The salivary proteome of *Tetranychus urticae*

*Key to its polyphagous nature?*

Jonckheere, W.S.A.

**Publication date**
2018

**Document Version**
Other version

**License**
Other

**Citation for published version (APA):**

**General rights**
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
A putative salivary effector family of *Tetranychus urticae* shows host plant dependent expression

Partly redrafted after:

**Host plant dependent expression of salivary proteins**

5.0. **Abstract**

The salivary protein repertoire released by the herbivorous pest *Tetranychus urticae* is assumed to hold key to its success on diverse crops. We report on a spider mite-specific protein family which is expanded in *T. urticae*. The encoding genes have an expression pattern restricted to the anterior podocephalic glands while peptide fragments were found in *T. urticae’s* secretome, supporting the salivary nature of these proteins. As peptide fragments were identified in a host-dependent manner, we designated this family as the ‘Secreted HOst-responsive protein of Tetranychidae’ (SHOT) family. The proteins were divided in three groups based on sequence similarity. Unlike *TuSHOT3* genes, *TuSHOT1 & 2* genes were highly expressed when feeding on a subset of Fabaceae, while expression was depleted on other hosts. *TuSHOT1 & 2* expression was induced within 24 hours after certain host transfers, pointing towards transcriptional plasticity rather than selection as the cause. Transfer from an ‘inducer’ to a ‘non-inducer’ plant was associated with slow yet strong downregulation of *TuSHOT1 & 2*, occurring over generations rather than hours. This asymmetric on/off regulation points towards host-specific effects of SHOT proteins, which is further supported by the diversity of SHOT genes identified in Tetranychidae with a distinct host repertoire. Preliminary experiments for SHOT functional validation were performed, including yeast two-hybrid studies, which revealed the interaction between *TuSHOT2b* and MUT9-like kinases in *Arabidopsis*. 
5.1. Introduction

Being sessile organisms, plants use a wide range of strategies to ward off arthropod herbivores. These include constitutive chemical and physical defenses which discourage feeding and are present irrespective of the presence of herbivores, as well as induced defenses which are specifically activated upon attack (Howe and Jander, 2008; Wu and Baldwin, 2010). The latter include the production of toxic and anti-digestive compounds, but also the attraction of natural enemies of the herbivore via induced volatiles. Furthermore, resources can be re-allocated towards non-attacked 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 have evolved host-specific adaptations to deal with specific defenses (Nosil, 2002; Ali and Agrawal, 2012; Barrett and Heil, 2012; Kant et al., 2015). On the other end of the diet breadth spectrum are the polyphagous herbivorous 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 1000 different plant species belonging to more than 100 different families (Migeon and Dorkeld, 2006-2017; Grbić et al., 2011). 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 (Grbić et al., 2011; Santamaría et al., 2012; Wybouw et al., 2012; Dermauw et al., 2013a,b; Van Leeuwen and Dermauw, 2016). 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 intraspecific genetic variability (Magalhaes et al., 2007) and it has been proposed that local populations form host races that do not perform equally well on all potential host plant species (Gotoh et al., 1993; Navajas, 1998; Agrawal et al., 2002; Díaz-Riquelme et al., 2016).

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

Wybouw et al., 2015) were shown to suppress these defenses thereby increasing their fitness (Ataide et al., 2016).

Recognition of an invading organism is key for the plant to initiate a successful defense response. Dodging such recognition has been well studied for plant-pathogen interaction, 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’ (DAMPs), 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 herbivorespecific molecules, known as ‘pathogen-’ or ‘herbivore-associated molecular patterns’ (PAMPs or HAMPs), also proposed to be detected by PRRs, are leading to PAMP/HAMP-triggered immunity (PTI or HTI) (Jones and Dangl 2006, Erb et al., 2012, Pieterse et al., 2012). However, pathogens and herbivores evolved so-called effector molecules (often proteins) that enable them to evade PTI/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 (ETS). 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/HTI. This in its turn drives the evolution of alternative effectors that yet again can suppress host resistance (Jones and Dangl, 2006; Barrett and Heil, 2012; Erb et al., 2012; Win et al., 2012; Lapin and Van den Ackerveken, 2013). While the molecular details of this so-called ‘Zigzag model’ have been worked out for pathogens, it most likely applies for herbivores as well (e.g., Jing et al., 2017; Mondal, 2017). However, the extent of the parallels is still a matter of intensive research (Kaloshian and Walling, 2015).

As the saliva of arthropod herbivores forms an important source of molecules that could trigger or manipulate plant defense responses (e.g., Musser et al., 2002; Mutti et al., 2008; Hogenhout and Bos, 2011), 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 (Villarroel 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, Mitchum et al., 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 (SHOT) appears to be unique for spider mites, is expanded in *T. urticae* compared to close relatives and the corresponding genes show rapid changes in host plant dependent expression. A next step was to investigate SHOT protein function. Yeast two-hybrid (Y2H) screens were performed to identify their host interaction partners, yielding an insight in the cellular processes they target.

### 5.2. Material and Methods

#### 5.2.1. Mite lines

For peptidomics analyses, *T. urticae* mites (London strain) were transferred from bean (*Phaseolus vulgaris* cv. ‘Prelude’, Fabaceae) to new host plants as described earlier (Wybouw et al., 2015; Jonckheere et al., 2016 (Chapter 3)). In short, four replicate lines on soy bean (*Glycine max* cv. ‘Merlin’, Fabaceae) and tomato (*Solanum lycopersicum* cv. ‘Moneymaker’, Solanaceae) were established. The spider mites were maintained on their host plants during five generations for *G. max* (Jonckheere et al., 2016 (Chapter 3)), and for over 30 generations for *S. lycopersicum* (Wybouw et al., 2015).

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

Collection of spider mite saliva was done as in Jonckheere et al. (2016) (Chapter 3). Briefly, mites on *P. vulgaris*, *G. max* and *S. lycopersicum* were
collected (mites originating from the same replicated host plant line were pooled) and transferred to feeding arenas containing holidic artificial diet encapsulated in Parafilm. The spider mites were allowed to feed for 24 hours, whereupon the remaining diet, containing saliva, was collected. Samples were stored at -80°C until sufficient material (about 0.3 ml, 10–15 diet hemispheres/sample) was collected for nano-LC-MS/MS analysis. Samples (0.2–0.3 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 spindown filter (Merck Chemicals, Belgium). The flow-through, containing the peptides, was collected and a solid phase extraction was performed on a Pierce C18 spin column (ThermoFisher Scientific, Belgium), according to the manufacturers protocol. The eluate was collected, dried in a vacuum centrifuge and stored at -80°C until analysis. Liquid chromatography mass spectrometric (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 HPLC-grade water and 5 μl was subsequently loaded on the trapping column (PepMap C18, 300 μm x 20 mm, ThermoFisher Scientific) with an isocratic flow of 2% acetonitrile in water with 0.1% formic acid at a flow rate of 5 μl min⁻¹. After 2 min, the column-switching valve was switched, placing the pre-column online with the analytical capillary column, an Acclaim PepMap RSLC C18 (2 μm, 100 Å, 50 μm i.d. x 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 where a full scan spectrum (350–2000 m/z, resolution 60,000) was followed by a maximum of twenty HCD tandem mass spectra (100 to 2000 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.

5.2.3. 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)
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 11th 2016) and the common Repository of Adventitious Proteins (cRAP) database (available at http://www.thegpm.org/crap/; 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), Pyrolidone from Q (-17.026549 Da), Oxidation of M (+15.994915 Da), variable modifications during refinement procedure: Acetylation of protein N-term (+42.010565 Da), Pyroli done from E (-18.010565 Da), Pyrolidone from Q (-17.026549 Da), Pyroli done 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 mass spectrometry 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.

