heterozygous parents, in this case a congeneric hybridization. A single locus, *Er2*, has been identified as a genetic source of WAA resistance in ‘Robusta 5’ progeny, and molecular markers linked to *Er2* have been found in G.87, but not G.935 (Fazio & Beers 2010). Profiling the G.87 and G.935 transcriptomes under biotic stress from WAA will provide a more robust molecular characterization of their immune responses. WAA perform almost identically on G.87 and G.16, despite the substantial differences in genetic backgrounds. The inverse relationship between genetic background and resistance between genotypes supports the notion that plant immune phenotypes are complex and multifaceted, and not dependent upon a sole gene-for-gene interaction. Using differential expression and coexpression network analyses to study the steady-state transcriptional patterns for uninfested plants for each genotype may lead to finding groups of functionally related genes that confer resistance to WAA. Resolving plant immune networks of non-model species can be challenging, so we complement the analyses in apple with WAA transcriptomics, including the differential expression of aphids feeding on different host genotypes during early colonization as well as effector prediction and characterization.

Apple transcriptome undergoes remodeling shortly after colonization

Aphid feeding on apple plants remodeled the host transcriptome with a total of 675 genes differentially expressed between all infested and control plants. Enriched bins included downregulated photosynthesis genes (Wilcoxon rank-sum identified 23 genes, P = 0.0002), and upregulated cell wall organization genes, specifically for suberin and cutin synthesis (Wilcoxon rank-sum identified 8 genes, P = 0.03). The three bins with the most DE genes were 'RNA biosynthesis' (64), 'protein biosynthesis' (44) and 'solute transport' (34) (Fig. 2). Plant transcriptional remodeling occurs shortly after aphid feeding in other plants, with comparable numbers of DE genes after 48 hours: 637 DE genes in tomato after potato aphid feeding (Coppola et al. 2013), and ~650 DE genes in maize after corn leaf aphid feeding (Tzin et al. 2015). Contrasting infested and control plants of G.16 resulted in 790 DE genes and a Wilcoxon rank-sum test identified enriched bins for cell wall organization, carbohydrate metabolism, auxin biosynthesis, and RNA biosynthesis. Only 15 and 5 DE genes were identified between infested and control plants of G.87 and G.935, respectively. As mentioned earlier, G.87 and G.935 both share similar genetic backgrounds compared to G.16, which may explain the similarities in expression patterns.

Primary metabolism.Photosynthesis

Photosynthetic downregulation is a common plant response to diverse forms of biotic stress ranging from viruses, bacteria, fungi, and arthropods (Bilgin et al. 2010). Across all infested and control plants, we found 23 DE genes involved in photosynthesis,

all of which were downregulated in infested plants (Fig. 2), including the downregulation of genes for three protein complexes that are required for the photosynthetic light reactions: photosystem II (PSII), cytochrome b_6f complex, and photosystem I (PSI). Researchers have shown that plant perception of PAMPs causes a reduced accumulation of photosynthetic proteins (Göhre et al. 2012), and induces defense gene expression through chloroplast-mediated retrograde Ca^{2+} signaling during PTI (Nomura et al. 2012). Photosynthetic inhibition is also triggered by effector recognition and has been shown to be essential for ETI (Su et al. 2018). Of the 675 DE genes, we identified 175 predicted to target the chloroplast, of which ~78% are downregulated. Global transcriptome studies have shown plant recognition of pathogens leads to rapid suppression of nuclear-encoded chloroplast targeted genes, especially downregulation of photosynthesis-related genes, which is followed by photosynthesis-derived ROS production and SA synthesis (de Torres Zabala et al. 2015; Lewis et al. 2015). We also observed the downregulation of a violaxanthin de-epoxidase gene, which is involved in non-photochemical quenching (NPQ), a process to protect the photosystems from high light stress by dissipating excess excitation energy as heat (Müller et al. 2001). ROS generation during PTI and ETI is linked to the downregulation of both photosynthesis and NPQ (Göhre et al. 2012; Su et al. 2018), possibly due to increased accumulation of excited electrons reducing O2 in place of proteins involved in electron transport and NPQ (Su et al. 2018). The downstream mechanisms following chloroplastic ROS production to promote defense gene expression are unclear, and the importance of chloroplast-mediated immune signaling for defense against insect herbivores is unknown. In our study, the

downregulation of photosynthesis-related genes shortly after aphid infestation possibly plays an important role in apple immune response to WAA and may be triggered by plant perception of WAA-derived HAMPs, or effectors.

Primary Metabolism.Growth

Suberin is a waxy polymer that forms a barrier between the environment and living plant tissue, and functions to prevent desiccation and protect against biotic attack (Graça 2015). We identified the upregulation of eight genes necessary for suberin synthesis and deposition after aphid infestation (Fig. 2). Although suberization of the cell wall has been shown to be a plant response to aphid feeding (Tzin et al. 2015), and may prevent further stylet penetration, it is unclear if suberin deposition is part of the general wound healing response caused by stylet piercing, or is elicited by direct recognition of aphid-derived molecules.

A clear trend of cell wall modifications has been demonstrated in other studies characterizing transcriptional and chemical responses to aphid feeding (Hohenstein et al. 2019; Rasool et al. 2017; Tetreault et al. 2019; Tzin et al. 2015). Plant cell walls contain biopolymers in the form of celluloses, hemicelluloses, pectins, lignins, suberins, and a large variety of cell wall proteins. During cell expansion, a complex of modifying proteins alter cell wall constituents to loosen the cell wall. Once a cell wall is loosened, turgor pressure generated from the central vacuole can then allow the cell to expand (Perrot-Rechenmann 2010). The cell wall is dynamic under various conditions and

transcriptional control of cell wall factors is crucial for cell development in the face of biotic stress (Houston et al. 2016). Our analysis shows that genes involved in cell wall loosening are largely upregulated after aphid feeding including 2 alpha-class expansins, a leucine-rich repeat extensin, 2 pectin methylesterases, and one polygalacturonase. Furthermore, genes involved in cell wall strengthening processes were generally downregulated including 2 lignin lacasses, 1 coniferin beta-glucosidase, 2 cellulose synthase catalytic components, and 1 xylan O-acetyltransferase. We hypothesize these changes to the cell wall are in favor of the insect, to 1.) facilitate stylet penetration through the apoplast towards sieve elements, and 2.) initiate gall formation, as increased cell size is characteristic of WAA galls (Staniland 1924).

The proliferation of xylem tissue is another key feature of WAA gall morphology and is a visible symptom under light microscopy (Brown et al. 1991). Vascularization is thought to be a general requirement for tumor formation in both plants and animals, as a high influx of nutrients is necessary for tumor enlargement (Ullrich & Aloni 2000). In the case of the crown galls induced by the model plant pathogen, *Rhizobium radiobacter*, informally known as *Agrobacterium tumefaciens*, neovascularization mediated by altered plant growth regulators is necessary for tumorigenesis (Ullrich & Aloni 2000). Xylem differentiation is known to be promoted by extracellular proteoglycans called xylogen-like arabinogalactan proteins (*XYLP*), named after the first of its kind to be discovered in the plant, *Zinnia elegans* L. (Motose et al. 2004) We report the upregulation of 5 *XYLP* genes induced by aphid feeding, which suggests WAA may be activating xylem

differentiation to initiate gall formation. The mechanisms underlying these changes are unknown, but auxin and cytokinin signaling is hypothesized to play a role in xylogenesis.

Plant defense response.Immune receptors

Similar to PTI, during herbivore-triggered immunity (HTI), plants undergo major transcriptional changes due to signal transductions leading to the activation of transcription factors that promote defense-related gene expression (Foyer et al. 2014; Jaouannet et al. 2014). WAA feeding induced the upregulation of genes associated with biotic stress signaling and response, including nine receptor-like kinases/proteins (RLKs/RLPs) involved in PAMP recognition and signal transduction, two nucleotide-binding leucine-rich repeat (NLR) receptors, five genes related to calcium signaling, four peroxidases, one NADPH-oxidase, and 38 transcription factors. The upregulation of many of these genes resembles a PTI or PTI-like response. For example, we found two PRRs that are predicted homologs of PATTERN□TRIGGERED IMMUNITY (PTI) COMPROMISED RECEPTOR□LIKE CYTOPLASMIC KINASE 1 (PCRK1), which is known to initiate PTI in Arabidopsis when elicited by bacterial PAMPs (Sreekanta et al. 2015). We also detected the upregulation of three predicted PRRs in the form of RLKs, an LRR-XII RLK, and two DUF26-containing RLKs that have been implicated in biotic stress responses. In Arabidopsis, both LRR-XII and DUF26-containing RLK transcript abundances were increased after exposure to the bacterial PAMP, flg22, the bacterial effector, HrpZ, and the plant pathogenic mold, *Phytophthora infestans*, while DUF26-containing RLK transcript abundance was also increased following exposure to the plant

bacterial pathogen, *Pseudomonas syringae (Lehti-Shiu et al. 2009)*. A meta-analysis that characterized the transcriptional responses of Arabidopsis challenged with various phloem-feeding insects with data from six published studies showed the aphid-induced transcripts, LRR-XII and DUF26-containing RLKs were significantly over-represented (Foyer et al. 2014). The authors propose that aphid-derived molecules are not directly recognized by the receptors, but rather aphids are indirectly detected through elicitation of PTI by bacteria present as symbionts or in the gut. This hypothesis requires further testing and could significantly alter how we understand and study aphid-recognition and downstream signaling. In support of this hypothesis, one particular bacterial protein, the GroEL chaperonin, from the aphid obligate endosymbiont, *Buchnera aphidicola*, was identified in the saliva of potato aphids, *Macrosiphum euphorbiae*, and *in planta* overexpression of GroEL as well as external application of the purified protein results in reduced aphid fecundity and increased expression of several PTI-associated genes (Chaudhary et al. 2014). We also found upregulation of two NLR receptors, which are also known as resistance genes or *R* genes, because they are effector-activated and initiate ETI, which is characterized as a robust immunity (Cui et al. 2015). Several studies have characterized two resistance genes, *Mi-1.2* in tomato, and the *Vat* gene in melon that both code for NLRs and confer resistance to the potato aphid and cotton aphid, respectively (Dogimont et al. 2014; Rossi et al. 1998). Because even a single NLR gene can confer resistance an aphid species, further characterizing these genes may have significant implications for apple rootstock breeders searching for effective sources of resistance.

