



OPEN

## Novel insights into plant defensin ingestion induced metabolic responses in the polyphagous insect pest *Helicoverpa armigera*

Javed A. Mulla & Vaijayanti A. Tamhane  

Lepidopteran insect pest *Helicoverpa armigera* is one of the most destructive pests of crop plants and several biotechnological approaches are being developed for its control. Plant defensins are small cationic and cysteine-rich peptides that play a role in plant defense. Ingestion of a defensin from *Capsicum annuum* (CanDef-20) induced a dose-dependent reduction in larval and pupal mass, delayed metamorphosis and also severely reduced fecundity and fertility in *H. armigera*. To understand the molecular mechanisms of CanDef-20 ingestion-mediated antibiosis in *H. armigera* larvae, a comparative transcriptomics analysis was carried out. Predominant downregulation of GOs represents serine-type endopeptidases, structural constituents of ribosomes and integral membrane components and differential upregulation of ATP binding, nucleus and translation, while up-regulation of nucleic acid binding represented by transposable elements, were detected. Different isoforms of lipase, serine endopeptidase, glutathione S-transferase, cadherin, alkaline phosphatase and aminopeptidases were found to be upregulated as a compensatory response to CanDef-20 ingestion. In vitro enzyme assays and qPCR analysis of some representative genes associated with vital cellular processes like metamorphosis, food digestion and gut membrane indicated adaptive differential regulations in CanDef-20 fed *H. armigera* larvae. We conclude that CanDef-20 ingestion affects insect metabolism in a number of ways through its interaction with cell membrane, enzymes, cytoplasmic proteins and triggering transposon mobilization which are linked to growth retardation and adaptive strategies in *H. armigera*.

Insect pests lead to substantial yield losses in crop plants either by direct damage or by the spreading of diseases. Of the number of insect pests that threaten and attack crop plants *Helicoverpa armigera* is polyphagous and the most devastating<sup>1</sup>. Control measures like use of various pesticides and cry transgenic crop-based approaches have been used globally, though *H. armigera* evolves resistance<sup>2</sup>, leading to failure of these methods necessitating development of novel biological approaches for sustainable and environment-friendly pest control.

Polyphagous insects have developed multiple resistance mechanisms like the production of glutathione S-transferase, glucose oxidase, overexpression of insensitive protease and cytochrome P450 mono-oxygenase to cope with plant defenses<sup>3</sup>. The molecular mechanisms of resistance of Bt transgenics have been studied in recent years. The mutation in cadherin and ABCC2 transporter gene present in the brush border epithelium of *H. armigera* and *H. virescens* resulted in the resistance against Bt toxin<sup>4</sup>. In addition to this, altered expression of alkaline phosphatase (ALP)<sup>5</sup>, aminopeptidase N (APN)<sup>6</sup> and regulatory events in mitogen-activated protein kinase (MAPK)<sup>7</sup> were the resistance mechanisms used by lepidopteran insects against Bt toxin. These studies provide evidence of notable diversity and plasticity in insect metabolism to counter plant defenses.

Plants have evolved sophisticated regulatory networks to present constitutive and induced defense responses against herbivore attacks<sup>3</sup>. Induction of jasmonate (JA) and salicylic acid (SA)<sup>4</sup> pathways followed by the production of secondary metabolites, proteinase inhibitor (PIs)<sup>5</sup>, and antimicrobial peptides (AMPs) determine a pest-specific plant defense response. Effect of AMPs on insects is not clearly deciphered, and it is known that there is a much lower probability to generate resistance to AMPs in bacteria than to antibiotics<sup>8</sup>. Therefore, studying

Department of Biotechnology (Jointly Merged With Institute of Bioinformatics and Biotechnology (IBB)), Savitribai Phule Pune University, Pune, Maharashtra 411007, India. ✉email: vatamhane@unipune.ac.in; vatamhane@gmail.com

the insect's counter defense mechanism against plant defense molecules like defensin peptides, will help us in pyramiding defense molecules for better insect pest control<sup>9</sup>.

AMPs like defensins were less explored in the context of plant–insect *H. armigera* interactions though their role as antifungal and antibacterial peptides is well studied via their interactions with the cell membranes<sup>10</sup>, cytoplasmic proteins<sup>11</sup> and transmembrane channels<sup>12</sup>. Plant defensins are cationic peptides with ~ 45–50 amino acids with a conserved three-dimensional structure representing cysteine-stabilized  $\alpha\beta$  (CS $\alpha\beta$ ) fold, four to five disulfide bonds and  $\gamma$ -core motif<sup>13</sup>—the major determinant conferring antifungal and antibacterial activity<sup>14</sup>. Some studies confirm the direct cell membrane binding activity of  $\gamma$ -core motif<sup>15,16</sup>. In our previous work, we have discussed the growth inhibitory potential of *Capsicum annuum* defensin (CanDef-20) on *H. armigera*<sup>17</sup>. CanDef-20 caused around 15% retardation in larval mass and a 12% reduction in pupal mass though the mechanisms through which CanDef-20 induces physiological and morphological implications in the insect pest remained unclear. The present study attempts to decipher the mechanism of action of CanDef-20 in *H. armigera* through a transcriptomics approach.

In this study, we aimed (i) to understand dose-dependent effects of CanDef-20 ingestion in *H. armigera* (ii) elucidate physiological and molecular responses of *H. armigera* fed with recombinant CanDef-20 using a comparative de novo transcriptomics approach and (iii) to understand the correlations in gene expression levels with protein/enzyme activity and adaptive responses of *H. armigera* towards the ingestion of plant defensin CanDef-20. Understanding the role and mechanism of action of plant defensins in insect pests will be a step forward in designing sustainable pest control measures for agriculture.

## Materials and methods

**CanDef-20 amplification and source of insect culture.** The *C. annuum* (local variety Phule Jyoti) flower tissues were used for amplification and cloning of CanDef-20 gene as described previously<sup>17</sup>. The *C. annuum* seeds were procured from the local market and grown in a polyhouse. The molecular biology experiments like amplification of the defensin gene, cloning of genes and recombinant expression of proteins was performed after approval from the Institutional Bio-Safety Committee (IBSC-SPPU) as per the guidelines from Department of Biotechnology, Ministry of Science and Technology, Government of India <http://dbtbiosafety.nic.in/> or Guidelines and Handbook.

The *H. armigera* eggs were obtained from the National Bureau of Agricultural Important Resources (NBAIR), Bangalore. Eggs were allowed to hatch and neonates were transferred on an artificial diet as described earlier<sup>17</sup>. Only insects with one generation raised in the laboratory were used for bioassays with CanDef-20.

**Cloning and heterologous expression of CanDefensin (CanDef-20) gene.** The expression of recombinant CanDef-20 protein and dot blot assay were performed as described previously<sup>17</sup>. For protein expression, the *Pichia pastoris* strain with gene constructs (CanDef-20) or the empty vector (EV) in pPICZ alpha vector were used in buffered methanol complex (BMMY) medium (10 × Yeast Nitrogen Base and buffered with 1 M potassium phosphate buffer pH 6.0) (Hi-media, India). The recombinant proteins were obtained from culture supernatant after 90% ammonium sulphate precipitation and dialysis. Protein samples were subjected to SDS-PAGE using 15% (w/v) polyacrylamide gels<sup>18</sup>. The recombinant proteins expressed using this system carried a C-terminal 6X His and myc tag.

**Dot blot assay of CanDef-20.** The recombinant CanDef-20 was confirmed by dot blot assay using rabbit anti-myc primary antibody (Abcam, USA) directed against c-myc tag of the recombinantly expressed CanDef-20. Anti-rabbit secondary antibody was used from Western Blot Development kit (Bangalore genie, India). Finally, the spots were developed by incubating the membrane in BCIP/NBT solution.

***H. armigera* bioassay.** *H. armigera* larval bioassay was performed by using second day neonates maintained on an artificial diet (AD). The AD used in the feeding assay was made up of major and minor components containing ascorbic acid, sorbic acid, yeast extract, cholesterol, vitamins, methyl hydroxy benzoate and chickpea flour which are mixed with agar–agar boiled in water to form a semisolid feed<sup>19</sup>. Total 25 neonates were individually released on AD incorporated with 15.5 and 31.25  $\mu\text{g/ml}$  recombinant CanDef-20 and it was considered as set-1 in a separate small container with diameter of 3 cm. Set-2 with 50 neonates was parallelly run and consists of 62.25, 125, 250  $\mu\text{g/ml}$  recombinant CanDef-20. The AD (control) and EV *P. pastoris* expressed proteins (EV control) were individually run for respective sets. Larval mass, development, mortality, pupal mass, fecundity and fertility was recorded for up to 22 days during the bioassay. All the bioassay sets were maintained at 30 °C with 70%  $\pm$  10% humidity with a photoperiod of 16 h. Larval and pupal mass was recorded periodically. Pupae formed were separated in beakers and were monitored for emergence of moths. The moths were fed with 10% sugar solution and allowed to mate in 8:4 and 12:6 (female: male ratio) for set-1 and 2 respectively in a cage (30 × 20 × 15 cm) with a detachable mesh cloth on lid for egg collection. The number of eggs and neonates hatched were recorded.

**RNA isolation and Illumina sequencing.** Insect samples (fed with 125  $\mu\text{g/ml}$  recombinant CanDef-20 and EV control incorporated AD) were used for RNA isolation by using TruSeq RNA Sample Prep Kit v2 as per the manufacturer's protocol. Libraries were generated using TruSeq DNA PCR-Free kit following the manufacturer's recommendations. Sequencing was performed using Illumina HiSeq. 2000 system (Illumina, San Diego, CA, USA) with paired-end reads strategy. The BCL (base calls) binary is converted into FASTQ by using illumina package bcl2fastq (v1.8.4).

**De novo transcriptome assembly and functional annotation of genes.** The raw reads were checked using Fastqc tool<sup>20</sup> and Cutadapt tool<sup>21</sup> was used for removal of adaptors and low-quality sequences. Also, it was found that sequence files have an average per base Phred score > 30. Normalized reads were assembled into longer fragments (contigs) using Trinity v2.0.6 software<sup>22</sup>. Assembled transcripts were searched for coding transcript by using transdecoder tool<sup>23</sup>. These assembled transcripts were further searched for the orf finding using transdecoder program and the completeness of the transcript. All the protein coding sequences were searched for further annotation using insect uniprot protein database with an e-value cutoff < 1e<sup>-10</sup>. Some differentially expressed uncharacterized genes were further identified by using Uniprot id and BLAST analysis. For further evaluation of the assembly and annotation completeness, BUSCO (Benchmarking Universal Single Copy Orthologs, version 5.4.3) analysis was performed by comparing with arthropod lineage in default settings (<http://busco.ezlab.org/>).

**Analysis of differentially expressed genes.** To compare gene expression profiles across the CanDef-20 and EV control fed *H. armigera*, clean reads of each library were mapped to assembly using RSEM software<sup>24</sup>, and abundance estimation was performed in the Trinity package. Expression levels of the transcripts were calculated based on the FPKM normalization method using the edgeR package<sup>25</sup>. Identification of differentially expressed genes (DEGs) was carried out using EV control libraries as a reference. The FPKM values of genes detected in both CanDef-20 fed (treated) and EV control were used to derive ratio (treated/ control), to which a log<sub>2</sub> scale was applied to get fold change values. Transcripts with log<sub>2</sub>FC ≥ ±2 with P-value ≤ 0.05 and FDR ≤ 0.05 were filtered out as the statistically most significant DEGs. To obtain more DEGs and uniquely expressed genes from CanDef-20 fed larvae, transcripts with log<sub>2</sub>FC ≥ +0.8 and ≤ -0.8 ratios were also identified. Transcripts with log<sub>2</sub>FC ≥ +0.8 were considered upregulated in CanDef-20 fed larvae, whereas transcripts with fold change values (≤ -0.8) were considered downregulated. Finally, GO enrichment analysis was performed for DEG sets using manual method and FUNC software<sup>26</sup>.