### 5.2.4 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 non-decoy (true) and not matching to the cRAP sequences (see Table S5.1). 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 analy-
sis, 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 signal peptide (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 Peptide Prediction with Abundance (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. *Tetranychus urticae* proteins without their SP were used as input for PPA, using default settings. Lastly, a previously published dataset of differentially expressed genes between dissected *T. urticae* proterosoma’s – containing the salivary glands – and intact females (Jonckheere et al., 2016) (Chapter 3) was mined for genes encoding proteins from which peptides were identified.

5.2.5. 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 http://bioinformatics.psb.ugent.be/orcae/overview/Tetur; Grbić et al., 2011), *T. evansi* and *T. lintearius* (unpublished data) and the transcriptomes of *Panonychus citri* and *P. ulmi* (Niu et al., 2012; Bajda et al., 2015). 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: 1st of February 2017). Protein sequence alignments were performed using the online version of MUSCLE (Edgar 2004) with 1000 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, 2014) with 1000 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., USA). 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 time points of transfer to tomato (Wybouw et al., 2015) and in mite lines adapted to different host plants (Jonckheere et al., 2016) was mined.

5.2.6. Whole-mount *in situ* hybridization

Primers for whole-mount *in situ* hybridization (ISH) probe templates were designed using Primer3 (Rozen and Skaletsky, 1999), with a preferred amplicon length of about 300 nucleotides (TABLE S5.2). The ISH procedure was performed as described earlier (Jonckheere et al., 2016) (Chapters 2 and 3).

5.2.7. RT-qPCR

Total RNA was isolated from 100-150 adult female spider mites using the RNeasy Mini Kit (Qiagen, Belgium) and DNA was removed using the Turbo-DNA-free Kit (Thermo Fisher Scientific, Belgium). 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, Belgium) combined with MxPro Software (Agilent Technologies). For each experiment, three biological replicates were used, each analyzed in two technical replicates. *Tetranychus urticae* (London strain) adapted to and maintained on *P. vulgaris* served as the reference condition in all qPCR experiments. The primers used for qPCR are listed in TABLE S5.2. 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+).
5.2.8. 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: the *T. urticae* London strain established on *P. vulgaris* and strains derived from this strain after long term maintenance (>10 generations) on three other plant species, being *S. lycopersicum*, barrel medic (*Medicago truncatula*, ecotype A17, Fabaceae) and cotton (*Gossypium hirsutum*, Malvaceae). Furthermore, three wild *T. urticae* strains were collected (Belgium, East Flanders province, 2011) from cherry (*Prunus avium* cv. ‘Kordia’, Rosaceae), elderberry (*Sambucus nigra*, Adoxaceae) and cucumber (*Cucumis sativus*, 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.

5.2.9. Short term expression modulation after a host transfer

Host transfer experiments were performed by transferring mites collected from the *S. lycopersicum* adapted strain to *P. vulgaris* plants (100-150 adult females, triplicate per time point). After 3, 6, 12 and 24 h, mites were collected again from the *P. vulgaris* plants for gene expression analysis. Mites maintained on *S. lycopersicum* were used for expression analysis at 0 hours. *Tetranychus urticae* adapted to and maintained on *P. 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: (1) sweet pea (*Lathyrus odoratus* cv. ‘Dorothy Eckford’, Fabaceae), (2) pea (*Pisum sativum* cv. ‘Ceresa’, Fabaceae), (3) *Mimosa* spp. (Fabaceae), (4) gorse (*Ulex europaeus*, Fabaceae), (5) peanut (*Arachis hypogaea*, Fabaceae), (6) lupin (*Lupinus angustifolius*, Fabaceae), (7) soybean (*Glycine max* cv. ‘Merlin’, Fabaceae), (8) maize (*Zea mays*, Poaceae), (9) gerbera daisy (*Gerbera jamesonii*, Asteraceae), (10) thale cress (*Arabidopsis thaliana*, ecotype Col-0, Brassicaceae) and (11) tobacco (*Nicotiana benthamiana*, Solanaceae). The experimental setup was as described above, with the only exception that gene expression analysis was performed 24 h after transfer.

5.2.10. Yeast two-hybrid screens of *SHOT* proteins

Three different yeast two-hybrid (Y2H) screens were performed to elucidate putative plant targets of spider mite *SHOT* proteins. The *SHOT* pro-
teins, lacking their native SP, were used as a ‘bait’ for interacting plant ‘prey’ proteins. (1) For a SHOT1 & 2 member of *T. urticae* (*TuSHOT2b*), an *A. thaliana* seedling (1 week old plants) cDNA library was used. (2) For a SHOT3 member of *T. urticae* (*TuSHOT3b*), a cDNA library from mite-infested *S. lycopersicum* leaves was screened. Both of these ‘ULTImate Y2H SCREEN’ analyses were performed by Hybrigenics Services (Paris, France). Additionally, (3) a SHOT3 member of *T. evansi* (*TeSHOT3b*) was submitted to a Y2H screen, using a tomato library. This screen was performed by Keygene (Wageningen, the Netherlands).

### 5.2.11. Evaluation of SHOT effector action in mite performance assays

*Tetur03g03680* (SHOT1b) and *tetur11g01360* (SHOT3a) were amplified without native SP. Att-B gateway recombination sequences and a simple Kozak sequence with initiation codon (CACCAGT) were added during two successive PCR reactions (gene specific primers, and primers with overlap sequences, Table 5.1), taking care to adjust for the reading frame. Next, PCR products were cloned into vector pDONR221 (Gateway system, ThermoFisher Scientific) and the genes were eventually transferred to the binary vector pK7WGF2 (Karimi et al., 2002). In this vector, the gene of interest is under the control of a CaMV 35S promotor and positioned downstream of eGFP. The destination vectors were transformed into the GV3101 strain of *Agrobacterium tumefaciens* by electroporation.

#### Table 5.1. Primers used for amplification of SHOT vector inserts

<table>
<thead>
<tr>
<th></th>
<th>Forward primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tetur11g01360_PCR1</td>
<td>gatgaacaacccgaactttttta aacg</td>
<td>ttattgtatatgtgagtctacc gaag</td>
</tr>
<tr>
<td>Tetur03g03690_PCR1</td>
<td>aatgaacaatcaccgcacttcg</td>
<td>ttaaccgaacatcttaccagcgac</td>
</tr>
<tr>
<td>Tetur11g01360_PCR2</td>
<td>AAAAAGCAGGCTTCA CCATGgatgaacaacccgaac</td>
<td>CAAGAAAGCTGGG GATTTatattgtatgttggag</td>
</tr>
<tr>
<td>Tetur03g03690_PCR2</td>
<td>AAAAAGCAGGCTTCA CCATGaatgaacaatcacc</td>
<td>CAAGAAAGCTGGG TTtaaccgaacatcttacc</td>
</tr>
<tr>
<td>Attb_PCR3</td>
<td>GGGGACAAGTTTGTGTA CAAAAGACAGGCT</td>
<td>GGGGACAAGTTTGTGTA CAAAAGACAGGCT</td>
</tr>
</tbody>
</table>

*Italicized* letter T: nucleotide to correct for reading frame

**Bold:** kozak with start codon

Fragments with lower-case nucleotides overlap with each other in the next PCR, as well as underlined nucleotide fragments.
For the *A. tumefaciens* transient transformation assays (ATTAs), agrobacteria of strain GV3101 harboring the SHOT1b:pK7WGF2 or SHOT3a:pK7WGF2 vectors, empty pK7WGF2 vector, or pBIN61 vector carrying the silencing inhibitor p19 (Voinnet et al., 2003) were grown at 28°C in 2 ml of YEB medium supplemented with the appropriate antibiotics. After 24 h, 25 ml YEB medium was inoculated with 50 μl pre-culture supplemented with the appropriate antibiotics, acetosyringone and MES. Cultures were grown for about 24 h until the optical density at 600 nm (OD$_{600}$) reached 0.6. The bacteria were subsequently harvested by centrifuging for 15 minutes at 5000 g. Next, cells were re-suspended in MMAi medium (2% sucrose, 10mM MES, 0.2 mM acetosyringone) to a final OD$_{600}$ of 0.6 and incubated at room temperature for 2 hours.

