

# SCIENTIFIC REPORTS



OPEN

## Transcriptional Analysis of The Adaptive Digestive System of The Migratory Locust in Response to Plant Defensive Protease Inhibitors

Jornt Spit<sup>1,\*</sup>, Michiel Holtof<sup>1,\*</sup>, Liesbet Badisco<sup>1</sup>, Lucia Vergauwen<sup>2,3</sup>, Elise Vogel<sup>1</sup>, Dries Knapen<sup>2,3</sup> & Jozef Vanden Broeck<sup>1</sup>

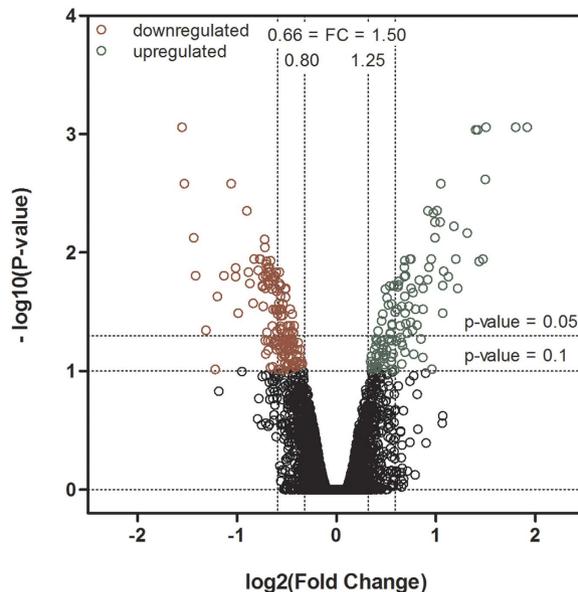
Received: 06 June 2016  
Accepted: 05 August 2016  
Published: 01 September 2016

Herbivorous insects evolved adaptive mechanisms to compensate for the presence of plant defensive protease inhibitors (PI) in their food. The underlying regulatory mechanisms of these compensatory responses remain largely elusive. In the current study, we investigated the initiation of this adaptive response in the migratory locust, *Locusta migratoria*, via microarray analysis of gut tissues. Four hours after dietary uptake of PIs, 114 and 150 transcripts were respectively found up- or downregulated. The results suggest a quick trade-off between compensating for potential loss of digestive activity on the one hand, and stress tolerance, defense, and structural integrity of the gut on the other hand. We additionally addressed the role of a group of related upregulated hexamerin-like proteins in the PI-induced response. Simultaneous knockdown of corresponding transcripts by means of RNA interference resulted in a reduced capacity of the locust nymphs to cope with the effects of PI. Moreover, since insect hexamerins have been shown to bind Juvenile Hormone (JH), we also investigated the effect of JH on the proteolytic digestion in *L. migratoria*. Our results indicate that JH has a stimulatory effect on the expression of three homologous chymotrypsin genes, while knocking down the JH receptor (methoprene tolerant) led to opposite effects.

Plants and insects have continuously interacted with each other in a co-evolutionary relationship. While plants can benefit from the pollinator activities of several insect species, herbivorous insects use their host plants as main feeding source. The insect midgut shows a high flexibility, which allows them to swiftly adapt to varying nutrient levels in their diet<sup>1</sup>. A remarkable case of this high flexibility is found in the insect's adaptive mechanisms to malnutritious factors. The constant threat of severe damage by feeding insects has led to the evolution of a wide range of specific protective strategies in the plants, among which secondary metabolites, such as protease inhibitors (PI)<sup>2</sup>. These PI primarily act in the digestive tract of the insect by targeting proteolytic digestive enzymes, or interfere with the structural properties or protective functions of the gut<sup>3</sup>. Resulting amino acid deficiencies will eventually manifest themselves in a limitation of the insect's growth, development, fecundity, and can ultimately lead to an increased mortality<sup>3</sup>. Since their discovery, PI have unquestionably received a lot of attention in agricultural sciences, considering their insecticidal potential<sup>4,5</sup>.

However, several important pest insects have evolved mechanisms to compensate for the antimetabolic effects of PI, with attenuation, or a complete lack of developmental defects as a result<sup>5</sup>. Over the past decades, the number of published studies describing these different adaptive strategies has grown. These can generally be divided into two major categories: increased expression of active proteases, outnumbering the PI in the gut lumen, and/or expression of PI-insensitive proteases, which are not a target of the ingested PI<sup>4,5</sup>. Both strategies have been observed in many different insect species when confronted with PI<sup>6–11</sup>. Nevertheless, while much of the current literature focuses on how insects counteract the presence of PI, the current understanding of the initiation and

<sup>1</sup>Department of Animal Physiology and Neurobiology, Zoological Institute KU Leuven, Naamsestraat 59, B-3000 Leuven, Belgium. <sup>2</sup>Systemic Physiological and Ecotoxicological Research (SPHERE), Department of Biology, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerpen, Belgium. <sup>3</sup>Zebrafishlab, Veterinary Physiology and Biochemistry, Department of Veterinary Sciences, University of Antwerp, Universiteitsplein 1, B-2610 Wilrijk, Belgium. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.V.B. (email: jozef.vandenbroeck@bio.kuleuven.be)



**Figure 1. Volcano plot of the retrieved microarray data, plotting the negative  $\log^{10}$  of the adjusted P-value against the  $\log^2$  of the fold change (FC).** This means that all transcripts with a FC < 1 have negative X-values and all transcripts with positive X-values have a FC > 1. The higher on the Y-axis, the lower the corresponding adjusted P-value. All values with a FC < 0.8, and P-value < 0.1 are indicated in red and are considered significantly down-regulated. All values with a FC > 1.25 and P-value < 0.1 are indicated in green and are considered upregulated.

regulation of these mechanisms still remains surprisingly unclear. This comprehension, however, could prove to be essential for the development of more effective PI-based strategies for pest control.

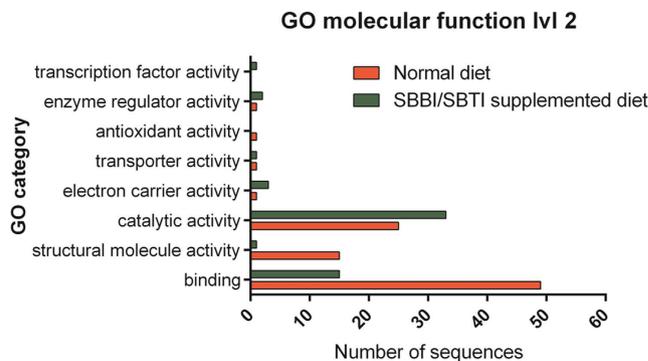
The migratory locust, *Locusta migratoria*, is considered among the most important agricultural pest insects, endangering an enormous acreage of cultivated areas worldwide each year<sup>12</sup>. Furthermore, its relatively large size makes it a highly suitable insect model for physiological research. Initial proteolytic digestion in this species is mainly steered by trypsins and chymotrypsins, complemented with a small activity of cysteine proteases<sup>13</sup>. We previously have shown that *L. migratoria* possesses a clear PI-based compensatory response<sup>13</sup>. It responds to the dietary uptake of soybean Bowman Birk (SBBI) and soybean trypsin inhibitors (SBTI) by inducing the expression of serine proteases active in the digestive system, while no effect on larval development could be observed<sup>13</sup>. Interestingly, we observed increased expression of serine proteases as quickly as four hours after the ingestion of the inhibitors, demonstrating the high flexibility of this response<sup>13</sup>. Like in other insects, the underlying regulatory mechanisms of this adaptation are however still unknown. Up to now, only few studies have tried to identify broader transcriptional changes associated with this phenomenon<sup>8,14–18</sup>. However, these studies mostly investigated long-term changes after dietary PI uptake and did not focus on the initial driving forces regulating the response.

Therefore, in this study, we conducted a microarray study to specifically observe short-term transcriptional changes in the gut and gastric caeca of the migratory locust at the previously observed critical time point of four hours after SBBI and SBTI ingestion<sup>13</sup>. Detailed analyses of differentially expressed transcripts showed a notable upregulation of several transcripts encoding hexamerin-like proteins, which are known to be juvenile hormone binding proteins (JHBPs)<sup>19</sup>. Interestingly, one other study performed in the Colorado potato beetle, *Leptinotarsa decemlineata*, also demonstrated an induced expression of several JHBPs after feeding on potato plants with induced defences<sup>16</sup>. These findings could suggest a more general role for JHBPs or JH in the regulation of the digestive process in insects. We therefore further pursued this idea. Knockdown studies by means of RNAi helped us reveal, for the first time, a possible regulatory mechanism involved in the PI-induced defence in this locust species. In addition, we were able to demonstrate a stimulatory effect of JH signalling on a specific group of serine proteases, further hinting towards a conserved role for this hormone in the regulation of digestion.

## Results and Discussion

### Two hundred and sixty-four differentially expressed transcripts were identified 4h after SBBI and SBTI ingestion.

Microarray slides were developed containing 35869 unique single probed transcripts (Spit *et al.*<sup>20</sup>). Transcript levels of combined midgut and caeca samples originating from 5th instar nymphs, 4 hours after exposure to a diet containing SBBI and SBTI (1%), were compared to those from control locusts. These plant derived inhibitors are well described and capable of inhibiting both insect trypsins and chymotrypsins, which constitute up to 90% of the total proteolytic activity in the digestive system of *L. migratoria*<sup>21–23</sup>. The administered dose of 1% of total protein content was previously shown to be successful in eliciting a quick response in the locust gut<sup>13</sup>. Figure 1 shows a volcano plot representing the results. Plotting the negative  $\log_{10}$  function of the adjusted P-value against the  $\log_2$  function of the fold change, visualises the up- and downregulated



**Figure 2.** GO term classification of differentially expressed annotated transcripts under GO category ‘Molecular function’ level2. Comparison of transcript numbers associated with different GO terms that are more abundant in control locusts (red) or locusts fed on a diet containing PI (green).

transcripts as a response to PI uptake. All transcripts with increasing or decreasing fold changes of at least 25% at an adjusted P-value cut-off of  $p < 0.1$  were considered differentially expressed as result of the treatment. In total, 114 and 150 transcripts were identified as respectively up- or down-regulated. A higher number of downregulated genes is not surprising, as it has been previously noticed that the expression of several genes is suppressed rather than induced under environmental stress<sup>24</sup>.