**Quantitative real-time PCR (RT-qPCR) validation.** For RNA extraction, whole larvae pulverized in liquid nitrogen (50 mg) were used from CanDef-20 ingesting 4<sup>th</sup> instar (125 and 250 µg/ml fed *H. armigera* larvae), EV control and AD fed. These were different insect samples from those used for sequencing but treated under the similar experimental conditions. cDNA was synthesized from 1000 ng of RNA using a verso cDNA Reverse Transcription Kit (ThermoFisher, USA) following the manufacturer's instructions. Primers were designed using NCBI Primer (Supplementary Table S3) and validated by analysis of their PCR amplification efficiencies (E) and correlation coefficients (R<sup>2</sup>), before the real time analysis. The reaction mixture consists of 1 µl of 0.5-fold diluted cDNA template, 5-µL 2 × Power SYBR<sup>®</sup> Green PCR Master Mix (Applied Biosystems, Lithuania), 1 µl of each gene-specific primer, and ddH<sub>2</sub>O to make up the volume. The reaction conditions were as follows: 50 °C for 2 min, 95 °C for 2 min, 40 cycles of 95 °C for 10 s, 60 °C for 30 s; with melt curve 5 s at 95 °C, 65–95 °C, 0.5 °C increase/cycle. The 2<sup>-ΔΔCt</sup> method was used to detect relative fold change in quantitative real-time PCR and *H. armigera* actin gene was used to normalize the data.

**Enzymatic analysis.** The *H. armigera* larvae fed on AD and AD incorporated with 125 µg/ml CanDef-20 and EV proteins respectively were used for detection of total enzyme activities of amylase, protease, lipase, glutathione S-transferase (GST), aminopeptidase and alkaline phosphatase enzymes. Briefly, 50 mg of whole larvae pulverized in liquid nitrogen was homogenized in 200 µl of 0.2 M glycine–NaOH buffer, pH 10.0 and kept at 4 °C for 2 h. The suspension was centrifuged at 4 °C for 20 min at 10,000 rpm and the resulting supernatant was used as *H. armigera* enzyme preparation (HEP). The total protein content of HEP was calculated by Bradford's method and it was stored at -20 °C till further use.

**Amylase enzyme assay.** The amylase activity was assayed by using the 3,5-Dinitrosalicylic acid (DNSA) method<sup>27</sup>. The assay mixture composed of 125 µl of 0.02 M sodium phosphate buffer pH 6.9 containing 6 mM sodium chloride and 40 µg of HEP, which was incubated at 37 °C for 15 min. This enzymatic reaction was terminated with 1.0 ml DNSA reagent, followed by incubation in boiling water bath for 5 min, cooling to room temperature and diluting the reactions 10 times with water and the absorbance measured at 540 nm. The free sugars released at the end of the amylase enzyme activity were estimated using a standard maltose graph.

**Protease enzyme assay.** The estimation of the total protease enzyme activity was determined by using Azocasein as a substrate<sup>19,28</sup> with 40 µg of HEP. Briefly, 40 µg of HEP enzyme was added into 200 µl of 1% azocasein (prepared in 0.2 M glycine–NaOH, pH 10.0) and incubated at 37 °C for 30 min. The reaction was stopped by the addition of 300 µl of 5% trichloroacetic acid. After centrifugation at 12,000 rpm for 10 min, an equal volume of 1 M NaOH was added into the supernatant. The absorbance was measured at 450 nm and one proteinase unit was calculated as the amount of enzyme that increased the absorbance by 1.0 OD under the given assay conditions. The activity of protease was expressed as units/µg/min/ml protein.

**Glutathione S transferase (GST) enzyme assay.** The GST activity was estimated by using 1-Chloro-2,4-dinitrobenzene (CDNB) (Sigma, Missouri, USA) as the substrate, following the method described earlier<sup>29</sup>. Reaction mixture (1000 µl) contained phosphate buffer pH=7 (980 µl), reduced glutathione (GSH) 200 mM (10 µl), CDNB 100 mM (10 µl), and HEP 40 µg. Buffer is added in control reactions instead of the enzyme. Absorption for the reaction mixtures were acquired on Shimadzu UV-1800 spectrophotometer (Cole-Parmer India). The reactions were carried out in 1 ml quartz cuvettes (Shilpent Quartz Cuvette, France) and absorbance

at 340 nm was recorded at intervals of 30 s for 15 min at 25 °C. All reactions were done in triplicates and GST specific activity was calculated as  $\mu\text{M}/\text{ml}/\text{min}$  by using  $9.6 \text{ mM}^{-1}$  as an extinction coefficient for CDNB-GSH conjugate at 340 nm.

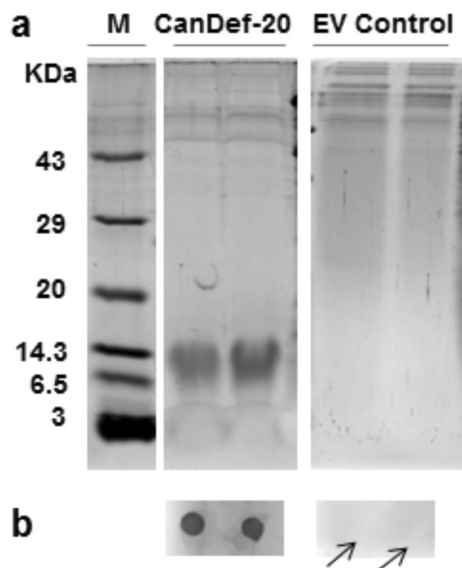
**Lipase enzyme assay.** Total lipase activity was estimated using the p-Nitrophenyl palmitate (pNPP, Sigma-Aldrich) assay<sup>30</sup>. Solution A consists 0.1 g gum arabica and 0.4 mL Triton X-100 (Hi-media, India) dissolved in 90 ml of distilled water and solution B contained 30 mg pNPP dissolved in 10 ml isopropanol<sup>31</sup>. The complete substrate solution was made by adding 9.5 mL of solution A to 0.5 ml of solution B drop-wise with constant stirring to obtain an emulsion that was stable for two hours. Using this substrate, the lipase activity was determined. 40  $\mu\text{g}$  of HEP was added in the substrate solution and it was incubated at 37°C for 30 min. The reaction was stopped by adding 1% sodium carbonate ( $\text{Na}_2\text{CO}_3$ ) and absorbance of the samples was measured at 410 nm.

**Alkaline phosphatase (ALP) and aminopeptidase (APN) enzyme assays.** p-nitrophenyl phosphate and leucine-p-nitroanilide (Sigma-Aldrich) were used as substrates respectively for detection of specific ALP and APN enzymatic activities as per protocol<sup>32</sup>. 40  $\mu\text{g}$  of HEP were mixed with ALP buffer [0.5 mM  $\text{MgCl}_2$ , 100 mM Tris/HCl (pH 9.5)] consisting 1.25 mM of p-nitrophenyl phosphate or APN buffer [0.2 M Tris/HCl (pH 8), 0.25 M NaCl] with 1 mM of leucine-p-nitroanilide. ALP enzyme activity was checked as the change in the absorbance at 450 nm for 3 min at 25 °C while APN enzyme activity was measured at 410 nm by checking rate of generation of p-nitroaniline.

**Statistical analysis.** The Student's t-test was used for statistical analysis and the level of significance was indicated as \* $P < 0.05$ , \*\* $P < 0.001$  and \*\*\* $P < 0.0001$  by comparing with readings of EV control. The graphs were plotted by using an online demo version of Graphpad prism and Microsoft excel. The model image was created with Biorender.com.

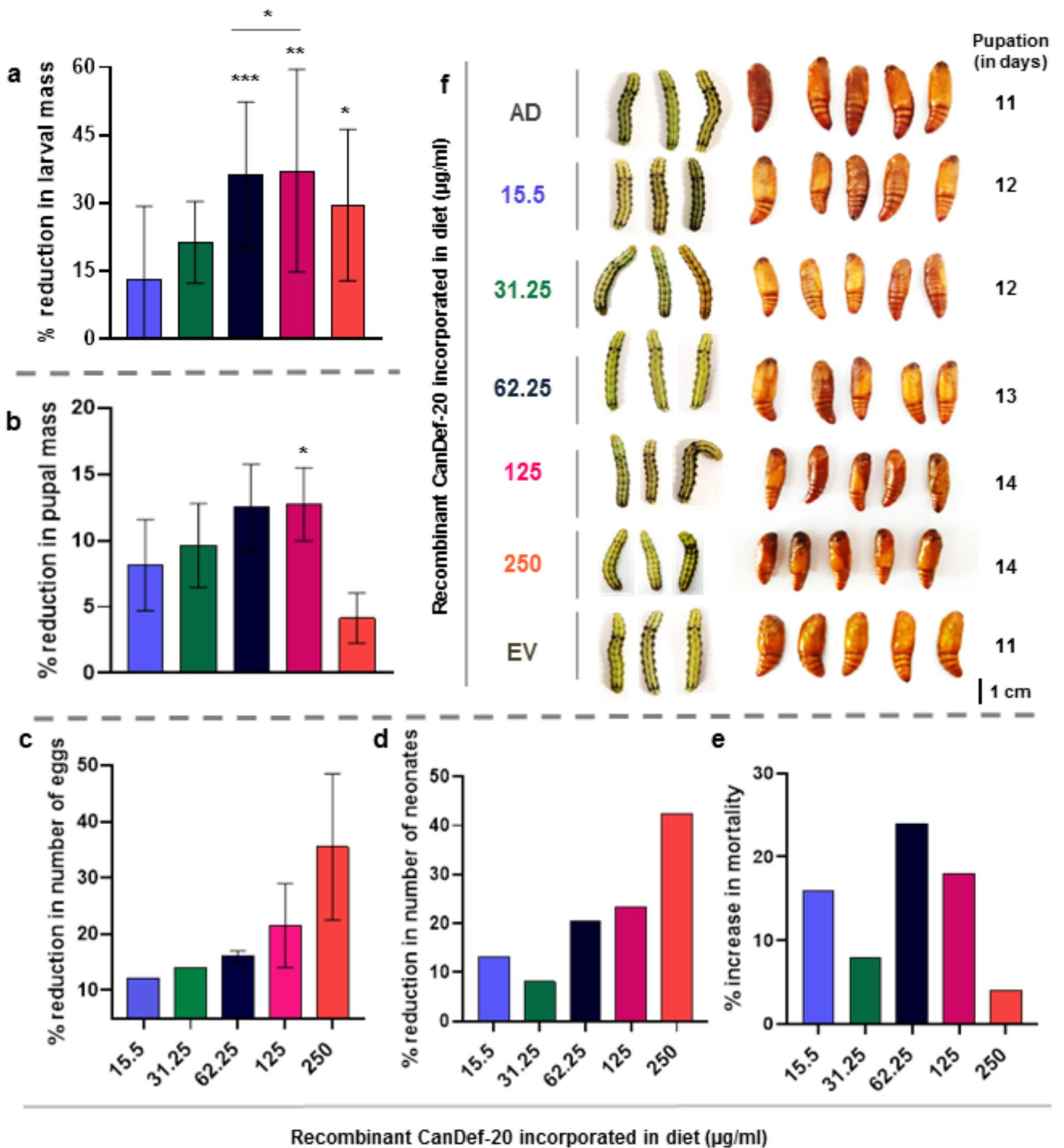
## Results

**Effect of CanDef-20 ingestion on growth and development of *H. armigera*.** Recombinantly expressed CanDef-20 protein showed a molecular mass of ~ 14 kDa and their dot blot assay with anti-myc antibodies confirmed the presence and expression of the recombinant proteins (Fig. 1a,b). The original images of gels and dot blot are given in Supplementary Fig. S1. The effect of CanDef-20 on the development of *H. armigera* was evaluated by measuring the larval mass, pupal mass, fertility and fecundity in insects fed on various doses of the recombinant protein incorporated in artificial diet (15.5, 31.25, 62.25, 125 and 250  $\mu\text{g}/\text{ml}$ ). *H. armigera* larvae ingesting recombinant CanDef-20 displayed a dose dependent effect on larval mass as compared to the larvae ingesting control diets (EV and AD fed) (Supplementary Fig. S2). Maximum reduction in the mass of larvae (37%) was observed when the larvae were fed 125  $\mu\text{g}/\text{ml}$  of CanDef-20 containing diet (Fig. 2a). Similarly, reduction in pupal mass (4% to 12%) was also observed in the larvae fed with CanDef-20 (Fig. 2b and Supplementary Fig. S2). Interestingly, upon ingestion of 250  $\mu\text{g}/\text{ml}$  CanDef-20, *H. armigera* larvae indicated a slight yet significant increase in larval and pupal mass compared to larvae ingesting 125  $\mu\text{g}/\text{ml}$  CanDef-20, though



**Figure 1.** The SDS-PAGE and dot blot analysis of recombinant CanDef-20 and EV proteins: (a) CanDef-20, EV protein and protein molecular weight ladder (lane M, Genei Laboratories Private Limited) were resolved on 15% SDS-PAGE and visualized by Coomassie blue staining. (b) The rabbit anti-myc primary antibody was used against myc epitope of pPICZ-alpha A vector in dot blot analysis. Secondary antibody provided with Western Blot Development kit (Bangalore genie, India) was used to visualise myc-tagged primary antibody complex.





