A Gene Family Coding for Salivary Proteins (SHOT) of the Polyphagous Spider Mite *Tetranychus urticae* Exhibits Fast Host-Dependent Transcriptional Plasticity

Wim Jonckheere,^{1,2} Wannes Dermauw,^{1,†} Mousaalreza Khalighi,¹ Nena Pavlidi,² Wim Reubens,¹ Geert Baggerman,^{3,4} Luc Tirry,¹ Gerben Menschaert,⁵ Merijn R. Kant,⁶ Bartel Vanholme,^{7,8} and Thomas Van Leeuwen^{1,2,†}

¹Laboratory of Agrozoology, Department of Crop Protection, Faculty of Bioscience Engineering, Ghent University, Coupure links 653, 9000 Ghent, Belgium; ²Department of Evolutionary Biology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 904, 1098 XH, Amsterdam, The Netherlands; ³Center for Proteomics (CFP), University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium; ⁴Flemish Institute for Technological Research (VITO), Boeretang 200, 2400 Mol, Belgium; ⁵Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University; ⁶Department of Population Biology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam; ⁷Department of Plant Biotechnology and Bioinformatics, Ghent University, Technologiepark 927, 9052 Gent, Belgium; and ⁸Centre for Plant Systems Biology, VIB, Technologiepark 927, 9052 Gent, Belgium

Accepted 13 September 2017.

The salivary protein repertoire released by the herbivorous pest *Tetranychus urticae* is assumed to hold keys to its success on diverse crops. We report on a spider mite-specific protein family that is expanded in *T. urticae*. The encoding genes have an expression pattern restricted to the anterior podocephalic glands, while peptide fragments were found in the *T. urticae* secretome, supporting the salivary nature of these proteins. As peptide fragments were identified in a host-dependent manner, we designated this family as the SHOT (secreted host-responsive protein of Tetranychidae) family. The proteins were divided in three groups based on sequence similarity. Unlike *TuSHOT3* genes, *TuSHOT1* and *TuSHOT2* genes were highly expressed when feeding on a subset of family Fabaceae, while expression was depleted on other hosts. *TuSHOT1* and *TuSHOT2* expression was induced within 24 h after certain host transfers, pointing toward transcriptional plasticity rather than selection as the cause. Transfer from an ‘inducer’ to a ‘noninducer’ plant was associated with slow yet strong down-regulation of *TuSHOT1* and *TuSHOT2*, occurring over generations rather than hours. This asymmetric on and off regulation points toward host-specific effects of SHOT proteins, which is further supported by the diversity of SHOT genes identified in Tetranychidae with a distinct host repertoire.

Being sessile organisms, plants use a wide range of strategies to ward off arthropod herbivores. These include constitutive chemical and physical defenses that discourage feeding and are

present irrespective of the presence of herbivores as well as induced defenses that are specifically activated upon attack (Howe and Jander 2008; Wu and Baldwin 2010). The latter include the production of toxic and antidiigestive compounds, but also, the attraction of natural enemies of the herbivore via induced volatiles. Furthermore, resources can be reallocated toward nonattacked tissues to minimize the negative fitness consequences of tissue loss (Howe and Jander 2008). Because of these defenses, plants can be challenging to feed on and many herbivores became monophagous specialists that have evolved host-specific adaptations to deal with specific defenses (Ali and Agrawal 2012; Barrett and Heil 2012; Kant et al. 2015; Nosil 2002). On the other end of the diet breadth spectrum are the polyphagous species, of which the two-spotted spider mite *Tetranychus urticae* (Acari: Tetranychidae) is a representative. This cosmopolitan pest has been recorded to feed on over 1,000 different plant species belonging to more than 100 different families (Grbić et al. 2011; Spider Mites Web). The exceptional polyphagous nature of *T. urticae* has been linked to the expansion of several gene families implicated in digestion, detoxification, and transport of xenobiotics (Dermauw et al. 2013a and b; Grbić et al. 2011; Santamaría et al. 2012; Van Leeuwen and Dermauw 2016; Wybouw et al. 2012). Despite being extremely polyphagous as a species, *T. urticae* does not necessarily consist of generalist individuals (Fox and Morrow 1981; Kant et al. 2008). It has high intra-specific genetic variability (Magalhães et al. 2007) and it has been proposed that local populations form host races that do not perform equally well on all potential host species (Agrawal et al. 2002; Díaz-Riquelme et al. 2016; Gotoh et al. 1993; Navajas 1998).

Although spider mites have acquired an impressive arsenal of adaptations, plant defenses mainly orchestrated by the phytohormones jasmonic acid and salicylic acid still prevent mites from achieving their maximal reproductive performance (Alba et al. 2015; Ament et al. 2004; Kant et al. 2008; Li et al. 2002, 2004). Manipulation of plant defenses offers a major fitness advantage to the pest. Spider mite species like *T. evansi* (Alba et al. 2015; Sarmiento et al. 2011), *T. ludeni* (Godinho et al.

Wim Jonckheere and Wannes Dermauw contributed equally.

[†]Corresponding authors: Thomas Van Leeuwen; E-mail: thomas.vanleeuwen@ugent.be and Wannes Dermauw; E-mail: wannes.dermauw@ugent.be

*The e-Xtra logo stands for “electronic extra” and indicates that one supplementary figure, seven supplementary tables, and one dataset are published online.

2015), and some genotypes of *T. urticae* (Alba et al. 2015; Kant et al. 2008; Wybouw et al. 2015) were shown to suppress these defenses, thereby increasing their fitness (Ataide et al. 2016).

Recognition of an invading organism is a key for the plant to initiate a successful defense response. Dodging such recognition has been well-studied for plant-pathogen interactions, yet appears to apply to arthropod herbivores as well. Tissue damage caused by pathogens or herbivores produces endogenous molecules referred to as damage-associated molecular patterns, which are proposed to be recognized by pattern recognition receptors (PRRs) of the plant and initiate defense responses (Erb et al. 2012; Pieterse et al. 2012). In addition, this general response can be tailored to the attacker when conserved patterns of pathogen- or herbivore-specific molecules, known as pathogen- or herbivore-associated molecular patterns (PAMPs or HAMPs), also proposed to be detected by PRRs, are leading to PAMP or HAMP-triggered immunity (PTI or HTI) (Erb et al. 2012; Jones and Dangl 2006; Pieterse et al. 2012). However, pathogens and herbivores evolved so-called effector molecules (often proteins) that enable them to evade PTI or HTI of some of their host plants (Deslandes and Rivas 2012; Kant et al. 2015; Lo Presti et al. 2015). Interference of pathogen effectors with PTI leads to effector-triggered susceptibility. In turn, some plant varieties have evolved special sensory proteins referred to as R proteins, which recognize these effectors, activating effector-triggered immunity (ETI) as a by-pass of suppressed PTI and HTI. This, in its turn, drives the evolution of alternative effectors that yet again can suppress host resistance (Barrett and Heil 2012; Erb et al. 2012; Jones and Dangl 2006; Lapin and Van den Ackerveken 2013; Win et al. 2012). While the molecular details of this so-called zigzag model have been worked out for pathogens, it most likely applies for herbivores as well (Jing et al. 2017; Mondal 2017). However, the extent of the parallels is still a matter of intensive research (Kaloshian and Walling 2016).

As the saliva of arthropod herbivores forms an important source of molecules that could trigger or manipulate plant defense responses (Hogenhout and Bos 2011; Musser et al. 2002; Mutti et al. 2008), we partly elucidated the salivary protein repertoire of *T. urticae* by means of comparative proteomics and transcriptomics (Jonckheere et al. 2016). Besides the finding of proteins presumably involved in digestion and detoxification, a large collection of proteins with unknown functions was identified. Some of these proteins inhibited plant defenses and enhanced mite reproduction, in line with a potential role as effectors (Villaruel et al. 2016). However, the experimental setup of the proteomics analysis did not allow for the identification of endogenous peptides within *T. urticae* saliva (Jonckheere et al. 2016). Yet, such peptides could play, like mature proteins, distinct roles in the plant-mite interaction, since peptide signaling regulates several aspects of plant growth and development and endogenous plant peptide elicitors have a role in regulating plant immunity to both pathogens and herbivores (Yamaguchi and Huffaker 2011). Furthermore, plant-parasitic nematodes secrete peptides mimicking plant hormone-like peptides to establish a successful interaction with its host (Mitchum et al. 2012, 2013). In addition, arthropods are shown to use peptide toxins to subdue their prey (Fletcher et al. 1997). It is, therefore, relevant to look for endogenous peptides in the saliva of *T. urticae*, as spider mites might potentially employ salivary peptides to manipulate plant responses to obtain maximal fitness.

In the current study, we examined the peptide fraction of *T. urticae* saliva and focused on a putative effector family. This protein family appears to be unique for spider mites, is expanded in *T. urticae* compared with close relatives, and the corresponding genes show rapid changes in host-dependent expression.

RESULTS

Peptidomics analysis of *T. urticae* saliva.

T. urticae lines propagated for at least five generations on different host plants (*Phaseolus vulgaris* [bean], *Glycine max* [soybean], and *Solanum lycopersicum* [tomato]) were allowed to feed on an artificial diet encapsulated in parafilm. After 24 h, the artificial diet was collected for each plant-adapted mite line and the peptide content was analyzed using nano-liquid chromatography tandem mass spectrometry (LC-MS/MS). The detected peptides were pooled into one dataset, retaining host-specific information. Parent protein identity was inferred from its daughter peptides (Codrea and Nahnsen 2016). This protein identification was conducted against a concatenated target/decoy version of the *T. urticae* protein database and the common Repository of Adventitious Proteins (cRAP) database (discussed below).

In total, 279 peptides were identified, of which 216, 116, and 133 were found in the *Phaseolus vulgaris*, *G. max* and *S. lycopersicum* samples, respectively (Fig. 1; Supplementary Table S1). Identified peptides typically had an arginine (R) or lysine (K) at their C-terminus (210 of 279 peptides) or were preceded by an R or K in the sequence of the proteins from which they originate (51 of 279). Considering that miscleavage of tryptic peptide bonds is a common phenomenon (Siepen et al. 2007), it can be assumed that the identified peptides originated from tryptic cleavage. The identified daughter peptides mapped to a total of 40 different parent proteins distributed over the *Phaseolus vulgaris* (36 proteins), *G. max* (32 proteins), and *S. lycopersicum* (32 proteins) samples. Twenty-six of these proteins (65%) were also identified in a previous study in which the salivary protein repertoire of *T. urticae* was studied using a proteomic approach (Jonckheere et al. 2016). Conserved domain database searches were used to find additional information on the function of the 40 identified proteins. Five, two, and six peptides belonged to either proteinase inhibitors (tetur22g00290, tetur40g00392, tetur95g00060, and tetur95g00080), serine proteases (tetur12g03940, tetur12g03950, tetur30g01440, tetur66g00060, and tetur66g00070), and proteins with an Armadillo-fold (InterPro domain IPR016024: tetur31g00630 and tetur49g00080), respectively. No function could be allocated to the other 29 proteins for which peptides were identified (266 of 279 peptides).