To confirm expression results from the microarray, several transcripts from the top right and top left corner of the volcano plot were selected, representing strongly up- or downregulated transcripts, respectively. In total, 7 transcripts were tested with RT-qPCR: LMC\_002270, LMC\_003930, LMC\_000698, LMC\_000083, LMC\_003601, LMC\_003606, and LMC\_004513. Transcript levels were measured in a new group of 5<sup>th</sup> larval locusts that were fed a diet containing SBBI and SBTI, and with the exception of LMC\_000083, expression levels did not differ significantly from the fold changes as observed in the microarray analysis (Supp. Fig. S1).

**Functional annotation of SBBI- and SBTI-responsive genes from *L. migratoria*.** Sequences of differentially regulated transcripts were functionally annotated using gene ontology (GO), and the InterProScan (IPS) tool that is integrated in the Blast2GO bioinformatics platform<sup>25,26</sup>. Out of 264, 95 transcripts did not produce any significant BLAST hit. 113 of the remaining 169 sequences with at least one hit were ultimately assigned GO terms. Comparing the main GO ontology domain ‘Molecular function’ (lvl2) shows that most transcripts are putatively involved in ‘catalytic activity’, ‘binding’ or ‘structural molecule activity’, with the latter two more dominantly present in insects that received control diet (Fig. 2). Annotation numbers resulting from the IPS are comparable. 133 sequences could be assigned to an InterPro family. From these sequences, 68 belonged to transcripts that were upregulated, and 65 to transcripts that were downregulated. Tables 1 and 2 show a summary of the results for respectively the up- and the down-regulated transcripts. A list containing all up- and down-regulated transcript IDs with their corresponding LogFC, calculated fold change, and putative sequence description is provided as supplementary data (Table S2).

**SBBI and SBTI ingestion quickly induces changes in the digestive enzyme profile.** Immediately apparent is the large number of upregulated sequences involved in protein digestion. Several transcripts are predicted S1 serine endoproteases, while others correspond to putative carboxypeptidases. Our data demonstrate that *L. migratoria* counteracts the malnutritious effects of SBBI and SBTI in the gut by inducing the expression of serine endoproteases, thereby supporting our previous research<sup>13</sup>. In addition, the observed elevated transcript levels for carboxypeptidases, suggest partial use of this class of exopeptidases to overcome the inhibitory effect of the ingested endoprotease inhibitors. This is the first time that this type of compensation is observed in *L. migratoria*. Partial reliance on insensitive proteases as counter defence however has already been observed in different other insects, including: *Helicoverpa armigera*, *Spodoptera exigua*, *Callosobruchus maculatus*, *Leptinotarsa decemlineata* and *Schistocerca gregaria*<sup>9,14,16,27,28</sup>. Our findings confirm that regulating the expression of proteolytic enzymes is a common strategy applied by insects to evade loss of proteolytic activity caused by the presence of PI in the gut. Surprisingly, in contrast to the observed trend of general upregulation of digestive proteases, one aminopeptidase and a dipeptidase sequence were found downregulated in response to dietary PI uptake (Table 2). The underlying reason for this is unclear. In some insects there exists a trade-off between the expression of proteases susceptible to the ingested PI and insensitive proteases<sup>3,9</sup>. However, the used protease inhibitors, SBBI and SBTI, are not capable of inhibiting aminopeptidases or dipeptidases. A possible explanation could be that the expression of these proteases is influenced by the levels of free amino acids present in the gut. When these levels are affected by the presence of PI, downregulation of these enzymes might occur.

The locusts’ exposure to PI also results in an upregulation of a group of enzymes involved in lipid metabolism and carbohydrate hydrolysis (Table 1). This could be consistent with an exploitation of alternative sources of energy to compensate for a loss in energy normally derived from protein digestion<sup>14,18</sup>. On the other hand, a  $\beta$ -glucosidase (encoded by 3 transcripts), two hexokinases and cellulase were found downregulated under the same conditions. Interestingly, the simultaneous downregulation of  $\beta$ -glucosidase and cellulase, while other carbohydrate hydrolases were upregulated, was also observed in microarray studies from *L. decemlineata*<sup>16</sup>, and

Putative function	InterPro <sup>§</sup>	# Seq <sup>¶</sup>	Description	Fold Change <sup>¶</sup>
Storage protein/JH-binding	IPR013788	21	Hemocyanin/hexamerin	1.4–2.8
	IPR010562	2	Haemolymph juvenile hormone binding	1.4–1.5
Protein digestion	IPR001254	16	Serine protease family S1	1.4–2.0
	IPR000834	5	Peptidase M14, carboxypeptidase A	1.3–1.6
Carbohydrate digestion	IPR001360	1	$\beta$ -glucosidase	1.6
	IPR000933	1	$\alpha$ -L-fucosidases	1.7
	IPR001139	1	Glucosylceramidase	1.5
Lipid digestion	IPR000734	1	Lipase	1.6
Detox/Stress response	IPR001128	3	Cytochrome P450	1.3–2.0
	IPR002018	2	Carboxylesterase, type B	1.4–1.6
Peritrophic membrane	IPR002557	7	Chitin binding domain	2.0–2.1
Other	IPR001071	1	$\alpha$ -tocopherol transport	1.6
	IPR000863	1	Sulfotransferase domain	1.5
	IPR012337	1	Ribonuclease H-like domain	1.4
	IPR001888	2	Transposase/methyltransferase	1.4
	IPR002110	1	Ankyrin repeat	1.4
	IPR011042	1	Six-bladed beta-propeller, TolB-like	1.4
	IPR008037	1	Proteinase inhibitor I19, pacifastin	1.3

**Table 1. Summary of InterPro families from upregulated transcripts that were annotated in *L. migratoria* after PI ingestion.** <sup>§</sup>InterPro domain ID based on the InterPro database scan<sup>69</sup>. <sup>¶</sup>Number of transcript sequence occurrences corresponding to the indicated InterPro ID. <sup>¶</sup>When multiple transcripts are identified the range between the minimal and maximal fold change is presented.

*C. maculatus*<sup>14</sup>. In addition, a low-density lipoprotein (LDL) receptor and an apo-lipoprotein (lipocalin) were also downregulated. Both are involved in the trafficking of lipids to energy demanding tissues or the fat body, suggesting less energy is available for redirecting or building reserves. Other enzymes including short- and long chain dehydrogenase, nucleoside hydrolase and carbonic anhydrase were also downregulated, indicating a metabolic adjustment to use energy for other tasks, like the observed enhanced proteolytic digestion.

*Structural integrity and defences are in a trade-off with compensatory expression of digestive enzymes.* Another intriguing observation is that, after being confronted with SBBI and SBTI in their gut, locusts respond by suppression of transcripts encoding proteins involved in both muscle and cytoskeleton structure as well as cell adhesion (Table 2). This suggests a weakening of the structural integrity of the insect gut in response to PI uptake. However, 4 hours after PI ingestion, structural changes were not yet visible. It could certainly be of interest to further investigate if the muscle structure of the gut indeed is compromised at later stages. Changes in the ultrastructure of the midgut epithelium have been observed in longer term studies on the PI compensatory responses in insects. In the black field cricket, *Teleogryllus commodus*, reduction of midgut wall depth, vacuolization of epithelial cells, swelling of the microvilli and eventual rupture of epithelial cells was observed after consecutive days of ingesting PI<sup>29</sup>. In *Drosophila melanogaster*, a remarkable reduction in the length of midgut microvilli was also observed in SBBI-fed larvae<sup>30</sup>. Other studies investigating morphological changes in response to PI ingestion are lacking, but these results suggest that structural changes may indeed coincide with PI ingestion in insects.

Interestingly, reduced transcript levels for several defence- and stress response related genes are also observed (Table 2). Among these are transcripts encoding C-type lectin proteins, which mainly function in innate immunity as microbial pattern recognition molecules<sup>31</sup>, while haem peroxidases and cytochrome P450-like enzymes are generally known to be important during oxidative stress. Heat shock protein HSP20 was also found downregulated. In insects, HSP20 has been reported to play critical roles during thermal adaptation<sup>32,33</sup>. In addition, carboxylesterase encoding transcripts were downregulated, which were previously shown to be involved in the xenobiotic detoxification in *L. migratoria*<sup>34</sup>. Combined, these results indicate that, while executing the counter defence against the dietary challenge presented by PI, locusts may be more susceptible for the risks of other environmental factors. In contrast, another carboxylesterase B-like sequence, and some additional cytochrome P450 encoding transcripts were found upregulated, which could indicate a more specific role for these proteins during PI-compensation in the locust gut. While structural properties of the gut appear negatively affected, chitin binding mucin-like peritrophins were upregulated after PI ingestion. These proteins are part of the peritrophic membrane (PM) that surrounds the food bolus in insects, protecting the midgut epithelium from microorganisms and mechanical damage. This suggests a strengthening of the PM after PI detection in the gut. Similar results were obtained in corn rootworm larvae during scN inhibitor challenge<sup>35</sup>. It was suggested that a decrease in protein degradation efficiency could lead to a more abrasive food bolus, so a more rapid turnover of the PM is required to protect the midgut epithelium<sup>35</sup>.