**Figure 2.** *H. armigera* growth and development in CanDef-20 feeding assay as compared to control (AD) and EV control diet: (a, b) The % reductions in the larval and pupal mass of *H. armigera* are shown by bar graph respectively. (c) The effect of CanDef-20 feeding on fecundity of *H. armigera* larvae is shown by % decrease by respective treatments. (d) The number of neonates hatched per 1,000 eggs was recorded for each treatment. (e) The mortality induced by different concentrations of CanDef-20 in *H. armigera* larvae. (f) The morphological variations in the late fifth instar *H. armigera* larvae and their pupae indicated the pronounced effect of CanDef-20 on larval and pupal growth. The % reduction in larval and pupal mass values are significantly different from EV control at \* for  $P < 0.05$ , \*\* for  $P < 0.001$  and \*\*\* for  $P < 0.0001$ , respectively.

the mass was still lower than EV control ingesting larvae. Additionally, around 48–72 h delay in pupation was observed in insects fed with 125 and 250 µg/ml CanDef-20. The egg laying and egg hatching were also affected in *H. armigera* moths raised from larvae fed with higher concentrations of CanDef-20. 12–21% reduction was observed in total number of eggs laid by *H. armigera* moths raised from CanDef-20 fed larvae (Fig. 2c). The

egg fertility was seen to be reduced from 13 to 68%. 250 µg/ml CanDef-20 feeding resulted in lowest number (440) of neonates emerging per 1,000 eggs (Fig. 2d). Larvae feeding on 62.25 µg/ml CanDef-20 diets showed high mortality (24%) amongst all tested concentrations (Fig. 2e). CanDef-20 ingestion induced larval and pupal mass reduction was also evident from the comparative morphologies of late fifth instar *H. armigera* larvae and their pupae (Fig. 2f). The 24–72 h of delayed pupation was observed in the larvae fed with CanDef-20. The day where ~70% larvae entered to pupation was showed in Fig. 2f. The delayed metamorphosis effects of CanDef-20 on the larvae and pupae of *H. armigera* are also evident in a separate larval bioassay with 150 µg/ml and 300 µg/ml CanDef-20 (Supplementary Fig. S2).

**Comparative transcriptome analysis of *H. armigera* larvae fed with CanDef-20 and EV incorporated diets.** The feeding experiments demonstrated a dose-dependent reduction in larval and pupal mass of *H. armigera* when ingested up to 125 µg/ml CanDef-20 containing diet. The 4<sup>th</sup> instar *H. armigera* larvae fed for 9 days on 125 µg/ml CanDef-20 or fed on EV protein containing diet were used for a comparative transcriptomic analysis using RNA-seq.

Illumina sequencing of *H. armigera* libraries yielded 1091.8 and 1278.1 million read pairs for CanDef-20 and EV diet fed larvae respectively. The de novo assembly analysis of clean reads resulted in a total of 1,57,768 transcripts of which 65,535 assembled transcripts (41%) were >200 bp in length and were annotated with UniProtKB insect data resulting in 13,779 transcript annotations (Table 1). Out of the 13,779 annotated transcripts 6982 showed >75% query coverage and these were used for further comparative transcriptome analysis of DEGs between CanDef-20 and EV fed *H. armigera* larvae. Of the 6982 genes, 47% were found to align with genus *Heliothis* and 6.19% aligned with genus *Helicoverpa*. Applying the criteria of  $\log_2FC \geq \pm 2$  with  $P\text{-value} \leq 0.05$  and  $FDR \leq 0.05$  to the 13,779 transcripts resulted in detection of 2012 transcripts of which 56 transcripts were found upregulated and 529 were downregulated.

Applying the criteria of  $\log_2FC \geq +0.8$  and  $\leq -0.8$  ratios to 6982 transcripts with >75% query coverage, 2327 (33.32%) genes were found to be downregulated and 659 (9.4%) genes were found to be upregulated upon CanDef-20 feeding. Additionally 1679 genes were found to be uniquely expressed in CanDef-20 fed larvae and 1545 genes were uniquely expressed in EV control fed larvae.

The BUSCO analysis was applied on all assembled transcripts and it detected 19% complete (154 Complete and single-copy BUSCOs, 39 Complete and duplicated BUSCOs), while 24% of as fragmented (243 BUSCOs) and 57% are missing (577 BUSCOs) out of 1013 single copy orthologs for arthropods (Supplementary Table S4).

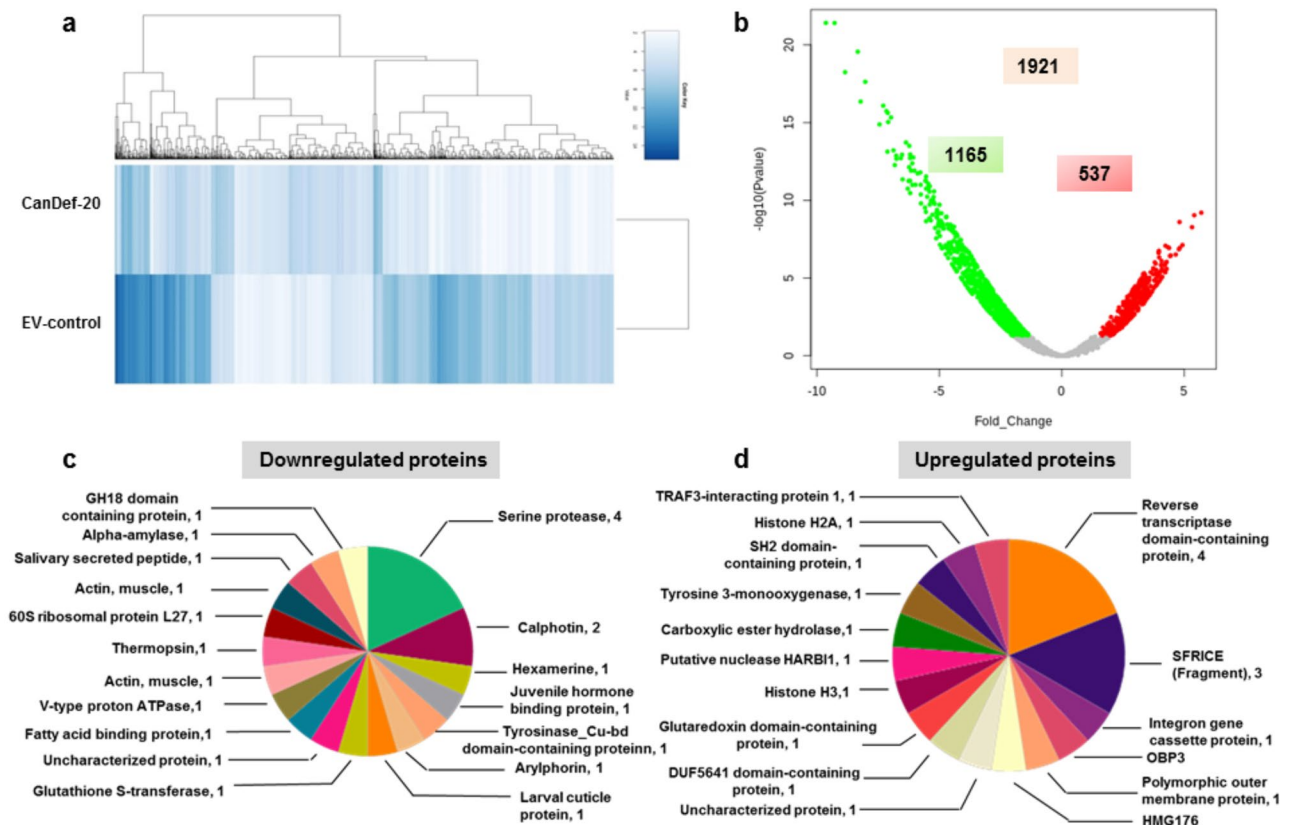
Samples		Raw reads		Clean reads	
CanDef-20 Fed		10,918,133		10,593,388	
EV Control Fed		12,781,063		12,436,430	
Total assembled transcripts		157,768			
Number of Unigenes		142,916			
Unigenes (>200 nucleotide)		65,535			
Transcripts annotated (with insect NR database)		13,779			
Transcripts with best query coverage $\geq 75\%$		6982			
log <sub>2</sub> FC $\geq \pm 2$ with $P\text{-value} \leq 0.05$ and $FDR \leq 0.05$ (2012 transcripts)					
Upregulated transcripts		56			
Downregulated transcripts		529			
log <sub>2</sub> FC $\geq +0.8$ ratio and $\leq -0.8$ ratio (treated to control) (6982 transcripts)					
Upregulated transcripts		659			
Downregulated transcripts		2327			
Unique in CanDef-20 fed (treated)		1679			
Unique in EV control fed (only found in control)		1545			
Transcript ( $P\text{ value} \leq 0.05$ )		1921			
Statistics related to Transcript				Statistics related to Gene	
Contig	N10:	864	Contig	N10:	690
Contig	N20:	559	Contig	N20:	483
Contig	N30:	437	Contig	N30:	397
Contig	N40:	370	Contig	N40:	344
Contig	N50:	324	Contig	N50:	308
Median (contig length)	275		Median (contig length)	271	
Average	337.4		Average	320.28	
Total assembled bases	53,230,840		Total assembled bases	45,772,751	

**Table 1.** Transcriptomic assembly statistics of *H. armigera* larvae fed with CanDef-20.

**Differential gene expression analysis in CanDef-20 fed *H. armigera*.** The overall gene expression pattern in the control and CanDef-20 fed larvae was determined by hierarchical clustering and results indicated a prominent downregulation of most genes in CanDef-20 fed *H. armigera* (Fig. 3a). The volcano plot indicated significant differential regulation of 1702 out of 1921 transcripts with  $P$ -value  $\leq 0.05$  (Fig. 3b; Table 2). The analysis with  $\geq$  two fold criteria with  $P$ -value  $\geq 0.05$  and  $FDR \geq 0.05$  showed juvenile-hormone binding protein (JHBP), hexamerin, serine protease, arylphorin and calphotin transcripts represented amongst the highly downregulated transcripts while reverse transcriptase domain containing protein, putative nuclease HARBI1, carboxylic ester hydrolase and tyrosine 3-monooxygenase were represented amongst the highly upregulated transcripts (Fig. 3c and 3d; Supplementary Fig. S3; Supplementary Table S1). Also, juvenile hormone-suppressible protein, serine protease and some isoforms of glutathione S-transferase were uniquely found in EV control larvae with very high FPKM values ranging from 6054–1386 whereas some isoforms of arginine kinase, endonuclease-reverse transcriptase and serine protease were uniquely detected in CanDef-20 fed larvae and had high FPKM values ranging from 211–83. A single transcript of GST was upregulated (TRINITY\_DN83033\_c9\_g1\_i1 with FPKM 2.44) and other isoform of GST (TRINITY\_DN82857\_c2\_g2\_i2 with FPKM 8.44) was also found amongst genes uniquely expressed in CanDef-20 fed *H. armigera* larvae.

A comparative GO enrichment for differentially expressed genes with  $\log_2FC \geq +0.8$  and  $\leq -0.8$  ratios criteria was carried out and top10 GOs are represented (Figs. 4 and 5). Majority of top10 GOs from the differentially expressed set were represented by downregulated genes in *H. armigera* upon CanDef-20 ingestion (except the GO nucleic acid binding, which was upregulated) (Fig. 4); while the uniquely expressed set was represented by majority of genes induced by CanDef-20 (Fig. 5). Integral component of membrane, ATP binding and translation were found to be most prominent differentially regulated GOs in the CanDef-20 fed larvae (Fig. 4a). The subcategories of top most differentially regulated GOs are represented in Fig. 4b. GO ATP binding subcategorized into ATP dependent enzymatic processes like protein kinase, proteasome activating ATPase and helicase; while the GO integral component of membrane included sub GOs like mitochondrial membrane associated genes, signal peptidase complex, proton transporting V-type ATPase, plasma membrane and microtubule tethering complex. Uniquely expressed GOs showed prominence of integral component of membrane, ATP binding and carbohydrate metabolic process (Fig. 5a and b).

Some prominent GOs namely kinases, ATPases, serine-endopeptidases, integral membrane components, lipases and transposons—appeared in differentially expressed as well as uniquely expressed sets. Different isoforms of the above mentioned genes were detected and found to have variable expression levels as represented



**Figure 3.** Transcriptomic analysis: (a) Hierarchical clustering analysis of differentially expressed genes between CanDef-20 fed and EV control fed *H. armigera* larvae. (b) Volcano plot is showing statistically significant gene expression. The down and up regulation was denoted by green and red dots respectively. (c) and (d) Top 25 down and upregulated DEG's are shown by pie chart respectively. The number of transcripts appeared for particular gene is denoted.