Except for tetur18g00530, all proteins from which peptides were identified had a predicted signal peptide (SP) for secretion and 34 of the 40 proteins (85%) were predicted to be targeted toward the extracellular space (Supplementary Table S2). In addition, most of the corresponding genes (33 of 40) were, based on an RNA-seq expression analysis, significantly over-expressed in the proterostoma—the body region that harbors the salivary glands—relative to the entire mite body (Benjamini-Hochberg adjusted *P* value ≤ 0.05 and fold change (FC) ≥ 8). Together these data support the salivary nature of the identified proteins.

About 50% (149 of 279) of the identified peptides mapped to only two protein groups (tetur29g01360; tetur07g00160 and tetur07g01660) that were present in all host-specific samples. Eight peptides belonged to four *T. urticae* proteins (tetur22g00290, tetur30g01440, tetur31g00830, and tetur49g00080) that were exclusively found in the sample of mites that had been feeding from *S. lycopersicum* (Fig. 1). Remarkably, about 20% (49 of 279) of the identified peptides mapped to a group of related parent proteins (tetur03g03620, tetur03g03670, tetur03g03680, tetur03g03690, tetur03g03700, tetur03g03730, tetur03g10083, tetur03g10093, tetur11g01360, tetur11g06390, and tetur11g06400). These proteins and homologs thereof are here designated to the SHOT (secreted host-responsive protein in Tetranychidae) family.

T. urticae SHOT family analysis.

tBLASTn searches resulted in the identification of 13 *SHOT* genes in the *T. urticae* genome (available at the Online Resource for Community Annotation of Eukaryotes website [OrcAE]) (Grbić et al. 2011), while four and two were found in

the draft genomes of *T. lintearius* and *T. evansi*, respectively (Supplementary Table S3; Supplementary Dataset S1). The *Panonychus ulmi* and *Panonychus citri* transcriptomes harbored one and three *SHOT* transcripts, respectively, i.e., contig 00033 (*Panonychus ulmi* [Bajda et al. 2015]) and contigs 183, 22851, and 29686 (*Panonychus citri* [Bajda et al. 2015]). BLASTp and tBLASTn searches against the nonredundant protein database (E-value < E^{-5}) and Acari genomes, respectively, using intact *Tetranychidae* *SHOT* members as a query did not yield hits with proteins of nontetranychid species.

Genes lacking a proper start codon or containing a frameshift (*T. urticae*: 1, *T. lintearius*: 2, *Panonychus citri*: 2) were considered as pseudogenes or fragments and were not included for further analysis. A maximum-likelihood phylogenetic analysis of *SHOT* protein sequences (Fig. 2) revealed that the *SHOT* family could be divided into three different clusters: a cluster containing tetur03g03620, tetur03g03680, tetur03g03700, tetur03g03740, tetur03g10083, and TISHOT1a (referred to as *SHOT1*), a cluster containing tetur03g03670, tetur03g03690, tetur03g03730 tetur03g10093 and TISHOT2a (referred to as *SHOT2*), and a third cluster containing tetur11g01360, tetur11g06390, tetur11g06400, TeSHOT3a, TeSHOT3b, *Panonychus ulmi* contig 00033 and *Panonychus citri* contig 183 (referred to as *SHOT3*).

Intact *T. urticae*, *T. lintearius*, and *T. evansi* *SHOT* genes are single exon genes and, in *T. urticae*, *SHOT3* genes are located in an 11-kb region of scaffold 11 (671,365 to 682,501 bp), while *TuSHOT1* and *TuSHOT2* genes are in close proximity of each other on scaffold 3 (1,642,274 to 1,703,045 bp). Except for tetur03g03620, *TuSHOT1* and *TuSHOT2* genes occur as consecutive pairs (tandem duplications) of one *SHOT1* gene and a neighboring *SHOT2* gene that is located on the opposite strand (1,682,835 to 1,703,045 bp).

All intact *SHOT* members have a predicted SP for secretion and are enriched in glycine, alanine, and serine (on average 17.7, 17.6, and 11.1%, respectively [Supplementary Table S4]). In contrast to *Panonychus* *SHOT* members, *Tetranychus* *SHOTs* possess either a proline-glutamic acid, proline-glutamine, or proline-aspartate motif at the N-terminus (Supplementary Fig. S1). The average length of *SHOT1*, *SHOT2*, and *SHOT3* proteins is 324, 275, and 319 amino acids, respectively. For those *TuSHOT* genes of which expression data of developmental stages was available, the average expression of intact *SHOT* genes across all feeding stages (larva, nymph, adult) relative to the expression in the embryo stage was four- to 18-fold higher than the overall genome-wide average of this expression ratio (average genome-wide ratio of 2.2 [Jonckheere et al. 2016]). Except for tetur03g03620, all intact *SHOT* genes were significantly overexpressed (fold changes ranging from 17 to 40) in the proterostoma relative to the entire mite body. Interestingly, the peptides originating from *SHOT1* and *SHOT2* proteins (41 of 42) were almost uniquely identified in either the *Phaseolus vulgaris* or *G. max* samples (Fig. 1A), while four out of seven *SHOT3* peptides were only found in *S. lycopersicum* samples.

Fig. 1. Peptides identified in an artificial diet probed by *Tetranychus urticae* lines adapted to different host plants. **A**, A Venn diagram depicting the overlap in peptides identified in the different samples. The number of unique proteins from which these peptides could originate is shown between brackets. All but one (41 of 42) of the peptides originating from *SHOT1* and *SHOT2* proteins were identified in the *Glycine max* or *Phaseolus vulgaris* samples exclusively (shaded in gray). **B**, Heatmap of normalized intensity ($-\log_{10}$ transformed) of identified peptides from *SHOT* proteins. The Euclidean distance metric and Ward's method were used for clustering of columns. *SHOT1* and *SHOT2* peptides that were in silico predicted to have the highest confidence to be detected (Peptide Prediction with Abundance analysis [Muntel et al. 2015]) are marked with one asterisk, while variants of these peptides are marked with two asterisks.

We also looked at precursor peak (MS1) intensities of peptides, which can be used for label-free relative protein quantification (Argentini et al. 2016) (discussed below). SHOT1 and SHOT2 peptides flanking the predicted SP (with amino acid sequences NEQSPELNQLQGNVHGR, NEESPELNQLQSNVHGR, NEQSPELNQLQSNVHGR, and NEQSPDFDLQSNVHGR) were among the top ten peptides with the highest MS1 intensity in both the *Phaseolus vulgaris* and *G. max* samples, while a similar peptide of SHOT3 proteins (DEQPELSLNQKVHGR) was among the top ten peptides with the highest MS1 intensity in the *S. lycopersicum* sample (Fig. 1B). This indicates that the expression of the *SHOT* genes is host plant-dependent.

Salivary gland-specific expression of *SHOT* genes.

The spatial expression pattern of three *T. urticae* *SHOT* genes was investigated using whole-mount in situ hybridization (ISH). Three *SHOT* genes were selected as representative for the different subfamilies (*tetur03g03690* for *TuSHOT2b*, *tetur03g03700* for

TuSHOT1c, and *tetur11g01360* for *TuSHOT3a*). The genes were shown to be expressed in the anterior podocephalic glands of *T. urticae* (Fig. 3). Simultaneously performed ISH experiments with sense control probes did not show localized expression, supporting specific staining of the antisense probes.

SHOT gene expression on various hosts.

The expression level of *SHOT* genes of *T. urticae* feeding on different hosts was studied using reverse transcription-quantitative polymerase chain reaction (RT-qPCR). Due to strong gene sequence similarity within the *SHOT* family, three nonspecific primer sets were used during qPCR analysis, each amplifying a group of *SHOT* genes (SHOT1_q_Fw and Rv primers amplifying a region of *tetur03g03680*, *tetur03g03700*, *tetur03g03740*, and *tetur03g10083* [belonging to the *SHOT1* cluster], SHOT2_q_Fw and Rv primers amplifying *tetur03g03670*, *tetur03g03690*, *tetur03g03730*, and *tetur03g10093* [belonging to the *SHOT2* cluster], and SHOT3_q_Fw and Rv amplifying *tetur11g01360*,

Fig. 2. Phylogeny of the *SHOT* family and expression levels of *Tetranychus urticae* *SHOT* genes in different mite lines. **A**, Midpoint-rooted maximum-likelihood tree of *SHOT* proteins from *T. urticae* (*TuSHOT*), *T. evansi* (*TeSHOT*), *T. lintearius* (*TISHOT*), *Panonychus citri* (*PcSHOT*), and *Panonychus ulmi* (*PuSHOT*). Three groups can be distinguished: *SHOT1*, *SHOT2*, and *SHOT3*. Branch length represents the substitutions per site, while numbers above the branches show bootstrap support based on 1,000 bootstrap replicates. **B**, Top, Heatmap of cyanine intensities of *TuSHOT* genes in strains adapted to *Phaseolus vulgaris*, *Glycine max*, *Solanum lycopersicum*, *Gossypium hirsutum*, or *Zea mays* (Jonckheere et al. 2016). Bottom, Heatmap of $\log_2$ fold changes of *TuSHOT* genes after host shifts between *Phaseolus vulgaris* and *S. lycopersicum* (B_24hT, mites from the London reference strain on *Phaseolus vulgaris* transferred to *S. lycopersicum* for 24 h; B_30GT, mites from the *Phaseolus vulgaris* strain grown on *S. lycopersicum* for 30 generations; B_30GT_2GT, mites from the *Phaseolus vulgaris* strain grown on *S. lycopersicum* for 30 generations and transferred back to *Phaseolus vulgaris* for two generations (Wybouw et al. 2015; The Dryad Digital Repository).

