Research Article

Received: 20 October 2019 Revised: 6 January 2020 Accepted article published: 24 January 2020 Published online in Wiley Online Library: 3 March 2020

(wileyonlinelibrary.com) DOI 10.1002/ps.5761

NPC1b as a novel target in controlling the cotton bollworm, *Helicoverpa armigera*

Jin-Cheng Zheng,[a,b,c†] Xiao-Rong Yue,[a,b†] Wen-Qing Kuang,[a,b] Sa-Li Li,[a,b] Rui Tang,[a,b] Zhan-Feng Zhang,[a,b] Anwar Kurban,[d] Saif-Ur-Rehman,[e] Chaoyang Zhao,[f] Tong-Xian Liu[a,b] and Xiangfeng Jing[a,b*]

Abstract

BACKGROUND: Insects cannot synthesize sterols and must acquire them from food. The mechanisms underlying how insects uptake dietary sterols are largely unknown except that NPC1b, an integral membrane protein, has been shown to be responsible for dietary cholesterol uptake in *Drosophila melanogaster*. However, whether *NPC1b* orthologs in other insect species, particularly the economically important pests, function similarly remains to be determined.

RESULTS: In this study, we characterized the function of *NPC1b* in *Helicoverpa armigera*, a global pest that causes severe yield losses to many important crops. Limiting dietary cholesterol uptake to insects significantly inhibited food ingestion and weight gain. Compared to the wild-type *H. armigera*, the CRISPR/Cas9-edited *NPC1b* mutant larvae were incapable of getting adequate cholesterol and died in their early life stage. Gene expression profile and *in situ* hybridization analyses indicated that *NPC1b* was mainly expressed in the midgut where dietary cholesterol was absorbed. Expression of *NPC1b* was also correlated with the feeding life stages and was especially upregulated during early larval instars. Protein-ligand docking and sequence similarity analyses further demonstrated that NPC1b proteins of lepidopteran insects shared a relatively conserved cholesterol binding region, NPC1b_NTD, which, however, was highly divergent from bees-derived sequences.

CONCLUSION: *NPC1b* was crucial for dietary cholesterol uptake and growth of *H. armigera*, and therefore could serve as an insecticide target for the development of a novel pest-management approach to control this economically significant insect pest with little off-target effect on bees and sterol-autotrophic animals.

Supporting information may be found in the online version of this article.

Keywords: *Helicoverpa armigera*; cholesterol absorption; *NPC1b*; CRISPR/Cas9

1 INTRODUCTION

The cotton bollworm, *Helicoverpa armigera*, is an important agricultural pest that mainly attacks reproductive organs of economically important crops such as cotton, tomato, legume, sorghum, maize, and fruit trees.[1,2] This species is one of the most widely distributed insect pests and causes substantial yield losses in Africa, Asia, Australia, and Europe. Recently, *H. armigera* was reported to be expanding its range in South and North America.[3] The successful management of this pest heavily relies on pesticides and transgenic crops expressing toxins from the bacterium *Bacillus thuringiensis*. However, with the escalating resistance evolved by insects and the increasing environmental concerns from overuse, particularly the negative impacts on non-target organisms, people are continuously losing the control power over this pest.[4–6] It is hence in an urgent need to discover environment-friendly agents that efficiently act on novel molecular targets of insects with high specificity.

Cholesterol and its derivatives are essential for all eukaryotes as both a vital structural component of cellular membrane and a precursor to important biomolecules.[7,8] However, unlike mammals, insects and other arthropods cannot synthesize sterols *de novo*,[9] and thus typically obtain them from food. Previous studies have

* *Correspondence to: X Jing, Key Laboratory of Integrated Pest Management on the Loess Plateau of Ministry of Agriculture, College of Plant Protection, Northwest A&F University, Yangling, Shaanxi 712100, China. E-mail: jxf_zb@sina.cn*