Plant defense response.Signal transduction

Mitogen-activated protein kinase (MAPK) signaling cascades are involved in many aspects of plant function, and in plant-pathogen interactions, they are also known to operate during PTI and ETI signaling (Rasmussen et al. 2012). Pathogen recognition by plant immune receptors activates MAPK cascades, which leads to the activation of transcription factors to promote the expression of defense-related genes (Rasmussen et al. 2012). We identified one aphid-induced MAPK and one mitogen-activated kinase kinase also known as MAP3K or MEKK. The signaling events following aphid recognition are poorly characterized, but it is hypothesized that MAPK cascades essential for PTI are also important for defense signaling during HTI. For example, the silencing of one or a few MAPKs in tomato resulted in reduced *R* gene-mediated (*Mi-1)* aphid resistance in tomato (Li et al. 2006). Furthermore, MAPKs play an important role in hormone-mediated defense gene regulation against chewing insect-herbivores (Hettenhausen et al. 2015). Taken together, it is likely that MAPK cascades function similarly during HTI as they do in PTI. However, if PRRs are detecting bacterial associates of aphids rather than the aphids directly, a congruent hypothesis postulates that the following MAPK cascades are a part of PTI and is triggered upon recognition of bacteria-derived molecules.

G.16 decreased auxin biosynthesis in response to WAA feeding

Plant auxin response

After WAA feeding, observed changes in gene expression were more substantial in G.16 compared to G.87 and G.935. Contrasting samples of infested and control G.16 plants resulted in 790 DE genes. Wilcoxon rank-sum tests identified concentrations of DE genes involved in cell wall organization and auxin biosynthesis. Similar to the findings across all aphid vs control plants, cell wall loosening proteins were upregulated in aphid infested samples, including 2 alpha-class expansins and 1 pectin methylesterase, which might suggest host cells are expanding in aphid colonized tissue.

In G.16, we observed the downregulation of 5 genes that code for indole-3-acetamide hydrolases, which convert indole-3 acetamide (IAM) into indole-3 acetic acid (IAA), the most common form of auxin (Mano & Nemoto 2012). Auxin modulates defense responses to pathogens, and host repression of auxin signaling increases resistance against pathogens (Kunkel & Harper 2018; Navarro et al. 2006). Pathogens have evolved to enhance auxin signaling to antagonize host immune pathways as well as stimulate the growth of infested tissue (Kunkel & Harper 2018). For example, several gall-forming bacteria, including *A. tumefaciens*, increase auxin synthesis to promote cell expansion and division (Duca et al. 2014). Modification of auxin signaling also appears to be important in insect gall formation, as feeding by galling insects has been shown to elevate auxin levels in host plants, however, the mechanisms are unknown (Byers, John Allen Brewer, J. Wayne Denna, Donald W. 1976; Mapes & Davies 2001; Tooker & De

Moraes 2011). Manipulation of auxin homeostasis by pathogens has been well studied but remains elusive in plant-aphid interactions. A study with *Aphis gossypii* (cotton-melon aphid) and *Cucumis melo* (melon), has demonstrated repression of auxin signaling via miRNAs is associated with the *Vat* resistance gene. Furthermore, chemically inhibiting the auxin receptor, transport inhibitor response□1 (TIR□1), in susceptible plants reduced aphid fecundity, a characteristic of *Vat* resistance (Sattar et al. 2016). Decreased IAA synthesis in G.16 may be part of the immune response to WAA, but downstream changes to the defensive gene regulatory network will need to be further studied.

Apple genotypes vary in constitutive expression of defense-related genes

Studying the differences in constitutively expressed genes between the three genotypes is important for understanding how each genotype responds to biotic stress. Contrasts between G.87 vs G.935, G.16 vs G.87, and G.16 vs G.935 resulted in 2118, 5102, and 4290 DE genes, respectively. Notably, the constitutive expression profile for G.16 is highly dissimilar to G.935 and G.87, which is reflective of the differences in their genetic backgrounds. Of the G.87 vs G.935 DE genes, there were zero enriched bins. Of the DE genes between G.87 vs G.16 and G.16 vs G.935, enriched bins included cell wall organization, glycosyltransferases, oxidoreductases, phosphorylation, protein modification, and enzyme classification. Cell wall organization genes were generally upregulated in G.87, and glycosyltransferases were generally upregulated in G.16, while

the other enriched bins had mixes of up and downregulated genes (Figure 4). As mentioned earlier, WAA feeding induced greater transcriptional changes in G.16 compared to the other two genotypes, which corresponds to the greater differences in constitutively expressed genes in G.16 compared to G.935 and G.87.

Aphid salivary glands encode plant manipulating effector genes

Effector characterization is critical for understanding insect-induced plant responses, and omics-scale approaches have revealed large arsenals of candidate effectors, many of which contain uncharacterized protein domains, and high rates of positive selection and gene duplications (Bos et al. 2010; Boulain et al. 2018; Thorpe et al. 2016; Zhao et al. 2015). After differential expression analysis between WAA SG and WB samples we identified 5,377 DE transcripts at the 'isoform' level. Of the SG-upregulated transcripts, we obtained 2,311 protein sequences with both start and stop codons predicted from the Transdecoder-produced peptide fasta file. Differential expression analysis of WAA across different host genotypes resulted in 26, 17 and 4 DE genes for the contrasts G.935 vs G.16, G.87 vs G.16, and G.935 vs G.87, respectively. Our effector prediction pipeline identified 390 candidate effectors, of which 281 are only conserved in insect-herbivore lineages. Of the total 390 candidate effectors, 140 show sequence similarity to proteins in the NCBI nr database at an e-value of 0.1 or lower, of which 45 are uncharacterized proteins. Lastly, the BLASTP analysis showed 22 candidate effectors are aphid-specific, 20 of which are uncharacterized (Supplementary

Table 1). The high proportion of uncharacterized candidate effectors with herbivore-specific BLAST hits is a sign of their novelty and suggests they likely evolved functions important for the interaction with host plants. Based on the effector prediction in this study and in other aphid species, it is clear that aphids deploy a large arsenal of diverse salivary proteins into host tissues (Bos et al. 2010; Thorpe et al. 2016), and may work in a combined effort to overcome plant defenses and establish long term colonization, similar to plant pathogen effector arsenals (Toruño et al. 2016).

Enzymes have been previously identified in aphid saliva, and are thought to be involved in detoxification of plant allelochemicals, and initiate digestion (Carolan et al. 2009; Miles 1999). By leveraging the NCBI nr database and BLASTP we assigned tentative functions to 95 candidate effectors including various enzymes such as glycoside hydrolases (GHs) (8), peptidases (6), peroxidases (5), lipases (4), and several other enzymes (Supplementary Table 1). In plant-pathogens, multiple classes of enzymes have been previously identified as factors promoting pathogenesis (Kubicek et al. 2014; Lebeda et al. 2001). For example, GHs constitute a diverse family of cellulases, and in plants, most GHs are involved in the metabolism of cell wall polysaccharides (Minic 2008), and in plant-pathogenic bacteria and fungi, GHs are secreted to degrade polysaccharide chains in plant cell walls (Hématy et al. 2009; Tomme et al. 1995). GHs have also been found in several herbivorous insect lineages, including aphids (Calderón-Cortés et al. 2012; Eyun et al. 2014; Wybouw et al. 2016). Proteomic examinations of aphid saliva have identified GHs, which are suggested to function in the degradation of

plant cell walls to support intercellular stylet penetration (Harmel et al. 2008; Miles 1999; Rao et al. 2013). The secretion of GHs in WAA salivary glands and the transcriptional upregulation of cell wall loosening molecules in apple upon WAA feeding, suggests a possible multi-pronged offensive to weaken the plant cell wall and manipulate the apple phenotype.

Another group of proteins present in the WAA candidate effector repertoire includes peroxidases, which catalyze electron transfer from hydrogen peroxide (H2O2) to an electron acceptor and are involved in diverse biological processes. In plant-pathogenic fungi, peroxidases scavenge ROS produced by the plant as part of the antioxidant defense system essential for successful colonization (Huang et al. 2011; Mir et al. 2015; Missall et al. 2004). Similar to pathogen infestation, aphid feeding also causes ROS burst as part of the plant defense response (De Ilarduya et al. 2003; Lei & Zhu-Salzman 2015; Rasool et al. 2017), and peroxidases have been previously reported in aphid saliva (Chaudhary et al. 2015; Cherqui & Tjallingii 2000; Miles & Peng 1989). As mentioned earlier, WAA infestation resulted in apples' downregulation of photosynthesis and NPQ related genes, which is linked to H2O2 and superoxide generation, two major ROS (Serrano et al. 2016; Su et al. 2018), and upregulation of an NADPH-oxidase, which catalyzes the production of superoxide (Panday et al. 2015). Analogous to fungal antioxidant defenses, the salivary peroxidases we identified in WAA may function to counter ROS burst by scavenging H2O2.