The agrobacteria harboring the SHOT1b:pK7WGF2, SHOT3a: pK7WGF2 or empty pK7WGF2 vector were co-infiltrated at a 1:1 ratio with the 35S-p19 agrobacteria into the abaxial side of *N. benthamiana* leaves using a needleless syringe. Two days after infiltration, the leaves were cut from the plants and the trichomes on the adaxial side were removed by rubbing with wet paper tissue. Next, leaf disks (diameter of 20 mm) were punched out of the leaves and 60 disks transformed with each construct (eGFP-SHOT1b, eGFP-SHOT3a or eGFP) were distributed, adaxial side up, on wet cotton wool (semi-randomly divided over several plates). One female adult mite (*T. urticae*, London strain on *P. vulgaris*), 2-4 h since turning adult, was added to each leaf disk. Disks with mites were placed in an incubator (Panasonic MLR-352H) at 26°C, 16:8 L:D and 60% RH. The eGFP fluorescence of additional control leaf discs was monitored to ascertain transgene expression throughout the experiment. Eggs were counted (blind) 3 days after transfer of mites to the leaf disks. Data analysis of egg number was performed in Rstudio (Team 2015) using the Kruskal-Wallis rank sum test. Kruscalmc test, using R package pgirmess, was used to determine significant differences among treatments.

5.2.12. Transient expression of *SHOT* genes in *S. lycopersicum* and the effect on JA and SA levels

*Agrobacteria* (strain ID1249s) harboring SHOT1b:pK7WGF2, SHOT3c:pK7WGF2 or the empty pK7WGF2 vector were infiltrated in *S. lycopersicum* (5 weeks old) leaflets using a needleless syringe. For each plant, three leaflets were infiltrated. Additional control plants remained untreated.
or were infiltrated with *Agrobacterium*-free infiltration medium. Per time point and per treatment, four plants were used for hormone data collection, while one extra plant per treatment was used for evaluation of eGFP-expression.

After two and five days, leaflets were collected (without petiolule), pooled per plant as one biological replicate, flash frozen in liquid nitrogen and stored at -80°C. Next, plant hormone analysis was performed using the procedure of Wu et al. (2007), with some minor modifications as described by Alba et al. (2015) (see also Chapter 2). Phytohormone titers were compared across treatments per time point using Anova with Tukey HSD test (R-studio; Team 2015).

### 5.3. Results

#### 5.3.1. Peptidomics analysis of *T. urticae* saliva secreted into artificial diet

*T. urticae* lines propagated for at least 5 generations on different host plants (*P. vulgaris*, *G. max* and *S. lycopersicum*) 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-LC-MS/MS. The detected salivary peptides from all host plant adapted mite samples were pooled into one dataset, retaining host specific information. In total, 279 peptides were identified, of which 216, 116 and 133 were found in the *P. vulgaris*, *G. max* and *S. lycopersicum* samples, respectively (Figure 5.1, Table S5.3). Identified peptides typically had an arginine (R) or lysine (K) at their C-terminus (210/279 peptides), or were preceded by an R or K in the sequence of the proteins from which they originate (51/279). Taking into account 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 peptides mapped to a total of 40 different proteins distributed over the *P. 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) (Chapter 3). 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,
Host plant dependent expression of salivary proteins

A

S. lycopersicum
P. vulgaris
G. max
1−2
2−3
3−4
4−5
not detected

-log10(normalized intensity)

SHOT1
SHOT2
SHOT3

B

NEQSPENLQGNVHGR*
NEQSPENLQSNVHGR*
NEESPEPNLQSNVHGR*
NEQSPENLQSNVHGR**
NEQSPENLQSNVHGR**
GIGGVANQADAEIK
GIGGVANQADAEIK
IGSAADKKFDALSGAVGK
AGMAESGKAQKADAGK
LASAGNVAAGAVFESNAMK
LASAGNVAAGAVFESNAMK
EQSPELNSQNHGR**
NEQSPENLQSNVHGR**
IASNVVSSVASAGK
ANMASAGKAGQKMAN
AGEMANSGKAAAGDAAG
NEQSPENLQSNVHGR**
EESPEPNLQSNVHGR**
NEQSPDFDLQSNVHGR*
NEQSPDFDLQSNVHGR**
EESPEPNLQSNVHGR**
NEQSPDFDLQSNVHGR**
SPDFDLQSNVHGR**
VVCNEQSPDFDLQSNVHGR**
VVCNEQSPDFDLQSNVHGR**
PDFSLSQNSVHGR**
VADVTQSGVQVAGK
VADVTQSGVQVAGK
GQADVSSQINGVK
VAGVQDAVTSVQWQAGK
AKGQGVDSQKTNGLGK
VQMDGQKAVGSGGKV
GQGNSQKTNGLGK
VQMDGQKAVGSGGKV
VAGVQAVTQSVQWQAGK
VAGVQAVTQSVQWQAGK
VAGVQAVTQSVQWQAGK
AFLAVVAGK
AFLAVVAGK
GQADVSSQINGVK
VAGVQAVTQSVQWQAGK
AKGQGVDSQKTNGLGK
VQMDGQKAVGSGGKV
GQGNSQKTNGLGK
VQMDGQKAVGSGGKV
VAGVQAVTQSVQWQAGK
VAGVQAVTQSVQWQAGK
VAGVQAVTQSVQWQAGK
AFLAVVAGK
AFLAVVAGK
AFLAVVAGK
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 (266/279 peptides).

With the exception of tetur18g00530, all proteins from which peptides were identified, had a predicted SP for secretion and 34 of the 40 proteins (85%) were predicted to be targeted towards the extracellular space (TABLE S5.4). In addition, most of the corresponding genes (33/40) were, based on an RNA-seq expression analysis, significantly upregulated in the proterosoma – the body region which harbors the salivary glands – relative to the entire mite body (Benjamini-Hochberg adjusted \( p \leq 0.05 \) and FC \( \geq 8 \)) (TABLE S5.4). Together these data support the salivary nature of the identified proteins.

About 50% (149/279) of the identified peptides mapped to only two protein groups (tetur29g01360; tetur07g00160 & tetur07g01660) that were present in all host specific samples (TABLES S5.3 and S5.4). 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* (FIGURE 5.1, TABLE S5.4). Remarkably, about 20% (49/279) of the identified peptides mapped to a group of related proteins (tetur03g03620, tetur03g03670, tetur03g03680, tetur03g03690, tetur03g03700, tetur03g03730, tetur03g10083, tetur03g10093, tetur11g01360, tetur11g06390 and tetur11g06400). These

**FIGURE 5.1.** Peptides identified in artificial diet probed by *Tetranychus urticae* lines adapted to different host plants. (A) Venn diagram depicting the overlap between *T. urticae* peptides identified in secretions from mites adapted to *G. max*, *P. vulgaris* and *S. lycopersicum*. The number of unique proteins from which these peptides could originate is shown between brackets. All but one (41/42) peptides originating from SHOT1 & 2 proteins were identified in the *G. max* or *P. vulgaris* samples exclusively (shaded in gray). (B) Heatmap of normalized intensity (\(-\log_{10}\) transformed) of identified peptides from SHOTs. The Euclidean distance metric and Ward’s method were used for clustering of columns. SHOT1 and SHOT2 peptide groups were sorted based on decreasing order of intensity values of the *P. vulgaris* sample while those of the SHOT3 peptide group were sorted based on decreasing order of intensity values of the *S. lycopersicum* sample. Those SHOT1 & 2 peptides that were *in silico* predicted (PPA; Muntel et al., 2015) to have the highest confidence to be detected are marked with one asterisk. Those peptides marked with two asterisks represent variants of these SHOT1 & 2 peptides.
proteins and homologs thereof are here designated to the ‘SHOT’ (Secreted Host-responsive protein in Tetranychidae) family.