Sr. No	Transcript id	Fold change	-(log P value)	Matched uniprot id	Protein names	Organism
<b>Upregulated proteins</b>						
1	TRINITY_DN83412_c6_g2_i1	5.694941	9.201742	A0A2H1WG43	SFRICE_037529	<i>S. frugiperda</i> (Fall armyworm)
2	TRINITY_DN82568_c3_g1_i3	5.408217	9.038418	P18488	Homeotic protein empty spiracles	<i>D. melanogaster</i> (Fruit fly)
3	TRINITY_DN83065_c3_g5_i1	5.318082	8.268862	A0A2R7VU22	Uncharacterized protein	<i>O. fasciatus</i> (Large milkweed bug)
4	TRINITY_DN82238_c11_g4_i1	4.925622	7.120954	A0A1Y1L125	Uncharacterized protein Nucleoporin NUP188 homolog	<i>P. pyralis</i> (Eastern firefly)
5	TRINITY_DN82995_c2_g1_i1	4.815952	6.935561	A0A1W4WR17	Twitchin	<i>A. planipennis</i> (Emerald ash borer)
6	TRINITY_DN82945_c1_g1_i1	4.806776	8.599529	A0A151J635	Uncharacterized protein CCHC-type domain-containing protein	<i>T. cornetzi</i>
7	TRINITY_DN82860_c12_g1_i1	4.781217	6.848041	A0A2A4JYZ1	Uncharacterized protein Reverse transcriptase domain-containing protein	<i>H. virescens</i> (Tobacco budworm moth)
8	TRINITY_DN83005_c2_g1_i8	4.652604	6.511229	A0A2H1VQ04	SFRICE_039558	<i>S. frugiperda</i> (Fall armyworm)
9	TRINITY_DN82539_c5_g2_i2	4.444028	6.456275	A0A2A4IVF1	Uncharacterized protein Superoxide dismutase	<i>H. virescens</i> (Tobacco budworm moth)
10	TRINITY_DN81078_c1_g1_i2	4.426557	6.427306	A0A182IYI1	Uncharacterized protein	<i>A. atroparvus</i>
<b>Downregulated proteins</b>						
1	TRINITY_DN83198_c0_g1_i2	-9.6387	21.40195	A0A2A4JMW0	Uncharacterized protein-basic juvenile hormone-suppressible protein 1	<i>H. virescens</i> (Tobacco budworm moth)
2	TRINITY_DN82746_c0_g1_i1	-9.2829	21.3964	Q68YP2	Hexamerine	<i>H. armigera</i> (Cotton bollworm)
3	TRINITY_DN83202_c0_g3_i3	-8.8531	18.23386	A0A0K8SDP7	Uncharacterized protein-Calphotin	<i>L. hesperus</i> (Western plant bug)
4	TRINITY_DN83234_c0_g1_i6	-8.3346	19.55163	A0A2A4J4S9	Uncharacterized protein arylphorin subunit beta-like	<i>H. virescens</i> (Tobacco budworm moth)
5	TRINITY_DN83234_c0_g1_i2	-8.2174	16.34541	G3LF43	Arylphorin	<i>H. armigera</i> (Cotton bollworm)
6	TRINITY_DN82063_c0_g1_i6	-8.02709	17.61496	A0A0K8T7D1	Uncharacterized protein Actin cytoskeleton-regulatory complex protein PAN1-like	<i>L. hesperus</i> (Western plant bug)
7	TRINITY_DN83386_c1_g3_i11	-7.4470	14.87827	O18447	Serine protease (Trypsin-like protease)	<i>H. armigera</i> (Cotton bollworm)
8	TRINITY_DN82805_c0_g1_i4	-7.2939	16.0792	O18439	Diverged serine protease	<i>H. armigera</i> (Cotton bollworm)
9	TRINITY_DN82625_c1_g1_i1	-7.1640	15.72952	O02443	Larval cuticle protein 1	<i>H. armigera</i> (Cotton bollworm)
10	TRINITY_DN83068_c1_g1_i1	-7.1219	13.12375	A0A2A4J8G1	Uncharacterized protein chitinase-like protein EN03 isoform X4	<i>H. virescens</i> (Tobacco budworm moth)

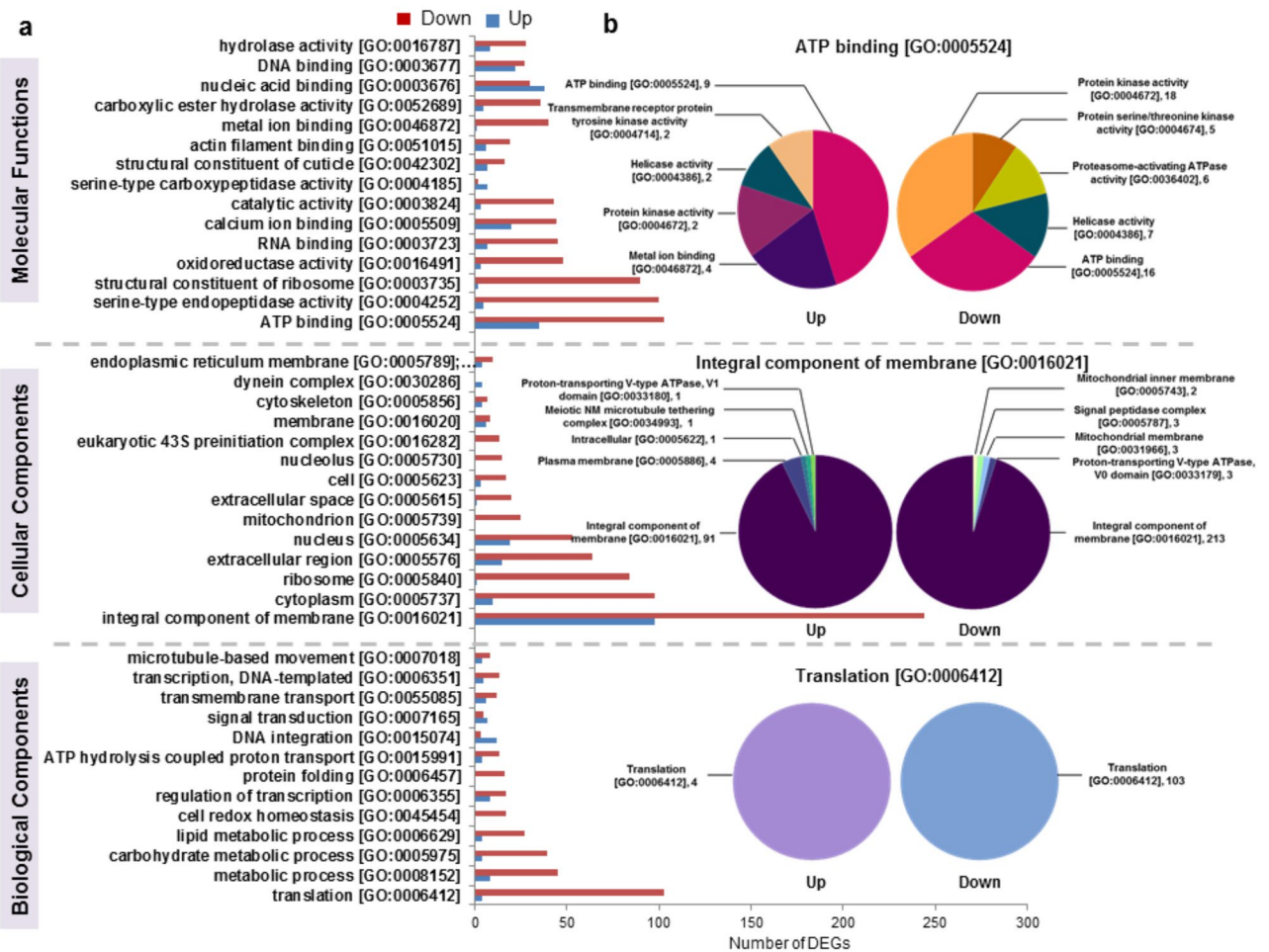
**Table 2.** Top 10 upregulated and downregulated proteins in CanDef-20 fed insect shown in volcano plot.

by the heatmaps (Fig. 6a and b; Supplementary Table S2). Downregulated isoforms of genes were balanced by few upregulated isoforms indicating a CanDef-20 induced compensatory response in *H. armigera*. For example, in GO transposable elements, isoforms were found to be highly upregulated, while in GO lipase interestingly only three isoforms were found to be upregulated (LP1, TLP and LP2) and > 15 downregulated. LP1 and TLP are not characterized from *H. armigera* (Supplementary Table S2), though they showed homology with lipase-1 (NM\_001043501.1) and triacylglycerol lipase (XM\_038014482.1) respectively found in *Bombyx mori*. The third (LP2) upregulated lipase has homology with *H. armigera* triacyl glycerol lipases (XM\_047184139.1).

**CanDef-20 fed *H. armigera* larvae have higher activities for selected enzymes.** HEP from CanDef-20 fed and EV control protein fed *H. armigera* was analyzed for detection of amylase, protease, lipase, GST, aminopeptidase and alkaline phosphatase activities (Fig. 7). A significant increase in the total enzyme activities of all these enzymes were found in the larvae fed on CanDef-20. Amylase, protease, GST and aminopeptidase showed an increase of around 50–63% whereas lipases and alkaline phosphatases showed 35% and 24% increase respectively in CanDef-20 fed larvae. The RNA seq data had indicated a prominent downregulation of most of the isoforms of genes for these enzymes in response to Candef-20 feeding (Fig. 6). However, a few specific enzyme isoform coding genes were found to be upregulated for all the selected enzymes (except amylases). These selectively upregulated enzyme isoforms coding genes would have most prominently contributed to the increase in net in vitro activity of these enzymes.

**Validation of selected differentially regulated *H. armigera* genes upon CanDef-20 ingestion by real-time quantitative reverse transcription PCR (RT-qPCR).** Relative gene expression levels of JHBP, hexamerin, calphotin, kinases, ATPases, integral components of membrane, lipases, serine endopeptidases and transposons were studied from CanDef-20 fed *H. armigera* larvae (at 125 µg/ml and 250 µg/ml) and EV control by RT-qPCR (details of primer are given in Supplementary Table S3). The gene expression of JHBP, hexamerin and calphotin were found to be downregulated in *H. armigera* upon CanDef-20 ingestion and this is consistent with its RNA-seq data (Fig. 8a).





**Figure 4.** Gene Ontology enrichment of DEGs of CanDef-20 fed insects compared to EV control: (a) Gene Ontology (GO) classifications of transcripts into three categories like molecular function, cellular component and biological process. The X-axis corresponds to the number of DEGs appeared in analysis while the Y-axis represents different GOs. Red and blue bars are showing down and up regulation respectively. (b) The sub categorization of top most GO is further analyzed and shown in the form of pie chart. The name and occurrence of sub GO is denoted.

The isoforms of arginine kinase, pyruvate kinase, V-type ATPase and its subunit C, were down regulated and this is consistent with its RNA-seq data, but adenylate kinase and V-type ATPase subunit B showed contrasting expression levels in qPCR versus RNA-seq analysis (Fig. 8a,b).

Gene expression levels of isoforms of membrane proteins like, aminopeptidase, alkaline phosphatase 2 and mucin were also down regulated and it is consistent with RNA-seq data (Fig. 8c).