For insect herbivore research, effector investigation has been given a high-priority in recent years (Basu et al. 2018; Kaloshian et al. 2014; Stuart 2015). Studies using molecular genetics and functional genomics methods have led the way for effector identification and determining their impact on aphid fitness (Kaloshian et al. 2014). Using BLASTP, we identified 3 WAA candidate effectors with sequence similarity to three previously characterized aphid effectors: *C002* (e-value = 1.97E-50, logFC in SG=8.3), *SHP* (e-value = 1.79E-53, logFC in SG=6.5), and *Armet* (e-value = 4.35E-95, logFC in SG=1.7) (Table 3). RNAi knockdown and *in planta* overexpression of these effectors have proven useful techniques to determine their importance for aphid fitness and have advanced our hypotheses of their functions. The first aphid effector identified was in the pea aphid, *Acyrthosiphon pisum,* and named C002. RNAi-based knockdown of pea aphid *C002* transcripts led to reduced phloem feeding and higher rates of mortality on fava bean (Mutti et al. 2006; Mutti et al. 2008). Additionally, *in planta* overexpression of the cloned *MPC002* from green peach aphid, *Myzus persicae*, in tobacco resulted in increased fecundity of *M. persicae (Bos et al. 2010)*. The aphid effector, Armet, is a highly conserved gene, apparently distributed across metazoa (Wang et al. 2015). In pea aphids, RNAi knockdown of Armet reduced phloem feeding and lifespan (Wang et al. 2015). Lastly, the effector, SHP (structural sheath protein), is thought to be secreted along the stylet path towards the phloem and polymerizes the gelling saliva into a protective sheath around the stylet (Will et al. 2012). In pea aphids and grain aphids, *Sitobion avenae*, RNAi knockdown of SHP disrupted sheath formation, affected aphid feeding ability and reduced aphid fecundity (Abdellatef et al. 2015; Will & Vilcinskas

2015). The exact functions of these effectors remain unknown, but because *C002*, *SHP* and *Armet* protein sequences are conserved across distantly related subfamilies within Aphididae, they likely function in similar biological processes. However, proper characterization of WAA candidate effectors will be necessary to validate their functions and importance in WAA-apple interaction. Coupling effector silencing techniques with transcriptomic profiling of plants may lead to more refined hypotheses of how aphid salivary proteins alter plant immune responses at the transcriptional level.

CONCLUSION

The molecular dialogue between plants and aphids is analogous in many ways to plant-pathogen interactions. However, our current understanding of the plant immune network and aphid effector activity remains fragmented. In recent years, transcriptomic analyses of aphid infested plants have identified hundreds of differentially expressed genes under varying conditions of susceptible and resistant host genotypes, including genes involved in signal transduction, primary metabolism, growth, and defense. These studies also reinforce the point that plant immune networks are complex, and generally, resistance is not attributable to the expression of a few genes, but rather a myriad of genes that are expressed at different stages of infestation as part of the immune response to aphids. Studies on aphid saliva have shown that aphids deliver hundreds of effectors into their hosts, suspected to alter plant processes to suppress immunity and enhance nutrition. Here, by profiling the transcriptomes of resistant and susceptible apple

rootstock genotypes during early woolly apple aphid infestation and characterizing the genes present in salivary glands, we find further support for the pathogen paradigm in aphid-plant interactions. Our data show substantial chloroplastic changes at the transcriptional level including photosynthetic and NPQ downregulation, which are linked to ROS-mediated defense signaling. We also find major upregulation of signal transduction components that resemble a PTI/PTI-like response, such as pattern recognition receptors, MAPKs, ROS generators, *R* genes, and many transcription factors. In addition to defense-related gene expression, we find woolly apple aphid feeding induces transcriptional changes possibly linked to gall induction, such as xylogen-like arabinogalactan proteins, which induce xylem differentiation, and upregulation of proteins involved in cell wall loosening, which suggests cellular expansion. These findings may lead to improved hypotheses of the molecular mechanisms underlying aphid-plant interactions, woolly apple aphid gall induction, and aid in the development of viable woolly apple aphid resistant apple rootstocks.

LITERATURE CITED

Abdellatef, E. et al., 2015. Silencing the expression of the salivary sheath protein causes transgenerational feeding suppression in the aphid *Sitobion avenae*. *Plant Biotechnology Journal*, 13(6), pp.849–857.

Anders, S., Pyl, P.T. & Huber, W., 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2), pp.166–169.

Andrews, S., 2010. *FastQC: a quality control tool for high throughput sequence data*, Available at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Ateyyat, M.A. & Al-Antary, T.M., 2009. Susceptibility of nine apple cultivars to woolly apple aphid, *Eriosoma lanigerum* (Homoptera: Aphididae) in Jordan. *International Journal of Pest Management*, 55(1), pp.79–84.

Baker, A.C., 1915. The woolly apple aphid. *United States Department of Agriculture, Washington, DC.*, Rept. No. 101, 1-55.

Basu, S., Varsani, S. & Louis, J., 2018. Altering plant defenses: herbivore-associated molecular patterns and effector arsenal of chewing herbivores. *Molecular Plant-Microbe Interactions: MPMI*, 31(1), pp.13–21.

Bilgin, D.D. et al., 2010. Biotic stress globally downregulates photosynthesis genes. *Plant, Cell & Environment*, 33(10), pp.1597–1613.

Blackman, R.L. & Eastop, V.F., 2000. *Aphids on the world's crops: an identification and information guide*, John Wiley & Sons Ltd.

Blackman, R.L. & Eastop, V.F., 1994. *Aphids on the world's trees: an identification and information guide*, Cab International.

Blommers, L.H.M., 1994. Integrated pest management in european apple orchards. *Annual Review of Entomology*, 39(1), pp.213–241.

Bos, J.I.B. et al., 2010. A functional genomics approach identifies candidate effectors from the aphid species *Myzus persicae* (green peach aphid). *PLoS Genetics*. Available at: http://dx.doi.org/10.1371/journal.pgen.1001216.

Boulain, H. et al., 2018. Fast evolution and lineage-specific gene family expansions of aphid salivary effectors driven by interactions with host-plants. *Genome Biology and Evolution*, 10(6), pp.1554–1572.

Brown, M.W., Glenn, D.M. & Wisniewski, M.E., 1991. Functional and anatomical disruption of apple roots by the woolly apple aphid (Homoptera: Aphididae). *Journal of Economic Entomology*, 84(6), pp.1823–1826.

Bus, V.G.M. et al., 2010. Genome mapping of an apple scab, a powdery mildew and a woolly apple aphid resistance gene from open-pollinated mildew immune selection. *Tree Genetics & Genomes*, 6(3), pp.477–487.

Bus, V.G.M. et al., 2008. Genome mapping of three major resistance genes to woolly apple aphid (*Eriosoma lanigerum* Hausm.). *Tree Genetics & Genomes*, 4(2), pp.223–236.

Byers, John Allen Brewer, J. Wayne Denna, Donald W., 1976. Plant growth hormones in pinyon insect galls. *Marcellia*, 39, pp.125–134.

Calderón-Cortés, N. et al., 2012. Endogenous plant cell wall digestion: A key mechanism in insect evolution. *Annual Review of Ecology, Evolution, and Systematics*, 43(1), pp.45–71.

Camacho, C. et al., 2009. BLAST+: architecture and applications. *BMC Bioinformatics*, 10(421), p.1.

Carolan, J.C. et al., 2011. Predicted effector molecules in the salivary secretome of the pea aphid (*Acyrthosiphon pisum*): a dual transcriptomic/proteomic approach. *Journal of Proteome Research*, 10(4), pp.1505–1518.

Carolan, J.C. et al., 2009. The secreted salivary proteome of the pea aphid *Acyrthosiphon pisum* characterised by mass spectrometry. *Proteomics*, 9(9), pp.2457–2467.

Chaudhary, R. et al., 2014. GroEL from the endosymbiont *Buchnera aphidicola* betrays the aphid by triggering plant defense. *Proceedings of the National Academy of Sciences of the United States of America*, 111(24), pp.8919–8924.

Chaudhary, R. et al., 2015. Potato aphid salivary proteome: enhanced salivation using resorcinol and identification of aphid phosphoproteins. *Journal of Proteome Research*, 14(4), pp.1762–1778.

Chen, M.S. et al., 2008. Analysis of transcripts and proteins expressed in the salivary glands of Hessian fly (*Mayetiola destructor*) larvae. *Journal of Insect Physiology*, 54(1), pp.1–16.

Cherqui, A. & Tjallingii, W.F., 2000. Salivary proteins of aphids, a pilot study on identification, separation and immunolocalisation. *Journal of Insect Physiology*, 46(8), pp.1177–1186.

Cohen, H. et al., 1996. Susceptibility of the woolly apple aphid parasitoid, *Aphelinus mali* (Hym.: Aphelinidae), to common pesticides used in apple orchards in Israel. *Entomophaga*, 41(2), pp.225–233.

Coppola, V. et al., 2013. Transcriptomic and proteomic analysis of a compatible tomato-aphid interaction reveals a predominant salicylic acid-dependent plant response. *BMC Genomics*, 14(1), p.1.

Crane, M.B. et al., 1937. Studies on the resistance and immunity of apples to the woolly aphis, *Eriosoma lanigerum* (Hausm.). *Journal of Pomology and Horticultural Science*, 14(2), pp.137–163.

Cui, H., Tsuda, K. & Parker, J.E., 2015. Effector-triggered immunity: from pathogen perception to robust defense. *Annual Review of Plant Biology*, 66, pp.487–511.

Cummins, J.N., Forsline, P.L. & Mackenzie, J.D., 1981. Woolly apple aphid (*Eriosoma lanigerum*) colonization on *Malus* cultivars. *Journal-American Society for Horticultural Science*, 106(1), pp.26–30.

De Ilarduya, O.M., Xie, Q. & Kaloshian, I., 2003. Aphid-induced defense responses in *Mi-1*-mediated compatible and incompatible tomato interactions. *Molecular Plant-Microbe Interactions: MPMI*, 16(8), pp.699–708.

Dobin, A. et al., 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics,* 29(1), pp.15–21.

Dodds, P.N. & Rathjen, J.P., 2010. Plant immunity: towards an integrated view of plant–pathogen interactions. *Nature Reviews Genetics*, 11(8), pp.539–548.

Dogimont, C. et al., 2014. The Vat locus encodes for a CC-NBS-LRR protein that confers resistance to *Aphis gossypii* infestation and *A. gossypii*-mediated virus resistance. *The Plant Journal: For Cell and Molecular Biology*, 80(6), pp.993–1004.