5.3.2. *T. urticae* SHOT family analysis

tBLASTn searches resulted in the identification of 13 SHOT genes in the *T. urticae* genome, while four and two were found in the draft genomes of *T. lintearius* and *T. evansi*, respectively (TABLE S5.5, DATA S1). The *P. ulmi* and *P. citri* transcriptomes harbored one and three SHOT transcript(s) respectively (contig 00033 [*P. ulmi*; Bajda et al. (2015); contig 183, 22851 and 29686, *P. citri*; Bajda et al. (2015); see DATA S1]). BLASTp and tBLASTn searches against the non-redundant protein database (E-value < E^-5) and Acari genomes, respectively, using intact *Tetranychidae* SHOT members as query did not yield hits with proteins of non-tetranychid species (TABLE S5.5).

Genes lacking a proper start codon or containing a frameshift (*T. urticae*: 1, *T. lintearius*: 2, *P. citri*: 2) were considered as pseudogenes/fragments and were not included for further analysis. A maximum-likelihood phylogenetic analysis of SHOT protein sequences (FIGURE 5.2) revealed that the SHOT family could be divided into three different clusters: a cluster containing tetur03g03620, tetur03g03680, tetur03g03700, tetur03g03740, tetur03g10083 and *Tl* shot1a (referred to as SHOT1), a cluster containing tetur03g03670, tetur03g03690, tetur03g03690, tetur03g03730 tetur03g10093 and *Tl* shot2a (referred to as SHOT2), and a third cluster containing tetur11g01360, tetur11g06390, tetur11g06400, *Tl* shot3a, *Tl* shot3b, *P. ulmi* contig 00033 and *P. 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 a 11 kb region of scaffold 11 (671,365 bp – 682,501 bp), while *T. urticae* SHOT1 & 2 genes are in close proximity of each other on scaffold 3 (1,642,274 bp – 1,703,045 bp). Except for tetur03g03620, *T. urticae* SHOT1 & 2 genes occur as consecutive pairs (tandem duplications) of one SHOT1 gene and a neighboring SHOT2 gene that is located on the opposite strand (1,682,835bp – 1,703,045 bp).

All intact SHOT members have a predicted SP for secretion (TABLE S5.5) and are enriched in glycine, alanine and serine (on average 17.7, 17.6 and 11.1%, respectively, TABLE S5.5). In contrast to *Panonychus* SHOT members, *Tetranychus* SHOTs possess either a proline-glutamic acid, proline-glutamine or proline-aspartate motif at the N-terminus (FIGURE S5.1). The average length of SHOT1, SHOT2 and SHOT3 proteins is 324, 275 and 319 amino
acids, respectively. For those SHOT genes of which expression data of developmental stages was available, the average expression of intact *T. urticae* SHOT genes across all the feeding stages (larva, nymph, adult) relative to the expression in the embryo stage was 4- to 18-fold higher than the overall genome-wide average of this expression ratio (average genome-wide ratio of 2.2; Chapter 3; Jonckheere et al., 2016). Except for *tetu03g03620*, all intact SHOT genes were significantly upregulated (fold changes ranging from 17 to 40) in the proterosoma relative to the entire mite body (Table S5.5). Interestingly, the peptides originating from SHOT1 & 2 proteins were almost uniquely identified in either the *P. vulgaris* or *G. max* samples (41/42) (Figure 5.1A), whereas four out of seven SHOT3 peptides were only found

**FIGURE 5.2.** 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* (*TlSHOT*), *P. citri* (*PcSHOT*) and *P. 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 1000 bootstrap replicates. (B) (Top) Heatmap of cyanine intensities of putative *T. urticae* SHOT genes in different host plant adapted mite lines (*P. vulgaris*, *G. max*, *S. lycopersicum*, *G. hirsutum*, and *Z. mays*; Jonckheere et al., 2016). (Bottom) Heatmap of log2FCs of SHOT genes in *T. urticae* after host shifts between *P. vulgaris* and *S. lycopersicum* (B_24hT, mites from the London reference strain on *P. vulgaris* transferred to *S. lycopersicum* for 24h; B_30GT, mites from the *P. vulgaris* strain grown on *S. lycopersicum* for 30 generations; B_30GT_2GT, mites from the *P. vulgaris* strain grown on *S. lycopersicum* for 30 generations and transferred back to *P. vulgaris* for two generations (Wybouw et al., 2015).
Host plant dependent expression of salivary proteins

in *S. lycopersicum* samples. SHOT1 & 2 peptides flanking the predicted signal peptide (with aa sequences ‘NEQSPELNLQGNVHGR’, ‘NEESPELNLQSNVHGR’, ‘NEQSPELNLQSNVHGR’ and ‘NEQSPDFDLQSNVHGR’, see Figure S5.1) were among the top ten of peptides with the highest MS1 intensity in both the *P. vulgaris* and *G. max* samples, while a similar peptide of SHOT3 proteins (‘DEQPELSLNGKVHGR’) was among the top ten of peptides with the highest MS1 intensity in the *S. lycopersicum* sample (Figure 5.1B, Table S5.3). This indicates that the expression of the SHOT genes is host plant dependent.

5.3.3 Confirmation of salivary gland-specific expression using whole-mount in situ hybridization

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/TuSHOT2b, tetur03g03700/TuSHOT1c, tetur11g01360/TuSHOT3a*). The genes were shown to be expressed in the anterior podocephalic glands of *T. urticae* (Figure 5.3). Simultaneously performed ISH experiments with sense control probes did not show localized expression, supporting specific staining of the anti-sense probes.

**Figure 5.3. Expression localization of Tetranychus urticae SHOT genes.** Whole-mount *in situ* hybridization reveals specific expression of SHOT genes in the anterior podocephalic glands of *T. urticae*. ISH reaction product is visible as a red signal on the confocal fluorescence images, while the mite body emits green autofluorescence. 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) Lateral view of mite from ISH experiment using an antisense *SHOT3a* probe. (B) Dorsal view of mite from ISH experiment using an antisense *SHOT2b* probe. Scale bar in each panel represents 100 μm.
5.3.4. SHOT gene expression after long term feeding on different hosts

The expression level of *SHOT* genes of *T. urticae* feeding on different hosts was studied using RT-qPCR. Due to strong gene sequence similarity within the *SHOT* family, three non-specific primers sets were used during qPCR analysis, each amplifying a group of *SHOT* genes. 

- **SHOT1_q_Fw/Rv** primers amplifying a region of *tetu03g03680, tetu03g03700, tetu03g03740* and *tetu03g10083* (belonging to the *SHOT1* cluster),
- **SHOT2_q_Fw/Rv** primers amplifying *tetu03g03670, tetu03g03690, tetu03g03730* and *tetu03g10093* (belonging to the *SHOT2* cluster) and
- **SHOT3_q_Fw/Rv** primers amplifying a region of *tetu03g03680, tetu03g03700, tetu03g03740* and *tetu03g10083* (belonging to the *SHOT1* cluster),

![Figure 5.4](image-url)