*Juvenile hormone binding proteins (JHBPs) are upregulated.* Among the most highly upregulated transcripts is a large group of sequences encoding putative hexamerin-like proteins. Hexamerins are high molecular weight proteins comprised of six identical or similar subunits. They are related to hemocyanin, but lost their ability to

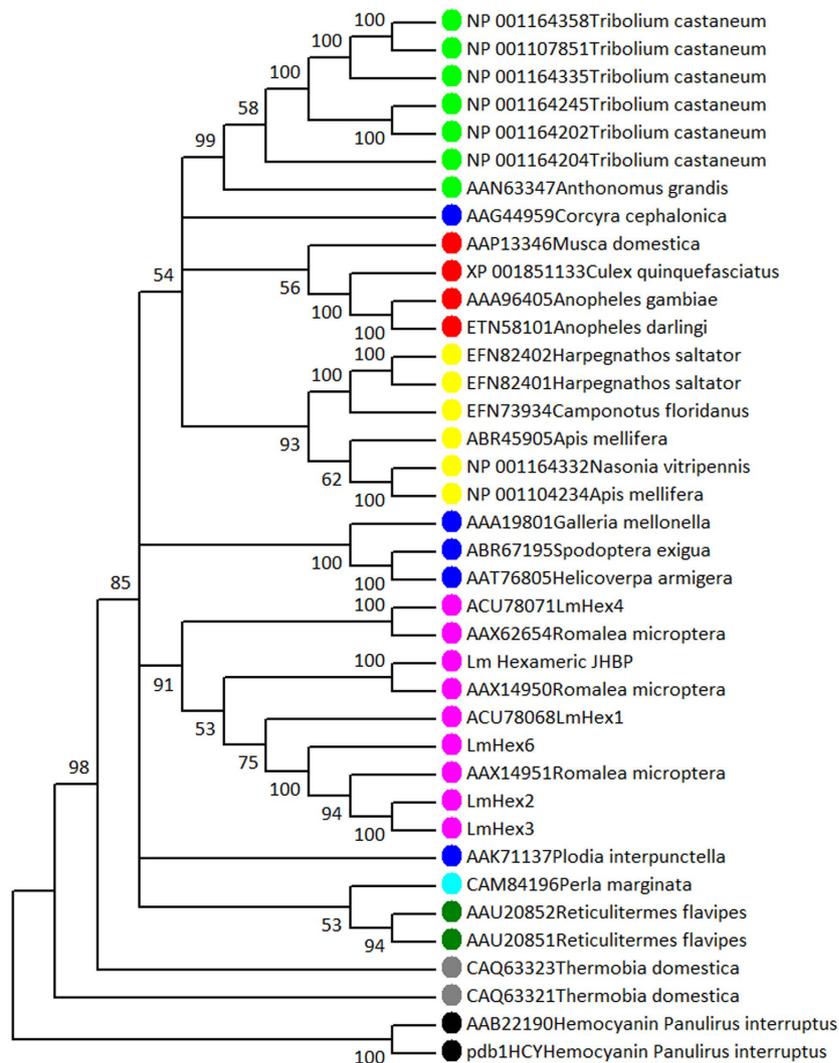
Putative function	InterPro <sup>§</sup>	# Seq <sup>¶</sup>	Description	Fold Change <sup>¶</sup>
Structural	IPR004000	15	Actin-related protein	0.6–0.7
	IPR002928	6	Myosin	0.7–0.8
	IPR001781	4	Zinc finger, LIM-type	0.6–0.8
	IPR002035	3	von Willebrand factor, type A	0.4–0.7
	IPR013098	2	Immunoglobulin I-set domain	0.7–0.8
	IPR027707	2	Troponin T	0.6–0.7
	IPR001715	1	Calponin homology domain	0.7
	IPR000301	1	Tetraspanin	0.8
Defense	IPR001152	1	Thymosin $\beta$ -4	0.8
	IPR001304	3	C-type lectin	0.4–0.6
Carbohydrate metabolism	IPR001360	3	$\beta$ -glucosidase	0.4–0.6
	IPR001312	2	Hexokinase	0.5–0.7
	IPR001701	1	Cellulase	0.7
Lipid metabolism	IPR000566	1	Lipocalin	0.7
	IPR002172	1	Low-density lipoprotein (LDL) receptor class A	0.8
Other metabolism	IPR002198	1	Short-chain dehydrogenase	0.7
	IPR002085	1	Long-chain alcohol dehydrogenase	0.7
	IPR023186	1	Inosine/uridine-nucleoside hydrolase	0.8
	IPR001148	1	Alpha carbonic anhydrase	0.7
	IPR008597	1	Destabilase	0.6
Protein digestion	IPR000994	1	Peptidase M24, aminopeptidase	0.7
	IPR008257	1	Peptidase M19, dipeptidase	0.6
Detox/Stress response	IPR002018	3	Carboxylesterase, type B	0.6–0.7
	IPR002007	1	Haem peroxidase	0.7
	IPR002403	1	Cytochrome P450, E-class, group IV	0.7
	IPR008978	1	Heat shock HSP20-like chaperone	0.6
Other	IPR002048	1	EF-hand domain	0.7
	IPR011011	1	Zinc finger, FYVE/PHD-type	0.7
	IPR001253	1	Translation initiation factor 1A (eIF-1A)	0.6
	IPR020309	1	Uncharacterized protein family, CD034/YQF4	0.7
	IPR004119	1	Protein of unknown function DUF227	0.5
	IPR012464	1	Protein of unknown function DUF1676	0.6

**Table 2. Summary of InterPro families from downregulated transcripts that were annotated in *L. migratoria* after PI ingestion.** <sup>§</sup>InterPro domain ID based on the InterPro database scan<sup>69</sup>. <sup>¶</sup>Number of transcript sequence occurrences corresponding to the indicated InterPro ID. <sup>¶</sup>When multiple transcripts are identified the range between the minimal and maximal fold change is presented.

bind oxygen. Hexamerins are best known as hemolymph storage proteins providing amino acids and energy during non-feeding periods<sup>36</sup>. More recently, additional functions for this protein family were suggested. It could be demonstrated that hexamerins are phase-dependently expressed in both *L. migratoria*<sup>37</sup> and *S. gregaria*<sup>38</sup>. Moreover, two hexamerin genes were found to participate in the JH-dependent regulation of caste polyphenism in termites<sup>39</sup>, suggesting their importance in regulating phenotypic plasticity. Apart from their function as storage proteins, in *L. migratoria*, hexamerins were also reported as juvenile hormone binding proteins (JHBPs) that function in the regulation of JH levels<sup>19</sup>. Interestingly, a family of low molecular weight JHBPs was also found upregulated in a study of the response of *L. decemlineata* to induced potato plant defences<sup>16</sup>. These findings prompted us to perform a comprehensive search in the Supplementary Data of related microarray studies, revealing several genes involved in JH production and degradation, e.g. JH acid methyltransferase, JH esterase and JH epoxide hydrolase, were affected during PI compensatory responses in *C. maculatus* and *H. virescens*<sup>14,15</sup>. Combined, these results indicate that JH and/or JHBPs are potentially more directly involved in the regulation of digestion and the proteolytic compensation after dietary exposure to antinutritional components. This idea is further supported by a growing number of reports on JH regulating the expression of certain proteolytic enzymes in the midgut<sup>40–44</sup>.

**Hexamerin knockdown perturbs the PI compensatory response.** The remarkable consensus between some of the results from our microarray study and those from other studies on insects fed with different kinds of PI could suggest that similar regulatory networks are present in different orders of insects. One of the most intriguing results from these studies is the differential expression of JHBP family members, suggesting an important role in the regulation of the PI-induced compensatory response.

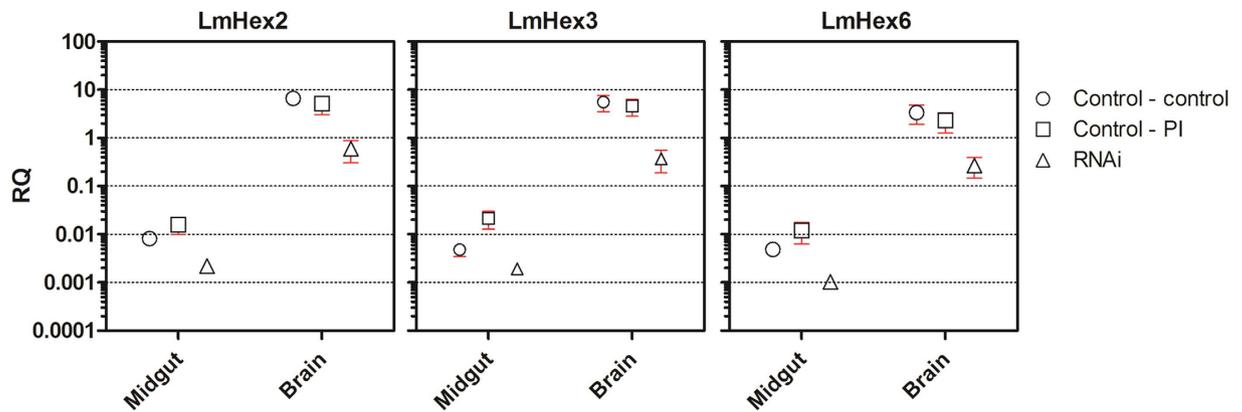
The microarray analyses revealed a set of 21 differentially expressed transcripts encoding hexamerin-like proteins in the larval gut after ingestion of SBBI and SBTI. Many of these transcripts appeared closely related or



**Figure 3. Maximum Likelihood tree of insect hexamerin-like proteins.** Protein sequences of hemocyanin from *Panulirus interruptus* were used as outgroup to root the tree. The tree was inferred from 100 bootstrap replicates and bootstrap values for every node are presented. Accession numbers and species name are depicted. Sequences from Orthoptera are represented in purple, Coleoptera in light green, Lepidoptera in dark blue, Diptera in red, Hymenoptera in yellow, Plecoptera in light blue, Isoptera in dark green, Zygentoma in grey.

identical to each other and showed substantial similarity (>98% identity). Based on sequence identity of (overlapping) nucleotide sequences, the transcripts could be further assembled into three distinct full length contig sequences. For all three contigs a coding region from start to stop codon was present, together with 5' and 3' UTRs (Supp. Fig. S3). For *L. migratoria*, five sequences encoding hexamerin-like proteins and one so-called hexameric JHBP can be found in GenBank. Contig 2 was identical to hexamerin-like protein 2 (*LmHex 2*; [GenBank: FJ609739.1]), while contig 3 was identical to hexamerin-like protein 3 (*LmHex 3*; [GenBank: FJ609740.1]). The sequence from contig 1 was not identical to any known hexamerin sequence from *L. migratoria* and was therefore designated as *L. migratoria* hexamerin-like protein 6 (*LmHex6*). Its sequence was confirmed by PCR and uploaded to Genbank (accession number: BK009413). Phylogenetic analysis shows that the hexamerin-like sequences from *Locusta* and other Orthoptera form a monophyletic group within the insect hexamerins, which was supported by high bootstrap values (Fig. 3). The sequence for *LmHex5* was not included in this analysis since only a partial sequence for this protein was present in GenBank. The amino acid sequence of *LmHex2* was most similar to *LmHex3* (96%), while they both share 75% identity with *LmHex6*. The identity with, and between, other hexamerin-like proteins from *L. migratoria* varies between 35% and 45%, suggesting *LmHex2*, *LmHex3*, and *LmHex6* are highly related genes (Supp. Table S4).