† *These two authors contributed equally to this work.*

a *State Key Laboratory of Crop Stress Biology for Arid Areas, College of Plant Protection, Northwest A&F University, Yangling, Shaanxi, China*

b *Key Laboratory of Integrated Pest Management on the Loess Plateau of Ministry of Agriculture, College of Plant Protection, Northwest A&F University, Yangling, Shaanxi, China*

c *Basic Medical School, Guizhou University of Traditional Chinese Medicine, Guiyang, Guizhou, China*

d *College of Agronomy, Xinjiang Agricultural University, Urumqi, Xinjiang, China*

e *College of Plant Protection, Northwest A&F University, Yangling, Shaanxi, China*

f *College of Natural and Agricultural Sciences, University of California, Riverside, CA, USA*

demonstrated that interfering with dietary cholesterol usage delays insect growth and increases insect mortality.[10–19]

The uptake of cholesterol in animals requires these molecules to move across intestinal epithelium to reach various internal tissues. This absorbing process includes two steps: (i) dietary cholesterol transportation through the apical membrane of enterocytes into cells, and (ii) intracellular cholesterol trafficking through the cytoplasm to the basal membrane where cholesterol is loaded into lipophorin.[8] Sterol carrier protein-2 (SCP-2), a soluble protein abundant in enterocyte cytoplasm, has been shown to play an important role in the second step.[13,20–22] Inhibitors targeting SCP-2 significantly restricted dietary cholesterol trafficking, and showed high toxicity to insect pests and low toxicity to mammals that can synthesize their own cholesterol.[14–16,18,22] Comparably, Niemann–Pick Type C1 L1 (NPC1L1) protein was found to function in the first step of the cholesterol absorbing process in mammals, through mediating the transportation of dietary cholesterol from gut lumen into enterocytes across the apical membrane of cells.[23,24] In parallel, the *NPC1L1* ortholog identified in *D. melanogaster*, named *NPC1b*, was also shown to be essential for dietary cholesterol uptake.[25] These orthologous proteins, NPC1L1 and NPC1b, are both integral membrane proteins and typically comprise three domains, the N-terminal domain (NTD), the sterol sensing domain (SSD), and the Patched domain, among which NTD specifically binds cholesterol to mediate its transportation.[26,27] Given the structural and functional conservation shared by the *D. melanogaster* NPC1b and mammal NPC1L1 proteins, it is reasonable to hypothesize that employing *NPC1b* orthologs for cholesterol uptake is a conserved mechanism among insects, despite that some insect species may lack this gene and have evolved other mechanisms as substitutes.[27]

To test this hypothesis and to discover a novel insecticide target, we here investigated the molecular and functional characteristics of *NPC1b* in *H. armigera*. Results indicated that the temporal and spatial expression of *NPC1b* was linked to dietary sterol uptake. Loss-of-function mutation of *NPC1b* in *H. armigera* generated by the CRISPR/Cas9 genome editing technology disrupted insect growth and resulted in lethality. In addition, we found that the cholesterol-binding domain, NTD, of NPC1b was conserved in caterpillar species but was divergent from the NTD domain in bees. This work, to our knowledge, represents the second example to reveal the indispensable role that NPC1b plays in cholesterol uptake in insects and will shed light on the development of novel pest management strategies.

2 MATERIALS AND METHODS

2.1 Experimental animals

Helicoverpa armigera pupae were purchased from Keyun Industry Co., Ltd (Jiyuan, Henan, China). The newly emerged adults were allowed to mate and produce eggs. The rearing conditions were set as 27 °C/25 °C [light/dark (L/D)] with a 16 h L/8 h D photoperiod.