Domes, K., Scheu, S. & Maraun, M., 2007. Resources and sex: Soil re-colonization by sexual and parthenogenetic oribatid mites. *Pedobiologia*, 51(1), pp.1–11.

Duca, D. et al., 2014. Indole-3-acetic acid in plant–microbe interactions. *Antonie van Leeuwenhoek*, 106(1), pp.85–125.

Elzinga, D.A., De Vos, M. & Jander, G., 2014. Suppression of plant defenses by a *Myzus persicae* (green peach aphid) salivary effector protein. *Molecular Plant-Microbe Interactions: MPMI*, 27(7), pp.747–756.

Eyun, S.-I. et al., 2014. Molecular evolution of glycoside hydrolase genes in the Western corn rootworm (*Diabrotica virgifera virgifera*). *PloS One*, 9(4), p. e94052.

Fazio, G. & Beers, B., 2010. Wooly Apple Aphid resistance in advanced rootstock selections, *Washington Tree Fruit Research Commission*.

Fazio, G., Robinson, T.L. & Aldwinckle, H.S., 2015. The Geneva apple rootstock breeding program. *Plant Breeding Reviews*, 39, pp.379–424.

Fernandes, G.W. & Negreiros, D., 2001. The occurrence and effectiveness of hypersensitive reaction against galling herbivores across host taxa. *Ecological Entomology*, 26(1), pp.46–55.

Foyer, C.H., Verrall, S.R. & Hancock, R.D., 2014. Systematic analysis of phloem-feeding insect-induced transcriptional reprogramming in *Arabidopsis* highlights common features and reveals distinct responses to specialist and generalist insects. *Journal of Experimental Botany*, 66(2), pp.495–512.

Gallego Romero, I. et al., 2014. RNA-seq: impact of RNA degradation on transcript quantification. *BMC Biology*, 12, p.42.

Gilabert, A. et al., 2009. Climate and agricultural context shape reproductive mode variation in an aphid crop pest. *Molecular Ecology*, 18(14), pp.3050–3061.

Giliomee, J.H., Strydom, D.K. & Van Zyl, H.J., 1968. Northern Spy, Merton and Malling-Merton rootstocks susceptible to woolly aphid, *Eriosoma lanigerum*, in the Western Cape. *South African Journal of Agricultural Science*, 11, pp.183–186.

Gill, H.K. & Garg, H., 2014. Pesticides: environmental impacts and management strategies. In *Pesticides-Toxic Aspects*. Available at: DOI: 10.5772/57399

Giron, D. et al., 2016. Insect-induced effects on plants and possible effectors used by galling and leaf-mining insects to manipulate their host-plant. *Journal of Insect Physiology*, 84, pp.70–89.

Göhre, V. et al., 2012. Molecular crosstalk between PAMP-triggered immunity and photosynthesis. *Molecular Plant-Microbe Interactions: MPMI*, 25(8), pp.1083–1092.

Gontijo, L.M., Cockfield, S.D. & Beers, E.H., 2012. Natural enemies of woolly apple aphid (Hemiptera: Aphididae) in Washington state. *Environmental Entomology*, 41(6), pp.1364–1371.

Grabherr, M.G. et al., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29(7), pp.644–652.

Graça, J., 2015. Suberin: the biopolyester at the frontier of plants. *Frontiers in Chemistry*, 3, p.62.

Harmel, N. et al., 2008. Identification of aphid salivary proteins: a proteomic investigation of *Myzus persicae*. *Insect Molecular Biology*, 17(2), pp.165–174.

Heil, M. & Land, W.G., 2014. Danger signals--damaged-self recognition across the tree of life. *Frontiers in Plant Science*, 5, p.578.

Hématy, K., Cherk, C. & Somerville, S., 2009. Host–pathogen warfare at the plant cell wall. *Current Opinion in Plant Biology*, 12(4), pp.406–413.

Hettenhausen, C., Schuman, M.C. & Wu, J., 2015. MAPK signaling: a key element in plant defense response to insects. *Insect Science*, 22(2), pp.157–164.

Hoffmann, A.A. et al., 2008. A high incidence of parthenogenesis in agricultural pests. *Proceedings of the Royal Society B: Biological Sciences*, 275(1650), pp.2473–2481.

Hogenhout, S.A. & Bos, J.I.B., 2011. Effector proteins that modulate plant--insect interactions. *Current Opinion in Plant Biology*, 14(4), pp.422–428.

Hohenstein, J.D. et al., 2019. Transcriptional and chemical changes in soybean leaves in response to long-term aphid colonization. *Frontiers in Plant Science*, 10, p.310.

Houston, K. et al., 2016. The plant cell wall: a complex and dynamic structure as revealed by the responses of genes under stress conditions. *Frontiers in Plant Science*, 7, p.984.

Huang, K. et al., 2011. HYR1-mediated detoxification of reactive oxygen species is required for full virulence in the rice blast fungus. *PLoS Pathogens*, 7(4), p. e1001335.

Jaouannet, M. et al., 2014. Plant immunity in plant--aphid interactions. *Frontiers in Plant Science*, 5, p.663.

Jensen, P.J. et al., 2014. Mapping in an apple (*Malus* x *domestica*) F1segregating population based on physical clustering of differentially expressed genes. *BMC Genomics*, 15, p.261.

Jones, J.D.G. & Dangl, J.L., 2006. The plant immune system. *Nature*, 444(7117), pp.323–329.

Kaloshian, I. & Walling, L.L., 2014. Hemipteran and dipteran pests: effectors and plant host immune regulators. *Journal of Integrative Plant Biology*, 58(4), pp.350–361.

Kassambara, A. et al., 2017. survminer: Drawing Survival Curves using’ggplot2'. *R package version 0. 3*, 1.

King, G.J. et al., 1991. The “European Apple Genome Mapping Project”-developing a strategy for mapping genes coding for agronomic characters in tree species. *Euphytica/ Netherlands Journal of Plant Breeding*, 56(1), pp.89–94.

Klepikova, A.V. et al., 2017. Effect of method of deduplication on estimation of differential gene expression using RNA-seq. *PeerJ*, 5, p.e3091.

Knight, R.L. et al., 1962. The inheritance of resistance to woolly aphid, *Eriosoma lanigerum* (HSMNN.), in the apple. *The Journal of Horticultural Science*, 37(3), pp.207–218.

Kubicek, C.P., Starr, T.L. & Glass, N.L., 2014. Plant cell wall–degrading enzymes and their secretion in plant-pathogenic fungi. *Annual Review of Phytopathology*, 52(1), pp.427–451.

Kunkel, B.N. & Harper, C.P., 2018. The roles of auxin during interactions between bacterial plant pathogens and their hosts. *Journal of Experimental Botany*, 69(2), pp.245–254.

Kuśnierczyk, A. et al., 2008. Towards global understanding of plant defence against aphids--timing and dynamics of early *Arabidopsis* defence responses to cabbage aphid (*Brevicoryne brassicae*) attack. *Plant, Cell & Environment*, 31(8), pp.1097–1115.

Lavandero, B. et al., 2011. Estimating gene flow between refuges and crops: a case study of the biological control of *Eriosoma lanigerum* by *Aphelinus mali* in apple orchards. *Plos One*. Available at: http://dx.doi.org/10.1371/journal.pone.0026694.

Lavandero, B. et al., 2009. Isolation and characterization of microsatellite loci from the woolly apple aphid *Eriosoma lanigerum* (Hemiptera: Aphididae: Eriosomatinae). *Molecular Ecology Resources*, 9(1), pp.302–304.

Lavandero, B. et al., 2009. Landscape composition modulates population genetic structure of *Eriosoma lanigerum* (Hausmann) on *Malus domestica* Borkh in central Chile. *Bulletin of Entomological Research*, 99(01), p.97.

Law, C.W. et al., 2014. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology*, 15(2), p.29.

Lebeda, A. et al., 2001. The role of enzymes in plant-fungal pathogens interactions / Die Rolle der Enzyme in den Beziehungen zwischen Pflanzen und pilzlichen Erregern. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz / Journal of Plant Diseases and Protection*, 108(1), pp.89–111.

Lehti-Shiu, M.D. et al., 2009. Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant Physiology*, 150(1), pp.12–26.

Lei, J. & Zhu-Salzman, K., 2015. Enhanced aphid detoxification when confronted by a host with elevated ROS production. *Plant Signaling & Behavior*, 10(4), p.e1010936.

Lewis, L.A. et al., 2015. Transcriptional dynamics driving MAMP-triggered immunity and pathogen effector-mediated immunosuppression in *Arabidopsis* leaves following infection with *Pseudomonas syringae* pv tomato DC3000. *The Plant Cell*, 27(11), pp.3038–3064.

Li, Q. et al., 2006. *Mi-1*-Mediated aphid resistance involves salicylic acid and mitogen-activated protein kinase signaling cascades. *Molecular Plant-Microbe Interactions: MPMI*, 19(6), pp.655–664.

Li, W. & Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22(13), pp.1658–1659.

Mano, Y. & Nemoto, K., 2012. The pathway of auxin biosynthesis in plants. *Journal of Experimental Botany*, 63(8), pp.2853–2872.

Mapes, C.C. & Davies, P.J., 2001. Indole-3-acetic acid and ball gall development on *Solidago altissima*. *The New Phytologist*, 151(1), pp.195–202.

Miles, P.W., 1999. Aphid saliva. *Biological Reviews of the Cambridge Philosophical Society*, 74(1), pp.41–85.

Miles, P.W. & Peng, Z., 1989. Studies on the salivary physiology of plant bugs: detoxification of phytochemicals by the salivary peroxidase of aphids. *Journal of Insect Physiology*, 35(11), pp.865–872.

Miller, A.D. et al., 2012. Population genetics of the wheat curl mite (*Aceria tosichella* Keifer) in Australia: implications for the management of wheat pathogens. *Bulletin of Entomological Research*, 102(2), pp.199–212.

Minic, Z., 2008. Physiological roles of plant glycoside hydrolases. *Planta*, 227(4), pp.723–740.