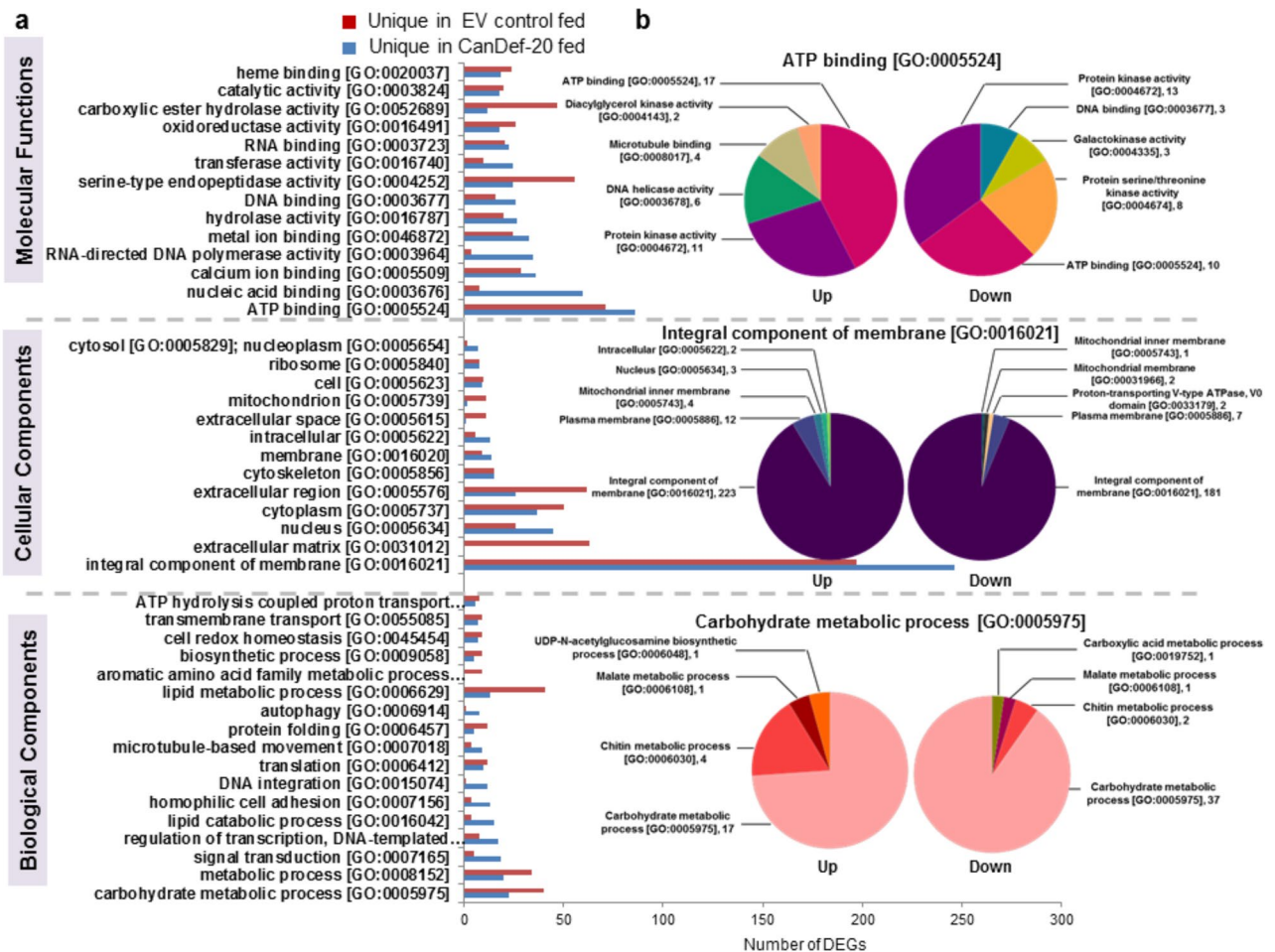
Different isoforms of lipases (lipase, N lipase and LP2), serine protease and trypsin were found to be up or down regulated and that is consistent with their respective RNA-Seq data (Fig. 8d and e).

Similarly, ATP synthase also showed down regulation of gene expression correlating with the RNA-seq data (Fig. 8f). qPCR analysis showed contrasting expression pattern of all the tested isoforms of endonuclease reverse transcriptase genes (ERT4, 12 and 2) with its RNA-seq data respectively (Fig. 8g).

## Discussion

Plants employ arsenal of biomolecules to pose a defense response against pests and pathogens, some of these molecules are target specific while some are broad acting. Plant defensins are best known for their antifungal properties<sup>33</sup> though they also possess insect growth retardation potential. In the present study ingestion of a defensin from *C. annuum* (recombinant CanDef-20) led to a dose dependent negative effect on the growth and development of the insect pest *H. armigera*, similar to lepidopteran and coleopteran insect antibiosis reported earlier<sup>34,35</sup>. *H. armigera* fed very high concentrations of CanDef-20 (250 µg/ml) deviated from the trend of dose dependent larval mass reduction, most likely due to the defensin induced adaptive molecular responses (discussed further).

Though insect growth retardation is a prominent effect due to feeding on defensin containing diet, the molecular mechanisms involved in plant defensin-mediated metabolic changes in insects like delayed life stages, growth retardation, mass reduction and loss of fecundity remain largely unknown. The comparative transcriptomics

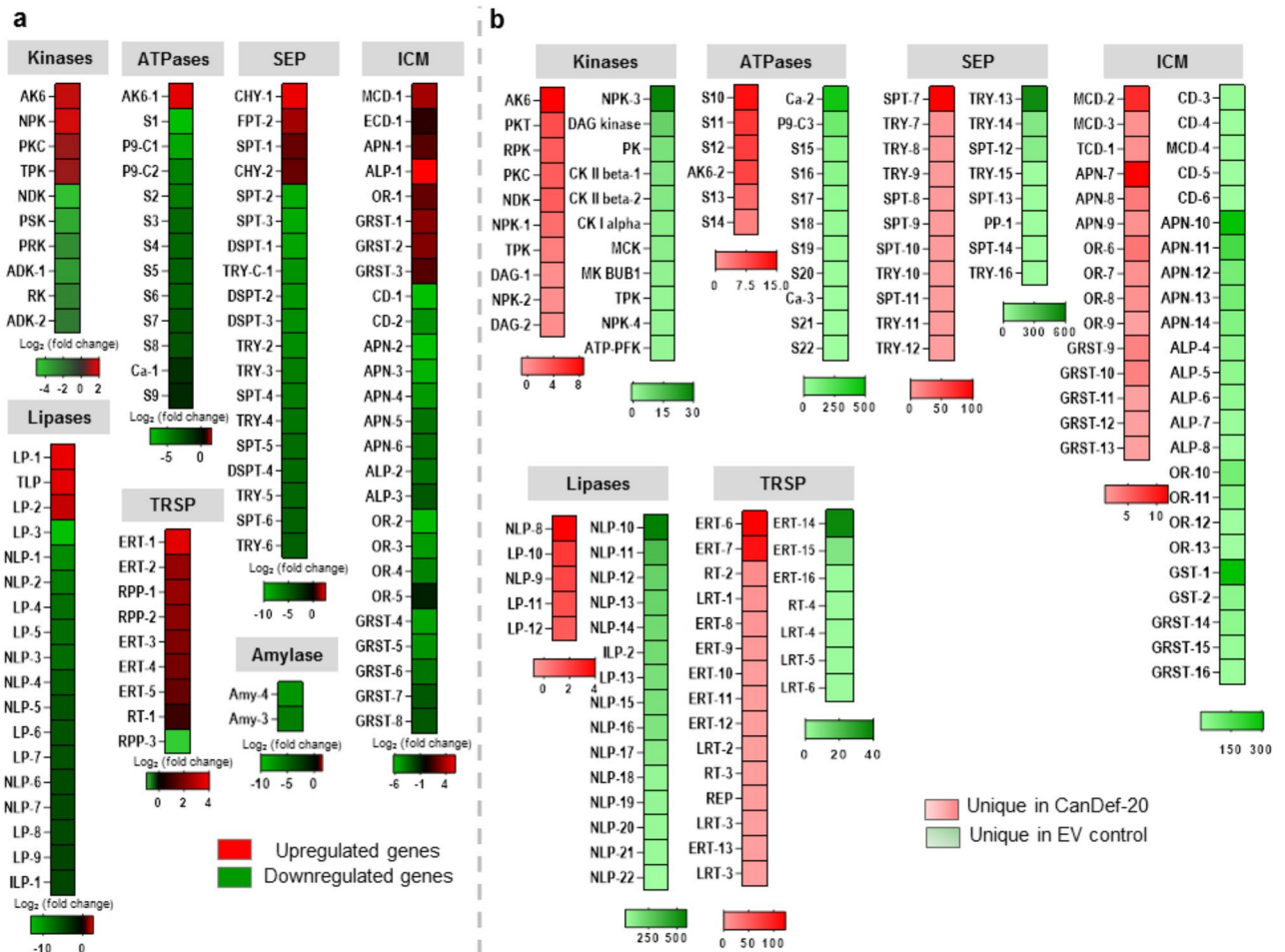


**Figure 5.** Gene Ontology enrichment of DEGs of uniquely expressed in CanDef-20 and EV fed insects: (a) Gene Ontology (GO) classifications of transcripts into three categories like molecular function, cellular component and biological process. The X-axis corresponds to the number of DEGs appeared in analysis while the Y-axis represents different GOs. Red and blue bars were showing unique expression in EV control and CanDef-20 fed respectively. (b) The sub categorization of top most GO is further analyzed and shown in the form of pie chart. The name and occurrence of sub GO is denoted.

revealed changes in pathways and processes induced by CanDef-20 ingestion in *H. armigera* larvae. AMPs like defensins are known to interact with different cellular components like cell membrane, cytoplasmic proteins and they can also enter the nucleus and interact with DNA<sup>36–39</sup>. Ingestion of CanDef-20 peptide mediates multiple effects on *H. armigera* metabolism through defensin–secreted protein/enzyme interaction(s), defensin–membrane interactions and defensin-induced regulation of target gene expression (Fig. 9).

One of the known modes for plant defensin activity on insects is in the form of inhibition of their digestive enzymes like amylases and proteinases<sup>17,40–42</sup>. In the present study, CanDef-20 fed larvae showed higher gut amylase and proteinase activities. Though, the transcriptomics data revealed a prominent downregulation of alpha-amylase and serine endopeptidase (serine proteases and trypsin) transcripts and notable upregulation of a few isoforms of chymotrypsin and serine protease. This indicated compensatory metabolic adjustments made by *H. armigera* for balancing digestion in presence of CanDef-20. Adjustments in the expression of proteinase isoforms at gene and transcript levels in insects are common upon exposure to proteinase inhibitors<sup>43,44</sup>. Similarly, gut amylase enzymes modulate their expression in response to ingestion of plant amylase inhibitor<sup>45</sup>.

The insect lipase inhibitory potential of plant defensins has not been reported earlier, though we noted a downregulation of most *H. armigera* lipase transcripts (neutral lipase, lipases and triacylglycerol lipases) upon CanDef-20 ingestion. Also, a simultaneous net increase in lipase activity in CanDef-20 fed larvae was observed, which can be attributed to the compensatory upregulated lipase transcript isoforms LP1, LP2 and TLP (classified under carboxylic ester hydrolase activity [GO:0052689]). Carboxylic ester hydrolases are also involved in the detoxification of several insecticides imparting resistance development in insects<sup>46</sup>. It is also known that different isoforms of lipase enzymes are needed for lipid mobilization in insect under starvation stress<sup>47</sup>. TAG lipases and a phospholipase gets upregulated during starvation in *D. melanogaster* and it may be involved in lipid mobilization<sup>48,49</sup>. Thus CanDef-20 ingestion seems to induce starvation-like stress and toxin stress in *H. armigera*, which further manifests into differential lipase expression and carboxylic ester hydrolase mediated lipid breakdown.



**Figure 6.** Differential and unique expression pattern of some selected gene categories: (a) CanDef-20 induced differential regulation of *H. armigera* genes from selected GOs is represented by the heatmaps. Different gene isoforms of kinases, ATPases, serine-endopeptidases (SEP), integral membrane components (ICM), lipases and transposons (TRSP) are shown on the basis of  $\log_2$  fold change. The down and up regulation is shown by green and red colored gradient respectively. (b) The unique expression pattern of different gene isoforms of kinases, ATPases, serine-endopeptidases (SEP), integral membrane components (ICM), lipases and transposons (TRSP) are detected in CanDef-20 and EV control fed larvae separately and represented by the heatmaps. The unique expression pattern in CanDef-20 fed larvae is shown by red gradient. The green gradient is used to show unique expression in EV fed larvae.