In order to more closely evaluate the function of these hexamerin-like proteins, the transcript levels for *LmHex2*, *LmHex3*, and *LmHex6* were determined in 11 different adult locust tissues under normal dietary conditions (Supp. Fig. S5). Expression of all three was highest in the fat body, the flight muscles, the central nervous system and the gonads. High transcript levels in the gonads suggest that hexamerins may supply amino acids for gonad development and gamete production in *L. migratoria* adults, while expression in the flight muscles

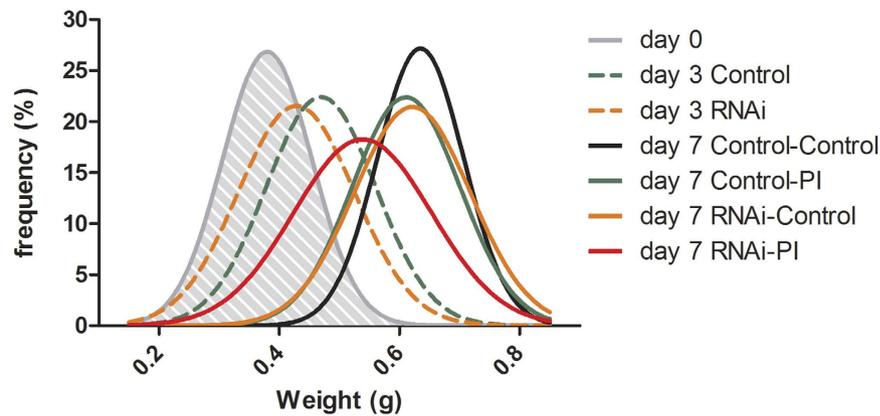


**Figure 4. Mean relative transcript levels for hexamerin-like proteins after knockdown in combination with protease inhibitor feeding.** Relative transcript levels in midguts and brains were examined for 5<sup>th</sup> larval locusts that received dsRNA for *gfp* as a control and were subsequently fed a control diet (open circles), or a diet containing SBBI and SBTI (open squares), and were compared to locusts that received a combination of dsRNA for *LmHex2*, 3 and 6 (open triangles). Hexamerin expression data for RNAi–control and RNAi–PI data were combined since no difference in transcript levels could be detected between these two groups. Relative transcript levels are normalized against two reference genes, *rp49* and *rps13*. Axes are scaled logarithmically, and each horizontal dotted line represents a 10 fold difference in expression. RQ = relative quantity. Means  $\pm$  SEM are represented (n = 5 pools of >5 individuals). Statistical differences were determined by ANOVA. For both brain and midgut, transcript levels were significantly reduced by RNAi ( $p < 0.01$ ).

supports the idea that hexamerins could be used for flight activity<sup>45</sup>. Alternatively, expression of hexamerin-like proteins in these tissues might also be related to their JH binding capacity and influence the transport of JH in or out of these tissues, thereby regulating reproductive and developmental physiology. Surprisingly, in sharp contrast with the high expression level in these tissues, transcript levels of *LmHex2*, *LmHex3* and *LmHex6* were very low in the foregut, midgut, hindgut, gastric caeca and Malpighian tubules, which are all associated with the alimentary tract. However, when fed with PI rich diet, the transcript levels of hexamerins were altered in the gut. No clearly detectable changes could be observed in the brain (Fig. 4). These results could hint towards a more specified role of hexamerins in the gut during variable dietary conditions.

To further determine the possible role of these proteins in the flexible regulation of digestion, we performed a simultaneous knockdown of the three hexamerin-like sequences. Seven days after injection of dsRNA, the transcript levels of *LmHex2*, *LmHex3* and *LmHex6* were reduced with respectively 95%, 93% and 94% in the locusts' brains (Fig. 4). It must be noted that, while a very high knockdown efficiency was achieved, brain transcript levels after knockdown were still remarkably higher than the transcript levels in the gut of control animals. Nevertheless, while transcript levels in the gut were already low in control animals, they could still be significantly reduced through dsRNA injection (73%, 76% and 88%). To investigate the role of the knocked down genes in the PI-induced compensatory response, experimental locusts were fed with either a PI rich diet or a control diet. These conditions were compared with the same dietary conditions in control locusts. Individual locusts were weighed throughout the experiment. In this set-up, failure in the compensatory response would manifest in differences in weight gain. All groups significantly gained weight after three and seven days. Mean weight was significantly different after 3 days for control versus RNAi treated locusts (T-test;  $p < 0.05$ ). Nevertheless, RNAi treated animals fed with a control diet regained their normal growth rate and at the end of the experiment their weight was no longer statistically different from control animals grown on a control or PI-rich diet (Fig. 5). In contrast, knockdown of two important storage proteins, SeHex and SeSp1 in *Spodoptera exigua* resulted in increased mortality, suggested to be caused by a lack of nutritional reserves<sup>46</sup>. However, in this study, we only targeted the group of hexamerins that were upregulated in the gut after PI ingestion. Additional storage proteins or hexamerins exist<sup>19</sup>, which could explain the lack of noticeable developmental defects in the PI-untreated knockdown condition. Mean weight of insects feeding for 4 days on PI-enriched diet also did not differ significantly from control animals, confirming swift adaptation of *L. migratoria* to SBBI and SBTI. Opposed to the other conditions, the growth of locusts that received *hexamerin*-dsRNA in combination with SBBI and SBTI slowed down, and at the end of the experiment the mean weight of this group was significantly lower compared to all other groups (ANOVA;  $p < 0.01$ ).

To understand the underlying reason for the observed delay in growth, the transcript levels of several serine protease genes that were previously shown to be upregulated in *L. migratoria* nymphs after prolonged exposure to SBBI and SBTI uptake were analyzed after day 7 (Fig. 6). With the exception of *LmChy1* [GenBank: BK008825], our results confirmed upregulation of serine proteases after PI ingestion, be it slightly less pronounced than previously reported<sup>13</sup>. Interestingly, transcript levels of *LmChy1*, *LmChy2* [GenBank: BK008826] and *LmChy4* [GenBank: BK008828] are lower in locusts that were additionally fed a diet containing PI (Fig. 6). Transcript levels of *LmChy3* [GenBank: BK008827], *LmTry1B* [GenBank: BK008820], *LmTry2A* [GenBank: BK008822] and *LmTry2B* [GenBank: BK008823] are not affected by the treatment, and the usual upregulation of these genes after PI ingestion is still observed after PI ingestion of RNAi animals. Remarkably, our previous work showed that



**Figure 5. Frequency distribution of weight from control and experimental 5<sup>th</sup> instar locust larvae at different time points.** At day 0, half the locust population was injected with dsRNA for *LmHex2*, *LmHex3* and *LmHex6*, or dsRNA for *gfp* as a control. After three days half the locusts from each group received a diet containing SBBI and SBTI, while the other half received a control diet supplemented with BSA. Weight gain was recorded during the experiment and is presented as frequency distributions for each subsequent subpopulation. Each final group contained at least 30 individuals, both male and female. Statistical differences in weight (ANOVA) are further discussed in the text.