Mir, A.A. et al., 2015. Systematic characterization of the peroxidase gene family provides new insights into fungal pathogenicity in *Magnaporthe oryzae*. *Scientific Reports*, 5, p.11831.

Missall, T.A., Pusateri, M.E. & Lodge, J.K., 2004. Thiol peroxidase is critical for virulence and resistance to nitric oxide and peroxide in the fungal pathogen, *Cryptococcus neoformans*. *Molecular Microbiology*, 51(5), pp.1447–1458.

Motose, H., Sugiyama, M. & Fukuda, H., 2004. A proteoglycan mediates inductive interaction during plant vascular development. *Nature*, 429(6994), pp.873–878.

Müller, P., Li, X.P. & Niyogi, K.K., 2001. Non-photochemical quenching. A response to excess light energy. *Plant Physiology*, 125(4), pp.1558–1566.

Mutti, N.S. et al., 2008. A protein from the salivary glands of the pea aphid, *Acyrthosiphon pisum*, is essential in feeding on a host plant. *Proceedings of the National Academy of Sciences of the United States of America*, 105(29), pp.9965–9969.

Mutti, N.S. et al., 2006. RNAi knockdown of a salivary transcript leading to lethality in the pea aphid, *Acyrthosiphon pisum*. *Journal of Insect Science*, 6, pp.1–7.

Nabity, P.D., 2016. Insect-induced plant phenotypes: revealing mechanisms through comparative genomics of galling insects and their hosts. *American Journal of Botany*, 103(6), pp.979–981.

Navarro, L. et al., 2006. A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science*, 312(5772), pp.436–439.

Nicholas, A.H., Spooner-Hart, R.N. & Vickers, R.A., 2005. Abundance and natural control of the woolly aphid *Eriosoma lanigerum* in an Australian apple orchard IPM program. *Biocontrol*, 50(2), pp.271–291.

Niu, L. et al., 2018. Dynamic transcriptomes of resistant and susceptible peach lines after infestation by green peach aphids (*Myzus persicae* Sülzer) reveal defence responses controlled by the Rm3 locus. *BMC Genomics*, 19(1), p.846.

Nomura, H. et al., 2012. Chloroplast-mediated activation of plant immune signalling in *Arabidopsis*. *Nature Communications*, 3, p.926.

Norelli, J.L. et al., 2003. Resistance of Geneva and other apple rootstocks to *Erwinia amylovora*. *Plant Disease*, 87(1), pp.26–32.

Ogawa, K. & Miura, T., 2014. Aphid polyphenisms: trans-generational developmental regulation through viviparity. *Frontiers in Physiology*, 5, p.1.

Painter, R.H., 1951. *Insect resistance in crop plants*, The Macmillan Company New York.

Panday, A. et al., 2015. NADPH oxidases: an overview from structure to innate immunity-associated pathologies. *Cellular & Molecular Immunology*, 12(1), pp.5–23.

Patch, E.M., 1916. Elm leaf rosette and woolly aphid of the apple. *Technical Bulletin. Maine Agricultural and Forest Experiment Station*, Bulletin 2.

Patro, R. et al., 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, 14(4), pp.417–419.

Penman, D.R. & Chapman, R.B., 1980. Woolly apple aphid outbreak following use of fenvalerate in apples in Canterbury, New Zealand. *Journal of Economic Entomology*, 73(1), pp.49–51.

Perrot-Rechenmann, C., 2010. Cellular responses to auxin: division versus expansion. *Cold Spring Harbor Perspectives in Biology*, 2(5), p.a001446.

Pringle, K.L. & Heunis, J.M., 2001. Woolly apple aphid in South Africa: biology, seasonal cycles, damage and control. *Deciduous Fruit Grower*, 51(4), pp.22–36.

Rao, S.A.K., Carolan, J.C. & Wilkinson, T.L., 2013. Proteomic profiling of cereal aphid saliva reveals both ubiquitous and adaptive secreted proteins. *PloS One*, 8(2), p.e57413.

Rasmussen, M.W. et al., 2012. MAP kinase cascades in *Arabidopsis* innate immunity. *Frontiers in Plant Science*, 3, p.169.

Rasool, B. et al., 2017. Redox control of aphid resistance through altered cell wall composition and nutritional quality. *Plant Physiology*, 175(1), pp.259–271.

Robinson, T. et al., 2002. The Geneva series of apple rootstocks from Cornell: performance, disease resistance, and commercialization. *XXVI International Horticultural Congress: Genetics and Breeding of Tree Fruits and Nuts 622*. pp. 513–520.

Rock, G.C. & Zeiger, D.C., 1974. Woolly apple aphid infests Malling and Malling-Merton rootstocks in propagation beds in North Carolina. *Journal of Economic Entomology*, 67(1), pp.137–138.

Rodriguez, P.A. & Bos, J.I.B., 2013. Toward understanding the role of aphid effectors in plant infestation. *Molecular Plant-Microbe Interactions: MPMI*, 26(1), pp.25–30.

Rogers, D.J., Sharma, N. & Stretton, D.C., 2011. Toxicity of pesticides to *Aphelinus mali*, the parasitoid of woolly apple aphid. *New Zealand Plant Protection*, 64, p.235.

Rossi, M. et al., 1998. The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. *Proceedings of the National Academy of Sciences of the United States of America*, 95(17), pp.9750–9754.

Sandanayaka, W.R.M. et al., 2003. Characteristics associated with woolly apple aphid *Eriosoma lanigerum*, resistance of three apple rootstocks. *Entomologia Experimentalis et Applicata*, 109(1), pp.63–72.

Sandanayaka, W.R.M. & Bus, V.G.M., 2005. Evidence of sexual reproduction of woolly apple aphid, *Eriosoma lanigerum*, in New Zealand. *Journal of Insect Science*, 5(1), p.27.

Sandanayaka, W.R.M., Bus, V.G.M. & Connolly, P., 2005. Mechanisms of woolly aphid [*Eriosoma lanigerum* (Hausm.)] resistance in apple. *Journal of Applied Entomology*, 129(9-10), pp.534–541.

Sandström, J., Telang, A. & Moran, N.A., 2000. Nutritional enhancement of host plants by aphids — a comparison of three aphid species on grasses. *Journal of Insect Physiology*, 46(1), pp.33–40.

Sano, M. & Akimoto, S.I., 2011. Morphological phylogeny of gall-forming aphids of the tribe Eriosomatini (Aphididae: Eriosomatinae). *Systematic Entomology*, 36(4), pp.607–627.

Sattar, S., Addo-Quaye, C. & Thompson, G.A., 2016. miRNA-mediated auxin signalling repression during Vat-mediated aphid resistance in *Cucumis melo*. *Plant, Cell & Environment*, 39(6), pp.1216–1227.

Schwacke, R. et al., 2019. MapMan4: a refined protein classification and annotation framework applicable to multi-omics data analysis. *Molecular Plant*, 12(6), pp.879–892.

Serrano, I., Audran, C. & Rivas, S., 2016. Chloroplasts at work during plant innate immunity. *Journal of Experimental Botany*, 67(13), pp.3845–3854.

Shoene, W.J. & Underhill, G.W., 1935. Life history and migration of the apple woolly aphis. *Technical Bulletin. Virginia Agricultural Experiment Station*, 57.

Simão, F.A. et al., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19), pp.3210–3212.

Sreekanta, S. et al., 2015. The receptor-like cytoplasmic kinase PCRK 1 contributes to pattern-triggered immunity against *Pseudomonas syringae* in *Arabidopsis thaliana*. *The New Phytologist*, 207(1), pp.78–90.

Staniland, L.N., 1924. The immunity of apple stocks from attacks of woolly aphis (*Eriosoma lanigerum*, Hausmann). Part II. The causes of the relative resistance of the stocks. *Bulletin of Entomological Research*, 15(2), pp.157–170.

Stone, G.N. & Schönrogge, K., 2003. The adaptive significance of insect gall morphology. *Trends in Ecology and Evolution*, 18(10), pp.512–522.

Stuart, J., 2015. Insect effectors and gene-for-gene interactions with host plants. *Current Opinion in Insect Science*, 9, pp.56–61.

Su, J. et al., 2018. Active photosynthetic inhibition mediated by MPK3/MPK6 is critical to effector-triggered immunity. *PLoS Biology*, 16(5), p.e2004122.

Tetreault, H.M. et al., 2019. Global responses of resistant and susceptible sorghum (*Sorghum bicolor*) to sugarcane aphid (*Melanaphis sacchari*). *Frontiers in Plant Science*, 10, p.145.

Theobald, F.V., 1922. The woolly aphid of the apple and elm. *Journal of Pomology and Horticultural Science*, 2, pp.73–94.

Therneau, T.M. & Lumley, T., 2014. Package "survival." *Survival analysis Published on CRAN*. Available at: https://tbrieder.org/epidata/course_reading/e_therneau.pdf.

Thimm, O. et al., 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal: For Cell and Molecular Biology*, 37(6), pp.914–939.

Thorpe, P., Cock, P.J.A. & Bos, J., 2016. Comparative transcriptomics and proteomics of three different aphid species identifies core and diverse effector sets. *BMC Genomics*, 17(172), pp.1–18.

Timm, A.E., Pringle, K.L. & Warnich, L., 2005. Genetic diversity of woolly apple aphid *Eriosoma lanigerum* (Hemiptera: Aphididae) populations in the Western Cape, South Africa. *Bulletin of Entomological Research*, 95(3), pp.187–191.

Tjallingii, W.F. & Esch, T.H.H., 1993. Fine structure of aphid stylet routes in plant tissues in correlation with EPG signals. *Physiological Entomology*, 18(3), pp.317–328.

Tomme, P., Warren, R.A. & Gilkes, N.R., 1995. Cellulose hydrolysis by bacteria and fungi. *Advances in Microbial Physiology*, 37, pp.1–81.