Fig. 3. Expression localization of *Tetranychus urticae* *SHOT* genes. Whole-mount in situ hybridization reveals specific expression in the anterior podocephalic glands. All tested *SHOT* genes (*SHOT1c*, *SHOT2b*, and *SHOT3a*) show an identical expression pattern. No specific signal development was visible in negative control samples using sense probes. **A**, *SHOT3a* probe, lateral view. **B**, *SHOT2b* probe, dorsal view. Scale bar represents 100 μ m.

tetur11g06390, and *tetur11g06400* [belonging to the *SHOT3* cluster]) (Supplementary Table S5). *T. urticae* strains maintained on *Prunus avium*, *Cucumis sativus*, and *Sambucus nigra* showed a highly significant reduced expression of *SHOT1* and *SHOT2* genes, compared with the reference *T. urticae* London strain adapted to *Phaseolus vulgaris* (Fig. 4). This extremely reduced expression ranged from a log₂FC of -9.15 (95% confidence interval [CI]: -8.15 to -10.15) for the *SHOT2* gene group of mites adapted to *Cucumis sativus* to a log₂FC of -13.93 (95% CI: -13.21 to -14.66) for the *SHOT1* gene group of mites adapted to *Prunus avium*. In absolute numbers, this means expression dropped 568- to 15,608-fold compared with the expression in mites maintained on *Phaseolus vulgaris*. A similar observation was made for *T. urticae* (London strain) adapted to *S. lycopersicum* and *Gossypium hirsutum*. Also, in this case, *SHOT1* and *SHOT2* genes of mites feeding on these two host plants were expressed at extremely reduced levels. However, mites adapted to the fabacean host *Medicago truncatula* had *SHOT1* and *SHOT2* gene expression levels that were significantly higher than in the reference mite strain maintained on *Phaseolus vulgaris*. In short, when adapted to *Phaseolus vulgaris* or *Medicago truncatula*, *T. urticae* does express *SHOT1* and *SHOT2* genes at high levels while these genes are virtually nonexpressed when the mites were maintained on *Prunus avium*, *Gossypium hirsutum*, *Cucumis sativus*, *S. lycopersicum*, or *Sambucus nigra*. The expression levels of the *SHOT3* genes were slightly but significantly upregulated when feeding on all investigated host plants, relative to *Phaseolus vulgaris*, except for *Medicago truncatula* (Fig. 4). The variation in expression level for *SHOT3* genes was, however, less pronounced than the variation observed for the *SHOT1* and *SHOT2* genes. For example, the *SHOT3* genes of mites adapted to *Prunus avium* were most strongly upregulated, with an associated log₂FC of 3.18 (95% CI: 1.96 to 4.40).

Time dependency of *SHOT* gene expression induction.

Preliminary experiments showed that *SHOT1* and *SHOT2* gene expression levels could quickly increase when mites were transferred from noninducer plants such as *S. lycopersicum* to inducer plants such as *Phaseolus vulgaris*. Expressional changes after such host transfer were studied in closer detail over a 24-h time span. *T. urticae* mites adapted to *S. lycopersicum* were transferred to *Phaseolus vulgaris* plants and the expression of *SHOT* gene groups was monitored 3, 6, 12, and 24 h after transfer. Mites collected from the original host plant *S. lycopersicum* were used for expression analysis at 0 h. Mites adapted to and maintained on *Phaseolus vulgaris* served as a reference condition. The expression ratios are represented in Figure 5. While the expression ratios of *SHOT1* and *SHOT2* genes of *T. urticae* adapted to *S. lycopersicum* were extremely low compared with the *Phaseolus vulgaris* reference line (mean log₂FC of -12.35 at 0 h), mites transferred to *Phaseolus vulgaris* showed a severely reduced downregulation as soon as 3 h after transfer. Expression was still significantly lower than in mites from the reference strain (mean log₂FC of -4.08 at 3 h), although expression of both *SHOT1* and *SHOT2* groups increased about 300 times within 3 h after host transfer. The transcript levels increased further until, after 24 h on *Phaseolus vulgaris*, gene expression levels equaled or exceeded the expression level of reference mites adapted to and maintained on *Phaseolus vulgaris*. Expression of *SHOT3* genes of the *S. lycopersicum*-derived mites was significantly higher than the *Phaseolus vulgaris* reference strain at all time-points after transfer. As 24 h appeared to be sufficient for *T. urticae* to fully induce *SHOT1* and *SHOT2* gene expression after host transfer, this time scale was used for additional short-term expression analyses.

SHOT gene expression modulation after host transfer.

T. urticae mites from the *S. lycopersicum*-adapted strain were transferred to new host plants, and, 24 h after transfer, the

Fig. 4. Relative transcript abundance of *SHOT* groups in *Tetranychus urticae* strains adapted to different host plants. The different host-adapted *T. urticae* strains were originally derived from the London strain, except for three wild strains (indicated by °). Log₂-transformed data are presented as expression levels in the different strains, relative to the *Phaseolus vulgaris*-adapted strain. Error bars show the 95% confidence interval (CI), while asterisks indicate the significance level of the expression ratio of the host strains relative to the strain on *Phaseolus vulgaris* (one asterisk [*] indicates $P \leq 0.05$, two [**] $P \leq 0.01$, three [***] $P \leq 0.001$, and four [****] $P \leq 0.0001$).

expression levels of *SHOT* genes were evaluated (Fig. 6). The *SHOT3* genes were expressed at a slightly yet significant higher level (mean log₂FC over all host plants, 1.61 ± 0.37 [standard error of the mean]) when mites fed on all tested host plants,

relative to the *Phaseolus vulgaris*-adapted strain. For the *SHOT1* and *SHOT2* genes, different responses were noticeable. Genes of these subfamilies were either significantly down-regulated (*T. urticae* transferred to *Lathyrus odoratus*, *Mimosa*

Fig. 5. Relative transcript abundance of *SHOT* gene groups in *Tetranychus urticae* transferred from *Solanum lycopersicum* to *Phaseolus vulgaris* for increasing periods of time. Mites originating from a strain adapted to *S. lycopersicum* were transferred to *Phaseolus vulgaris* plants for 3, 6, 12, and 24 h. Mites maintained on *S. lycopersicum* were used for gene expression at 0 h. Log₂-transformed data are presented as gene expression levels in mites that had been transferred to *Phaseolus vulgaris* for the hours shown relative to *T. urticae* adapted to and maintained on *Phaseolus vulgaris*. Error bars show the 95% confidence interval (CI), while asterisks indicate the significance level (one asterisk [*] indicates $P \leq 0.05$, two [**] $P \leq 0.01$, three [***] $P \leq 0.001$, and four [****] $P \leq 0.0001$).

Fig. 6. Relative transcript abundance of *SHOT* gene groups in *Tetranychus urticae* transferred to different host plants for 24 h. Mites from the strain adapted to *Solanum lycopersicum* were transferred to a new host for 24 h. Fabacean species are marked above the bar chart and include several Papilionoideae (Fabeae and non-Fabeae) and one Mimosoideae species. Log₂-transformed data are presented as gene expression levels in *T. urticae* maintained on the new host for 24h, relative to *T. urticae* adapted to and maintained on *Phaseolus vulgaris*. Error bars show the 95% confidence interval (CI), while asterisks indicate the significance level (one asterisk [*] indicates $P \leq 0.05$, two [**] $P \leq 0.01$, three [***] $P \leq 0.001$, and four [****] $P \leq 0.0001$).

spp., *Arabidopsis thaliana*, *Gerbera jamesonii*, *Nicotiana benthamiana*, and *Zea mays*) or highly expressed (*T. urticae* transferred to *Arachis hypogaea*, *G. max*, *Lupinus angustifolius*, and *Ulex europaeus*). The mean relative expression levels of *SHOT1* and *SHOT2* genes of mites transferred to *Pisum sativum* were much lower than the reference strain (mean log₂FC: -10.17), yet, this was bordering significance due to high variation between biological replicates. Therefore, the experiment with *Pisum sativum* was repeated, with the same outcome (data not shown).

DISCUSSION

Peptidomics analysis of *T. urticae* saliva.

The *T. urticae* salivary proteome was previously analyzed by shotgun proteomics in which a tryptic digest was performed on the secreted salivary protein mixture prior to nano-LC-MS/MS analysis (Jonckheere et al. 2016). To determine the endogenous peptide content of spider mite saliva, we now complemented this proteomics analysis with a peptidomics analysis. Spider mite saliva was collected as by Jonckheere et al. (2016), but the samples were enriched in protein fragments or peptides smaller than 10 kDa and no tryptic digestion was applied. It should be noted, however, that peptides longer than 30 aa could not be detected by our approach, due to the technical limitations of bottom-up peptidomics (Steen and Mann 2004). In future experiments, this drawback might be circumvented by performing a proteomics analysis (i.e., proteolytic digestion with trypsin or another protease prior to sample analysis) on a sample enriched in peptide fragments smaller than 10 kDa.

Except for peptides mapping to tetu18g00530, all peptides identified in this study mapped to *T. urticae* proteins longer than 100 aa and almost all peptides showed signs of tryptic cleavage, indicating that the peptides identified in the artificial diet mainly comprise tryptic cleavage fragments of longer proteins instead of being discrete short open reading frame-encoded peptides. As peptidomics approaches do not involve in vitro digestion of the sample prior to analysis, endogenously secreted tryptic proteases most likely digested the secreted proteins in the saliva or artificial diet. Strikingly, a previous proteomics analysis (Jonckheere et al. 2016) and, to a lesser extent, current study revealed the presence of endogenous trypsins (serine proteases with InterPro domain IPR033116 and IPR018114) in *T. urticae* saliva, suggesting that tryptic peptides could be generated in situ by endogenous enzymes.

SHOTs are produced in the salivary glands and are secreted in a host-dependent manner.

The peptidomics approach to identify salivary proteins revealed a novel gene family, designated the SHOT family. *SHOT* genes were expressed at considerably higher levels in the salivary gland region (proterosoma) compared with the entire spider mite body (Jonckheere et al. 2016), while whole-mount in situ hybridizations confirmed that *SHOT* genes are expressed in the salivary glands of *T. urticae*, more specifically in the anterior podocephalic glands (Fig. 3). Hence, *SHOT* genes can be considered to encode genuine salivary proteins. Interestingly, SHOTs were secreted in a host plant-dependent manner. For example, peptides belonging to proteins of the *SHOT1* and *SHOT2* group were almost exclusively found in artificial diet enriched with saliva of mites adapted to the fabacean plants *Phaseolus vulgaris* or *G. max* (Fig. 1A and B), whereas *SHOT3* peptides were identified in diet after feeding of all the host plant-adapted mite lines (Fig. 1B). Remarkably, N-terminal peptides of mature SHOTs had a very high MS1 intensity compared with peptides matching to other regions of mature SHOTs (Fig. 1). Although tempting to link the abundance of

these N-terminal SHOT peptides to a biological function of these peptides per se, these peptide fragments also had the highest likelihood to be detected by MS (peptide prediction with abundance (PPA) analysis [Supplementary Table S6; discussed below]), and their high intensity is probably due to their specific physicochemical properties.

Induction of *SHOT* expression by host transfer.