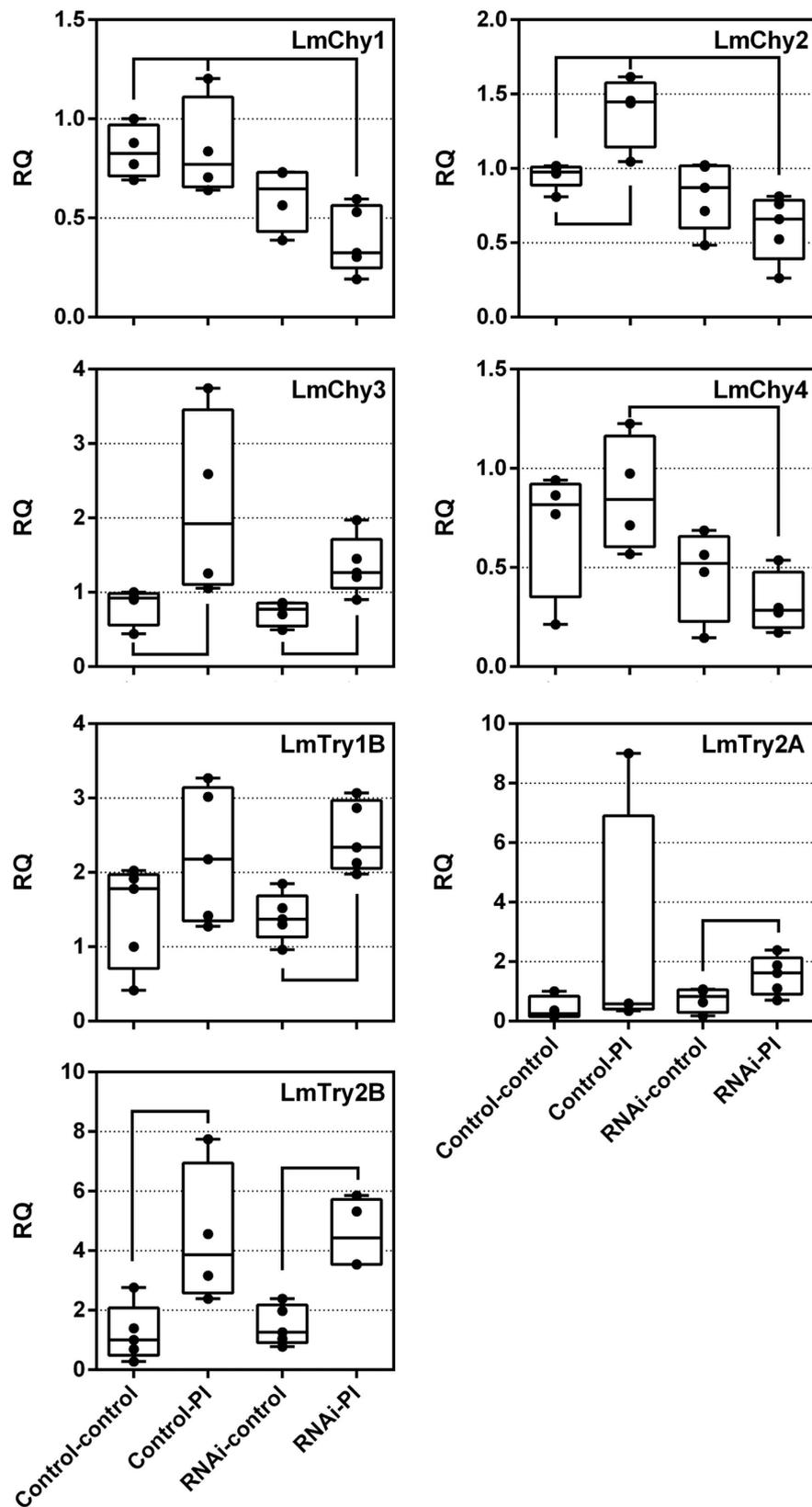
*LmChy1*, *LmChy2* and *LmChy4* tightly cluster in a phylogenetic analysis of *L. migratoria* serine proteases<sup>13</sup>. This could indicate that their expression might be under the same regulatory control mechanism. Lowered expression of several proteases, normally contributing significantly to the proteolytic digestion, could help explain the reduced weight gain observed in RNAi animals fed with the PI-containing diet. Nonetheless, upregulation of some proteases was retained, suggesting that sufficient proteolysis remained to ensure the insects' survival, albeit at a slightly slower growth rate.

#### Juvenile hormone signalling has an effect on the expression of *LmChy1*, *LmChy2* and *LmChy4*.

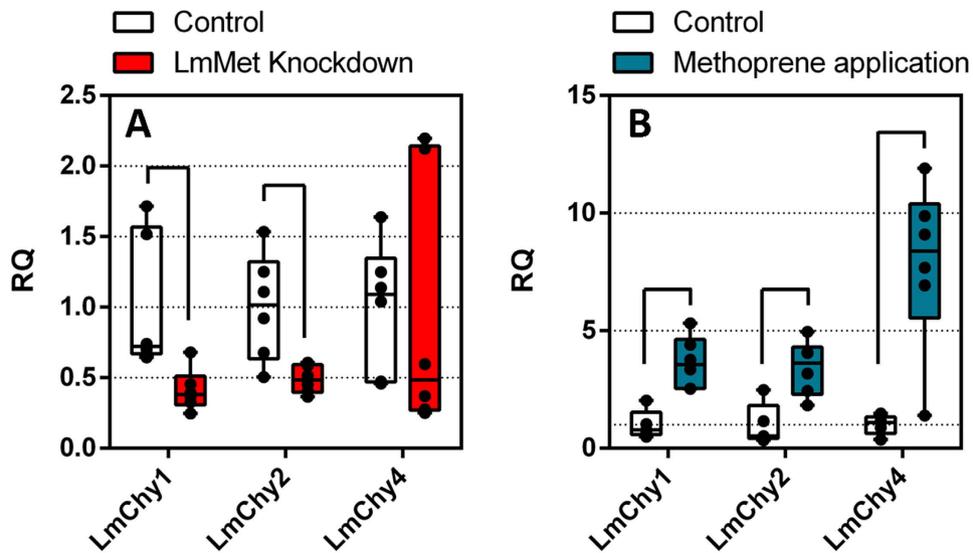
The reported ability of hexamerins to bind and transport JH in the hemolymph<sup>19</sup>, and the effects of a knockdown on *LmChy1*, *LmChy2*, and *LmChy4* expression during a PI-induced compensatory response led us to further investigate the possible role of JH in the regulation of the proteolytic digestion. We approached this hypothesis in two different ways. First, by interfering with JH signalling, through the knockdown of the sole known JH receptor, methoprene tolerant (*LmMet*)<sup>47</sup>. Second, by stimulating JH signalling through topical application of methoprene, which acts as JH analogue.

An RNAi-mediated knockdown of the *Met* gene [GenBank: KF471131.1] was performed using an identical dsRNA construct as reported by Guo *et al.*<sup>48</sup>. An average knockdown efficiency of 50% could be observed in the brain. However, reaching similar knockdown values in the midgut proved difficult (Supp. Fig. S6). Nevertheless, effects on protease expression were still clearly detectable. Figure 7A shows that the knockdown of *LmMet* results in significantly decreased transcript levels of the chymotrypsin genes, *LmChy1*, *LmChy2* (MWW;  $p < 0.01$ ), while transcript levels of *LmChy4* show a strong trend in the same direction. Interestingly, this is the same group of sequences that was affected by the hexamerin knockdown. No changes in transcript levels of any other serine proteases were detected (Supp. Fig. S7). Treatment of locust nymphs with methoprene results in the opposite effect. As shown in Fig. 7B, the transcript levels of *LmChy1*, *LmChy2* and *LmChy4* are clearly induced in the methoprene treated locusts compared to the control group (MWW;  $p < 0.01$ ). Other serine proteases were again not affected by the treatment (Supp. Fig. S7). In addition, we also observed a significant stimulation in the expression of the hexamerin genes, *LmHex 2*, *LmHex 3*, and *LmHex 6* in the midgut, confirming the successful application of methoprene and subsequent stimulation of JH signalling in the insects (Supp. Fig. S7). The transcript levels of *LmMet* did not differ.

Juvenile hormone is one of the most studied insect hormones. It has several well-described functions throughout the lifespan of all insects, but it is best known for its role as regulator of larval moulting and reproductive maturation<sup>49</sup>. In female insects, the cross talk between JH and feeding, mainly through the insulin and Target of rapamycin (TOR) nutrient sensing pathways, is vital for the production of yolk proteins in the fat body. The underlying mechanisms of this interaction however still remain largely elusive<sup>50</sup>. Although these findings indicate a link between JH and feeding, the direct involvement of JH in the regulation of the digestive system hasn't been studied in great detail yet. There exist scarce reports on certain digestive proteases seemingly under the regulatory control of JH<sup>41–44,51</sup>. However, the best-studied case of JH involvement in the digestive system of an insect species comes from the mosquito, *Aedes aegypti*. It is well established that the ingestion of a blood meal in the female mosquito takes place in two distinct phases. A first stage when a small amount of early trypsins initiates the proteolytic digestion is always followed by a second stage when much larger amounts of late trypsins continue this digestion. Studies have shown that the transcription of the early trypsins is under strict regulatory control of JH<sup>52</sup>. Based on feeding behaviour, insects can roughly be subdivided into two major categories. There are the non-continuous feeders (predators and hematophagous insects, such as mosquitoes), and the more continuous feeders (e.g. Coleoptera, Lepidoptera, Orthoptera). In general, it is believed that the digestive process in



**Figure 6. Relative transcript levels of selected proteases after knockdown of *LmHex2*, *LmHex3*, and *LmHex6*, plus PI ingestion.** Transcript levels in the midgut were normalized against two reference genes, *rp49* and *rps13*. Box plots are presented based on five pools of over five individuals. RQ = relative quantity. Relevant statistically significant differences are visualized by connecting black lines ( $p < 0.05$ , one way ANOVA).



**Figure 7.** Relative transcript levels of *LmChy1*, *LmChy2* and *LmChy4* after *LmMet* knockdown (A) and after topical application of methoprene (B). Transcript levels in the midgut were normalized against two reference genes, *rp49* and *rps13*. Box plots are presented based on six pools of at least four individuals. RQ = relative quantity. Significant differences are visualized by connecting black lines ( $p < 0.05$ , MWW).

continuous feeders is more or less constitutive, whereas in discontinuous feeders, the digestive system is more tightly regulated<sup>53,54</sup>. Interestingly, phylogenetic analyses have shown a clear divergence of insect trypsin and chymotrypsins. Dipteran early trypsin appear to be more related to an ancient trypsin molecule, while late trypsin show much more sequence resemblance with insect chymotrypsins<sup>13,55</sup>. In addition, *LmChy1*, *LmChy2* and *LmChy4* also appear to have evolved from a trypsin ancestor<sup>13</sup>. The present data therefore indicate that the digestive system of continuous feeders, like locusts, could also be regulated up to a certain level by JH, and that this function of JH on certain digestive enzymes could be present in a much broader group of insects. This leads us to hypothesize an evolutionary conserved role of JH in the digestive system in insects, but has to be further investigated in more species.