Tooker, J.F. & De Moraes, C.M., 2011. Feeding by a gall-inducing caterpillar species alters levels of indole-3-acetic and abscisic acid in *Solidago altissima* (Asteraceae) stems. *Arthropod-Plant Interactions*, 5(2), pp.115–124.

de Torres Zabala, M. et al., 2015. Chloroplasts play a central role in plant defence and are targeted by pathogen effectors. *Nature Plants*, 1, p.15074.

Toruño, T.Y., Stergiopoulos, I. & Coaker, G., 2016. Plant-pathogen effectors: cellular probes interfering with plant defenses in spatial and temporal manners. *Annual Review of Phytopathology*, 54, pp.419–441.

Tzin, V. et al., 2015. Dynamic maize responses to aphid feeding are revealed by a time series of transcriptomic and metabolomic assays. *Plant Physiology*, 169(3), pp.1727–1743.

Ullrich, C.I. & Aloni, R., 2000. Vascularization is a general requirement for growth of plant and animal tumours. *Journal of Experimental Botany*, 51(353), pp.1951–1960.

Usadel, B. et al., 2005. Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. *Plant Physiology*, 138(3), pp.1195–1204.

Valenzuela, I. et al., 2010. Patterns of genetic variation and host adaptation in an invasive population of *Rhopalosiphum padi* (Hemiptera: Aphididae). *Annals of the Entomological Society of America*, 103(6), pp.886–897.

Wang, W. et al., 2015. Armet is an effector protein mediating aphid–plant interactions. *The FASEB Journal*, 29(5), pp.2032–2045.

Wearing, C.H., Attfield, B.A. & Colhoun, K., 2010. Biological control of woolly apple aphid, *Eriosoma lanigerum* (Hausmann), during transition to integrated fruit production for pipfruit in Central Otago, New Zealand. *New Zealand Journal of Crop and Horticultural Science*, 38(4), pp.255–273.

Will, T. et al., 2012. Aphid gel saliva: sheath structure, protein composition and secretory dependence on stylet-tip milieu. *PloS One*, 7(10), p.e46903.

Will, T. et al., 2007. Molecular sabotage of plant defense by aphid saliva. *Proceedings of the National Academy of Sciences of the United States of America*, 104(25), pp.10536–10541.

Will, T. & Vilcinskas, A., 2015. The structural sheath protein of aphids is required for phloem feeding. *Insect Biochemistry and Molecular Biology*, 57, pp.34–40.

Woldemariam, M.G., Baldwin, I.T. & Galis, I., 2011. Transcriptional regulation of plant inducible defenses against herbivores: a mini-review. *Journal of Plant Interactions*, 6(2-3), pp.113–119.

Wybouw, N. et al., 2016. Horizontal gene transfer contributes to the evolution of arthropod herbivory. *Genome Biology and Evolution*, 8(6), pp.1785–1801.

Zhao, C. et al., 2015. A massive expansion of effector genes underlies gall-formation in the wheat pest *Mayetiola destructor*. *Current Biology*, 25, pp.613–620.

Zipfel, C., 2014. Plant pattern-recognition receptors. *Trends in Immunology*, 35(7), pp.345–351.

Züst, T. & Agrawal, A.A., 2016. Mechanisms and evolution of plant resistance to aphids. *Nature Plants*, 2, p.15206.

TABLES AND FIGURES

Apple RNAseq libraries: read statistics and metabolism gene counts CHO = carbohydrate, AA = amino acid, ^ = outlier from MDS analysis * = low RNA concentration and undetected RIN score (NA)						not deduplicated						deduplicated					
						genes present		metabolism genes present				genes present		metabolism genes present			
sample	genotype	GC (%)	sequences (millions)	duplicates (%)	uniquely mapped reads (millions)	total	total (%)	CHO	AA	Lipid	Total	total	total (%)	CHO	AA	Lipid	Total
1CA	16	55	23.1	89.3	6.4	28059	53.2	321	265	505	1091	28058	53.2	320	265	503	1088
4CA	16	54	25.8	88.7	12.9	31012	58.8	333	277	530	1140	31003	58.8	332	277	526	1135
25WA	16	54	24.9	89.8	12.8	31014	58.8	337	276	532	1145	31010	58.8	334	276	531	1141
37CA^	16	54	22.4	89.0	11.4	28643	54.3	327	272	506	1105	28636	54.3	326	272	505	1103
40WA	16	54	25.3	89.4	12.5	30377	57.6	334	274	519	1127	30375	57.6	333	274	515	1122
46CA	16	55	23.1	90.9	11.5	29826	56.6	333	274	515	1122	29816	56.5	332	274	512	1118
48CA	16	55	23.4	90.4	11.9	29624	56.2	325	274	510	1109	29623	56.2	324	274	509	1107
18c*	16	54	21.5	89.3	11.3	28888	54.8	331	273	513	1117	28885	54.8	330	273	512	1115
29c	16	54	22.2	89.8	11.6	29749	56.4	330	275	521	1126	29745	56.4	328	275	518	1121
43c	16	53	21.0	84.3	11.8	30454	57.7	328	273	502	1103	30449	57.7	325	273	499	1097
50c	16	53	21.0	82.2	12.4	30820	58.4	334	274	521	1129	30815	58.4	332	274	519	1125
51c*^	16	54	23.8	87.6	13.2	28655	54.3	321	265	489	1075	28640	54.3	320	265	485	1070
2CA	87	54	24.3	89.7	12.5	30284	57.4	334	271	527	1132	30282	57.4	333	271	526	1130
6CA	87	54	22.0	87.4	11.8	31188	59.1	336	277	538	1151	31187	59.1	336	277	536	1149
15CA	87	54	21.9	87.0	11.6	31444	59.6	338	278	539	1155	31442	59.6	337	278	538	1153
16WA	87	53	22.0	81.4	12.8	32729	62.1	343	278	544	1165	32725	62.0	341	278	543	1162
20WA	87	52	23.9	79.3	14.4	32585	61.8	340	277	545	1162	32581	61.8	339	277	542	1158
27CA	87	53	21.2	81.1	12.4	33276	63.1	340	278	556	1174	33272	63.1	340	278	552	1170
33WA	87	53	24.4	83.6	13.0	32403	61.4	338	277	542	1157	32403	61.4	338	277	540	1155
34WA	87	53	22.1	86.2	12.1	31084	58.9	337	275	528	1140	31083	58.9	336	275	525	1136
36WA	87	52	22.4	74.8	14.2	33496	63.5	342	280	558	1180	33493	63.5	340	280	555	1175
39CA*	87	54	21.5	87.6	11.2	29940	56.8	332	274	520	1126	29938	56.8	332	274	520	1126
42WA	87	53	25.4	85.8	13.0	32221	61.1	339	276	544	1159	32217	61.1	338	276	541	1155
10c	87	53	24.5	84.5	13.8	30820	58.4	335	274	532	1141	32099	60.9	339	275	544	1158
22c	87	52	24.4	79.9	14.8	33374	63.3	344	280	546	1170	33373	63.3	341	280	543	1164
24c	87	53	21.2	79.1	12.9	33038	62.6	344	281	545	1170	33032	62.6	340	281	541	1162
41c	87	53	21.7	81.9	12.6	32215	61.1	337	278	546	1161	32207	61.1	334	278	543	1155
45c	87	53	20.4	81.5	12.0	31940	60.6	341	278	541	1160	31930	60.5	338	278	538	1154
13WA^	935	52	21.4	78.1	13.3	31846	60.4	334	279	534	1147	31844	60.4	333	279	531	1143
14WA	935	53	21.2	82.1	12.2	32004	60.7	336	276	537	1149	31999	60.7	336	276	534	1146
19CA	935	53	21.9	84.9	11.9	32302	61.2	339	274	538	1151	32301	61.2	337	274	536	1147
21CA	935	53	25.1	83.5	14.1	33548	63.6	341	278	547	1166	33546	63.6	339	278	545	1162
23WA	935	51	23.8	74.1	15.2	33990	64.4	342	281	549	1172	33986	64.4	340	281	546	1167
32CA	935	53	23.7	79.7	13.8	32384	61.4	335	277	525	1137	32381	61.4	335	277	523	1135
44CA	935	54	21.7	88.2	11.4	31143	59.0	332	276	524	1132	31141	59.0	331	276	522	1129
5c	935	53	24.5	86.8	13.2	30820	58.4	335	274	532	1141	30819	58.4	334	274	528	1136
26c	935	51	25.7	75.1	16.6	34888	66.1	347	280	555	1182	34886	66.1	344	280	552	1176
31CA	935	52	24.5	79.6	14.4	34102	64.7	342	280	547	1169	34099	64.7	341	280	544	1165
35c	935	50	22.5	70.6	15.6	33621	63.7	341	278	544	1163	33617	63.7	338	278	539	1155
49c^	935	52	22.4	79.1	14.4	29249	55.5	322	278	494	1094	29240	55.4	321	278	492	1091
Mean		**53.1**	**23.0**	**83.8**	**12.8**	**31476**	**59.7**	**336**	**276**	**531**	**1142**	**31505**	**59.7**	**334**	**276**	**529**	**1139**

Table 1. RNAseq data for 40 apple libraries: read quality, GC content, duplication percentage, reads mapped and total genes present in metabolism bins for deduplicated and non-deduplicated datasets.