Gene-expression analysis revealed that *SHOT1* and *SHOT2* genes were highly expressed in *T. urticae* lines adapted to *Phaseolus vulgaris* or *G. max*, while, on average, being much lower expressed in lines adapted to *S. lycopersicum*, *Gossypium hirsutum*, or *Zea mays*. *SHOT3* genes, on the other hand, did not show such extreme variation in expression levels (Fig. 2B) (Jonckheere et al. 2016). To test whether the host-dependent expression of the *SHOT1* and *SHOT2* genes was the result of adaptation (between generations) or transcriptional plasticity (within a single generation), we performed a time course experiment (Fig. 5). Since 24 h sufficed to fully induce *SHOT1* and *SHOT2* expression, selection could be excluded as cause for this differential expression. In an additional set of experiments, mites were transferred from the noninducer *S. lycopersicum* to several other hosts for 24 h (Fig. 6). Combining all host-dependent gene expression data, it was found that the *SHOT1* and *SHOT2* genes are highly expressed when *T. urticae* is feeding on a subset of fabacean hosts, more specifically, subfamily Papilionoideae, with the exclusion of the Fabaeae tribe, and that this induction occurs rapidly in all cases. Interestingly, host transfer of mites in the opposite direction (from inducer to noninducer host) is associated with a relatively slow downregulation of *SHOT1* and *SHOT2* genes (Fig. 2A) (Wybouw et al. 2015). Such slow downregulation could be determined by the timing of food digestion or excretion, as inducing plant compounds may be detained in the mite body. On the other hand, it may suggest that *SHOT1* and *SHOT2* upregulation when feeding on a subset of Papilionoideae hosts is much more important to the mite than downregulation on non-Papilionoideae hosts and point toward a Papilionoideae-specific plant trait, such as a metabolite that triggers expression. Notably, transcriptional plasticity in a different group of mite effector genes was also observed in response to competition (Schimmel et al. 2017), reinforcing the notion that spider mites might be able to customize the release of salivary (effector) proteins depending to the circumstances.

The transcriptional plasticity of genes like *SHOT1* and *SHOT2* suggests that a sensory system may be involved, allowing *T. urticae* to detect environmental cues such as those from host plants and to customize salivary gland gene expression accordingly. Recently, it has been shown that the chemosensory perception of *T. urticae* is well-developed and consists of an extraordinarily large number of chemosensory receptors in lineage-specific proliferations (Ngoc et al. 2016). A potential factor determining *SHOT1* and *SHOT2* induction may be secondary metabolites. Isoflavonoids, for example, are predominantly found in subfamily Papilionoideae (Wink 2013), including those used in this study, but non-Fabaeae Papilionoideae (such as *Arachis hypogaea*, *G. max*, *Lupinus angustifolius*, *Ulex europaeus*) specific isoflavonoids have, to our knowledge, not yet been reported. Feeding on *Mimosa* spp., *Lathyrus odoratus*, and *Pisum sativum* did result in variable *SHOT1* and *SHOT2* expression levels between biological replicates (Fig. 6, note the rather large 95% CI), while the variability on the other tested plant species was much smaller. Maintaining replicates of these plants under different controlled conditions or subjecting them to different stresses prior to mite transfer and expression analysis might point toward the plant traits responsible for the variability in *SHOT1* and *SHOT2* gene expression.

Host-dependent expression of salivary genes has also been described in aphids and caterpillars (Afshar et al. 2010; Eyres et al. 2016; Lu et al. 2016; Pan et al. 2015). *ACYPI006346* expression, for example, differed between *Acyrtosiphon pisum* colonies feeding on three different host plants (Pan et al. 2015). However, these three aphid lines were each adapted to one of the three hosts, and the observed differential expression does therefore not necessarily reflect a plastic response as, alternatively, it may be the result of plant-dependent selection. Finally, differential expression of genes, including several putative salivary genes, of aphids from within a single generation on different hosts has been observed as well (Eyres et al. 2016; Lu et al. 2016). However, the rapid and strong upregulation in response to a specific set of closely related plant species is, to our knowledge, unprecedented.

SHOT genes and host plant range.

Despite the polyphagous nature of *T. urticae*, *Phaseolus vulgaris* is one of its most suitable hosts (Yano et al. 1998) and fabacean plants, in general, appear to be readily accepted as a host by *T. urticae* (van den Boom et al. 2003). Furthermore, the plant family with the most species on which *T. urticae* has been recorded is the Fabaceae family (recorded on 119 Fabaceae species) (Spider Mites Web). The diversification of SHOTs into multimember subfamilies (Fig. 2A) in *T. urticae* may be causally linked to the expanded host range of this species with its apparent preference for fabacean plants. The distribution of intact homologs across the related spider mites *T. lintearius*, *T. evansi*, *Panonychus citri*, and *Panonychus ulmi* supports this hypothesis.

First, *T. lintearius* has a narrow host range (six recorded host plants [Spider Mites Web]) and is considered to be an *Ulex* spp. (family Fabaceae, subfamily Papilionoideae) specialist. This close relative of *T. urticae* can typically be found on common gorse (*Ulex europaeus*) (Hill and O'Donnell 1991; Ireson et al. 2003; Norambuena et al. 2007), a plant we have shown to rapidly induce *SHOT1* and *SHOT2* expression in *T. urticae* (Fig. 6). *T. lintearius* was found to possess *SHOT1* and *SHOT2* genes (Fig. 2A), while the two identified *TISHOT3* genes appeared to be pseudogenes or fragments. Such pseudogenes can be regarded as 'fossilized footprints of past gene expression' and are relics of former genes that no longer possess biological functions (Podlaha and Zhang 2010). Likely, the common ancestor of *T. lintearius* and *T. urticae* possessed intact *SHOT3* genes, which were subsequently lost in *T. lintearius* when it specialized on its fabacean host *Ulex europaeus*.

Second, *T. evansi* is considered to be a specialist of family Solanaceae, although it has been recorded on other host plants as well (Navajas et al. 2013; Spider Mites Web). *T. urticae* does not express *SHOT1* and *SHOT2* genes when feeding on solanaceous species such as *S. lycopersicum* or *Nicotiana benthamiana* (strong downregulation relative to the reference inducer *Phaseolus vulgaris*) (Figs. 4, 5, and 6). Accordingly, members belonging to *SHOT1* and *SHOT2* appear to be absent in the *T. evansi* genome. However, *T. evansi* does possess two *SHOT3* genes (Fig. 2A). The *T. urticae* homologs of these *TeSHOT3* genes were expressed at high levels in *T. urticae* feeding on all tested host plants (Figs. 4 and 6).

Third, the oligophagous *Panonychus citri* has been recorded on 108 hosts belonging to different families, yet it can typically be found on citrus plants and fruit trees. As such, it is only rarely found on fabacean plants like *T. evansi* (Spider Mites Web; Takeyama et al. 2006). In line with this, *Panonychus citri* *SHOT1* and *SHOT2* genes could also not be identified in the transcriptome of this mite, yet, one gene belonging to the *SHOT3* cluster was found to be expressed.

Fourth and finally, a second *Panonychus* species, *Panonychus ulmi*, has been recorded on 141 different hosts, most of which

belong to the Rosaceae family and including only a few fabacean plants (Spider Mites Web). The only *SHOT* gene of *Panonychus ulmi* found to be expressed is a *SHOT3* type.

In summary, all tetranychid species for which genomic or transcriptomic resources were available possessed *SHOT* genes. *SHOT1* and *SHOT2* genes were only discovered in spider mites of which fabacean species are commonly reported as natural hosts, presumably reflecting a specialized function of these proteins in a subset of family Fabaceae. As stated by Vandermoten et al. (2014), polyphagous herbivores are exposed to a greater diversity of plant defenses, which may require a larger complement of salivary proteins. The expansion of the *SHOT* family in *T. urticae* may be an example of this hypothesis.

SHOTs as putative effectors.

In stylet feeding herbivores such as gall midges (Stuart et al. 2012; Zhao et al. 2015) and aphids (Carolan et al. 2011; Hogenhout and Bos 2011; Mondal 2017; Mutti et al. 2008), effectors are typically delivered inside the host via physical injection of salivary secretions. This injection is suggested to allow for immediate interaction of spider mite effectors with the damaged plant tissue (Bensoussan et al. 2016). We have shown that SHOTs are produced in the salivary glands of *T. urticae* and are injected into the mite diet, supporting a suggested role as effectors. While functional validation is required to determine if SHOTs alter host cell structure and function (Hogenhout and Bos 2011), there are additional indications that SHOTs may be effectors. Indeed, host-dependent expression levels, as reported here for *SHOT* genes, may imply that a protein serves as an effector, mediating herbivore-plant interaction (Pan et al. 2015). Further, *SHOT* genes are highly expressed in the feeding stages of *T. urticae* relative to the nonfeeding embryonic stage, pointing toward a feeding-related function. In addition, tetranychid SHOTs do not show homology (BLASTp E-value <10) with proteins of other organisms while also lacking known functional domains. Therefore, SHOTs are likely not involved in general functions such as digestion and detoxification but, rather, possess a unique function, potentially as an effector. The amino acid composition of *Tetranychus* SHOT (discussed above) is somehow reminiscent of *Mycobacterium tuberculosis* PE_PGRS effector proteins. These *Mycobacterium* proteins are characterized by a proline-glutamic acid motif at their N-terminus and are rich in glycine-alanine repeats (Deng et al. 2017; Tian and Jian-Ping 2010). Some PE_PGRS proteins in *M. tuberculosis* bind calcium (Yeruva et al. 2016), and if SHOTs would have a similar property, they could potentially interfere with the Ca²⁺ signaling of the host plant, inhibiting defense induction. Sabotage of calcium-dependent sieve-tube exclusion in attacked plants by salivary components that bind calcium has been described for phloem-feeding aphids (Will et al. 2007). However, spider mites do not feed from vascular sap and the sequence homology is low and any statement regarding possible function is rather speculative and should be the subject of further research.

The diversification of the *SHOT* gene family in Tetranychidae is correlated with the diversity and identity of the host plants of each of the respective mite species. The polyphagous spider mite *T. urticae* has intact *SHOT* genes belonging to all three subfamilies we identified, while the monophagous species *T. lintearius* and the oligophagous species *T. evansi*, *Panonychus citri*, and *Panonychus ulmi* possess either only intact *SHOT1* and *SHOT2* or only intact *SHOT3* genes. The highly inducible and fast regulation of *TuSHOT1* and *TuSHOT2* gene expression after transfer to particular hosts suggests they may play a central role in determining mite-plant compatibility. We formulated three hypotheses with respect to host plant compatibility. Our first hypothesis is that the *SHOT1* and *SHOT2*

proteins could facilitate colonization on all the host plants tested but could be needed in different quantities on different hosts, e.g., in higher quantities and for a longer period of time when feeding on a subset of family Fabaceae (subfamily Papilionoideae, excluding the Fabaeae tribe) compared with the other hosts. In that case, downregulation may reflect mostly a fine-tuning event that could also serve to save resources. This hypothesis is also in line with the observation that downregulation of *SHOT1* and *SHOT2* expression on non-Fabaceae spp. is relatively slow compared with the speed of upregulation under the reverse conditions. Such asymmetric on and off regulation suggests that the mite need to upregulate the *SHOT* genes on the subset family Fabaceae is much greater than its need to downregulate these on the alternative hosts. If so, it is possible that, on this subset of family Fabaceae, long-lasting high production of SHOT proteins is required for these to take effect, while, for the other hosts, such effects are reached either faster, at lower protein levels, or both. Such host-dependent fine-tuning of secreted effectors has also been suggested for microbial pathogens (Guyon et al. 2014). This hypothesis, however, contradicts the fact that tetranychid species such as *Panonychus* spp. and *T. evansi* do not have *SHOT1* and *SHOT2* genes while mainly feeding on nonfabacean hosts.