## Conclusions

Exploiting natural plant defence mechanisms, such as antinutritional PI, is an attractive option to generate future insect resistant crops. Nevertheless, several pest species possess a remarkably flexible digestive system, allowing them to circumvent the presence of PI in their gut and avoid developmental defects. Our previous research demonstrated that an increase in proteolytic enzyme expression could already be observed four hours after the dietary uptake of SBBI and SBTi (1%) in *L. migratoria*. Here, we therefore investigated large scale transcriptional changes that occur in the midgut and gastric caeca of fifth instar *L. migratoria* nymphs, four hours after dietary uptake of SBBI and SBTi. In total, the levels of 264 transcripts were found to be altered in these conditions. The results indicate that fewer resources will be invested in immunity and structural integrity of the gut, to be able to compensate for the presence of the PI in the intestinal tract, through adjusting the digestive enzyme profile. Moreover, a group of related upregulated transcripts suggests hexamerin-like proteins likely fulfil an important role in the compensation process. The specific function of hexamerin-like proteins as JHBPs led us to further examine the possible role of hexamerins and JH in the regulation of proteolytic digestion in the migratory locust. Our combined results clearly suggest a role for both hexamerins and JH in the regulation of the expression of a specific set of proteolytic enzyme encoding genes in the migratory locust. Under regular dietary conditions, JH signalling (via the methoprene tolerant receptor) elicits a significant stimulatory effect on the expression of a specific group of chymotrypsin-like serine proteases: *LmChy1*, *LmChy2* and *LmChy4*. However, when their diet contains elevated levels of PI, more proteases are needed. One plausible strategy to meet these demands relies on a JH-stimulated overexpression of digestive proteases. In order to capture more JH in the digestive tract, hexamerins (or other JHBPs) may act as specific carriers of the hormone. Hence, we hypothesize that after PI uptake, transcript levels of hexamerins are swiftly upregulated in the gut to assist the JH-stimulated expression of certain proteases.

## Materials and Methods

**Rearing of animals and sample collection.** Locusts (*Locusta migratoria*) were reared under crowded conditions with controlled temperature of 32 °C, light (14 h photoperiod) and relative humidity (40–60%). Locusts were fed daily with grass. For experiments, nymphs were developmentally synchronized on the day of the 5th larval molt (Day 0). For extraction of total RNA, tissues were dissected in *L. migratoria* Ringer's solution (composition: 9.82 g/l NaCl; 0.48 g/l KCl; 0.19 g/l NaH<sub>2</sub>PO<sub>4</sub>; 0.25 g/l NaHCO<sub>3</sub>; 0.73 g/l MgCl<sub>2</sub>; 0.32 g/l CaCl<sub>2</sub>; pH 6.5) and immediately transferred to liquid nitrogen. Samples were stored at –80 °C until further processing.

**RNA extraction and cDNA synthesis.** The Lipid tissue extraction kit (Qiagen, Hilden, Germany) was utilized to extract RNA from dissected tissues. DNaseI treatment was performed to remove traces of genomic DNA contamination. Concentration of extracted RNA was assessed using a NanoDrop ND-1000 UV-VIS Spectro-photometer. Quality of the extracted RNA to be used in the microarray hybridizations was further assessed using the Agilent 2100 Bioanalyzer. In case of cDNA synthesis, equal quantities of RNA were used as template. Synthesis of cDNA was performed using the Superscript III reverse transcriptase kit (Invitrogen, Carlsbad, USA), and random hexamer primers (Invitrogen) and dNTPs (Roche, Basel, Switzerland), following the manufacturer's protocol.

**Feeding experiment microarray.** After two days of normal feeding, individual 5<sup>th</sup> instar locusts received a single meal of 1.5% agar based artificial diet containing 2.4% Wesson salt mixture, 0.5% linoleic acid, 0.6% cholesterol, 18.8% Vanderzant vitamin mixture, 50% cellulose, 14% dextrin, 8.1% caseine, 2.8% peptone and 2.8% albumin. One group received this meal complemented with plant derived PI (SBBI; Soybean Bowman Birk Inhibitor, and SBTI; Soybean Trypsin Inhibitor: 1% of total protein content; as previously described<sup>13</sup>), while another group received this meal complemented with BSA to achieve equal total protein contents in the diets. 4 hours after ingestion of the meal, midgut and caeca samples were dissected for RNA extraction. Tissues for RNA extraction were dissected in 3 pools of at least 5 individuals for every condition. For confirmation of the microarray results by qPCR, a separate experiment was conducted in a similar setup.

**Labeling of samples.** Starting from total RNA, fluorescently labeled cRNA was generated by combining equal quantities of midgut and caeca RNA. Labeling of the samples was performed with the Quick Amp Labeling Kit (Agilent Technologies, Santa Clara, USA), according to the manufacturer's instructions. In brief, 1 µg of RNA from each sample was with a Cy5 (red) and Cy3 (green) fluorescent dye. mRNA was reverse transcribed to cDNA by using a poly(dT) primer that is coupled to an antisense T7 promoter. The reverse transcriptase also catalyzes synthesis of the second cDNA strand. Prior to cDNA synthesis, a Spike A or Spike B mix was added to all samples to be labeled with Cy3 or Cy5, respectively. These Spike mixes contain polyadenylated transcripts from the adenovirus E1A, that are premixed in various quantities and ratios. They are used as a control for the workflow. In a second step, the second cDNA strand functions as template for the synthesis of cRNA, for which fluorescently labeled dCTPs (Cy3 or Cy5) are provided. The fluorescently labeled samples were subsequently purified by means of the RNeasy Mini Kit (Qiagen), according to the manufacturer's instructions. For all samples, the yield and activity were higher than the minimum requirements of 825 ng total cRNA and 8 pmol Cy per µg cRNA, as measured with a NanoDrop spectro-photometer.

**Design of the microarray.** To reduce the number of transcripts to be analysed, we eliminated all transcripts that were expressed in neither brain- nor gut tissues, based on a self-self hybridisation protocol (Spit *et al.*<sup>20</sup>). From the 48802 unique ESTs available for *L. migratoria*, 12933 were omitted. This allowed the 35869 remaining transcripts to be spotted on a 4 × 44K Custom Gene Expression Microarray slide (Agilent technologies). For each of the resulting unique sequences, the previously determined best probe was spotted in addition to control features. The remaining spots were filled with random doubles to reach the platform maximum of 44K. For the analysis of three pooled control samples and three pooled experimental samples, an n + 2 A-optimal design was chosen (n = 6). This means that 8 hybridizations were performed in a balanced and even design<sup>56</sup>. The data files and platform information have been deposited in the GEO repository of NCBI, and are accessible under the accession number GSE79995.

**Hybridization, scanning and data analysis.** Hybridizations were performed at 65 °C for 17 h. Next, the arrays were washed: twice in GE Wash Buffer 1 (Agilent technologies) for 1 min each, in GE Wash Buffer 2 (Agilent technologies) at 37 °C for 1 min, in acetonitrile (Sigma Aldrich, Saint Louis, USA) for 1 min and in Stabilization and Drying Solution (Agilent technologies) for 30 sec. Subsequently, the microarray slides were scanned using a Genepix Personal 4100A confocal scanner (Molecular devices, Sunnyvale, USA) at a resolution of 5 µm and excitation wavelengths of 635 nm and 532 nm. The photomultiplier tube (PMT) voltages for separate wavelengths were adjusted to obtain an overall green/red ratio as close as possible to 1. Images were processed using GenePix Pro 6.0 software (Molecular devices) for spot identification and quantification of the fluorescent signal intensities. The data obtained from the eight hybridizations were used for statistical analysis by means of the R package limma<sup>57</sup>. Spots for which  $FG < BG + 2SD$  on all arrays, were deleted before analysis. Median intensities were background corrected using a normal-exponential convolution model and were normalized by Locally Weighed Scatterplot Smoothing (Loess)<sup>58</sup>. For each target transcript a linear model was fitted to the intensity ratios. Next, transcripts were ranked in order of evidence of differential expression using an empirical Bayes method and 'PI-fed vs control' contrasts were evaluated by the linear models<sup>59</sup>. False discovery rate (FDR) control was obtained by calculating adjusted p-values according to the Benjamini-Hochberg procedure and a cut-off was set at a FDR of  $p < 0.10$ <sup>60</sup>. Differentially regulated transcripts were functionally annotated using the InterProScan (IPS) tool that is integrated in the Blast2GO software<sup>25,26</sup>. In addition, Gene ontology (GO) annotation was performed using the Blast2GO software<sup>61</sup>. The best blast hit was used for sequence description.

**Transcript assembly and phylogenetic analysis of hexamerin-like proteins.** Transcripts putatively encoding hexamerin-like proteins that were identified from an upregulated set of sequences after PI feeding were assembled using DNA Baser software. Transcripts that showed at least 98% nucleotide identity were considered belonging to the same sequence. Signal peptide predictions were made using SignalP4.1<sup>62</sup> and multiple sequence alignments of insect hexamerin-like proteins were made using MAFFT alignment software<sup>63</sup>. Phylogenetic analysis was performed using MEGA version 6<sup>64</sup> with aligned amino acid sequences that were manually trimmed

to obtain regions with the highest homology. The evolutionary history was inferred by using the Maximum Likelihood method based on the Jones w/freq. model. The bootstrap consensus tree was inferred from 100 replicates. Positions containing gaps were treated as missing data and were eliminated. Protein sequences of hemocyanin from the decapod *Panulirus interruptus* were used to root the tree.