2.2 Insect performance on different concentrations of dietary cholesterol

An artificial diet developed for *Helicoverpa zea* was used and this diet contained trace level of sterols unless sterols are manually added.[28,29] *Helicoverpa armigera* hatchlings were reared on the diet containing cholesterol at a concentration of 1 mg/g, which resembles typical sterol content in plant tissues.[12] Upon molting into 4th instar, the insects were weighted (initial weight) and then randomly transferred onto one of the five diets which contained 0, 0.25 mg/g, 0.5 mg/g, 1 mg/g, and 2 mg/g of cholesterol, respectively. Thirty larvae were used for each treatment. Wet diet blocks were weighed prior to insect feeding and replaced every other day. Removed blocks were dried at 50 °C to constant mass and then weighed. To estimate the initial dry mass, wet mass and dry mass of five blocks were individually measured and a regression equation was generated to determine the water content of wet diets. Food consumed by each insect was calculated as the food dry mass change before and after being fed. After feeding for 5 days, the larvae were weighed (final weight) and food consumption was calculated. The relative growth rate (RGR) was individually determined by the formula: RGR = [$\log_e$(final weight/initial weight)]/number of days.[30]

2.3 Generation of *NPC1b* mutant using CRISPR/Cas9

Helicoverpa armigera NPC1b gene was previously identified,[27] and the genomic DNA was verified by Sanger sequencing (Supporting Information, Fig. S1, Table S1). We used long deletion knockout strategies to generate insect mutants through the co-injection of two guide-RNAs (sgRNAs) which target a 334-bp fragment of *NPC1b* (Fig. S2).[31] The specificity of sgRNA sequence was examined against *H. armigera* genome using the CRISPR/Cas system (Cas9/gRNA) off-targeter (CasOT, http://casot.cbi.pku.edu.cn), and the off-target sites with relatively high risk of mutations were experimentally evaluated.[32]

To synthesize *H. armigera NPC1b* sgRNAs, we initially generated their DNA templates using the GeneArt™ Precision gRNA Synthesis Kit (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's instruction. The DNA templates were then purified with the QIAquick PCR purification kit (QIAGEN, Hiden, Germany) and used to synthesize the sgRNAs with GeneArt™ Precision gRNA Synthesis Kit (Thermo Fisher Scientific). The sgRNAs were stored at −80 °C for further use.

Eggs laid within 2 h were fixed onto a microscope slide (24 mm, 50 mm) using a double-sided adhesive tape. A mixture containing two sgRNAs (100 ng/μL for each) and Cas9 protein (125 ng/μL, TrueCut Cas9 Protein v2, Thermo Fisher Scientific, Shanghai, China) was injected into each embryo using PV820 Pneumatic PicoPump (World Precision Instruments Inc., New Haven, CT, USA) equipped with Sutter Instrument 0.58 mm fire-polished glass capillary needles (Sutter BF-100-58-10). Eggs injected with Cas9 protein alone were used as control. Hatchling (G_0 generation) from these injected eggs were used for performance assay and cholesterol quantification.

2.4 Dissection and tissue collection

Head, epidermis and whole alimentary canal were dissected from 2-day-old 5th instar *H. armigera* in ice-cold phosphate-buffered saline (PBS) buffer (Beyotime, Haimen, China). The whole canal was separated into foregut (FG), the first-half midgut (MG1), the second-half midgut (MG2), Malpighian tubes (MT), and hindgut (HG) (Fig. S3(A)). All tissues were kept in RNA Stabilization Reagent (Qiagen, Hiden, Germany) until RNA extraction. Each biological replicate comprised material from two larvae for epidermis, four larvae for each part of the whole canal (except for Malpighian tubules which were from ten larvae) or ten larvae for head, and six biological replicates were prepared for each type of tissue.

To investigate the response of *NPC1b* to dietary cholesterol concentrations, five diets with different cholesterol concentrations, i.e. 0, 0.25 mg/g, 0.5 mg/g, 1 mg/g and 2 mg/g, were prepared. *Helicoverpa armigera* hatchlings were reared on the 1 mg/g diet

and then randomly assigned to each of these five diets upon reaching the 3rd instar. The guts were dissected from 2-day-old 3rd instar insects to quantify the gene expression. Eight biological replicates were collected for each treatment and each replicate comprised two guts.