**Figure 5.4.** Relative transcript abundance of *SHOT* gene groups in *Tetranychus urticae* strains adapted to different host plants. The different host-adapted mite strains originally derived from the *T. urticae* London strain, except for the three wild strains, which are marked with a degree symbol (°). Log2 transformed data are presented as gene expression levels in the different host-adapted *T. urticae* strains, relative to gene expression levels in *T. urticae* adapted to *P. vulgaris*. Three qPCR primer pairs amplified the three *SHOT* transcript groups (Table S5.1): *SHOT1, SHOT2* and *SHOT3*. 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 *P. vulgaris* (*p* ≤ 0.05; **p** ≤ 0.01; ***p*** ≤ 0.001; ****p*** ≤ 0.0001).
amplifying tetur11g01360, tetur11g06390 and tetur11g06400 (belonging to the SHOT3 cluster) (TABLE S5.2). Tetranychus urticae strains maintained on P. avium, C. sativus and S. nigra showed a highly significant reduced expression of SHOT1 & 2 genes, compared to the reference T. urticae London strain adapted to P. vulgaris (FIGURE 5.4). 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 C. sativus to a log2FC of -13.93 (95% CI: -13.21 to -14.66) for the SHOT1 gene group of mites adapted to P. avium. In absolute numbers, this means expression dropped 568 to 15,608 fold compared to the expression in mites maintained on P. vulgaris. A similar observation was made for T. urticae (London strain) adapted to S. lycopersicum and G. hirsutum. Also in this case SHOT1 & 2 genes of mites feeding on these two host plants were expressed at extremely reduced levels. However, mites adapted to the fabacean host M. truncatula had SHOT1 & 2 gene expression levels which were significantly higher than in the reference mite strain maintained on P. vulgaris. In short, when adapted to P. vulgaris or M. truncatula, T. urticae does express SHOT1 & 2 genes at high levels, while these genes are virtually non-expressed when the mites where maintained on P. avium, G. hirsutum, C. sativum, S. lycopersicum or S. nigra. The expression levels of the SHOT3 genes were slightly but significantly upregulated when feeding on all investigated host plants relative to P. vulgaris, except for M. truncatula (FIGURE 5.4). The variation in expression level for SHOT3 genes was, however, less pronounced than the variation observed for the SHOT1 & 2 genes. For example, the SHOT3 genes of mites adapted to P. avium were most strongly upregulated, with an associated log2FC of 3.18 (95% CI: 1.96 to 4.40).

5.3.5. SHOT gene expression during a 24 h time window after host transfer

Preliminary experiments showed that SHOT1 & 2 gene expression levels could quickly increase when mites were transferred from ‘non-inducer’ plants such as S. lycopersicum to ‘inducer’ plants such as P. vulgaris. Expressional changes after such host transfer were studied in closer detail over a 24 h time window. Tetranychus urticae mites adapted to S. lycopersicum were transferred to P. 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 *P. vulgaris* served as reference condition. The expression ratios are represented in Figure 5.5. While the expression ratios of *SHOT1* & 2 genes of *T. urticae* adapted to *S. lycopersicum* were extremely low compared to the *P. vulgaris* reference line (mean log₂FC of -12.35 at 0 h), mites transferred to *P. 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 lev-

![Figure 5.5](image_url)

**Figure 5.5.** Relative transcript abundance of *SHOT* gene groups in *T. urticae* transferred from *S. lycopersicum* to *P. vulgaris* for an increasing period of time. Mites originating from a strain adapted to *S. lycopersicum* were transferred to *P. 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 which had been transferred to *P. vulgaris* for x h, relative to gene expression levels in *T. urticae* adapted to and maintained on *P. vulgaris*. Three qPCR primer pairs amplified the three *SHOT* transcript groups (Table S5.1): *SHOT1*, *SHOT2* and *SHOT3*. 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 *P. vulgaris* (*p*≤0.05; **p**≤0.01; ***p**≤0.001; ****p**≤0.0001).
Host plant dependent expression of salivary proteins

eels increased further until, after 24 h on P. vulgaris, gene expression levels equaled or exceeded the expression level of reference mites adapted to and maintained on P. vulgaris. Expression of SHOT3 genes of the S. lycopersicum-derived mites was significantly higher than the P. vulgaris reference strain at all time points after transfer. As 24 h appeared to be sufficient for T. urticae to fully induce SHOT1 & 2 gene expression after host transfer, this time scale was used for additional short term expression analyses.

Figure 5.6. Relative transcript abundance of SHOT gene groups in Tetranychus urticae transferred to different host plants for 24 h. Mites from the strain adapted to S. lycopersicum were transferred to a new host for 24 h. Fabacean plant species (legumes) are marked above the bar chart, and include several Papilionoideae (Fabae and non-Fabae) and one Mimosoideae species. Log2 transformed data are presented as gene expression levels in T. urticae maintained on the new host for 24 h, relative to gene expression levels in T. urticae adapted to and maintained on P. vulgaris. Three qPCR primer pairs amplified the three SHOT transcript groups (Table S5.1): SHOT1, SHOT2 and SHOT3. Error bars show the 95% confidence interval (CI) while asterisks indicate the significance level of the expression ratio of the mites on the new host relative to the strain on P. vulgaris (*p≤0.05; **p≤0.01; ***p≤0.001; ****p≤0.0001).
5.3.6. **SHOT gene expression modulation after short term transfer to various hosts**

*T. urticae* mites from the *S. lycopersicum* adapted strain were transferred to new host plants, and 24 h after transfer the expression levels of **SHOT** genes were evaluated (Figure 5.6). The **SHOT3** genes were expressed at a slightly, yet significant, higher level (mean ± SEM log2FC over all host plants: 1.61 ± 0.37) when mites fed on all tested host plants, relative to the *P. vulgaris* adapted strain. For the **SHOT1 & 2** genes, different responses were noticeable. Genes of these subfamilies were either significantly downregulated (*T. urticae* transferred to *L. odoratus*, *Mimosa* spp., *A. thaliana*, *G. jamesonii*, *N. benthamiana* and *Z. mays*), or highly expressed (*T. urticae* transferred to *A. hypogaea*, *G. max*, *L. angustifolius* and *U. europaeus*). The mean relative expression levels of **SHOT1 & 2** genes of mites transferred to *P. sativum* were much lower than the reference strain (mean log2FC: -10.17), yet this was bordering significance due to high variation between biological replicates. The experiment with *P. sativum* was therefore repeated, yet with the same outcome (data not shown).

5.3.7. **Yeast two-hybrid screens of **SHOT** proteins**

To elucidate **SHOT** host plant interaction targets, three yeast two-hybrid screens were performed. (1) *TuSHOT2b* was found to interact with very high confidence with *A. thaliana* MUT9-like kinases (MLK), encoded by the genes *AT5G18190* (MLK1), *AT3G03940* (MLK2) and *AT3G13670* (MLK4). (2) For the *TuSHOT3b* Y2H screens, only one interacting *S. lycopersicum* protein had a good confidence score (Solyc07g053740.1.1, Ethylene-responsive transcription factor 4). However, this 'prey' was also shown to interact with the control validation bait, and therefore most likely represents a false positive interaction partner. In addition, two prey proteins showed a moderate confidence of interaction (Solyc07g053740.1.1, Storekeeper protein; and Solyc10g051340.1.1, Adenylyl cyclase-associated protein). (3) *TeSHOT3b* did interact with diverse *S. lycopersicum* proteins such as a LRR receptor-like serine/threonine-protein kinase, an isochorismatase hydrolase, a nodulin-like protein and phytoalexin deficient 4.

5.3.8. **Evaluation of **SHOT** effector action in mite performance assays**

Transient expression of **SHOT1b** and **SHOT3a** eGFP-fusion proteins in the non-leguminous plant *N. benthamiana* was performed, while the number
Host plant dependent expression of salivary proteins

of eggs produced during three days was used as a spider mite performance index (preliminary data) (FIGURE 5.7). The performance of \textit{T. urticae} feeding on \textit{N. benthamiana} leaf disks expressing eGFP-SHOT1b was significantly

![Figure 5.7](image-url)

**FIGURE 5.7.** Spider mite reproductive performance on \textit{Nicotiana benthamiana} transiently expressing \textit{SHOT} genes. The figure shows the average number of eggs laid by one adult \textit{T. urticae} female after three days on \textit{N. benthamiana} leaf discs producing SHOT-eGFP fusion proteins or eGFP as control (preliminary data). The error bars represent the standard error (SE), while different letters indicate statistical differences according to Kruskal-Wallis rank sum test ($p<0.05$).