Our second hypothesis is that these particular SHOT1 and SHOT2 proteins are involved in the establishment of a compatible interaction with the subset of family Fabaceae, on the one hand, but are detrimental for the establishment of a compatible interaction with non-Fabaceae spp. on the other. In that case, downregulation may be needed primarily to limit the ecological costs of the interaction. In aphids, for example, it is known that effectors can take effect in an opposite manner in different host plants (Kettles and Kaloshian 2016; Pitino and Hogenhout 2013; Rodriguez et al. 2017; Thorpe et al. 2016). For example, expression of the *Macrosiphum euphorbiae* effector Me47 in *Nicotiana benthamiana* enhanced reproductive performance of *Myzus persicae*, while delivery of Me47 into *Arabidopsis thaliana* reduced *M. persicae* reproductive performance (Kettles and Kaloshian 2016). In view of this second hypothesis, the high expression of *SHOT* genes when feeding on some Fabaceae spp. exclusively and their concomitant downregulation on the other plant species could indicate these other hosts respond to the SHOTs in an unfavorable manner, for example, by initiating ETI. However, ETI operates via specialized sensory proteins (called R proteins), which are usually species- or even cultivar-specific. Thus, one would expect SHOT-induced ETI to occur in a limited range of hosts—not in a broad range (Ercolano et al. 2012).

Our third hypothesis is that SHOT1 and SHOT2 proteins are involved in the establishment of a compatible interaction with a subset of family Fabaceae but do not affect the extent of compatibility with other hosts. For example, the SHOT1 and SHOT2 proteins could serve to interfere with family Fabaceae-specific recognition of spider mites and the subsequent defense activation. The transient downregulation observed after transfer from a fabacean (*Phaseolus vulgaris*) to a nonfabacean host (*S. lycopersicum*) is rather slow (Wybouw et al. 2015), which may suggest that the cost of expression may not be particularly high when feeding on non-Fabaceae spp. The downregulation in the mites feeding on non-Fabaceae spp. would most probably reflect a resource-saving mechanism. In view of this hypothesis, SHOT1 and SHOT2 proteins could thus be specialized for acting on a subset of family Fabaceae.

Conclusions.

A peptidomics analysis of spider mite saliva revealed that about 20% of all identified peptides mapped to a spider mite-specific protein family, which we named the SHOT family. SHOTs are encoded by genes specifically expressed in the

anterior podocephalic (salivary) glands of *T. urticae*. A subset of the SHOT family, SHOT1 and SHOT2, appeared to be secreted in a host-dependent manner. Furthermore, *SHOT1* and *SHOT2* genes were rapidly induced when mites were transferred to certain fabacean host plants. Together with the fact that the identity of *SHOT* members could be linked to the host plant repertoire in related Tetranychidae spp. points toward a host-specific function for these putative effectors. Interestingly, host transfers in the opposite direction (from *Phaseolus vulgaris* to *S. lycopersicum*) appear to result in a slow yet strong decrease of *SHOT1* and *SHOT2* gene expression, taking place over generations rather than hours. To our knowledge, SHOT1 and SHOT2 proteins are the first putative arthropod effectors for which asymmetric on and off regulation in a host plant-dependent manner, i.e., rapid upregulation on one host and slow downregulation on another has been demonstrated.

MATERIALS AND METHODS

Mite lines.

For peptidomics analyses, *T. urticae* mites (London strain) were transferred from bean (*Phaseolus vulgaris* cv. Prelude, family Fabaceae) to new host plants as described earlier (Jonckheere et al. 2016; Wybouw et al. 2015). In short, four replicate lines on soybean (*G. max* cv. Merlin, family Fabaceae) and tomato (*S. lycopersicum* cv. Moneymaker, family Solanaceae) were established. The spider mites were maintained on their host plants during five generations for *G. max* (Jonckheere et al. 2016) and for over 30 generations for *S. lycopersicum* (Wybouw et al. 2015).

Nano-LC-MS/MS analysis of secreted saliva.

Collection of spider mite saliva was done as described by Jonckheere et al. (2016). Briefly, mites on *Phaseolus vulgaris*, *G. max*, and *S. lycopersicum* were collected (mites originating from the same replicated host plant line were pooled) and were transferred to feeding arenas containing holidic artificial diet encapsulated in parafilm. The spider mites were allowed to feed for 24 h, whereupon the remaining diet, containing saliva, was collected. Samples were stored at -80°C until sufficient material (about 0.3 ml, 10 to 15 diet hemispheres per sample) was collected for nano-LC-MS/MS analysis. Samples (0.2 to 0.3 ml; two technical replicates of artificial diets enriched with saliva of mites maintained on three different host plant species) were filtered on a Millipore Amicon Ultra 0.5-ml 10K spin-down filter (Merck Chemicals). The flow-through containing the peptides was collected and a solid phase extraction was performed on a Pierce C18 spin column (ThermoFisher Scientific), according to the manufacturers protocol. The eluate was collected, was dried in a vacuum centrifuge, and was stored at -80°C until analysis. LC-MS analysis was performed on an Easy nLC 1000 (ThermoFisher Scientific) connected to a Q-Exactive plus Orbitrap mass spectrometer (ThermoFisher Scientific). The dried eluate was dissolved in 10 μl of 2% acetonitrile in high-pressure liquid chromatography-grade water and 5 μl was subsequently loaded on the trapping column (PepMap C18, 300 $\mu\text{m} \times 20$ mm, ThermoFisher Scientific) with an isocratic flow of 2% acetonitrile in water with 0.1% formic acid at a flow rate of 5 $\mu\text{l min}^{-1}$. After 2 min, the column-switching valve was switched, placing the precolumn online with the analytical capillary column, an Acclaim PepMap RSLC C18 (2 μm , 10.0 nm, 50 μm internal diameter $\times$ 15 cm, Thermo Scientific). Separation was conducted using a linear gradient from 2% acetonitrile in water, 0.1% formic acid to 40% acetonitrile in water, 0.1% formic acid, in 100 min. The flow rate was set at 300 nL min^{-1} . The Q-Exactive plus Orbitrap was set up in a data-dependent MS/MS mode in which a full scan spectrum (350 to 2,000 m/z , resolution 60,000)

was followed by a maximum of 20 higher energy collisional dissociation (HCD) tandem mass spectra (100 to 2,000 *m/z*). Peptide ions were selected as the 20 most intense peaks of the MS1 scan. The normalized collision energy used was 28% in HCD. We applied a dynamic exclusion list of 45 s.

Peptide identification.

Peak lists obtained from MS/MS spectra were identified using OMSSA version 2.1.9 (Geer et al. 2004), X!Tandem version X! Tandem Vengeance (2015.12.15.2) (Craig and Beavis 2004), and MS-GF+ version Beta (v10282) (Kim and Pevzner 2014). The search was conducted using SearchGUI version 3.2.9 (Vaudel et al. 2011). Protein identification was conducted against a concatenated target/decoy version of the *T. urticae* protein database holding 19,086 *T. urticae* target sequences (version of August 11, 2016, available at OrcAE [Grbić et al. 2011]) and the common cRAP database (116 sequences). The decoy sequences were created by reversing the target sequences in SearchGUI. The identification settings were as follows: no cleavage specificity; 10.0 ppm as MS1 and 10.0 ppm as MS2 tolerances; variable modifications: amidation of the peptide C-term (−0.984016 Da), pyroglutamine from Q (−17.026549 Da), oxidation of M (+15.994915 Da); variable modifications during refinement procedure: acetylation of protein N-term (+42.010565 Da), pyroglutamine from E (−18.010565 Da), pyroglutamine from Q (−17.026549 Da), pyroglutamine from carbamidomethylated C (−17.026549 Da). Peptide spectrum matches (PSMs) were inferred from the spectrum identification results, using PeptideShaker version 1.16.2 (Vaudel et al. 2015) PSMs were validated at a 1.0% false discovery rate (FDR) estimated using the decoy hit distribution. The MS data along with the identification results have been deposited to the ProteomeXchange Consortium (Vizcaino et al. 2014), via the PRIDE partner repository (Martens et al. 2005), with the dataset identifier PXD006385.

Peptidomics data filtering, quantification, and bioinformatics analyses.

First, for each LC-MS/MS replicate (two technical replicates for each host-specific sample) the modest Feature Finder (moFF) (Argentini et al. 2016) tool was used to extract the MS1 intensities from the Thermo raw files for the validated PSMs matching to nondecoy (true) sequences and not matching to cRAP sequences (Supplementary Table S7). moFF assigns an intensity to each PSM, based on the apex of the ion chromatogram that was extracted for each PSM (Argentini et al. 2016). The moFF program (moff.py) was run with the following settings: −tol 10 −rt_w 5 −rt_p 0.6. Those PSMs for which no intensities could be extracted (−1 values) were discarded from further analysis. Next, the normalized intensity value for each PSM was calculated by dividing the intensity value by the total sum of intensity values of all PSMs for each replicate. A MS1 intensity value was then calculated for each peptide by averaging the normalized intensity values across the different PSMs, matching with each peptide. Only those peptides with a MS1 intensity value in both technical replicates were used for further analysis, and the MS1 intensity for each peptide was averaged across the two technical replicates for each host plant-specific sample. Those proteins for which an averaged MS1 intensity across technical replicates could be derived for one of their peptides were subjected to several analyses. Proteins were first mined for conserved domains using CD-search (default E-value threshold E^{-2}) at the Conserved Domain database (Marchler-Bauer et al. 2011). Subsequently, they were screened (D cutoff of 0.450) for the presence of a SP, using SignalP 4.0 (Petersen et al. 2011), and their subcellular localization was predicted using YLoc-HighRes version ‘animals’ (Briesemeister

et al. 2010). The PPA tool was used to predict easily detectable peptides (peptides that are likely to provide the best detection sensitivity) from *T. urticae* proteins of which peptides were identified. *T. urticae* proteins without their SP were used as input for PPA, using default settings. Last, a previously published dataset of differentially expressed genes between dissected *T. urticae* proterosomes containing the salivary glands and intact females (Jonckheere et al. 2016) was mined for genes encoding proteins from which peptides were identified.

Homology searches and *in silico* expression analysis.

Homologous proteins were identified by tBLASTn searches (E-value threshold of E^{-5}) against the genome sequences of *T. urticae* (available at OrcAE [Grbić et al. 2011]), *T. evansi* and *T. lintearius* (unpublished data), and the transcriptomes of *Panonychus citri* and *Panonychus ulmi* (Bajda et al. 2015; Niu et al. 2012). In addition, BLASTp and tBLASTn searches (E-value threshold of E^{-5}) were performed against the non-redundant protein database and available Acari genomes at the NCBI website, respectively (date accessed: February 1, 2017). Protein sequence alignments were performed using the online version of MUSCLE (Edgar 2004) with 1,000 iterations. Model selection was done with ProtTest 3.4 (Darriba et al. 2011) and, according to the Akaike information criterion, WAG+G+F was optimal for the phylogenetic reconstruction. A maximum likelihood analysis was performed using RAXML v8 (Stamatakis 2014a) (Stamatakis 2014b) with 1,000 bootstrapping replicates. The resulting tree was midpoint-rooted and visualized using MEGA 6.0 (Tamura et al. 2013). Tree layout was edited with Adobe Illustrator software (Adobe Inc.). Expression levels of *T. urticae* genes were compared between feeding (mobile) *T. urticae* stages (i.e., larva, nymph, and adult) and the embryo stage (Grbić et al. 2011). Finally, an expression dataset of *T. urticae* genes across different timepoints of transfer to tomato (Wybouw et al. 2015) and in mite lines adapted to different host plants (Jonckheere et al. 2016) was mined.