To investigate the sites for sterol absorption, 2-day-old 5th instar insects were used. The whole alimentary canal connecting head capsule and anus was quickly dissected on ice with food residue in gut lumen removed. The whole gut was cut into four sections (FG, MG1, MG2, and HG) as described earlier and Malpighian tubules were removed. These tissues were subjected to gas chromatography mass spectrometry (GC-MS) analysis and four individual larvae were used. The relative cholesterol content in the four sections of a gut from one same individual insect was calculated relative to the section with highest cholesterol content (set as 100%).

To investigate gene expression at different life stages, two experiments were conducted. In one experiment, the insects were fed with the diet containing 1 mg/g of cholesterol and relative expression in eight life stages, e.g. 4-day-old eggs, hatchlings unfed for 6 h (i.e. no food was given), hatchlings fed on the diet for 6 h, 2-day-old larvae of 2nd, 3rd, 4th and 5th instars, and 4-day-old pupae were collected respectively for gene expression quantification. Eight biological replicates were prepared. In the other experiment, we used the diets containing different concentrations of cholesterol (0.5 mg/g, 1 mg/g, and 2 mg/g) to feed the insects and five different stages, e.g. 4-day-old eggs, hatchlings unfed for 6 h, hatchlings fed on the diet for 6 h, 2-day-old larvae of 5th instars, and 4-day-old pupae were collected. Four biological replicates were prepared.

2.5 Nucleic acid extraction and cDNA preparation

Total RNA was extracted using the M5 Universal RNA Mini Kit (Mei5 Biotechnology, Beijing, China) and genomic DNA was digested following manufacturer's protocols. Approximately 1 μg RNA was used to prepare cDNA by the PrimeScript™ RT reagent kit with gDNA Eraser (TaKaRa, Dalian, China) according to the manufacturer's protocol.

Extraction of insect genomic DNA was performed using the Universal DNA Mini Kit (Mei5 Biotechnology). Quantity and purity of extracted DNA were examined with a NanoDrop™ 2000 (Thermo Fisher Scientific, San Jose, CA, USA).

2.6 PCR primers and gene sequence verification

Primers used for polymerase chain reaction (PCR) analyses, including regular and quantitative, are listed in Table S1. DNA sequences of PCR products were validated by Sanger sequencing (TSINGKE, Beijing, China) when necessary.

2.7 Cholesterol quantification

We used a previously described protocol for cholesterol quantification.[11,33] Briefly, insect samples were freeze-dried and total sterols (free sterols and sterol esters) were extracted using a mixed solvent of chloroform and methanol. Cholestane was added into each sample as the internal standard. Subsequently, water was added, and the lower layer (chloroform part) was removed and concentrated. Cholesterol esters were hydrolyzed by incubating with 70% methanol [5% potassium hydroxide (KOH) *w/v*] and sterols were then extracted by hexane. The hexane solution was washed with 50% methanol/water to remove any basic residues. Sterols were derivatized to trimethylsilyl ether (TMS), and then identified and quantified by GC-MS using a Thermo trace 1310 GC coupled with an ISQ MS detector (Thermo Fisher Scientific) maintaining Agilent DB-17 column (Agilent Technologies, Palo Alto, CA, USA).

2.8 Quantitative real-time PCR (qRT-PCR) analysis

Each quantitative real-time PCR (qRT-PCR) reaction was conducted using a mixture of 10 μL real-time PCR super mix (Mei5 Biotechnology), 1 μL primer mix (5 μmol/L each primer), 2 μL cDNA and 7 μL doubly distilled water in QuantStudio 3 real-time PCR system (Thermo Fisher Scientific), and template-free samples were used as negative control. The PCR reaction was as follows: 95 °C for 1 min, and 40 cycles of 95 °C for 15 s, 60 °C for 15 s and 72 °C for 15 s. The denaturing step was added with 0.5 °C for 1 s from 65 °C to 95 °C to confirm single gene-specific peak. Two reference genes, *RPL13* (AY846882.1) and glyceraldehyde-3-phosphatededydrogenase (JF417983.1), were used to normalize gene expression across the different samples analyzed,[34] and a standard curve for each pair of primers was obtained to estimate primer efficiency. Three technical replicates were prepared for each biological sample.