**Synthesis of dsRNA for RNAi studies.** Double stranded RNAs for *gfp* (589 bp), *LmHex2* (715 bp), *LmHex3* (715 bp), *LmHex6* (735 bp) and *LmMet* (421 bp) were synthesized using the MEGAscript RNAi kit (Thermo Fisher Scientific, Waltham, USA), according to the manufacturer's instructions. Sense and antisense strands were synthesized from a PCR-generated DNA template containing T7 RNA Polymerase promoters on both 5' ends of the primers. Primer sequences are presented in Supp. Table S8. For *LmHex2* and *LmHex3* the DNA template could be generated in a single PCR reaction using the same primers. cDNA derived from fat body of fifth larval *L. migratoria* was used as a template for the PCR reactions. The PCR reaction was performed using REDTaq ready mix (Sigma) as a source of DNA polymerase, dNTPs and PCR buffer. Amplification of a PCR product of the desired length was checked by performing 1% gel electrophoresis. For the production of *gfp* dsRNA, an amplified *gfp* cDNA fragment was cloned in both sense and antisense direction in a TOPO 4.1 sequencing vector (Thermo Fisher Scientific), containing a T7 promoter site. Both RNA strands were then synthesized using T7 enzyme mix (Thermo Fisher Scientific) and sense and antisense strand were annealed subsequently. After dsRNA production, the DNA and ssRNA molecules were removed by nuclease treatment and dsRNA was further purified by solid phase adsorption purification, according to the manufacturer's instructions. The concentration of the produced dsRNA was estimated using a NanoDrop ND-1000 UV-VIS Spectro-photometer, and 1% agarose gel electrophoresis was performed to assess the integrity of the dsRNA.

**Simultaneous knockdown of hexamerin like sequences and PI feeding experiment.** At the first day after ecdysis (day 0), 130 5<sup>th</sup> instar nymphs were weighted and equally divided in two groups. One group was injected with 600 ng from a mixture of dsRNA for *LmHex2*, *LmHex3* and *Lmhex6*, while the other group was injected with 600 ng of dsRNA for *gfp* as a control. dsRNA was injected directly in the body cavity using a 710RN 100  $\mu$ l syringe (Hamilton). For three days locusts were fed daily with fresh grass. At day 3 locusts were weighted and both groups of nymphs were again equally subdivided into two groups each, which were transferred to an artificial diet, containing SBBI or SBTI (1%) or BSA, similar to the microarray feeding experiment. Consequently, these divisions resulted in four different treatments: Control-control, control-PI, RNAi-control and RNAi-PI, where the first part of the name refers to the dsRNA treatment and the second to the diet that was received starting from day 3. For the next 4 days locusts received freshly prepared artificial diet daily. At day 7 locusts were weighted and for each condition, midguts and brains were dissected for RNA extraction in 5 pools of at least 5 individuals.

**Met knockdown experiment.** At the first day after ecdysis (day 0), a total of 85 5<sup>th</sup> instar nymphs were divided in two groups. One group was injected with 600 ng of *LmMet* dsRNA, while the other group was injected with 600 ng of dsRNA for *gfp* as a control. Since initial knockdowns of the receptor proved difficult, locusts received a boost injection at day 2 to maximize the effect. During the experiment, locusts were fed *ad libitum* with grass. On day 4, midguts were collected 3 hours post feed for total RNA extraction. Tissues for RNA extraction were dissected in 6 pools of at least 5 individuals.

**Methoprene treatments.** At the first day after ecdysis (day 0), 60 male nymphs were divided in two groups. An experimental group of 30 nymphs was topically treated with the JH III hormone mimic, methoprene (Sigma Aldrich), dissolved in acetone, while a control group received only acetone. The locusts were treated with a daily dose of 100  $\mu$ g for four consecutive days. During the experiment, locusts were fed with grass *ad libitum*. On day 4, midguts were collected 3 hours post feed for total RNA extraction. Tissues for RNA extraction were dissected in 5 pools of at least 4 individuals.

**RT-qPCR.** Quantitative real time RT-PCR was performed according to MIQE requirements<sup>65</sup>. Primer express software (Applied Biosystems) was used to design primers. All primer sequences are displayed as Supp. Table S8. All primer pairs were validated with a standard curve based on a serial dilution of cDNA to determine the primer annealing efficiency. A dissociation protocol was performed to detect the presence of primer dimers and ensure production of a single PCR product. For all tested transcripts, only a single melting peak was found. qPCR reactions using a SYBR green protocol (Invitrogen) were performed in duplicate in 96 well plates according to the manufacturer's instructions, and were analyzed by the StepOne Plus System (ABI Prism, Applied Biosystems). Relative transcript levels were calculated using the  $\Delta\Delta$ Ct method<sup>66</sup>. To correct for sample to sample variation, expression was normalized against *rp49* and *rps13*, the two most stably expressed reference genes, as determined with geNorm<sup>67</sup>.

**Statistical analyses.** Weight gain and gene expression data were tested for normality by the D'Agostino-Pearson omnibus normality test<sup>68</sup>. In case of weight data, frequency distributions were calculated and non-linear Gaussian regression was used to fit the data. When data were distributed normally, significant differences between two groups were determined by student t-tests. Significant differences for multiple groups were determined by ANOVA. When data were not distributed normally, nonparametric alternatives were used for comparing two or multiple groups, i.e. Mann-Whitney-Wilcoxon (MWW) or Kruskal-Wallis tests, respectively. All statistical analyses were performed using GraphPad Prism software.

**Availability of data and materials.** The datasets supporting the conclusions of this article are available in the GEO repository of NCBI and are accessible under the accession number GSE79995, or can be found on the following website <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE79995>. In addition, several datasets supporting the conclusions of this article are included within the article and its additional files.

## References

- Huang, J. H., Jing, X. & Douglas, A. E. The multi-tasking gut epithelium of insects. *Insect Biochem. Mol. Biol.* **67**, 15–20 (2015).
- Mello, M. & Silva-Filho, M. Plant-insect interactions: an evolutionary arms race between two distinct defense mechanisms. *Brazilian J. Plant Physiol.* **14**, 71–81 (2002).
- Jongsma, M. A. & Bolter, C. The adaptation of insects to plant protease inhibitors. *J. Insect Physiol.* **43**, 885–895 (1997).
- Gatehouse, J. a. Prospects for using proteinase inhibitors to protect transgenic plants against attack by herbivorous insects. *Curr. Protein Pept. Sci.* **12**, 409–416 (2011).
- Zhu-Salzman, K. & Zeng, R. Insect response to plant defensive protease inhibitors. *Annu. Rev. Entomol.* **60**, 233–252 (2015).
- Gruden, K. *et al.* The cysteine protease activity of Colorado potato beetle (*Leptinotarsa decemlineata* Say) guts, which is insensitive to potato protease inhibitors, is inhibited by thyroglobulin type-1 domain inhibitors. *Insect Biochem. Mol. Biol.* **28**, 549–560 (1998).
- Markwick, N. P., Laing, W. A., Christeller, J. T., McHenry, J. Z. & Newton, M. R. Overproduction of digestive enzymes compensates for inhibitory effects of protease and alpha-amylase inhibitors fed to three species of leafrollers (Lepidoptera: Tortricidae). *J. Econ. Entomol.* **91**, 1265–1276 (1998).
- Moon, J., Salzman, R. a, Ahn, J.-E., Koiwa, H. & Zhu-Salzman, K. Transcriptional regulation in cowpea bruchid guts during adaptation to a plant defence protease inhibitor. *Insect Mol. Biol.* **13**, 283–291 (2004).
- Bown, D. P., Wilkinson, H. S. & Gatehouse, J. A. Differentially regulated inhibitor-sensitive and insensitive protease genes from the phytophagous insect pest, *Helicoverpa armigera*, are members of complex multigene families. *Insect Biochem. Mol. Biol.* **27**, 625–638 (1997).
- de Oliveira, C. F. R. *et al.* Insensitive trypsins are differentially transcribed during *Spodoptera frugiperda* adaptation against plant protease inhibitors. *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* **165**, 19–25 (2013).
- Bown, D. P., Wilkinson, H. S. & Gatehouse, J. A. Regulation of expression of genes encoding digestive proteases in the gut of a polyphagous lepidopteran larva in response to dietary protease inhibitors. *Physiol. Entomol.* **29**, 278–290 (2004).
- Chapuis, M.-P. *et al.* Outbreaks, gene flow and effective population size in the migratory locust, *Locusta migratoria*: a regional-scale comparative survey. *Mol. Ecol.* **18**, 792–800 (2009).
- Spit, J. *et al.* Effects of different dietary conditions on the expression of trypsin- and chymotrypsin-like protease genes in the digestive system of the migratory locust, *Locusta migratoria*. *Insect Biochem. Mol. Biol.* **48**, 100–109 (2014).
- Chi, Y. H. *et al.* Cowpea bruchid midgut transcriptome response to a soybean cystatin—costs and benefits of counter-defence. *Insect Mol. Biol.* **18**, 97–110 (2009).
- Govind, G. *et al.* Unbiased transcriptional comparisons of generalist and specialist herbivores feeding on progressively defenseless *Nicotiana attenuata* plants. *PLoS One* **5**, e8735 (2010).
- Petek, M. *et al.* a complex of genes involved in adaptation of *Leptinotarsa decemlineata* larvae to induced potato defense. *Arch. Insect Biochem. Physiol.* **79**, 153–181 (2012).
- Oppert, B., Elpidina, E. N., Toutges, M. & Mazumdar-Leighton, S. Microarray analysis reveals strategies of *Tribolium castaneum* larvae to compensate for cysteine and serine protease inhibitors. *Comp. Biochem. Physiol. Part D. Genomics Proteomics* **5**, 280–287 (2010).
- Kuwar, S. S., Pauchet, Y., Vogel, H. & Heckel, D. G. Adaptive regulation of digestive serine proteases in the larval midgut of *Helicoverpa armigera* in response to a plant protease inhibitor. *Insect Biochem. Mol. Biol.* **59**, 18–29 (2015).
- Braun, R. & Wyatt, G. R. Sequence of the Hexameric Juvenile Hormone-binding Protein from the Hemolymph of *Locusta migratoria*. *J. Biol. Chem.* **271**, 31756–31762 (1996).
- Spit, J., Badisco, L., Vergauwen, L., Knapen, D. & Vanden Broeck, J. Microarray-based annotation of the gut transcriptome of the migratory locust, *Locusta migratoria*. *Insect Mol. Biol.* **00** (2016). <http://www.ncbi.nlm.nih.gov/pubmed/27479692>
- Lam, W., Coast, G. M., Rayne, C., Lam, R., W., G. M. & Rayne, R. C. Characterisation of multiple trypsins from the midgut of *Locusta migratoria*. *Insect Biochem. Mol. Biol.* **30**, 85–94 (2000).
- Lam, W., Coast, G. M. & Rayne, R. C. Isolation and characterisation of two chymotrypsins from the midgut of *Locusta migratoria*. *Insect Biochem. Mol. Biol.* **29**, 653–660 (1999).
- Terra, W. R. & Ferreira, C. In *Insect Molecular Biology and Biochemistry* 365–418, doi: 10.1016/B978-0-12-384747-8.10011-X (2012).
- Hand, S. C. & Hardewig, I. Downregulation of cellular metabolism during environmental stress: mechanisms and implications. *Annu. Rev. Physiol.* **58**, 539–563 (1996).
- Conesa, A. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).
- Quevillon, E. *et al.* InterProScan: protein domains identifier. *Nucleic Acids Res.* **33**, W116–W120 (2005).
- Lara, P. *et al.* Adaptation of *Spodoptera exigua* (Lepidoptera: Noctuidae) to barley trypsin inhibitor BTI-CMe expressed in transgenic tobacco. *Transgenic Res.* **9**, 169–178 (2000).
- Spit, J. *et al.* Growth-inhibition effects of pacifastin-like peptides on a pest insect: the desert locust, *Schistocerca gregaria*. *Peptides* **34**, 251–257 (2012).
- Sutherland, P. W. *et al.* Ultrastructural changes to the midgut of the black field cricket (*Teleogryllus commodus*) following ingestion of potato protease inhibitor II. *J. Insect Physiol.* **48**, 327–336 (2002).
- Li, H.-M. *et al.* Bowman-Birk inhibitor affects pathways associated with energy metabolism in *Drosophila melanogaster*. *Insect Mol. Biol.* **19**, 303–313 (2010).
- Kanost, M. R., Jiang, H. & Yu, X.-Q. Innate immune responses of a lepidopteran insect, *Manduca sexta*. *Immunol. Rev.* **198**, 97–105 (2004).
- Huang, L.-H. & Kang, L. Cloning and interspecific altered expression of heat shock protein genes in two leafminer species in response to thermal stress. *Insect Mol. Biol.* **16**, 491–500 (2007).
- Shen, Y. *et al.* Cloning and expression analysis of six small heat shock protein genes in the common cutworm, *Spodoptera litura*. *J. Insect Physiol.* **57**, 908–914 (2011).
- Zhang, J. *et al.* Genomics-based approaches to screening carboxylesterase-like genes potentially involved in malathion resistance in oriental migratory locust (*Locusta migratoria manilensis*). *Pest Manag. Sci.* **67**, 183–190 (2011).
- Liu, Y., Salzman, R. a, Pankiw, T. & Zhu-Salzman, K. Transcriptional regulation in southern corn rootworm larvae challenged by soyacystatin N. *Insect Biochem. Mol. Biol.* **34**, 1069–1077 (2004).
- Burmester, T. Evolution and function of the insect hexamerins. *Eur. J. Entomol.* **96**, 213–225 (1999).
- Kang, L. *et al.* The analysis of large-scale gene expression correlated to the phase changes of the migratory locust. *Proc. Natl. Acad. Sci. USA* **101**, 17611–17615 (2004).