Whole-mount *in situ* hybridization.

Primers for whole-mount ISH probe templates were designed using Primer3 (Rozen and Skaletsky 1999), with a preferred amplicon length of about 300 nucleotides. The ISH procedure was performed as described earlier (Jonckheere et al. 2016).

RT-qPCR.

Total RNA was isolated from 100 to 150 adult female spider mites, using the RNeasy mini kit (Qiagen) and DNA was removed using the Turbo-DNA-free kit (Thermo Fisher Scientific). Next, 2 µg of RNA was used for cDNA synthesis using the Maxima first-strand cDNA synthesis kit (Thermo Fisher Scientific). The RT-qPCR reactions were conducted with the Maxima SYBR Green qPCR master mix (Thermo Fisher Scientific), using a Mx3005P qPCR system (Agilent Technologies) combined with MxPro software (Agilent Technologies). For each experiment, three biological replicates were used, each analyzed in two technical replicates. *T. urticae* (London strain) adapted to and maintained on *Phaseolus vulgaris* served as the reference condition in all qPCR experiments. Final melt curves and no-template controls ensured specific amplification and the absence of contamination, respectively. The raw quantification cycle (Cq) values were analyzed using qbase+ (Biogazelle, Belgium). Cq values were normalized against the housekeeping genes *ribosomal protein 49* (*Rp49*, *tetur18g03590*) and *ubiquitin C* (*UBQ*, *tetur03g06910*). The mean amplification efficiency of each reference target (*Rp49* and *UBQ*) and each target of interest was determined based on standard curves made from dilution series of pooled cDNA and was considered as acceptable (ranging from 2.01 to 2.05) for RT-qPCR. Unpaired

Student's *t* tests were used for statistical analysis, without assuming a normalized distribution (qbase+).

Expression levels in *T. urticae* after long-term feeding on different hosts.

In a first set of RT-qPCR experiments, four lab strains were used for expression analysis of *SHOT* genes, i.e., the *T. urticae* London strain established on *Phaseolus vulgaris* and strains derived from this strain after long-term maintenance (>10 generations) on three other plant species, namely, *S. lycopersicum*, barrel medic (*Medicago truncatula*, ecotype A17, family Fabaceae), and cotton (*Gossypium hirsutum*, family Malvaceae). Furthermore, three wild *T. urticae* strains were collected (East Flanders province, Belgium in 2011) from cherry (*Prunus avium* cv. Kordia, family Rosaceae), elderberry (*Sambucus nigra*, family Adoxaceae), and cucumber (*Cucumis sativus*, family Cucurbitaceae) plants (Wybouw et al. 2012). The wild strains were maintained under lab conditions on their original host plant species until used for expression analysis.

Short-term expression modulation after a host transfer.

Host transfer experiments were performed by transferring mites collected from the *S. lycopersicum*-adapted strain to *Phaseolus vulgaris* plants (100 to 150 adult females, triplicate per timepoint). After 3, 6, 12, and 24 h, mites were collected again from the *Phaseolus vulgaris* plants for gene expression analysis. Mites maintained on *S. lycopersicum* were used for expression analysis at 0 h. *T. urticae* adapted to and maintained on *Phaseolus vulgaris* served as reference condition in each qPCR experiment. In a follow-up experiment, different host plants were used as acceptor plants for the mite strain adapted to *S. lycopersicum*. These host plants were i) sweet pea (*Lathyrus odoratus* cv. Dorothy Eckford, family Fabaceae), ii) pea (*Pisum sativum* cv. Ceresa, family Fabaceae), iii) *Mimosa* spp. (family Fabaceae), iv) gorse (*Ulex europaeus*, family Fabaceae), v) peanut (*Arachis hypogaea*, family Fabaceae), vi) lupin (*Lupinus angustifolius*, family Fabaceae), vii) soybean (*G. max* cv. Merlin, family Fabaceae), viii) maize (*Zea mays*, family Poaceae), ix) gerbera daisy (*Gerbera jamesonii*, family Asteraceae), x) thale cress (*Arabidopsis thaliana*, ecotype Col-0, family Brassicaceae), and xi) tobacco (*Nicotiana benthamiana*, family Solanaceae). The experimental setup was as described above, with the only exception that gene expression analysis was performed 24 h after transfer.

ACKNOWLEDGMENTS

The authors thank V. Grbić and M. Grbić for providing access to the unpublished genomes of *T. evansi* and *T. lintearius*. W. Dermauw and G. Menschaert are postdoctoral fellows of the Research Foundation Flanders (FWO, Belgium). This project was supported by NWO (The Netherlands Organization for Scientific Research) under the FACCE-JPI ERA-NET Plus framework (GENOMITE, project ID 137), the FWO (grant G009312N to L. Tirry and T. Van Leeuwen and grant G053815N to L. Tirry, T. Van Leeuwen, and W. Dermauw) and The Special Research Fund (DOZA 01J131711 to W. Jonckheere). M. R. Kant was supported by NWO (STW-VIDI 13492).

LITERATURE CITED

Afshar, K., Dufresne, P. J., Pan, L., Merkx-Jacques, M., and Bede, J. C. 2010. Diet-specific salivary gene expression and glucose oxidase activity in *Spodoptera exigua* (Lepidoptera: Noctuidae) larvae. *J. Insect Physiol.* 56:1798-1806.

Agrawal, A. A., Vala, F., and Sabelis, M. W. 2002. Induction of preference and performance after acclimation to novel hosts in a phytophagous spider mite: Adaptive plasticity? *Am. Nat.* 159:553-565.

Alba, J. M., Schimmel, B. C. J., Glas, J. J., Ataíde, L. M. S., Pappas, M. L., Villarroel, C. A., Schuurink, R. C., Sabelis, M. W., and Kant, M. R. 2015. Spider mites suppress tomato defenses downstream of jasmonate

and salicylate independently of hormonal crosstalk. *New Phytol.* 205: 828-840.

Ali, J. G., and Agrawal, A. A. 2012. Specialist versus generalist insect herbivores and plant defense. *Trends Plant Sci.* 17:293-302.

Ament, K., Kant, M. R., Sabelis, M. W., Haring, M. A., and Schuurink, R. C. 2004. Jasmonic acid is a key regulator of spider mite-induced volatile terpenoid and methyl salicylate emission in tomato. *Plant Physiol.* 135:2025-2037.

Argentini, A., Goeminne, L. J., Verheggen, K., Hulstaert, N., Staes, A., Clement, L., and Martens, L. 2016. moFF: A robust and automated approach to extract peptide ion intensities. *Nat. Methods* 13:964-966.

Ataíde, L. M. S., Pappas, M. L., Schimmel, B. C. J., Lopez-Orenes, A., Alba, J. M., Duarte, M. V. A., Pallini, A., Schuurink, R. C., and Kant, M. R. 2016. Induced plant-defenses suppress herbivore reproduction but also constrain predation of their offspring. *Plant Sci.* 252:300-310.

Bajda, S., Dermauw, W., Greenhalgh, R., Nauen, R., Tirry, L., Clark, R. M., and Van Leeuwen, T. 2015. Transcriptome profiling of a spiroticlofen susceptible and resistant strain of the European red mite *Panonychus ulmi* using strand-specific RNA-seq. *BMC Genomics* 16:974.

Barrett, L. G., and Heil, M. 2012. Unifying concepts and mechanisms in the specificity of plant-enemy interactions. *Trends Plant Sci.* 17:282-292.

Bensoussan, N., Santamaria, M. E., Zhurov, V., Diaz, I., Grbić, M., and Grbić, V. 2016. Plant-herbivore interaction: Dissection of the cellular pattern of *Tetranychus urticae* feeding on the host plant. *Front. Plant Sci.* 7:1105.

Briesemeister, S., Rahnenführer, J., and Kohlbacher, O. 2010. YLoc—An interpretable web server for predicting subcellular localization. *Nucleic Acids Res.* 38:W497-W502.

Carolan, J. C., Caragea, D., Reardon, K. T., Mutti, N. S., Dittmer, N., Pappan, K., Cui, F., Castaneto, M., Poulain, J., Dossat, C., Tagu, D., Reese, J. C., Reeck, G. R., Wilkinson, T. L., and Edwards, O. R. 2011. Predicted effector molecules in the salivary secretome of the pea aphid (*Acyrtosiphon pisum*): A dual transcriptomic/proteomic approach. *J. Proteome Res.* 10:1505-1518.

Codrea, M. C., and Nahnsen, S. 2016. Platforms and pipelines for proteomics data analysis and management. Pages 203-215 in: *Modern Proteomics—Sample Preparation, Analysis and Practical Applications.*, H. Mirzaei and M. Carrasco, eds. Springer, Cham, Switzerland.

Craig, R., and Beavis, R. C. 2004. TANDEM: Matching proteins with tandem mass spectra. *Bioinformatics* 20:1466-1467.

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. 2011. ProtTest 3: Fast selection of best-fit models of protein evolution. *Bioinformatics* 27: 1164-1165.

Deng, W., Long, Q., Zeng, J., Li, P., Yang, W., Chen, X., and Xie, J. 2017. *Mycobacterium tuberculosis* PE_PGRS41 enhances the intracellular survival of *M. smegmatis* within macrophages via blocking innate immunity and inhibition of host defense. *Sci. Rep.* 7:46716.

Dermauw, W., Osborne, E. J., Clark, R. M., Grbić, M., Tirry, L., and Van Leeuwen, T. 2013a. A burst of ABC genes in the genome of the polyphagous spider mite *Tetranychus urticae*. *BMC Genomics* 14:317.

Dermauw, W., Wybouw, N., Rombauts, S., Menten, B., Vontas, J., Grbić, M., Clark, R. M., Feyereisen, R., and Van Leeuwen, T. 2013b. A link between host plant adaptation and pesticide resistance in the polyphagous spider mite *Tetranychus urticae*. *Proc. Natl. Acad. Sci. U.S.A.* 110: E113-E122.

Deslandes, L., and Rivas, S. 2012. Catch me if you can: Bacterial effectors and plant targets. *Trends Plant Sci.* 17:644-655.

Díaz-Riquelme, J., Zhurov, V., Ríoja, C., Pérez-Moreno, I., Torres-Pérez, R., Grimplet, J., Carbonell-Bejerano, P., Bajda, S., Van Leeuwen, T., Martínez-Zapater, J. M., Grbić, M., and Grbić, V. 2016. Comparative genome-wide transcriptome analysis of *Vitis vinifera* responses to adapted and non-adapted strains of two-spotted spider mite, *Tetranychus urticae*. *BMC Genomics* 17:74.

Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792-1797.

Erb, M., Meldau, S., and Howe, G. A. 2012. Role of phytohormones in insect-specific plant reactions. *Trends Plant Sci.* 17:250-259.

Ercoleano, M. R., Sanseverino, W., Carli, P., Ferriello, F., and Frusciante, L. 2012. Genetic and genomic approaches for R-gene mediated disease resistance in tomato: Retrospects and prospects. *Plant Cell Rep.* 31: 973-985.

Eyres, I., Jaquière, J., Sugio, A., Duvaux, L., Gharbi, K., Zhou, J.-J., Legeai, F., Nelson, M., Simon, J.-C., Smadja, C. M., Butlin, R., and Ferrari, J. 2016. Differential gene expression according to race and host plant in the pea aphid. *Mol. Ecol.* 25:4197-4215.

Fletcher, J. I., Smith, R., O'Donoghue, S. I., Nilges, M., Connor, M., Howden, M. E., Christie, M. J., and King, G. F. 1997. The structure of a novel insecticidal neurotoxin, ω -atracotoxin-HV1, from the venom of an Australian funnel web spider. *Nat. Struct. Biol.* 4:559-566.

- Fox, L. R., and Morrow, P. A. 1981. Specialization: Species property or local phenomenon? *Science* 211:887-893.
- Geer, L. Y., Markey, S. P., Kowalak, J. A., Wagner, L., Xu, M., Maynard, D. M., Yang, X., Shi, W., and Bryant, S. H. 2004. Open mass spectrometry search algorithm. *J. Proteome Res.* 3:958-964.
- Godinho, D. P., Janssen, A., Dias, T., Cruz, C., and Magalhães, S. 2015. Down-regulation of plant defence in a resident spider mite species and its effect upon con-and heterospecifics. *Oecologia* 180:161-167.
- Gotoh, T., Bruin, J., Sabelis, M., and Menken, S. 1993. Host race formation in *Tetranychus urticae*: Genetic differentiation, host plant preference, and mate choice in a tomato and a cucumber strain. *Entomol. Exp. Appl.* 68:171-178.
- Grbić, M., Van Leeuwen, T., Clark, R. M., Rombauts, S., Rouzé, P., Grbić, V., Osborne, E. J., Dermauw, W., Ngoc, P. C. T., Ortego, F., Hernández-Crespo, P., Diaz, I., Martínez, M., Navajas, M., Sucena, E., Magalhães, S., Nagy, L., Pace, R. M., Djuranović, S., Smagghe, G., Iga, M., Christiaens, O., Veenstra, J. A., Ewer, J., Villalobos, R. M., Hutter, J. L., Hudson, S. D., Velez, M., Yi, S. V., Zeng, J., Pires-daSilva, A., Roch, F., Cazaux, M., Navarro, M., Zhurov, V., Acevedo, G., Bjelica, A., Fawcett, J. A., Bonnet, E., Martens, C., Baele, G., Wissler, L., Sanchez-Rodriguez, A., Tirry, L., Blais, C., Demeestere, K., Henz, S. R., Gregory, T. R., Mathieu, J., Verdon, L., Farinelli, L., Schmutz, J., Lindquist, E., Feyereisen, R., and Van de Peer, Y. 2011. The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. *Nature* 479:487-492.
- Guyon, K., Balagué, C., Roby, D., and Raffaele, S. 2014. Secretome analysis reveals effector candidates associated with broad host range necrotrophy in the fungal plant pathogen *Sclerotinia sclerotiorum*. *BMC Genomics* 15:336.
- Hill, R., and O'Donnell, D. 1991. The host range of *Tetranychus lintearius* (Acarina: Tetranychidae). *Exp. Appl. Acarol.* 11:253-269.
- Hogenhout, S. A., and Bos, J. I. 2011. Effector proteins that modulate plant-insect interactions. *Curr. Opin. Plant Biol.* 14:422-428.
- Howe, G. A., and Jander, G. 2008. Plant immunity to insect herbivores. *Annu. Rev. Plant Biol.* 59:41-66.
- Ireson, J., Gourlay, A., Kwong, R., Holloway, R., and Chatterton, W. 2003. Host specificity, release, and establishment of the gorse spider mite, *Tetranychus lintearius* Dufour (Acarina: Tetranychidae), for the biological control of gorse, *Ulex europaeus* L. (Fabaceae), in Australia. *Biol. Control* 26:117-127.
- Jing, S., Zhao, Y., Du, B., Chen, R., Zhu, L., and He, G. 2017. Genomics of interaction between the brown planthopper and rice. *Curr. Opin. Insect Sci.* 19:82-87.
- Jonckheere, W., Dermauw, W., Zhurov, V., Wybouw, N., Van den Bulcke, J., Villarroel, C. A., Greenhalgh, R., Grbić, M., Schuurink, R. C., Tirry, L., Baggerman, G., Clark, R. M., Kant, M. R., Vanholme, B., Menschaert, G., and Van Leeuwen, T. 2016. The salivary protein repertoire of the polyphagous spider mite *Tetranychus urticae*: A quest for effectors. *Mol. Cell. Proteomics* 15:3594-3613.
- Jones, J. D., and Dangl, J. L. 2006. The plant immune system. *Nature* 444:323-329.
- Kaloshian, I., and Walling, L. L. 2016. Hemipteran and dipteran pests: Effectors and plant host immune regulators. *J. Integr. Plant Biol.* 58:350-361.
- Kant, M. R., Jonckheere, W., Knecht, B., Lemos, F., Liu, J., Schimmel, B. C. J., Villarroel, C. A., Ataíde, L. M. S., Dermauw, W., Glas, J. J., Egas, M., Janssen, A., Van Leeuwen, T., Schuurink, R. C., Sabelis, M. W., and Alba, J. M. 2015. Mechanisms and ecological consequences of plant defence induction and suppression in herbivore communities. *Ann. Bot. (Lond.)* 115:1015-1051.
- Kant, M. R., Sabelis, M. W., Haring, M. A., and Schuurink, R. C. 2008. Intraspecific variation in a generalist herbivore accounts for differential induction and impact of host plant defences. *Proc. Biol. Sci.* 275:443-452.
- Kettles, G. J., and Kaloshian, I. 2016. The potato aphid salivary effector Me47 is a glutathione-S-transferase involved in modifying plant responses to aphid infestation. *Front. Plant Sci.* 7:1142.
- Kim, S., and Pevzner, P. A. 2014. MS-GF+ makes progress towards a universal database search tool for proteomics. *Nat. Commun.* 5:5277.
- Lapin, D., and Van den Ackerveken, G. 2013. Susceptibility to plant disease: More than a failure of host immunity. *Trends Plant Sci.* 18:546-554.
- Li, C., Williams, M. M., Loh, Y.-T., Lee, G. I., and Howe, G. A. 2002. Resistance of cultivated tomato to cell content-feeding herbivores is regulated by the octadecanoid-signaling pathway. *Plant Physiol.* 130:494-503.
- Li, L., Zhao, Y., McCaig, B. C., Wingerd, B. A., Wang, J., Whalon, M. E., Pichersky, E., and Howe, G. A. 2004. The tomato homolog of CORONATINE-INSENSITIVE1 is required for the maternal control of seed maturation, jasmonate-signaled defense responses, and glandular trichome development. *Plant Cell* 16:126-143.
- Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S., and Kahmann, R. 2015. Fungal effectors and plant susceptibility. *Annu. Rev. Plant Biol.* 66:513-545.
- Lu, H., Yang, P., Xu, Y., Luo, L., Zhu, J., Cui, N., Kang, L., and Cui, F. 2016. Performances of survival, feeding behavior, and gene expression in aphids reveal their different fitness to host alteration. *Sci. Rep.* 6:19344.
- Magalhães, S., Fayard, J., Janssen, A., Carbonell, D., and Olivieri, I. 2007. Adaptation in a spider mite population after long-term evolution on a single host plant. *J. Evol. Biol.* 20:2016-2027.
- Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwatz, M., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Lu, F., Marchler, G. H., Mullokandov, M., Omelchenko, M. V., Robertson, C. L., Song, J. S., Thanki, N., Yamashita, R. A., Zhang, D., Zhang, N., Zheng, C., and Bryant, S. H. 2011. CDD: A conserved domain database for the functional annotation of proteins. *Nucleic Acids Res.* 39:D225-D229.
- Martens, L., Hermjakob, H., Jones, P., Adamski, M., Taylor, C., States, D., Gevaert, K., Vandekerckhove, J., and Apweiler, R. 2005. PRIDE: The proteomics identifications database. *Proteomics* 5:3537-3545.
- Mitchum, M. G., Hussey, R. S., Baum, T. J., Wang, X., Elling, A. A., Wubben, M., and Davis, E. L. 2013. Nematode effector proteins: An emerging paradigm of parasitism. *New Phytol.* 199:879-894.
- Mitchum, M. G., Wang, X., Wang, J., and Davis, E. L. 2012. Role of nematode peptides and other small molecules in plant parasitism. *Annu. Rev. Phytopathol.* 50:175-195.
- Mondal, H. A. 2017. Shaping the understanding of saliva-derived effectors towards aphid colony proliferation in host plant. *J. Plant Biol.* 60:103-115.
- Muntel, J., Boswell, S. A., Tang, S., Ahmed, S., Wapinski, I., Foley, G., Steen, H., and Springer, M. 2015. Abundance-based classifier for the prediction of mass spectrometric peptide detectability upon enrichment (PPA). *Mol. Cell. Proteomics* 14:430-440.
- Musser, R. O., Hum-Musser, S. M., Eichenseer, H., Peiffer, M., Ervin, G., Murphy, J. B., and Felton, G. W. 2002. Herbivory: Caterpillar saliva beats plant defences. *Nature* 416:599-600.
- Mutti, N. S., Louis, J., Pappan, L. K., Pappan, K., Begum, K., Chen, M.-S., Park, Y., Dittmer, N., Marshall, J., Reese, J. C., and Reeck, G. R. 2008. A protein from the salivary glands of the pea aphid, *Acyrtosiphon pisum*, is essential in feeding on a host plant. *Proc. Natl. Acad. Sci. U.S.A.* 105:9965-9969.
- Navajas, M. 1998. Host plant associations in the spider mite *Tetranychus urticae* (Acari: Tetranychidae): Insights from molecular phylogeography. *Exp. Appl. Acarol.* 22:201-214.
- Navajas, M., de Moraes, G. J., Auger, P., and Migeon, A. 2013. Review of the invasion of *Tetranychus evansi*: Biology, colonization pathways, potential expansion and prospects for biological control. *Exp. Appl. Acarol.* 59:43-65.
- Ngoc, P. C. T., Greenhalgh, R., Dermauw, W., Rombauts, S., Bajda, S., Zhurov, V., Grbić, M., Van de Peer, Y., Van Leeuwen, T., Rouzé, P., and Clark, R. M. 2016. Complex evolutionary dynamics of massively expanded chemosensory receptor families in an extreme generalist chelicerate herbivore. *Genome Biol. Evol.* 8:3323-3339.
- Niu, J. Z., Dou, W., Ding, T. B., Shen, G. M., Zhang, K., Smagghe, G., and Wang, J. J. 2012. Transcriptome analysis of the citrus red mite, *Panonychus citri*, and its gene expression by exposure to insecticide/acaricide. *Insect Mol. Biol.* 21:422-436.
- Norambuena, H., Martínez, G., Carrillo, R., and Neira, M. 2007. Host specificity and establishment of *Tetranychus lintearius* (Acari: Tetranychidae) for biological control of gorse, *Ulex europaeus* (Fabaceae) in Chile. *Biol. Control* 40:204-212.
- Nosil, P. 2002. Transition rates between specialization and generalization in phytophagous insects. *Evolution* 56:1701-1706.
- Pan, Y., Zhu, J., Luo, L., Kang, L., and Cui, F. 2015. High expression of a unique aphid protein in the salivary glands of *Acyrtosiphon pisum*. *Physiol. Mol. Plant Pathol.* 92:175-180.
- Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. 2011. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* 8:785-786.
- Pieterse, C. M., Van der Does, D., Zamioudis, C., Leon-Reyes, A., and Van Wees, S. C. 2012. Hormonal modulation of plant immunity. *Annu. Rev. Cell Dev. Biol.* 28:489-521.
- Pitino, M., and Hogenhout, S. A. 2013. Aphid protein effectors promote aphid colonization in a plant species-specific manner. *Mol. Plant-Microbe Interact* 26:130-139.