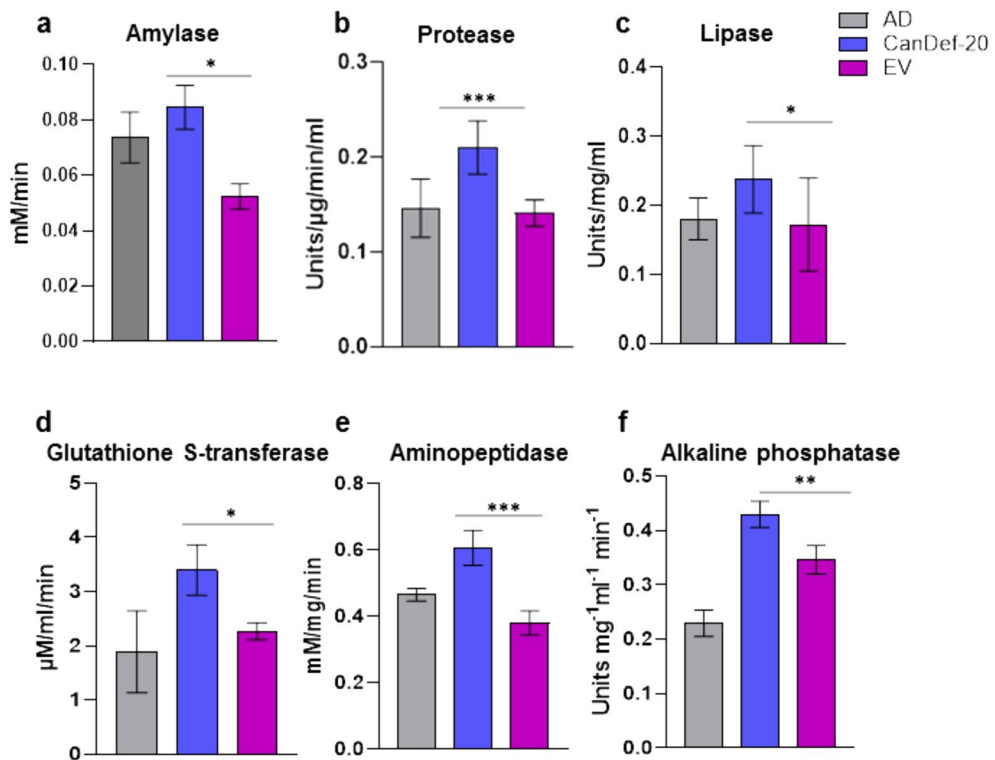
The available genomic and transcriptomic resources for *H. armigera* are not sufficient to map the range of transcripts expressed upon CanDef-20 ingestion as 64% of the differentially expressed transcripts identified in the present study are uncharacterized genes. The recently published *H. armigera* genome<sup>50</sup> resource may be useful to annotate some of the uncharacterized genes from the present study.

Recent studies have shown that defensins interact with microbial pathogens and/or tumour cell membranes by binding to specific phospholipids to cause membrane permeabilization leading to cell death<sup>38,51,52</sup>. The mechanism of membrane disruption varies as per the membrane composition and interacting defensin molecule(s). The fungal membrane interaction kinetics has been known for two plant defensins, (i) pea defensin PsD1 interacting with membranes containing the sphingolipid glucosylceramide (GluCer)<sup>53</sup> and (ii) MtDef4 from *Medicago truncatula* interacting with membrane phosphatidic acid (PA)<sup>15</sup>. We noted that integral components of the membrane (GO:0,016,021) is a highly downregulated GO in CanDef-20 fed *H. armigera* larvae, and it represents membrane proteins like V-type proton ATPases (V-ATPases), Cadherin, ALP and APN proteins.

V-ATPases transmembrane proton pumps which play important roles in numerous biological processes, such as protein degradation and synthesis, shifts in organelle pH, cell growth and cell autophagy<sup>54</sup>. CanDef-20 ingestion would lead to disturbance in cellular energy homeostasis in *H. armigera*, which can be correlated with the downregulation of V-ATPases. V-ATPases and its subunit were downregulated during starvation in tobacco hornworm *Manduca sexta*<sup>55</sup>. Interestingly, the RNAi targeting V-ATPase in *H. armigera* reduced larval mass and survival ability of the larvae<sup>56</sup>, correlating with the findings from the present work.

Cadherin, ALP and APN proteins are functional receptors for the Cry1A toxin in *Helicoverpa armigera* and their interaction with cry toxin mediates pore formation in the gut membrane leading to insect death<sup>5</sup>. Downregulation of these receptors and expression of their mutated isoforms leads to the development of resistance





**Figure 7.** Effect of CanDef-20 feeding on the enzyme activities: The *H. armigera* enzyme preparation (HEP) was used in the assays to detect (a) amylase, (b) protease, (c) lipase, (d) glutathione S-transferase (GST), (e) aminopeptidase and (f) alkaline phosphatase enzyme activities and is shown by bar graph. The enzyme activities from artificial diet (AD), recombinant CanDef-20 incorporated diet (CanDef-20) and empty vector expressed proteins (EV) were analyzed. The total enzyme activity values are significantly different from EV control at \* for  $P < 0.05$ , \*\* for  $P < 0.001$  and \*\*\* for  $P < 0.0001$  respectively.

against Cry1A toxin in lepidopteran insects<sup>4,57,58</sup>. CanDef-20 ingestion-induced downregulation of Cadherin, ALP and APN receptors in *H. armigera* is indicative of the toxic effect of CanDef-20. Additionally, transcripts of mutant cadherin and some isoforms of ALP and APN are seen to be upregulated in *H. armigera* upon CanDef-20 ingestion, revealing similarities in processes triggered by Cry1A toxin and CanDef-20. Similar simultaneous downregulation of APN1 and compensatory upregulation of APN6 is reported in the Cry1A resistant cabbage looper, *Trichoplusia ni*<sup>6</sup>.

GSTs are the other enzymes with xenobiotic detoxification potential<sup>59</sup>. The downregulation of several isoforms of GSTs and selective upregulation of two GST isoforms upon CanDef-20 ingestion in *H. armigera* is also indicating the toxic stress of CanDef-20 and corresponding adaptive response of the insect, as also represented by high GST protein activity (Fig. 7).

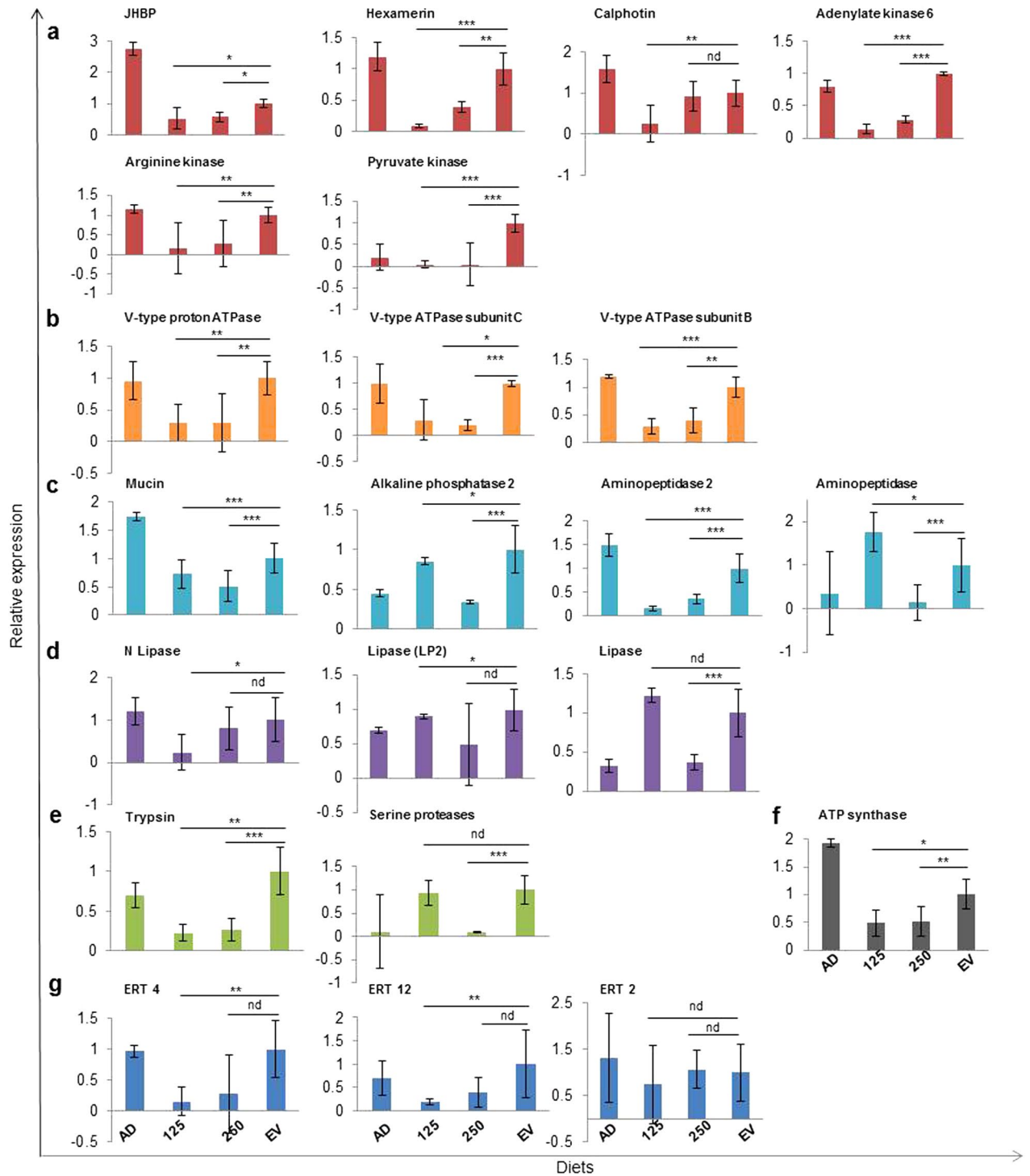
CanDef-20 ingestion leads to a change in the expression of some of the cytoplasmic enzymes like Arginine kinase (AK), Pyruvate kinase (PK), Superoxide demutase (SOD), Spermine oxidase and Chitin synthase (CHS).

AK is an invertebrate phosphagen kinase involved in the transport and buffering of energy by stabilizing the cellular ATP/ADP ratio<sup>60</sup>. RNA interference (RNAi) study targeting AK in coleopteran beetle *Phyllotreta striolata* resulted in mortality and loss in fecundity<sup>61</sup>. PK catalyzes phosphoenol pyruvate to pyruvate reaction of glycolysis, leading to the TCA cycle<sup>62</sup>. The inhibition of PK leads to decreased ATP levels and an imbalance in the energy metabolism in *Tribolium castaneum*<sup>63</sup>. Low levels of pyruvate in the pupal brains of *H. armigera* resulted in an extension of lifespan by initiating diapause<sup>64</sup>. The reduced larval and pupal mass along with the delayed pupation in CanDef-20 fed *H. armigera* may be correlated with the downregulation of AK and PK.

Plant defensins can cause cellular effects including generation of reactive oxygen species (ROS), which triggers secondary signaling and imbalance of ionic homeostasis, ultimately contributing to cell death<sup>65</sup>. SOD enzyme helps in combating ROS generation<sup>66</sup>. Plant defensins like RsAFP2 and Vu-Defr are also capable to induce ROS in *Candida albicans*<sup>67</sup> and *Leishmania amazonensis*<sup>68</sup> respectively. Upregulation of SOD upon CanDef-20 ingestion in *H. armigera* implies that defensin is generating oxidative stress in the larvae and the elevated SOD expression relieves ROS stress.