Legend: SG = salivary gland, WB = whole bodies, CHO = carbohydrate AA = amino acid

sample	tissue or host	GC (%)	sequences (millions)	duplication (%)	aligned (%)	aligned (millions)	metabolism genes present CHO	AA	Lipid	protein binding	peptidase activity	transporter activity
WAA_SG1	SG	40.5	45.6	87.8	94.7	43.2	100	39	229	1228	712	726
WAA_SG2	SG	37	40.5	87.55	95.3	38.5	108	37	231	1214	727	758
WAA_SG3	SG	40	34.8	87.3	93.9	32.6	92	36	217	1179	687	714
Average		39.2	40.3	87.6	94.6	38.1	100.0	37.3	225.7	1207.0	708.7	732.7
WAA_WB1	WB	38.5	63.5	82.25	93.6	59.5	111	47	253	1340	767	859
WAA_WB2	WB	38	54.4	81	92.2	50.1	112	47	256	1340	765	850
WAA_WB3	WB	37.5	54	81.65	92.8	50.1	109	47	252	1329	761	852
Average		38.0	57.3	81.6	92.9	53.2	110.7	47.0	253.7	1336.3	764.3	853.7
EL16_1	G16	45	21.4	79.65	94.4	20.1	116	41	258	1319	785	861
EL16_2	G16	44	21.9	78.55	94.3	20.7	127	46	261	1352	818	876
EL16_3	G16	51	21.6	87.9	95.0	20.5	104	35	230	1208	722	746
EL87_1	G87	40	20.7	67.2	93.4	19.3	128	49	277	1419	861	944
EL87_2	G87	38	22.1	75.05	90.9	20.1	123	47	266	1388	820	907
EL87_3	G87	50	22.3	88.25	95.4	21.2	106	37	240	1243	725	777
EL935_1	G935	41	18.7	71.75	92.9	17.3	125	48	271	1381	828	918
EL935_2	G935	39	19.6	70.75	93.0	18.2	123	49	277	1409	847	934
EL935_3	G935	40.5	21	71.8	92.5	19.4	124	51	276	1406	846	933
Mean		43.2	21.0	76.8	93.5	19.6	119.6	44.8	261.8	1347.2	805.8	877.3
Total Mean		**41.3**	**32.1**	**79.9**	**93.6**	**30.1**	**113.9**	**43.7**	**252.9**	**1317.0**	**778.1**	**843.7**

WAA Assembly Summary

assembled bases	GC (%)	total contigs	mean contig length	median contig length	N10	N50
209,786,507	30.8	268,921	780	397	5113	1381

Table 2. Aphid RNAseq data and de novo transcriptome assembly summary for 12 whole body, and 3 salivary gland libraries. Read quality, reads mapped and total genes present in several metabolism GO terms are included. Transcriptome assembly and clustering were performed with Trinity and CD-HIT.

Gene	*A. pisum* AphidBase Gene ID	*A. pisum* NCBI RefSeq protein ID	WAA Trinity ID	BLASTP E-value	Secretory	LogFC in salivary gland	Adj. P. Value	Orthology Level	Aphid Performance	References*
C002	ACYPI008617	XP_001948358.2	DN4506_c0_g1_i1	1.97E-50	Yes	8.3	4.00E-04	Aphididae	*In planta* overexpression increases fecundity & transcript knockdown reduces phloem feeding/survival/fecundity	[1, 2, 3, 4, 6, 11]
			DN20770_c0_g2_i4	3.23E-33	Yes	-0.04	0.96			
Me23	ACYPI002439	NP_001155475.1	DN20770_c0_g2_i2	3.13E-33	No	-0.71	0.09	Aphididae	*In planta* overexpression enhances fecundity	[5]
			DN6812_c0_g1_i16	1.72E-86	No	5.9	2.43E-05			
			DN6812_c0_g1_i21	5.67E-81	No	7.4	3.50E-03			
			DN6812_c0_g1_i3	1.79E-53	Yes	6.5	1.84E-07			
			DN6812_c0_g1_i7	9.63E-39	No	6	1.26E-05			
Shp	ACYPI009881	XP_001943898.1	DN6812_c0_g1_i13	7.77E-26	No	6.4	2.10E-07	Aphidomorpha	Transcript knockdown reduces fecundity	[7, 10]
			DN4653_c0_g3_i2	7.13E-86	Yes	-2.86	8.00E-04			
			DN11304_c0_g1_i1	5.09E-29	Yes	-0.5	0.5			
			DN4660_c0_g2_i3	5.05E-25	Yes	-1.07	0.23			
Mp10	ACYPI000097	NP_001119652.1	DN3464_c1_g1_i1	1.36E-07	Yes	-0.85	0.26	Insecta	*In planta* overexpression reduces fecundity	[3]
			DN2575_c7_g1_i1	2.44E-27	No	-3.27	1.00E-04			
			DN2761_c0_g1_i3	1.77E-14	Yes	-2.32	1.60E-03			
Mif1	ACYPI002465	XP_008184019.1	DN2761_c0_g1_i5	6.74E-13	No	-2.31	5.37E-05	Arthropoda	Transcript knockdown reduces fecundity	[8]
Armet	ACYPI008001	XP_001949541.1	DN5107_c1_g1_i1	4.35E-95	Yes	1.7	1.90E-03	Metazoa	Transcript knockdown reduces survival	[9]

*1 Mutti et al. 2006; 2 Mutti et al. 2008; 3 Bos et al. 2010; 4 Pitino et al. 2011; 5 Atamian et al. 2013; 6 Pitino and Hogenhout 2013; 7 Abdellatef et al. 2015; 8 Naessens et al. 2015; 9 Wang et al. 2015a; 10 Will and Vilcinskas 2015; 11 Zhang et al. 2015;

Candidate effector

Table 3. Woolly apple aphid transcripts orthologous to previously characterized aphid effectors. BLASTP against the NCBI non-redundant database was used to determine orthology of WAA transcripts to effectors in other aphid species. The effectors C002, Shp, and Armet are orthologous to three WAA candidate effectors.