![Figure 5.8](image-url)

**FIGURE 5.8.** Transient expression of \textit{TuSHOT} genes and the effect on phytohormones in \textit{Solanum lycopersicum} leaf discs. (A) Levels of jasmonic acid (JA) at 2 and 5 dpi. The error bars represent the standard error (SE) of the mean. No statistical differences (Tukey HSD test, $p<0.05$) could be found between different treatments at 2 dpi. There were significant differences between samples at 5 dpi, yet not between leaf discs producing eGFP-SHOT3a, eGFP-SHOT1b or eGFP. (B) Levels of salicylic acid (SA) at 2 and 5 dpi. The error bars represent the SE of the mean. No significant differences (Tukey HSD test, $p<0.05$) could be found between leaf discs producing eGFP-SHOT3a, eGFP-SHOT1b or eGFP at any time point. MC= medium control, UC= untreated control.
lower \((p = 0.03)\) than for mites feeding on eGFP expressing controls. Feeding on eGFP-SHOT3a transformed leaf disks did not result in a significantly altered performance \((p>0.05)\). Ectopic protein expression was evident from fluorescence microscopy investigation.

5.3.9. Transient expression of SHOT genes in \(S. lycopersicum\) and the effect on JA and SA levels

The JA and SA levels measured in transiently transformed \(S. lycopersicum\) leaflets is shown in Figure 5.8. Infiltration of \textit{Agrobacterium}-free medium (medium control, MC) did not lead to an augmented JA level, while it did result in a significant increase in the SA levels at both time points \((p<0.0001\) at 2 dpi; \(p<0.0001\) at 5 dpi). Infiltration with control \textit{Agrobacteria}, producing eGFP, resulted in an increase of JA levels at 5 dpi \((p = 0.014)\), but not yet at 2dpi. SA levels, however, were increased at both time points \((p<0.0000001\) at 2 dpi; \(p<0.00001\) at 5 dpi). No statistical difference in JA or SA levels between leaflets producing eGPF or salivary fusion proteins eGFP-SHOT1b or eGFP-SHOT3a could be observed at any time point.

5.4. DISCUSSION

5.4.1. Peptidomics analysis of spider mite saliva

The \textit{T. urticae} salivary proteome was previously analyzed by shotgun proteomics, where a tryptic digest was performed on the secreted salivary protein mixture prior to nano-LC-MS/MS analysis (Jonckheere et al., 2016) (Chapter 3). 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 in Jonckheere et al. (2016) (Chapter 3) but the samples were enriched in protein fragments/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 (see 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 tetur18g00530, all peptides that were identified in this study mapped to \textit{T. urticae} proteins longer than 100 aa and almost all peptides showed signs of tryptic cleavage, indicating that
Host plant dependent expression of salivary proteins

The *T. urticae* peptides identified in the artificial diet mainly comprise tryp tic 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 tryp tic proteases most likely digested the secreted proteins in the saliva or artificial diet. Strikingly, a previous proteomics analysis (Jonckheere et al., 2016) (Chapter 3) and, to a lesser extent, current study (see Table S5.2) 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.

5.4.2. 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 as SHOT family. Genes encoding SHOTs were considerably higher expressed in the salivary gland region (proterosoma), compared to the entire spider mite body (Jonckheere et al., 2016) (Chapter 3) and 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 (Figure 5.3). 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 & 2 group were almost exclusively found in artificial diet enriched with saliva of mites adapted to the legume plants *P. vulgaris* or *G. max* (Figure 5.1.A and B), whereas SHOT3 peptides were identified in diet after feeding of all the host plant-adapted mite lines (Figure 5.1.B). Remarkably, N-terminal peptides of mature SHOTs had a very high MS1 intensity compared to peptides matching to other regions of mature SHOTs (Figure 5.1, Table S5.3). 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 (PPA analysis, see Table S5.7), and their high intensity is probably due to their specific physicochemical properties.
5.4.3. The SHOT expression can be rapidly induced upon host transfer

Gene-expression analysis revealed that genes encoding SHOT1 & 2 proteins were highly expressed in *T. urticae* lines adapted to *P. vulgaris* or *G. max*, while on average being expressed 7-60x lower in lines adapted to *S. lycopersicum, G. hirsutum* or *Z. mays*. SHOT3 genes, on the other hand, did not show such extreme variation in expression levels (Figure 5.2.B) (Jonckheere et al., 2016). As both *P. vulgaris* and *G. max* belong to the Papilionoideae superfamily within the Fabaceae (legumes), the phytochemistry of both plants is quite comparable, relative to the chemistry of *S. lycopersicum, G. hirsutum* and *Z. mays*. Hence, the underlying phytochemistry might be reflected in a similar salivary protein repertoire when feeding either from *P. vulgaris* or *G. max*. The *T. urticae* lines used by Jonckheere et al. (2016) (Chapters 3 & 4) and used in this study for peptidomics analysis, were all derived from the London strain, of which the genome has been sequenced and annotated (Grbić et al., 2011).

To further survey host-dependent SHOT expression, we established additional adapted *T. urticae* lines derived from the London strain (this study). SHOT1 & 2 expression was found to be high in mites adapted to the legume host *M. truncatula*, yet expression was extremely low when feeding on the tested non-legume hosts. Wild *T. urticae* strains feeding on non-legume hosts were studied as well, and SHOT1 & 2 expression was always found to be very low relative to the London strain adapted to *P. vulgaris* (Figure 5.4). To test whether the host plant-dependent expression of the SHOT1 & 2 genes was the result of adaptation (between generations) or transcriptional plasticity (within a single generation), we performed a time course experiment. During this experiment, mites were transferred from *S. lycopersicum*, a host on which expression of SHOT1 & 2 genes was very low after adaptation, to the reference host *P. vulgaris*, on which expression of these genes is a 1000-fold higher. SHOT1 & 2 gene transcript levels were found to quickly increase during this time course experiment, and expression was equally high in mites transferred to *P. vulgaris* for 24 h as in mites that had been adapted to this plant for many generations. Since 24 h sufficed to fully induce SHOT1 & 2 expression, selection could be excluded as cause for this differential expression. In an additional set of experiments, mites were transferred from the ‘non-inducer’ *S. lycopersicum* to a number of other hosts for 24 h. Combining all host-dependent gene expression data, it is clear...
that the \textit{SHOT} 1 \& 2 genes are highly expressed when \textit{T. urticae} is feeding on a subset of legume hosts, more specifically 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 'non-inducer' host) is associated with a relatively slow downregulation of \textit{SHOT} 1 \& 2 genes. Indeed, when mites adapted to \textit{P. vulgaris} were transferred to \textit{S. lycopersicum} for 24 h (FIGURE 5.2.A) (Wybouw et al., 2015), \textit{SHOT} 1 \& 2 expression levels barely changed compared to mites adapted to \textit{P. vulgaris}. However, 30 generations after this transfer, members of the \textit{SHOT} 1 \& 2 groups were strongly downregulated (FIGURE 5.2.A). Such slow downregulation could be determined by the timing of food digestion and/or excretion, as inducing plant compounds may be detained in the mite body. On the other hand, it may suggest that \textit{SHOT} 1 \& 2 upregulation when feeding on a subset of Papilionoideae hosts is much more important to the mite than downregulation on non-Papilionoideae hosts and point towards a Papilionoideae-specific plant trait, such as a metabolite that triggers the 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 \textit{SHOT} 1 \& 2 suggests that a sensory system may be involved, allowing \textit{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 \textit{T. urticae} is well-developed and consists of an extraordinary large number of chemosensory receptors in lineage-specific proliferations (Ngoc et al., 2016). A potential factor determining the \textit{SHOT} 1 \& 2 induction in \textit{T. urticae} may be secondary metabolites. Isoflavonoids, for example, are predominantly found in the subfamily of Papilionoideae (Wink, 2013) including those used in this study, but non-Fabeae Papilionoideae specific isoflavonoids have not yet been reported. On the other hand, \textit{Mimosa}, \textit{Lathyrus} and \textit{Pisum} possess simple indoles, while all other tested Fabaceae hosts contain pyrrolizidines (Wink and Mohamed, 2003; Wink, 2013). The synthesis of these pyrrolizidine alkaloids is, however, not restricted to the Fabaceae (Ober and Kaltenegger, 2009) and does not explain the \textit{SHOT} 1 \& 2 gene expression differences between non-Fabeae Papilionoideae and all tested other plants. \textit{Tetranychus urticae} feeding
on *Mimosa* spp., *L. odoratus* and *P. sativum* did result in variable *SHOT1* & 2 expression levels between biological replicates (see the rather large 95% CI, FIGURE 5.6), 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 towards the plant traits responsible for the variability in *SHOT1* & 2 gene expression.