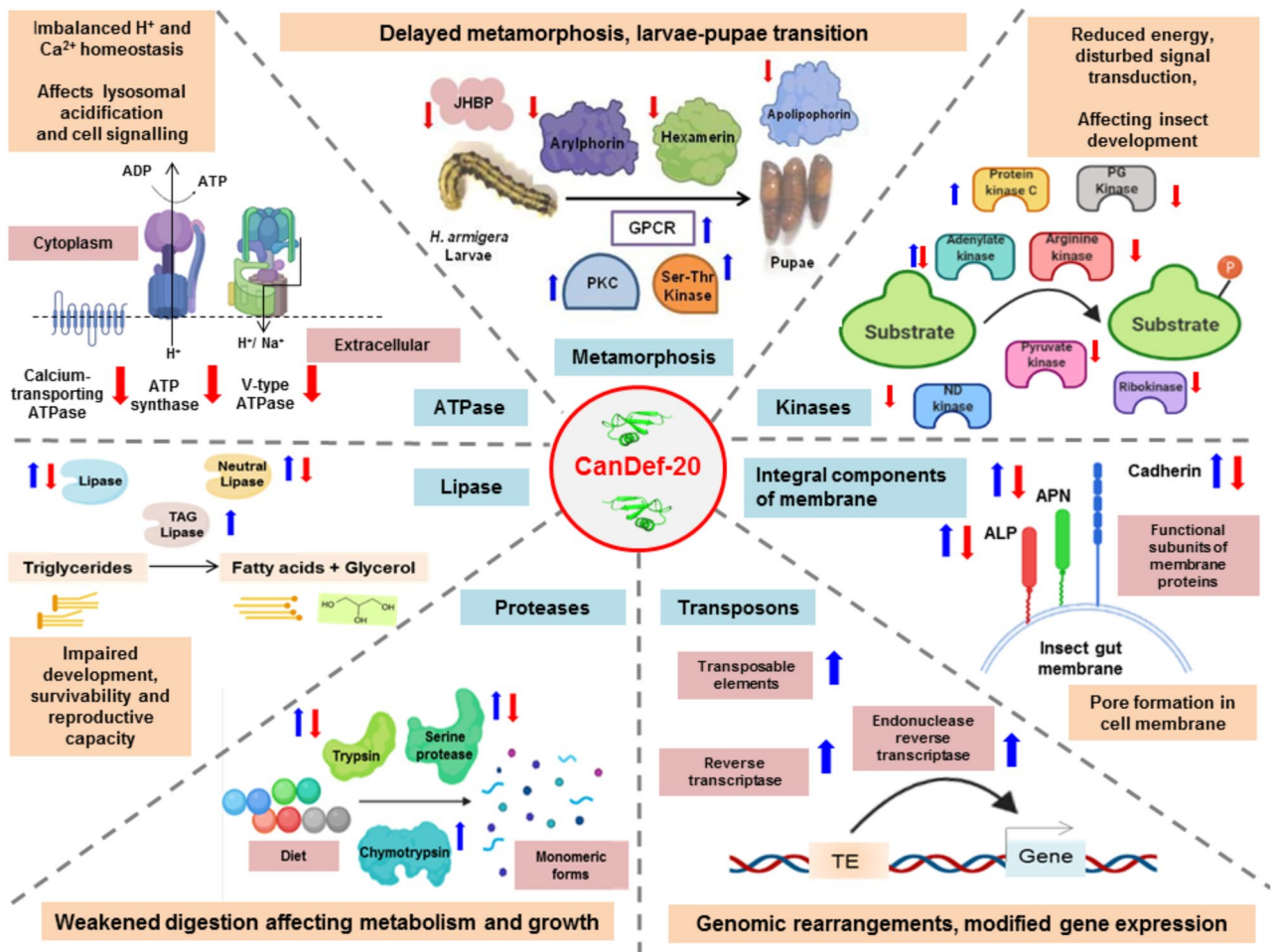
Interference in the activity of ribosomal machinery may impact the growth and development of insects<sup>69</sup>. In *Drosophila*, the transcript level of ribosomes and protein synthesis is downregulated under starvation of dietary amino acids<sup>70</sup>. The decrease in the transcript level of ribosomes is the result of nutrient stress posed by CanDef-20 feeding by *H. armigera*.





**Figure 8.** RNA-seq gene expression validation by qPCR analysis: qPCR analysis for selected genes is performed to validate the RNA-seq data. **(a)** Juvenile hormone binding protein (JHBP), hexamerin, calphotin, arginine kinase, pyruvate kinase, adenylate kinase 6, **(b)** V-type proton ATPase and its subunit C and B, **(c)** aminopeptidase 2, alkaline phosphatase 2 and aminopeptidase **(d)** N lipase, lipase (LP2), lipase, **(e)** trypsin, serine protease, **(f)** ATP synthase, **(g)** Endonuclease reverse transcriptase (ERT 4, ERT 12 and ERT 2). The relative gene expression values are significantly different from EV control at \* for  $P < 0.05$ , \*\* for  $P < 0.01$ , \*\*\* for  $P < 0.001$  and nd for no difference respectively.

CHS maintains the integrity of the chitin matrix and regulates larval to pupal metamorphosis, egg hatching and fecundity<sup>71</sup>. The silencing of CHS genes resulted in malformed phenotypes and increased mortality with



**Figure 9.** Model diagram representing metabolic reprogramming/pathways affected by CanDef-20 ingestion in *H. armigera*: Selected gene categories like transposons, kinase, integral components of membrane, serine endopeptidase, lipase, ATPase are shown and the role of key proteins involved in the process are highlighted. The down and upregulation of specific proteins and enzymes are shown by red and blue arrows respectively.

decreased molting rate in the nymph of *Locusta migratoria manilensis*<sup>72</sup> and egg laying in red mite *Panonychus citri*<sup>73</sup>. The extension of the larval period and decrease in egg laying may be a consequence of the downregulation of the CHS gene in CanDef-20 fed *H. armigera* larvae (Fig. 2).

Juvenile hormone binding protein (JHBP) in association with juvenile hormone (JH) and storage proteins like hexamerin, arylphorin, apolipoporphin are essential for multiple physiological processes in insects including larval development, metamorphosis, and adult reproduction<sup>74</sup>. It is reported that down regulation of JHBP leads to developmental delays in *H. armigera* larvae<sup>74</sup>. Also, hexamerin was found to be significantly down-regulated in pine sawyer beetle (*Monochamus alternatus*) when fed with insecticides (chloramine phosphorus)<sup>75</sup>, and arylphorin was down-regulated in Vip3Aa toxin-treated larvae of *Spodoptera exigua*<sup>76</sup>. Thus, down regulation of JHBP, hexamerin and arylphorin might have impact on metamorphosis of *H. armigera* larvae leading to a prolonged larval stage. However, counter upregulated transcripts of protein kinase, serine-threonine kinase and G-protein receptors helps in proceeding for pupal stage.

Endonuclease reverse transcriptase (retrotransposons) and transposable element (TEs) were the most represented upregulated transcripts in CanDef-20 fed *H. armigera* larvae. TEs by way of insertion mutagenesis and recombination near the functional genes contribute to the generation of novel mutations and are responsible to generate much of the genetic diversity that contributes to rapid evolution<sup>77</sup>. Insertion of TEs near detoxification genes like P450 conferred the adaptation and resistance to insecticides in *H. zea* moths and *D. melanogaster* flies<sup>78</sup>. The transposon insertion in gut membrane receptor cadherin genes leads to Bt resistance in pink bollworm<sup>79</sup>. We observed that feeding with CanDef-20 triggered the mobilization of RNA-type TEs like LINE jockey, LTR gypsy and LINE RTE in *H. armigera*. *H. virescens* showed upregulation of retrotransposons and DNA transposons in response to ingestion of glucosinolates from *Arabidopsis thaliana*<sup>80</sup>. Increased expression and transposition of TEs in response to CanDef-20 may generate heritable genetic diversity in *H. armigera* larvae and confer further selective advantages to the insects, which needs further investigations.

Disturbance in metabolism of larval stage due to the presence of inhibitors or toxins in the diet often offsets its life cycle. Our findings demonstrate that CanDef-20 ingestion exhibited larval and pupal growth inhibition in

*H. armigera* larvae. Comparative transcriptomics analysis indicated a downregulation and differential upregulation of genes involved in various metabolic pathways including energy homeostasis and enzymes related to food digestion, xenobiotic detoxification upon exposure to CanDef-20. We conclude that the retardation in the growth and development of *H. armigera* larvae ingesting CanDef-20 is mediated through its interactions with secreted proteins/enzymes, membrane components and regulation of target gene expression mediated via transcription factors and transposons. Our study provides a basis for future research on the potential practical applications of defensin peptide mediated control of the damaging pests like *H. armigera*.

## Data availability

The transcriptome assembly data generated during the current study is available at SRA (NCBI database) under the bioproject PRJNA869208. Additional data files on the transcriptome assembly are available at Mendeley data <https://doi.org/10.17632/gwjdhw6rw3.1>.

Received: 10 September 2022; Accepted: 1 February 2023

Published online: 23 February 2023

## References

- Stacke, R. F. *et al.* Damage assessment of *Helicoverpa armigera* (Lepidoptera: Noctuidae) in soybean reproductive stages. *Crop Prot.* **112**, 10–17 (2018).
- da Silva, F. R. *et al.* Comparative toxicity of *Helicoverpa armigera* and *Helicoverpa zea* (Lepidoptera: Noctuidae) to selected insecticides. *Insects* **11**, 431 (2020).
- Després, L., David, J. P. & Gallet, C. The evolutionary ecology of insect resistance to plant chemicals. *Trends Ecol. Evol.* **22**, 298–307 (2007).
- Gahan, L. J., Gould, F. & Heckel, D. G. Identification of a gene associated with Bt resistance in *Heliothis virescens*. *Science* **293**, 857–860 (2001).
- Zhao, M. *et al.* Functional roles of cadherin, aminopeptidase-N and alkaline phosphatase from *Helicoverpa armigera* (Hübner) in the action mechanism of *Bacillus thuringiensis* Cry2Aa. *Sci. Rep.* **7**, 1–9 (2017).
- Tiewisiri, K. & Wang, P. Differential alteration of two aminopeptidases N associated with resistance to *Bacillus thuringiensis* toxin Cry1Ac in cabbage looper. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 14037 (2011).
- Guo, Z. *et al.* MAPK signaling pathway alters expression of midgut ALP and ABCC genes and causes resistance to *Bacillus thuringiensis* Cry1Ac toxin in diamondback moth. *PLoS Genet.* **11**, 78 (2015).
- Yu, G., Baeder, D. Y., Regoes, R. R. & Rolff, J. Predicting drug resistance evolution: Insights from antimicrobial peptides and antibiotics. *Proc. R. Soc. B Biol. Sci.* **285**, 51 (2018).
- Bruce, T. J. A. Interplay between insects and plants: Dynamic and complex interactions that have coevolved over millions of years but act in milliseconds. *J. Exp. Bot.* **66**, 455–465 (2015).
- Bleackley, M. R. *et al.* *Nicotiana glauca* defensin chimeras reveal differences in the mechanism of fungal and tumor cell killing and an enhanced antifungal variant. *Antimicrob. Agents Chemother.* **60**, 6302 (2016).
- Mello, E. O. *et al.* Antifungal activity of PvD1 defensin involves plasma membrane permeabilization, inhibition of medium acidification, and induction of ROS in fungi cells. *Curr. Microbiol.* **62**, 1209–1217 (2011).
- Velivelli, S. L. S., Islam, K. T., Hobson, E. & Shah, D. M. Modes of action of a Bi-domain plant defensin MtDef5 against a bacterial pathogen *Xanthomonas campestris*. *Front. Microbiol.* **9**, 51 (2018).
- Tam, J., Wang, S., Wong, K. & Tan, W. Antimicrobial peptides from plants. *Pharmaceuticals* **8**, 711–757 (2015).
- Kovaleva, V., Bukhteeva, I., Kit, O. Y. & Nesmelova, I. V. Plant defensins from a structural perspective. *Int. J. Mol. Sci.* **21**, 5307 (2020).
- Sagaram, U. S. *et al.* Structural and functional studies of a phosphatidic acid-binding antifungal plant defensin MtDef4: Identification of an RGFRRR motif governing fungal cell entry. *PLoS ONE* **8**, e82485 (2013).
- de Mello, É. O., Taveira, G. B., de Carvalho, A. O. & Gomes, V. M. Improved smallest peptides based on positive charge increase of the  $\gamma$ -core motif from PvD 1 and their mechanism of action against *Candida* species. *Int. J. Nanomed.* **14**, 407–420 (2019).
- Mulla, J. A., Kibe, A. N., Deore, D. D., Jadhav, A. R. & Tamhane, V. A. Molecular characterization of diverse defensins ( $\gamma$ -thionins) from *Capsicum annuum* flowers and their effects on the insect pest *Helicoverpa armigera*. *Plant Gene* **26**, 100284 (2021).
- King, J. & Laemmli, U. K. Polypeptides of the tail fibres of bacteriophage T4. *J. Mol. Biol.* **62**, 465–477 (1971).
- Tamhane, V. A., Giri, A. P., Sainani, M. N. & Gupta, V. S. Diverse forms of Pin-II family proteinase inhibitors from *Capsicum annuum* adversely affect the growth and development of *Helicoverpa armigera*. *Gene* **403**, 29–38 (2007).
- Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. Available at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed 5th August 2022.
- Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
- Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
- TransDecoder - bio.tools. Available at: <https://bio.tools/TransDecoder>. Accessed 5th August 2022.
- RSEM. Available at: <https://deweylab.github.io/RSEM/README.html>. Accessed 5th August 2022.
- Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2009).
- Prüfer, K. *et al.* FUNC: A package for detecting significant associations between gene sets and ontological annotations. *BMC Bioinform.* **8**, 1–10 (2007).
- Bernfeld, P. Amylases,  $\alpha$  and  $\beta$ . *Methods Enzymol.* **1**, 149–158 (1955).
- Tamhane, V. A. *et al.* In vivo and in vitro effect of *Capsicum annuum* proteinase inhibitors on *Helicoverpa armigera* gut proteinases. *Biochim. Biophys. Acta Gen. Subj.* **1722**, 156–167 (2005).
- Labade, C. P., Jadhav, A. R., Ahire, M., Zinjarde, S. S. & Tamhane, V. A. Role of induced glutathione-S-transferase from *Helicoverpa armigera* (Lepidoptera: Noctuidae) HaGST-8 in detoxification of pesticides. *Ecotoxicol. Environ. Saf.* **147**, 612–621 (2018).
- Winkler, U. K. & Stuckmann, M. Glycogen, hyaluronate, and some other polysaccharides greatly enhance the formation of exolipase by *Serratia marcescens*. *J. Bacteriol.* **138**, 663–670 (1979).
- Sarate, P. J. *et al.* Developmental and digestive flexibilities in the midgut of a polyphagous pest, the cotton bollworm, *Helicoverpa armigera*. *J. Insect Sci.* **12**, 1–16 (2012).
- Arenas, I., Bravo, A., Soberón, M. & Gómez, I. Role of alkaline phosphatase from *Manduca sexta* in the mechanism of action of *Bacillus thuringiensis* Cry1Ab toxin. *J. Biol. Chem.* **285**, 12497 (2010).
- Lacerda, A. F., Vasconcelos, Á. A. R., Pelegrini, P. B. & de Grossi, S. M. F. Antifungal defensins and their role in plant defense. *Front. Microbiol.* **5**, 1–10 (2014).