38. Badisco, L. *et al.* Transcriptome analysis of the desert locust central nervous system: production and annotation of a *Schistocerca gregaria* EST database. *PLoS One* **6**, e17274 (2011).
39. Zhou, X., Tarver, M. R. & Scharf, M. E. Hexamerin-based regulation of juvenile hormone-dependent gene expression underlies phenotypic plasticity in a social insect. *Development* **134**, 601–610 (2007).
40. Chiou, S. J. *et al.* Cloning of the cDNA encoding Scg-SPRP, an unusual Ser-protease-related protein from vitellogenic female desert locusts (*Schistocerca gregaria*). *Insect Biochem. Mol. Biol.* **28**, 801–808 (1998).
41. Sui, Y.-P., Wang, J.-X. & Zhao, X.-F. The impacts of classical insect hormones on the expression profiles of a new digestive trypsin-like protease (TLP) from the cotton bollworm, *Helicoverpa armigera*. *Insect Mol. Biol.* **18**, 443–452 (2009).
42. Zhang, C. *et al.* A chymotrypsin-like serine protease cDNA involved in food protein digestion in the common cutworm, *Spodoptera litura*: Cloning, characterization, developmental and induced expression patterns, and localization. *J. Insect Physiol.* **56**, 788–799 (2010).
43. Bian, G., Raikhel, A. & Zhu, J. Characterization of a juvenile hormone regulated chymotrypsin like serine protease gene in aedes Aegypti mosquito. *Insect Biochem. Mol. Biol.* **38**, 190–200 (2008).
44. Cornette, R., Hayashi, Y., Koshikawa, S. & Miura, T. Differential gene expression in response to juvenile hormone analog treatment in the damp-wood termite *Hodotermopsis sjostedti* (Isoptera, Archotermopsidae). *J. Insect Physiol.* **59**, 509–518 (2013).
45. Cristino, A. S. *et al.* Organization, evolution and transcriptional profile of hexamerin genes of the parasitic wasp *Nasonia vitripennis* (Hymenoptera: Pteromalidae). *Insect Mol. Biol.* **19** Suppl 1, 137–146 (2010).
46. Tang, B., Wang, S. & Zhang, F. Two storage hexamerins from the beet armyworm *Spodoptera exigua*: cloning, characterization and the effect of gene silencing on survival. *BMC Mol. Biol.* **11**, 65 (2010).
47. Charles, J.-P. *et al.* Ligand-binding properties of a juvenile hormone receptor, Methoprene-tolerant. *Proc. Natl. Acad. Sci.* **108**, 21128–21133 (2011).
48. Guo, W. *et al.* Juvenile Hormone-Receptor Complex Acts on Mcm4 and Mcm7 to Promote Polyploidy and Vitellogenesis in the Migratory Locust. *PLoS Genet.* **10** (2014).
49. Simonet, G. *et al.* Neuroendocrinological and molecular aspects of insect reproduction. *J. Neuroendocrinol.* **16**, 649–659 (2004).
50. Badisco, L., Van Wielendaale, P. & Vanden Broeck, J. Eat to reproduce: a key role for the insulin signaling pathway in adult insects. *Front. Physiol.* **4**, 202 (2013).
51. Sindhu, A. & Nair, V. S. K. Influence of juvenile hormone analogue on food consumption and digestive enzyme activities in *Spodoptera mauritia* Boisdu. *Indian J. Exp. Biol.* **42**, 491–494 (2004).
52. Noriega, F. G. & Wells, M. A. A molecular view of trypsin synthesis in the midgut of *Aedes aegypti*. *J. Insect Physiol.* **45**, 613–620 (1999).
53. Lehane, M. J., Blakemore, D., Williams, S. & Moffatt, M. R. Regulation of digestive enzyme levels in insects. *Comp. Biochem. Physiol. Part B Biochem. Mol. Biol.* **110**, 285–289 (1995).
54. Lehane, M. & Billingsley, P. Biology of the insect midgut. *Parasitol. Today* **13**, 463 (1996).
55. Marshall, S. D. G. *et al.* Serine proteases identified from a *Costelytra zealandica* (White) (Coleoptera: Scarabaeidae) midgut EST library and their expression through insect development. *Insect Mol. Biol.* **17**, 247–259 (2008).
56. Knapen, D., Vergauwen, L., Laukens, K. & Blust, R. Best practices for hybridization design in two-colour microarray analysis. *Trends Biotechnol.* **27**, 406–414 (2009).
57. Smyth, G. K. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* 397–420, doi: 10.1007/0-387-29362-0\_23 (2005).
58. Smyth, G. K. & Speed, T. Normalization of cDNA microarray data. *Methods* **31**, 265–273 (2003).
59. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, Article3 (2004).
60. Benjamini, Y., Kenigsberg, E., Reiner, A. & Yekutieli, D. FDR adjustments of Microarray Experiments (FDR-AME). *math. tau. ac.il* 1–3 (2005).
61. The Gene Ontology Consortium. Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
62. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
63. Katoh, K., Kuma, K., Toh, H. & Miyata, T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511–518 (2005).
64. Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739 (2011).
65. Bustin, S. A. *et al.* MIQE précis: Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. *BMC Mol. Biol.* **11**, 74 (2010).
66. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the  $2^{-\Delta\Delta CT}$  Method. *Methods* **25**, 402–408 (2001).
67. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, research0034.1–research0034.11 (2002).
68. D'Agostino, R. B. In *Goodness-Of-Fit Techniques* 367–419 (1986).
69. Mitchell, A. *et al.* The InterPro protein families database: The classification resource after 15 years. *Nucleic Acids Res.* **43**, D213–D221 (2015).

## Acknowledgements

The authors are very grateful to Roger Jonckers, Evelien Herinckx, and all lab members who contributed to the maintenance of the locust colony. In addition, the authors would like to express their sincerest gratitude to Heleen Verlinden, Julie Toback, Rik Verdonck, Sven Zels and Senne Dillen for their help during dissections.

## Author Contributions

J.S., L.B., M.H., L.V., D.K. and J.V.B. designed the experiments. J.S., M.H. and E.V. performed the experiments. J.S., L.B. and M.H. analyzed the data. J.S., M.H. and J.V.B. wrote the paper. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Spit, J. *et al.* Transcriptional analysis of the adaptive digestive system of the migratory locust in response to plant defensive protease inhibitors. *Sci. Rep.* **6**, 32460; doi: 10.1038/srep32460 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016