Host-dependent expression levels of salivary genes have also been described in aphids and caterpillars (Afshar et al., 2010; Pan et al., 2015; Eyres et al., 2016; Lu et al., 2016). *ACYPI006346* expression, for example, differed between *Acyrthosiphon 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 host plants, 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.

### 5.4.4. The diversity of *SHOT* genes in Tetranychidae is linked to host plant range

Despite the polyphagous nature of *T. urticae, P. 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 most species on which *T. urticae* has been recorded is the Fabaceae family (recorded on 119 Fabaceae species) (Migeon and Dorkeld, 2006-2017). The diversification of *SHOTs* into multi-member subfamilies (FIGURE 5.2.A) in *T. urticae* may be causally linked to the expanded host range of this species with its apparent preference for legumes. The distribution of intact homologs across the related spider mites *T. lintearius, T. evansi, P. citri* and *P. ulmi* supports this hypothesis. (1) *T. lintearius* has a narrow host range (six recorded host plants; Migeon and Dorkeld 2006-2017) and is considered to be an *Ulex* spp. (Fabaceae, Papilionoideae) specialist. This close relative of *T. urticae* can typically be found on common gorse (*U. europaeus*) (Hill and O'Donnell 1991, Ireson et al., 2003, Norambuena et
Host plant dependent expression of salivary proteins

al., 2007), a plant we have shown to induce $SHOT_1 \& 2$ gene expression in $T. urticae$. Indeed, 24 h after transfer from the 'non-inducer plant' $S. lycopersicum$ to $U. europaeus$, expression levels of $TuSHOT_1 \& 2$ genes were in the same order of magnitude as on the reference 'inducer plant' $P. vulgaris$ (FIGURE 5.6). $Tetranychus lintearius$ was found to possess $SHOT_1 \& 2$ genes (FIGURE 5.2.A) whereas the two identified $TlSHOT_3$ genes appeared to be pseudogenes. 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, $T. lintearius'$ and $T. urticae'$ common ancestor possessed intact $SHOT_3$ genes, which were subsequently lost in $T. lintearius$ when it specialized on its fabacean host $U. europaeus$. (2) $T. evansi$ is considered to be a specialist of Solanaceae, although it has been recorded on other host plants as well (Migeon and Dorkeld 2006-2017, Navajas et al., 2013). $Tetranychus urticae$ does not express $SHOT_1 \& 2$ genes when feeding on solanaceous species such as $S. lycopersicum$ or $N. benthamiana$ (strong downregulation relative to the reference 'inducer plant' $P. vulgaris$; FIGURES 5.4, 5.5 & 5.6). Accordingly, members belonging to $SHOT_1 \& 2$ appear to be absent in the $T. evansi$ genome. However, $T. evansi$ does possess two $SHOT_3$ genes (FIGURE 5.2.A). The $T. urticae$ homologs of the $TeSHOT_3$ genes were expressed at high levels in $T. urticae$ feeding on all tested host plants (FIGURES 5.4 & 5.6). (3) In addition, the oligophagous $P. citri$ 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 $T. evansi$ (Takeyama et al., 2006; Migeon and Dorkeld, 2006-2017). In line with this, $SHOT_1 \& 2$ genes could also not be identified in the transcriptome of this mite, yet one gene belonging to the $SHOT_3$ cluster was found to be expressed. (4) Finally, a second $Panonychus$ species, $P. ulmi$, has been recorded on 141 different host plants, most of which belong to the Rosaceae and including few leguminous plants (Migeon and Dorkeld, 2006-2017). The only $SHOT$ gene of $P. ulmi$ found to be expressed is a $SHOT_3$ type.

In summary, all tetranychid species for which genomic or transcriptomic resources were available possessed $SHOT$ genes. $SHOT_1 \& 2$ 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 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.

5.4.5. SHOTs as putative effectors

In stylet feeding herbivores such as gall midges (Stuart et al., 2012; Zhao et al., 2015) and aphids (Mutti et al., 2008; Carolan et al., 2011; Hogenhout and Bos, 2011; Mondal, 2017), 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’s 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 non-feeding embryonic stage, pointing towards 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 (see section 3.2) 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 (Tian and Jian-ping, 2010; Deng et al., 2017). 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 SHOT genes belonging to all three subfamilies we identified, while the monophagous species *T. lintearius* and the oligophagous species *T. evansi, P. citri* and *P. ulmi* possess either only SHOT1 & 2 or only SHOT3 genes. The highly inducible and fast regulation of *TuSHOT1* & 2 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 & 2 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 Fabaceae (Papilionoideae excluding the Fabae tribe) compared to 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 & 2 expression on non-Fabaceae is relatively slow compared to the speed of upregulation under the reverse conditions. Such asymmetric on/off regulation suggests that the mite’s need to upregulate the SHOT genes on the subset Fabaceae is much greater than its need to downregulate these on the alternative hosts. If so, it is possible that on this subset of Fabaceae long-lasting high production of SHOT proteins is required for these to take effect, while for the other hosts such effects are reached faster and/or at lower protein levels. 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* sp. and *T. evansi* do not have SHOT1 & 2 genes while mainly feeding on non-fabacean hosts.

Our second hypothesis is that these particular SHOT1 & 2 proteins are involved in the establishment of a compatible interaction with the subset of Fabaceae on the one hand but are detrimental for the establishment of a compatible interaction with non-Fabaceae 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 (Pitino and Hogenhout, 2013; Kettles and Kaloshian, 2016; Thorpe et al., 2016; Rodriguez et al., 2017).
For example, expression of the *Macrosiphum euphorbiae* effector Me47 in *N. benthamiana* enhanced reproductive performance of *Myzus persicae* while delivery of Me47 into *A. 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 exclusively, and their concomitant down-regulation on the other plant species, could indicate these other hosts respond to the SHOTs in an unfavorable manner, for example by initiating effector-triggered immunity (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 & 2 proteins are involved in the establishment of a compatible interaction with a subset of Fabaceae but do not affect the extent of compatibility with other hosts. For example, the SHOT1 &2 proteins could serve to interfere with Fabaceae–specific recognition of spider mites and the subsequent defense activation. The transient downregulation observed after transfer from a fabacean (*P. vulgaris*) to a non-fabacean 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. The downregulation in the mites feeding on non-Fabaceae would most probably reflect a resource-saving mechanism. In view of this hypothesis, SHOT1 & 2 proteins could thus be specialized for acting on a subset of Fabaceae.

### 5.4.6. Evaluation of SHOT effector action

The ultimate proof of effector function is through functional validation. Yeast two-hybrid assays were performed, and *TuSHOT2b* was shown to interact with *Arabidopsis* serine/threonine protein kinases (MUT9-like kinases, MLKs). *Arabidopsis* mutants defective in the genes coding for these interaction partners, *At3g03940* and *At5g18190*, show pleiotropic phenotypes including dwarfism and hypersensitivity to osmotic/salt stress. It is suggested that these proteins are involved in chromatin reorganization (Wang et al., 2015). Interestingly, *mlk1,3,4* triple mutants exhibit stronger SA-induced defense marker gene expression, relative to wild type *Arabidopsis*. In addition, the oomycete effector HaRxL106 is suggested to manipulate the function of RCD1 (RADICAL-INDUCED CELL
Host plant dependent expression of salivary proteins

DEATH1), while these RCD1-type proteins interact with MLKs. As such, this oomycete effector suppresses Arabidopsis innate immunity (Wirthmueller et al., 2017). This study shows that MLKs may be relevant targets for effectors, such as TuSHOT proteins.