34. Chen, K. C., Lin, C. Y., Kuan, C. C., Sung, H. Y. & Chen, C. S. A novel defensin encoded by a mungbean cDNA exhibits insecticidal activity against bruchid. *J. Agric. Food Chem.* **50**, 7258–7263 (2002).
35. Lay, F. T., Brugliera, F. & Anderson, M. A. Isolation and properties of floral defensins from ornamental tobacco and petunia. *Plant Physiol.* **131**, 1283–1293 (2003).
36. Aerts, A. M. *et al.* The antifungal plant defensin HsAFP1 from *Heuchera sanguinea* induces apoptosis in *Candida albicans*. *Front. Microbiol.* **2**, 74 (2011).
37. Lee, J. *et al.* Coprisin-induced antifungal effects in *Candida albicans* correlate with apoptotic mechanisms. *Free Radic. Biol. Med.* **52**, 2302–2311 (2012).
38. Poon, I. K. H. *et al.* Phosphoinositide-mediated oligomerization of a defensin induces cell lysis. *Elife* **2014**, e01808 (2014).
39. Hayes, B. M. E., Bleackley, M. R., Anderson, M. A. & van der Weerden, N. L. The plant defensin NaD1 enters the cytoplasm of *Candida albicans* via endocytosis. *J. Fungi (Basel Switzerland)* **4**, 89 (2018).
40. Melo, F. R. *et al.* Inhibition of trypsin by cowpea thionin: Characterization, molecular modeling, and docking. *Proteins Struct. Funct. Genet.* **48**, 311–319 (2002).
41. Vijayan, S. *et al.* Enhanced antifungal and insect  $\alpha$ -amylase inhibitory activities of Alpha-TvD1, a peptide variant of *Tephrosia villosa* defensin (TvD1) generated through in vitro mutagenesis. *Peptides* **33**, 220–229 (2012).
42. Pelegrini, P. B. & Franco, O. L. Plant  $\gamma$ -thionins: Novel insights on the mechanism of action of a multi-functional class of defense proteins. *Int. J. Biochem. Cell Biol.* **37**, 2239–2253 (2005).
43. 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).
44. Lomate, P. R. *et al.* Integrated transcriptomic and proteomic analyses suggest the participation of endogenous protease inhibitors in the regulation of protease gene expression in *Helicoverpa armigera*. *Mol. Cell. Proteom.* **17**, 1324–1336 (2018).
45. Ashouri, S. & Farshbaf Pourabad, R. Regulation of gene expression encoding the digestive  $\alpha$ -amylase in the larvae of Colorado potato beetle, *Leptinotarsa decemlineata* (Say) in response to plant protein extracts. *Gene* **766**, 145159 (2021).
46. Qiao, C., Cui, F. & Yan, S. Structure, function and applications of carboxylesterases from insects for insecticide resistance. *Protein Pept. Lett.* **16**, 1181–1188 (2009).
47. Toprak, U. The role of peptide hormones in insect lipid metabolism. *Front. Physiol.* **11**, 434 (2020).
48. Vihervaara, T. & Puig, O. dFOXO regulates transcription of a drosophila acid lipase. *J. Mol. Biol.* **376**, 1215–1223 (2008).
49. Grönke, S. *et al.* Brummer lipase is an evolutionary conserved fat storage regulator in drosophila. *Cell Metab.* **1**, 323–330 (2005).
50. Zhang, J. *et al.* Population genomics provides insights into lineage divergence and local adaptation within the cotton bollworm. *Mol. Ecol. Resour.* **22**, 1875–1891 (2022).
51. Baxter, A. A. *et al.* The tomato defensin TPP3 binds phosphatidylinositol (4,5)-bisphosphate via a conserved dimeric cationic grip conformation to mediate cell lysis. *Mol. Cell. Biol.* **35**, 1964–1978 (2015).
52. Velivelli, S. L. S., Islam, K. T., Hobson, E. & Shah, D. M. Modes of action of a Bi-domain plant defensin MtDef5 against a bacterial pathogen *Xanthomonas campestris*. *Front. Microbiol.* **9**, 78 (2018).
53. De Medeiros, L. N. *et al.* Psd1 binding affinity toward fungal membrane components as assessed by SPR: The role of glucosylceramide in fungal recognition and entry. *Biopolymers* **102**, 456–464 (2014).
54. Harvey, W. R. & Wiczorek, H. Animal plasma membrane energization by chemiosmotic H<sup>+</sup> V-ATPases. *J. Exp. Biol.* **200**, 203–216 (1997).
55. Reineke, S., Wiczorek, H. & Merzendorfer, H. Expression of *Manduca sexta* V-ATPase genes mvB, mvG and mvd is regulated by ecdysteroids. *J. Exp. Biol.* **205**, 1059–1067 (2002).
56. Mao, J., Zhang, P., Liu, C. & Zeng, F. Co-silence of the coatomer  $\beta$  and v-ATPase A genes by siRNA feeding reduces larval survival rate and weight gain of cotton bollworm, *Helicoverpa armigera*. *Pestic. Biochem. Physiol.* **118**, 71–76 (2015).
57. Zhang, H. *et al.* Intra- and extracellular domains of the *Helicoverpa armigera* cadherin mediate Cry1Ac cytotoxicity. *Insect Biochem. Mol. Biol.* **86**, 41–49 (2017).
58. Jurat-Fuentes, J. L. & Adang, M. J. Characterization of a Cry1Ac-receptor alkaline phosphatase in susceptible and resistant *Heliothis virescens* larvae. *Eur. J. Biochem.* **271**, 3127–3135 (2004).
59. Sonu Koirala, B. K., Moural, T. & Zhu, F. Functional and structural diversity of insect glutathione S-transferases in xenobiotic adaptation. *Int. J. Biol. Sci.* **18**, 5713 (2022).
60. Ellington, W. R. Evolution and physiological roles of phosphagen systems. *Annu. Rev. Physiol.* **63**, 289–325 (2001).
61. Zhao, Y., Yang, G., Wang-Pruski, G. & You, M. *Phyllotreta striolata* (Coleoptera: Chrysomelidae): Arginine kinase cloning and RNAi-based pest control. *Science* **105**, 815–822 (2013).
62. Kayne, F. J. *Pyruvate Kinase*. *Enzymes* (Academic Press, London, 1973).
63. Xu, K. K., Pan, B. Y., Wang, Y. Y., Ren, Q. Q. & Li, C. Roles of the PTP61F gene in regulating energy metabolism of *Tribolium castaneum* (Coleoptera: Tenebrionidae). *Front. Physiol.* **11**, 1071 (2020).
64. Wang, T., Geng, S. L., Guan, Y. M. & Xu, W. H. Deacetylation of metabolic enzymes by Sirt2 modulates pyruvate homeostasis to extend insect lifespan. *Aging (Albany NY)* **10**, 1053 (2018).
65. Parisi, K. *et al.* The evolution, function and mechanisms of action for plant defensins. *Semin. Cell Dev. Biol.* **88**, 107–118 (2019).
66. Chaitanya, R. K., Shashank, K. & Sridevi, P. Oxidative Stress in Invertebrate Systems. in *Free Radicals and Diseases* (ed. Rizwan Ahmad) (IntechOpen, 2016). <https://doi.org/10.5772/64573>
67. Aerts, A. M. *et al.* The antifungal activity of RsAFP2, a plant defensin from *Raphanus sativus*, involves the induction of reactive oxygen species in *Candida albicans*. *J. Mol. Microbiol. Biotechnol.* **13**, 243–247 (2007).
68. Souza, G. S., de Carvalho, L. P., de Melo, E. J. T., Gomes, V. M. & de Carvalho, A. O. The toxic effect of Vu-Defra, a defensin from *Vigna unguiculata* seeds, on *Leishmania amazonensis* is associated with reactive oxygen species production, mitochondrial dysfunction, and plasma membrane perturbation. *Can. J. Microbiol.* **64**, 455–464 (2018).
69. Zhu, J. *et al.* Expression and RNA interference of ribosomal protein L5 gene in *Nilaparvata lugens* (Hemiptera: Delphacidae). *J. Insect Sci.* **17**, 895 (2017).
70. Li, L., Edgar, B. A. & Grewal, S. S. Nutritional control of gene expression in drosophila larvae via TOR, Myc and a novel cis-regulatory element. *BMC Cell Biol.* **11**, 1–13 (2010).
71. Merzendorfer, H. & Zimoch, L. Chitin metabolism in insects: Structure, function and regulation of chitin synthases and chitinases. *J. Exp. Biol.* **206**, 4393–4412 (2003).
72. Zhang, J. *et al.* Silencing of two alternative splicing-derived mRNA variants of chitin synthase 1 gene by RNAi is lethal to the oriental migratory locust, *Locusta migratoria manilensis* (Meyen). *Insect Biochem. Mol. Biol.* **40**, 824–833 (2010).
73. Ali, M. W. *et al.* RNA interference-based silencing of the chitin synthase 1 gene for reproductive and developmental disruptions in *Panonychus citri*. *Insects* **11**, 1–18 (2020).
74. Zhang, W., Liang, G., Ma, L., Jiang, T. & Xiao, H. Dissecting the role of juvenile hormone binding protein in response to hormone and starvation in the cotton bollworm, *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae). *J. Econ. Entomol.* **112**, 1411–1417 (2019).
75. Lin, T., Luo, L., Wei, C. & Wu, H. Microarray gene expression analysis of *Monochamus alternatus* (Coleoptera: Cerambycidae) after treatment with a sublethal dose of chloramine phosphorus. *Appl. Entomol. Zool.* **49**, 223–230 (2014).
76. Bel, Y., Jakubowska, A. K., Costa, J., Herrero, S. & Escriche, B. Comprehensive analysis of gene expression profiles of the beet armyworm *Spodoptera exigua* larvae challenged with *Bacillus thuringiensis* Vip3Aa toxin. *PLoS ONE* **8**, e81927 (2013).



77. Van De Lagemaat, L. N., Landry, J. R., Mager, D. L. & Medstrand, P. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet.* **19**, 530–536 (2003).
78. Chen, S. & Li, X. Transposable elements are enriched within or in close proximity to xenobiotic-metabolizing cytochrome P450 genes. *BMC Evol. Biol.* **7**, 46 (2007).
79. Wang, L. *et al.* Transposon insertion causes cadherin mis-splicing and confers resistance to Bt cotton in pink bollworm from China. *Sci. Rep.* **9**, 1–10 (2019).
80. Schweizer, F., Heidel-Fischer, H., Vogel, H. & Reymond, P. Arabidopsis glucosinolates trigger a contrasting transcriptomic response in a generalist and a specialist herbivore. *Insect Biochem. Mol. Biol.* **85**, 21–31 (2017).

## Acknowledgements

JM and VT thank Department of Science and Technology (DST), Science and Engineering Research Board (SERB), Government of India for providing the funding (CRG/2018/002324) and the Departmental Research and Development Program (DRDP), Institute of Bioinformatics and Biotechnology (IBB), Savitribai Phule Pune University (SPPU), Pune, Maharashtra, India for funding and supporting this work. We are thankful to the Head, Department of Microbiology for allowing us to use the qPCR machine. We are also thankful to Mr. Onkar Ghuge and Miss. Rutuja Shirke for helping in maintenance of insect culture and enzyme assay.

## Author contributions

J.M. – Experimentation, Data analysis, Figures and Manuscript writing. V.T. – Supervision, Conceptualization, Data analysis, Figures and Writing the manuscript, Project administration and Funding acquisition.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-29250-3>.

**Correspondence** and requests for materials should be addressed to V.A.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023