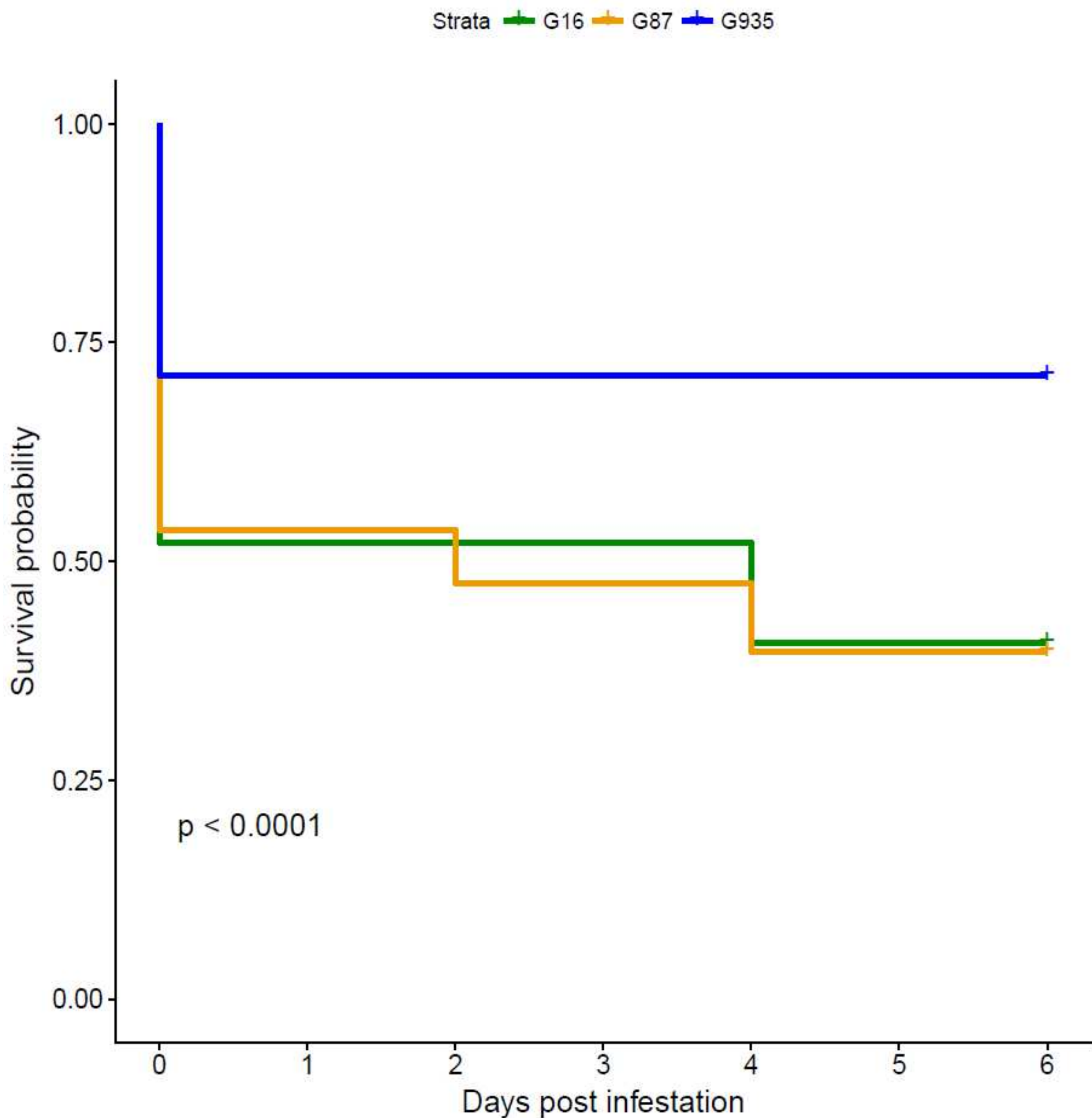
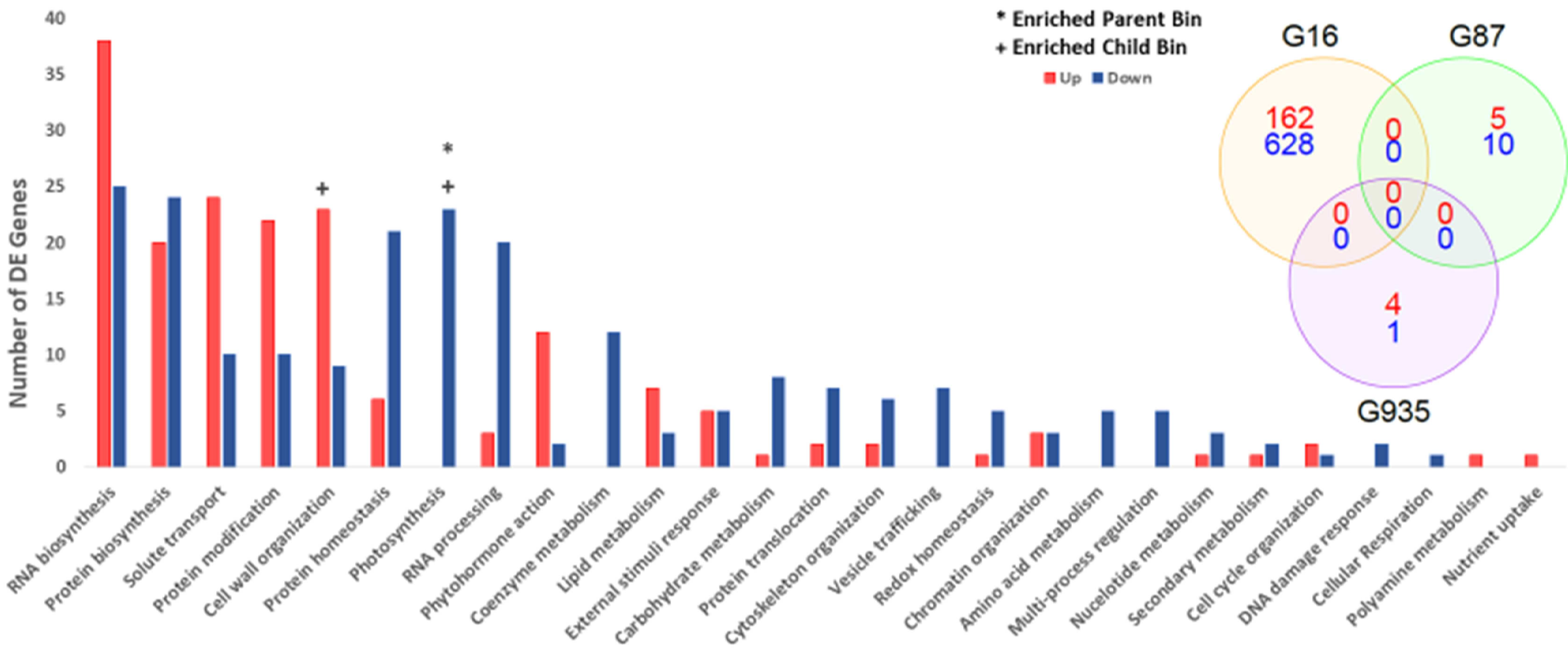
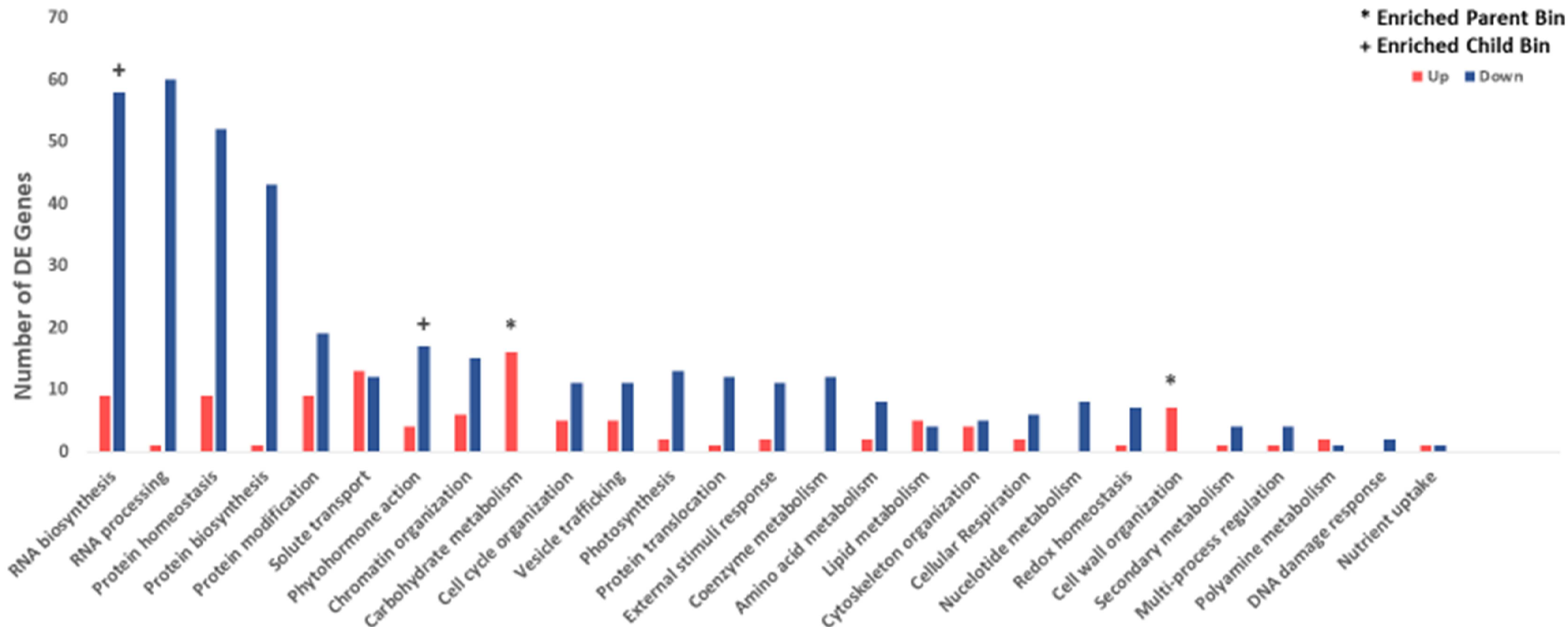


Figure 1. Aphid survival analysis on three host genotypes. Aphids performed equally well on G.16 and G.87, and significantly better on G.935. Host plants were challenged with WAA and aphids were counted 2, 4 and 6 DPI. Differences in rootstock resistances were determined with a Kaplan-Meier log-rank survival analysis. No aphids were found feeding on G.202 for the duration of the experiment, and therefore was omitted from the survival analysis.



All Aphid vs Control Wilcoxon-Rank Sum Enriched Bins	Total genes / Bin Total	Average logFC	p-value
Photosynthesis	23/354	-0.85	0.0002
Photophosphorylation	19/249	-0.90	0.0002
photosystem II	11/114	-0.88	0.043
Cell wall organisation.cutin and suberin	8/50	1.94	0.031

Figure 2. Venn diagram of up and downregulated genes after aphid feeding on three host genotypes and total up and downregulated genes across MapMan bins for all aphid-infested vs all control plants. Aphid feeding induced greater changes to host gene expression in G.16 vs G.87 and G.935. Aphid feeding also resulted in more downregulated genes (blue values) compared to upregulated genes (red values). An FDR threshold of <0.1 was used to determine significantly differentially expressed genes. In all aphid-infested vs control plants, 32428 total genes were tested for differential expression, and 675 were differentially expressed. Photosynthesis and cell wall organization bins were enriched as determined by Wilcoxon-rank sum tests.



G16 Aphid vs Control Wilcoxon-Rank Sum Enriched Bins	Total genes / Bin Total	Average logFC	p-value
Cell wall organisation	7/509	3.42	0.013
Carbohydrate metabolism	13/345	-1.46	0.036
RNA biosynthesis.organelle machinery	8/93	-1.62	0.048
Phytohormone action.auxin.biosynthesis.IAM pathway.indole-3-acetamide hydrolase	5/6	-2.47	0.036

Figure 3. Numbers of up and downregulated genes across MapMan bins for G.16 infested vs control plants. Cell wall organization, carbohydrate metabolism, RNA biosynthesis, and auxin synthesis bins were enriched as determined by Wilcoxon-rank sum tests. 32428 total genes were tested for differential expression, and 790 were differentially expressed at an FDR threshold of <0.1.

Control vs Control Wilcoxon Rank-Sum Enriched Bins			
G16 vs G935	total genes/ bin total	Average logFC	p-value
Protein modification	231/1615	-0.33	0.0001
Protein modification.phosphorylation	166/1224	0.61	2.38E-06
Enzyme classification	197/1362	0.86	0.002
Enzyme classification.EC_1 oxidoreductases	68/497	1.14	0.02
Enzyme classification.EC_2 transferases.EC_2.4 glycosyltransferase	26/170	2.38	0.0001
G87 vs G16			
Cell wall organisation	75/509	-1.41	1.94E-06
Cell wall organisation.hemicellulose	20/101	-1.82	0.044
Protein modification.phosphorylation	199/1224	-0.6	0.019
Enzyme classification.EC_2 transferases.EC_2.4 glycosyltransferase	26/170	2.24	0.0004

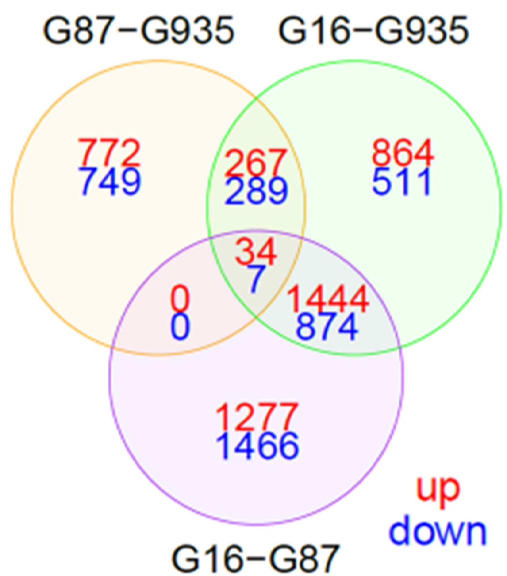


Figure 4. Differentially expressed genes across uninfested genotypes. Control plants of each genotype were contrasted to obtain profiles of constitutively expressed genes at FDR <0.1. G.16 constitutive expression was highly dissimilar to G.935 and G.87, in terms of numbers of DE genes and enriched bins, which included cell wall organization, protein modification, and enzyme classification.