2.9 Fluorescence *in situ* hybridization (FISH) of *NPC1b* mRNA

Guts dissected from 2-day-old 5th instar larvae were fixed in 4% (*w/v*) paraformaldehyde solution (Servicebio, Wuhan, China) at 4 °C, embedded in paraffin, and longitudinally sliced into 5 μm sagittal sections. After being incubated for 2 h at 62 °C, the sections were deparaffinized with xylene, rehydrated with gradient ethanol, and then boiled for 10 min for full antigen retrieval. Sections were then treated with 20 μg/mL proteinase K for 25 min at 37 °C, rinsed in 0.5 M PBS buffer, and incubated for 1 h at 37 °C. These pre-treated sections were then hybridized overnight with antisense DIG-dUTP-labeled *H. armigera NPC1b* RNA probes (503–540 nt) at 37 °C and those prepared without probes were used as control. All sections were washed in saline sodium citrate buffer, blocked in 1% bovine serum albumin (BSA) buffer for 30 min at room temperature, and incubated with 100-time diluted anti-mouse Alexa Fluor 488 (Jackson ImmunoResearch Laboratories, Inc.,West Grove, Pennsylvania, USA) for 50 min at 37 °C. Nuclei were stained with DAPI. The images were captured under Nikon 80i fluorescence microscope (Nikon Corporation, Tokyo, Japan).

2.10 NPC1b_NTD domain modeling and docking with cholesterol

Helicoverpa armigera NPC1b_NTD domain was predicted using the NPC1b protein sequence to search against the Pfam database (https://pfam.xfam.org/). An *in silico* model of *H. armigera* NPC1_NTD was constructed using human NPC1_NTD (PDB accession number: 3gkj) as template model in SWISS-MODEL (https://swissmodel.expasy.org/). The ligand–protein docking analysis was performed in Autodock v4.2 with AutoDockTools v1.5.6.[35] Cholesterol three-dimensional (3D) structure was downloaded from PubChem (https://pubchem.ncbi.nlm.nih.gov/) and docked to NPC1b_NTD using the Lamarckian genetic algorithm. The *x*, *y*, *z* coordinates of the center of the grid box around the protein were specified as 85, 83, and 92, respectively. The number of grid points was 40 × 40 × 50 with the grid point spacing as 0.375 Å. A total of 20 genetic algorithm runs were considered in each case with an initial population of 300 individuals and a maximum number of 5 000 000 energy evaluations. The resulting configurations that have less than or equal to 0.5 Å root mean square deviation

were clustered together and ranked by the energy. The best ligand–receptor orientation from the docked structures was chosen based on the lowest energy and the information in the docking orientation between 3gkj (NPC1_N) and its ligand. Docking complex was visualized in Discovery Studio 2017 R2 Client (Accelrys Software, Inc., San Diego, CA, USA).

2.11 Data analysis

R Version 3.61 was used to conduct survival analysis to compare the survival rate between the control and *NPC1b* mutants. Two sample comparison was performed using unpaired *t*-test. Analysis of variance (ANOVA) was used when multiple treatments were compared, followed by Tukey's *post hoc* test for pair-wise comparisons across all treatments. These analyses were conducted in SPSS 21 (SPSS Inc, Chicago, IL, USA).

3 RESULTS

3.1 Dietary sterol availability significantly effects insect performance

In this study, we initially performed a bioassay to assess the effect of sterol on *H. armigera* growth and development. Through feeding the 4th instar larvae with a variety of cholesterol concentrations, we observed that food intake and the RGR of the insects reduced with the decreasing amount of cholesterol added in the food (Fig. 1). Insects ingested significantly less amount of food (Fig. 1(A)) and grew slower (Fig. 1(B)) when sterol concentration reduced from 2.0 mg/g to 0 mg/g. For example, the larvae feeding on sterol-free diet only consumed 29% of food and the RGR was about half as those feeding on 0.25 mg/g of cholesterol, indicating that dietary sterol acquirement is very important to *H. armigera* larval growth and development.