The identified TuSHOT interaction partners in Arabidopsis may be the ‘intended’ effector target (i.e., a S gene products), which is the target to which the effector evolved to interact with in order to obtain a higher mite fitness. It may, however, also interact with the plant factors that induce a defense response, such as R gene products or associated decoy proteins. Indeed, when T. urticae feeds on a non-leguminous host such as Arabidopsis, the production of TuSHOT1 or TuSHOT2 proteins is downregulated (FIGURE 5.6), which suggests these proteins may have a negative effect on the fitness of mites feeding on these hosts. For the identification of the functionally most relevant interaction partner for T. urticae SHOT1 and SHOT2 proteins, a cDNA library of a fabacean host plant (e.g., M. truncatula) should be used. However, Y2H analyses on these libraries were not easily attainable.

Tetranychus urticae produced SHOT3 proteins when feeding on all studied plant species, although there is some variation in expression levels of the encoding genes. The plant species of the Y2H cDNA library likely is of minor importance. The Y2H screen of TuSHOT3b in S. lycopersicum did not seem to result in the identification of a biologically relevant plant target. TeSHOT3b on the other hand, interacted with several interesting S. lycopersicum proteins during the Y2H screens, though with low confidence. Amongst these identified proteins was a LRR receptor-like serine/threonine-protein kinase, for which a role in plant defense signaling is assumed (Afzal et al., 2008). Furthermore, interaction with an isochorismatase hydrolase occurred as well. This enzyme may inhibit the production of salicylic acid and has been proposed as plant-defense suppressor of the phytopathogenic fungus Verticillium dahlia (El-Bebany et al., 2010). Also, interaction between TeSHOTb3b and a nodulin-like protein occurred. Latter proteins are important for transport and for major aspects of plant development. Their activity has also been associated with the enhancement of pathogen fitness during host plant colonization (Denancé et al., 2014). Furthermore, interaction with PAD4 occurred, a lipase-like protein that is important in SA signaling (Jirage et al., 1999). While these plant interactors are plausible biological targets for spider mite salivary effectors, the in vivo relevance of
these interactions remains to be confirmed. Additional Y2H experiments are planned to further elucidate SHOT protein interaction partners.

The effect of transiently expressed TuSHOT1b and TuSHOT3a eGFP-fusion proteins in *N. benthamiana* leaves on mite reproduction was studied (preliminary data). Ectopic expression of a TuSHOT1b-eGFP led to a significant reduction ($p = 0.03$) of spider mite reproductive performance. Interestingly, genes coding for TuSHOT1 proteins are not expressed by *T. urticae* feeding *N. benthamiana* (see Figure 5.6). Presumably, TuSHOT1 proteins act as elicitors in *N. benthamiana*, and hypothetically also in most plants except for a subset of Fabaceae. The genes coding for TuSHOT3 are constitutively expressed in *T. urticae* adults, independent of the host plant, although some host-plant dependent expression was apparent (Figures 5.4 & 5.6). Ectopic expression of TuSHOT3 in *N. benthamiana* did not result in a significant change of spider mite performance. This may be caused by the fact that all mites, including those of the control group, are assumed to produce TuSHOT3 proteins themselves, potentially rendering the ectopically expressed TuSHOT3 redundant. The use of TuSHOT3 knock-down spider mites would provide a solution. However, this technique is not yet straightforward in spider mites, and would be hard to perform since three TuSHOT3 genes are involved in *T. urticae*. Another approach would be to use *T. evansi* as test organism. For example, ectopic expression of *T. urticae* TuSHOT1 & 2 in a leguminous plant such as *M. truncatula* could hypothetically enhance the performance of *T. evansi* (lab strains feeding on bean do exist), which does not possess SHOT1 & 2 genes. Also, considering the high number of putative salivary effector proteins of *T. urticae* (Chapter 3), the positive effect of each effector is potentially subtle and may be hard to detect using ATTA coupled to performance assays.

Transient expression of TuSHOT genes in *S. lycopersicum* did not result in detectable changes of leaf JA or SA levels. According to Alba et al. (2015), *S. lycopersicum* defense suppression by spider mites occurs downstream of JA and SA, and independent of the JA-SA antagonism. Potentially, most spider mite effectors therefore exert their function downstream of these phytohormones (Villarroel et al., 2016).

5.5. 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
Host plant dependent expression of salivary proteins

denominated 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 & 2, appeared to be secreted in a host-dependent manner. Furthermore, *SHOT1* & *2* 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, this points towards a host-specific function for these putative effectors. Interestingly, host transfers in the opposite direction (from *P. vulgaris* to *S. lycopersicum*) appear to result in a slow yet strong decrease of *SHOT1* & *2* gene-expression, taking place over generations rather than hours. To our knowledge, *SHOT1* & *2* proteins are the first putative arthropod effectors for which asymmetric on/off regulation in a host plant-dependent manner, i.e., rapid upregulation on one host, and slow downregulation on another, has been demonstrated. Additionally, we explored several techniques to gain insight into the molecular function of *SHOT* proteins. The results are promising, although further study is clearly needed. Preliminary data of transient expression experiments reveal that *SHOT* proteins may have an impact on spider mite performance. In addition, yeast two-hybrid assays revealed interacting plant proteins. However, care must be taken, since the interaction partners are expected to be highly host plant dependent.

Funding information

WD and GM are postdoctoral fellows of the Research Foundation Flanders (FWO, Belgium). This project was supported by NWO (The Netherlands) under the FACCE-JPI ERA-NET Plus framework (GENOMITE, project ID 137), the Research Foundation Flanders (FWO, Belgium) (grant G009312N to LT and TVL and grant G053815N to LT, TVL and WD) and The Special Research Fund (DOZA 01J131711 to WJ). MRK was supported by NWO (The Netherlands) (STW-VIDI 13492).

5.6. References


Host plant dependent expression of salivary proteins


Chapter 5


Host plant dependent expression of salivary proteins


Host plant dependent expression of salivary proteins


### 5.7. Supplementary Material

Supplementary information can be located at the MPMI website at https://doi.org/10.1094/MPMI-06-17-0139-R

**Figure S5.1.** Alignment of intact SHOTs. Intact SHOTs were aligned using MUSCLE (Edgar, 2004). The predicted signal peptides are indicated with an arrow, while those peptides that are adjacent to the SP, had a high MS1 intensity, and that were identified in this study, are framed with a blue rectangle. The proline-glutamic acid, proline-glutamine or proline-aspartate motif in *Tetranychus* SHOTs is indicated with an asterisk.

**Table S5.1.** Modest Feature Finder (MoFF) extracted MS1 intensities from the Thermo raw files for the validated PSMs matching to non-decoy (true) sequences [MS1 intensities from two technical LC-MS/MS replicates for each host plant specific sample (bean, soybean and tomato)].

**Table S5.2.** List of primers used in this study.

**Table S5.3.** Normalized MS1 intensities of peptides identified in artificial diet enriched with saliva of adult *T. urticae* females that were propagated for at least 5 generations on bean, soybean or tomato.

**Table S5.4.** Proteins of which peptides were identified (see Table S5.3) in artificial diet enriched with saliva of adult *T. urticae* females that were propagated for at least 5 generations on bean, soybean or tomato.

**Table S5.5.** Properties of intact members of the SHOT family.

**Table S5.6.** Amino acid composition of intact members of the SHOT family.

**Table S5.7.** PPA analysis of those proteins of which peptides were identified in this study (peptides were sorted based on decreasing detection probability).

**Data S5.1.** *SHOT* gene/transcript sequences (.txt format).