- Podlaha, O., and Zhang, J. 2010. Pseudogenes and their evolution. eLS. Published online.
- Rodriguez, P., Escudero-Martinez, C., and Bos, J. 2017. An aphid effector targets trafficking protein VPS52 in a host-specific manner to promote virulence. *Plant Physiol.* 173:1892-1903.
- Rozen, S., and Skaletsky, H. 1999. Primer3 on the WWW for general users and for biologist programmers. Pages 365-386 in: *Bioinformatics Methods and Protocols*, Vol. 132. Humana Press, Totowa, NJ, U.S.A.
- Santamaría, M. E., Hernández-Crespo, P., Ortego, F., Grbic, V., Grbic, M., Diaz, I., and Martinez, M. 2012. Cysteine peptidases and their inhibitors in *Tetranychus urticae*: A comparative genomic approach. *BMC Genomics* 13:307.
- Sarmento, R. A., Lemos, F., Bleeker, P. M., Schuurink, R. C., Pallini, A., Oliveira, M. G. A., Lima, E. R., Kant, M., Sabelis, M. W., and Janssen, A. 2011. A herbivore that manipulates plant defence. *Ecol. Lett.* 14: 229-236.
- Schimmel, B. C. J., Ataide, L. M. S., Chafi, R., Villarroel, C. A., Alba, J. M., Schuurink, R. C., and Kant, M. R. 2017. Overcompensation of herbivore reproduction through hyper-suppression of plant defenses in response to competition. *New Phytol.* 214:1688-1701.
- Siepen, J. A., Keevil, E. J., Knight, D., and Hubbard, S. J. 2007. Prediction of missed cleavage sites in tryptic peptides aids protein identification in proteomics. *J. Proteome Res.* 6:399-408.
- Stamatakis, A. 2014a. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312-1313.
- Stamatakis, A. 2014b. The RAxML v8. 2. X Manual. Heidelberg Institute for Theoretical Studies, Heidelberg, Germany.
- Steen, H., and Mann, M. 2004. The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* 5:699-711.
- Stuart, J. J., Chen, M.-S., Shukle, R., and Harris, M. O. 2012. Gall midges (Hessian flies) as plant pathogens. *Annu. Rev. Phytopathol.* 50:339-357.
- Takeyama, K., Mori, N., and Osakabe, M. 2006. Effect of cytochrome P450 inhibitor, piperonyl butoxide, on survival of *Panonychus citri* (McGregor) (Acari: Tetranychidae) on citrus leaves. *Appl. Entomol. Zool. (Jpn.)* 41: 487-491.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A., and Kumar, S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30:2725-2729.
- Thorpe, P., Cock, P. J. A., and Bos, J. 2016. Comparative transcriptomics and proteomics of three different aphid species identifies core and diverse effector sets. *BMC Genomics* 17:172.
- Tian, C., and Jian-Ping, X. 2010. Roles of PE_PGRS family in *Mycobacterium tuberculosis* pathogenesis and novel measures against tuberculosis. *Microb. Pathog.* 49:311-314.
- van den Boom, C. E. M., van Beek, T. A., and Dicke, M. 2003. Differences among plant species in acceptance by the spider mite *Tetranychus urticae* Koch. *J. Appl. Entomol.* 127:177-183.
- Van Leeuwen, T., and Dermauw, W. 2016. The molecular evolution of xenobiotic metabolism and resistance in chelicerate mites. *Annu. Rev. Entomol.* 61:475-498.
- Vandermoten, S., Harmel, N., Mazzucchelli, G., De Pauw, E., Haubruge, E., and Francis, F. 2014. Comparative analyses of salivary proteins from three aphid species. *Insect Mol. Biol.* 23:67-77.
- Vaudel, M., Barsnes, H., Berven, F. S., Sickmann, A., and Martens, L. 2011. SearchGUI: An open-source graphical user interface for simultaneous OMSSA and X!Tandem searches. *Proteomics* 11:996-999.
- Vaudel, M., Burkhart, J. M., Zahedi, R. P., Oveland, E., Berven, F. S., Sickmann, A., Martens, L., and Barsnes, H. 2015. PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nat. Biotechnol.* 33: 22-24.
- Villarroel, C. A., Jonckheere, W., Alba, J. M., Glas, J. J., Dermauw, W., Haring, M. A., Van Leeuwen, T., Schuurink, R. C., and Kant, M. R. 2016. Salivary proteins of spider mites suppress defenses in *Nicotiana benthamiana* and promote mite reproduction. *Plant J.* 86:119-131.
- Vizcaíno, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Ríos, D., Dienes, J. A., Sun, Z., Farrah, T., Bandeira, N., Binz, P. A., Xenarios, I., Eisenacher, M., Mayer, G., Gatto, L., Campos, A., Chalkley, R. J., Kraus, H. J., Albar, J. P., Martinez-Bartolomé, S., Apweiler, R., Omenn, G. S., Martens, L., Jones, A. R., and Hermjakob, H. 2014. ProteomeX-change provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* 32:223-226.
- Will, T., Tjallingii, W. F., Thönnessen, A., and van Bel, A. J. 2007. Molecular sabotage of plant defense by aphid saliva. *Proc. Natl. Acad. Sci. U.S.A.* 104:10536-10541.
- Win, J., Chaparro-Garcia, A., Belhaj, K., Saunders, D., Yoshida, K., Dong, S., Schornack, S., Zipfel, C., Robatzek, S., and Hogenhout, S. 2012. Effector biology of plant-associated organisms: concepts and perspectives. Pages 235-247 in: *Cold Spring Harbor Symposia on Quantitative Biology*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, U.S.A.
- Wink, M. 2013. Evolution of secondary metabolites in legumes (Fabaceae). *S. Afr. J. Bot.* 89:164-175.
- Wu, J., and Baldwin, I. T. 2010. New insights into plant responses to the attack from insect herbivores. *Annu. Rev. Genet.* 44:1-24.
- Wybouw, N., Balabanidou, V., Ballhorn, D. J., Dermauw, W., Grbić, M., Vontas, J., and Van Leeuwen, T. 2012. A horizontally transferred cyanase gene in the spider mite *Tetranychus urticae* is involved in cyanate metabolism and is differentially expressed upon host plant change. *Insect Biochem. Mol. Biol.* 42:881-889.
- Wybouw, N., Zhurov, V., Martel, C., Bruinsma, K. A., Hendrickx, F., Grbić, V., and Van Leeuwen, T. 2015. Adaptation of a polyphagous herbivore to a novel host plant extensively shapes the transcriptome of herbivore and host. *Mol. Ecol.* 24:4647-4663.
- Yamaguchi, Y., and Huffaker, A. 2011. Endogenous peptide elicitors in higher plants. *Curr. Opin. Plant Biol.* 14:351-357.
- Yano, S., Wakabayashi, M., Takabayashi, J., and Takafuji, A. 1998. Factors determining the host plant range of the phytophagous mite, *Tetranychus urticae* (Acari: Tetranychidae): A method for quantifying host plant acceptance. *Exp. Appl. Acarol.* 22:595-601.
- Yeruva, V. C., Kulkarni, A., Khandelwal, R., Sharma, Y., and Raghunand, T. R. 2016. The PE_PGRS proteins of *Mycobacterium tuberculosis* are Ca²⁺ binding mediators of host-pathogen interaction. *Biochemistry* 55: 4675-4687.
- Zhao, C., Escalante, L. N., Chen, H., Benatti, T. R., Qu, J., Chellapilla, S., Waterhouse, R. M., Wheeler, D., Andersson, M. N., Bao, R., Batterton, M., Behura, S. K., Blankenburg, K. P., Caragea, D., Carolan, J. C., Coyle, M., El-Bouhssini, M., Francisco, L., Friedrich, M., Gill, N., Grace, T., Grimmelikhuijzen, C. J., Han, Y., Hauser, F., Herndon, N., Holder, M., Ioannidis, P., Jackson, L., Javaid, M., Jhangiani, S. N., Johnson, A. J., Kalra, D., Korchina, V., Kovar, C. L., Lara, F., Lee, S. L., Liu, X., Löfstedt, C., Mata, R., Mathew, T., Muzny, D. M., Nagar, S., Nazareth, L. V., Okwuonu, G., Onger, F., Perales, L., Peterson, B. F., Pu, L. L., Robertson, H. M., Schermerhorn, B. J., Scherer, S. E., Shreve, J. T., Simmons, D., Subramanyam, S., Thornton, R. L., Xue, K., Weissenberger, G. M., Williams, C. E., Worley, K. C., Zhu, D., Zhu, Y., Harris, M. O., Shukle, R. H., Werren, J. H., Zdobnov, E. M., Chen, M. S., Brown, S. J., Stuart, J. J., and Richards, S. 2015. A massive expansion of effector genes underlies gall-formation in the wheat pest *Mayetiola destructor*. *Curr. Biol.* 25:613-620.

AUTHOR-RECOMMENDED INTERNET RESOURCES

Repository of Adventitious Proteins (cRAP) database:
<http://www.thegpm.org/crap>
 OrcaE (Online Resource for Community Annotation of Eukaryotes):
<http://bioinformatics.psb.ugent.be/orcae/overview/Tetur>
 Spider Mites Web: <http://www1.montpellier.inra.fr/cbgrp/spmweb>