3.2 *Helicoverpa armigera NPC1b* is critical for insect development and dietary cholesterol absorption

CRISPR/Cas9 was recently applied in lepidopteran species and we here used the dual sgRNA-directed CRISPR-Cas9 system to delete a long fragment in *NPC1b* gene from the *H. armigera* genome (Fig. S2).[1] In pilot experiments, we observed that approximately 58% of eggs injected by sgRNA/Cas9 hatched, which was slightly higher than Cas9-injected eggs (43%) but there was no significant difference (Fig. S4; $t = 2.078$, $P = 0.11$). We also observed that neonates hatched from CRISPR/Cas9-injected eggs were unable to molt to the 2nd instar, indicating that the editing effect of *NPC1b* is larval lethal. Thus, we injected sgRNA/Cas9 into 750 eggs and Cas9 into 640 eggs, respectively, to compare the genotypes and phenotypes of G_0 generation larvae. Out of these eggs, 433 (58%) hatched from the sgRNA/Cas9-injected eggs and 333 (52%) from Cas9-injected eggs, and randomly selected neonates were used for the following experiments.

Equal number of neonates ($n = 72$) from both sgRNA/Cas9-injected and Cas9-injected eggs were reared on 1 mg/g of cholesterol. *NPC1b* mutants only survived for up to 10 days and none of them molted to the 2nd instar, while the insects injected with only Cas9 protein can reached the 4th instar within 10 days (Fig. 2(A, B)). Similarly, *NPC1b* mutants cannot molt to the 2nd instar and more than 70% of the insects died within 10 days when a regular diet was used (Fig. S6). Genomic DNA was extracted from each of 20 sgRNA/Cas9-injected neonates among which 17 individuals exhibited approximately 300-bp-long deletions in the *NPC1b* gene examined by PCR amplification and agarose gel electrophoresis (primers: jianceNPC1b-F and jianceNPC1b-R; Table S1) and the other three were single-site mutations verified by Sanger sequencing (Fig. S2(C, D)). Moreover, the larger PCR band from the three random-selected individuals with the long fragment deletion also contained one site editing (Fig. S2(E)). Two sites in *H. armigera* genome containing only three mismatches to the on-target sites (CasOT, http://casot.cbi.pku.edu.cn) were considered high risk of off-target site mutation,[32] but we did not detect any mutation in these two sites in the randomly selected individuals (Fig. S5, Tables S1 and S2).

Cholesterol is a micro-nutrient and insects accumulate it mainly for structural purposes with a tiny portion serving as signaling molecules.[8] To compare cholesterol contents between the mutants and the control animals, we fed the mutants for 2 days and the control for 6 h after hatching to allow them to reach similar size ($t = 0.52$, $P = 0.66$), an indication that similar amount of food was assimilated and converted into body mass.[36] Both larvae were at the 1st instar. The genome-edited line contained significantly less amount of cholesterol than the control animals although the former were allowed to feed on the same diet for a longer period (Fig. 2(C); $t = 8.83$, $P < 0.0001$). Moreover, the relative amount of cholesterol per dry mass in the mutant larvae was only about half of that in the control animals (Fig. 2(D); $t = 5.39$, $P < 0.0001$), indicating that loss-of-function mutation of *NPC1b* severely impeded cholesterol uptake in the insects.

3.3 *NPC1b* expression was dramatically enriched in the midgut and regulated by dietary cholesterol concentrations

To profile *NPC1b* expression, different tissues were collected for quantitative PCR analysis. Results indicated that *NPC1b* expression was dramatically enriched in the midgut, especially the MG2, with a significantly higher level than other tissues tested (Fig. 3(A)). No

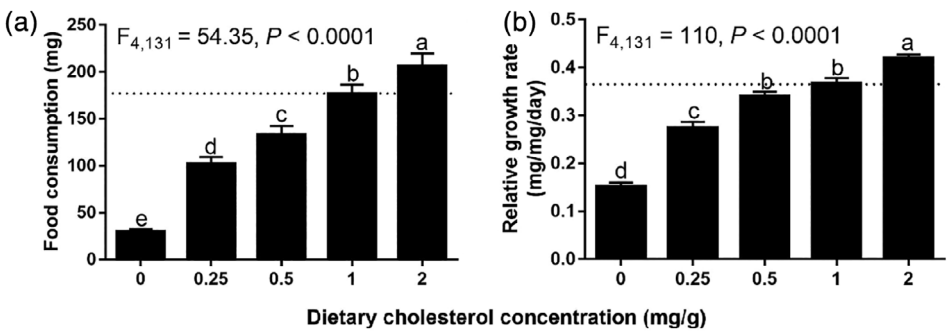

Figure 1. Impact of sterol intake on *H. armigera* larval performance. The 4th instar larvae tended to ingest more food (A) and grow faster (B) with the increasing dietary cholesterol concentrations. Data shown are mean ± standard error, and different letters indicate significant differences.

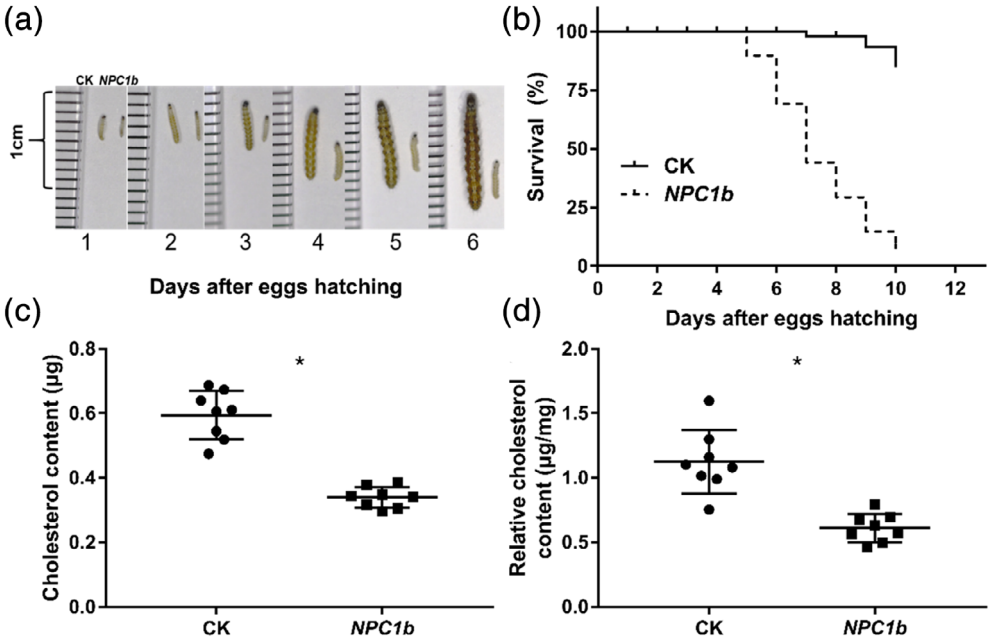

Figure 2. *NPC1b* gene is crucial for the growth and development of *Helicoverpa armigera* larvae and their cholesterol acquisition. Larval size (A), survival rate (B), total cholesterol amount (C), and relative cholesterol amount (D) between the *H. armigera NPC1b* mutants (denoted as *NPC1b*) and the control insects (denoted as CK) were compared. The *NPC1b* mutant line was generated using CRISP/Cas9 genome editing technology. In (A) and (B), hatchlings of the mutant line and the control were fed on diet containing 1 mg/g of cholesterol. Larval survival rate (B; $n = 72$, and shown as Kaplan–Meier curves) was significantly lower in the mutant line than the control determined by the log-rank test ($\chi^2 = 98.1$, $P < 0.001$). In (C) and (D), eight replicates, each containing 15 individuals, were prepared for the mutant line and the control, respectively. Cholesterol amount in *NPC1b* mutants was significantly less than that in the control (mean ± standard deviation; total cholesterol amount: $t = 8.83$, $P < 0.0001$; relative cholesterol amount: $t = 5.39$, $P < 0.0001$). *Indicates the significant difference between two treatments.

significant difference was found between other tissues. Interestingly, *NPC1b* expression tended to increase as the dietary cholesterol concentration decreased and it was highest when the insects fed on the diet containing no cholesterol (Fig. 3(B)).

To further visualize the localization of *NPC1b* transcripts in the gut, longitudinal gut sections were hybridized *in situ* with an antisense DIG-labeled *NPC1b* probe. Among the tissue collected, the midgut region appeared to be bound strongly by the probe, and the probe signal in HG was almost undetectable (Figs 3(C–H), S3(B, C)). Additionally, signal in the anterior midgut was relatively weaker than the other region of the midgut. In consistence with *NPC1b* expression profile, cholesterol content was generally more enriched in the midgut than the FG or the HG (Fig. 3(I)).

3.4 *NPC1b* gene expression varied at different stages of *H. armigera*

NPC1b expression was at the lowest in eggs and pupae, two non-feeding stages (Fig. 4(A)). In contrast, *NPC1b* expression was significantly higher in nearly all fed larvae (except for the 4th instar) than that in eggs and pupae. Moreover, *NPC1b* expression was significantly upregulated in the fed than the non-fed 6-h-old 1st instar larvae. The expression varied among different larval instars, which was significantly upregulated to the highest level in the 2nd instar among all stages tested and downregulated continuously in the 3rd and the 4th instar. After expressing at the lowest level in the 4th instar, *NPC1b* was upregulated again in the 5th instar.

We also tested the effect of dietary sterol concentration on *NPC1b* expression in different stages while only included eggs, the 1st instar unfed larvae, the 1st instar fed larvae, the 5th instar larvae, and pupae (Fig. 4(B)). Similarly, *NPC1b* expression level was the lowest in eggs and pupae. Larvae feeding on the diet containing 0.5 mg/g of cholesterol exhibited the similar expression pattern as those feeding on 1 mg/g of cholesterol (Table S3). In contrast, the expression in the larvae feeding on 2 mg/g of cholesterol was significantly downregulated compared to those feeding on the lower concentrations of cholesterol (i.e. 0.5 and 1 mg/g), in correlation with the earlier results that insects tend to respond to dietary cholesterol concentrations, depending on excessive or insufficient, by downregulating or upregulating the expression of *NPC1b*.

3.5 Cholesterol-binding sites in NPC1b was conserved among caterpillar species but very divergent from bees

To investigate the potential cholesterol-binding sites in NPC1b, we conducted an *in silico* docking analysis. NTD is the domain for cholesterol binding in NPC1 homologs, among which only two, human NPC1_NTD (3gkj) and human NPC1L1_NTD (3qnt), have been determined in crystal structures in their NTD domains.[37,38] Both proteins share approximately 30% identity to *H. armigera* NPC1b_NTD (Table S4), suggesting that they are equivalent in serving as the template for modeling the 3D structure of *H. armigera* NPC1b_NTD.[39] We chose NPC1_NTD as the template because it has a higher resolution (1.6 Å) and its crystallographic structure of protein/cholesterol complex is experimentally verified.

The 3D structure prediction demonstrated that NPC1b_NTD is composed of 13 α-helices and five β-sheets starting from residue 19 and extending to residue 268 (Figs 5(A), S7). Binding pocket residue analysis revealed that ten amino acid residues in the NTD domain are the putative cholesterol-interacting sites (Fig. 5 (B)). Half of these residues, i.e. Gln73, Phe98, Met105, Asn106 and Pro190, are conserved among lepidopteran species (Fig. 6). Notably, Gln73, which is the closest residue to cholesterol

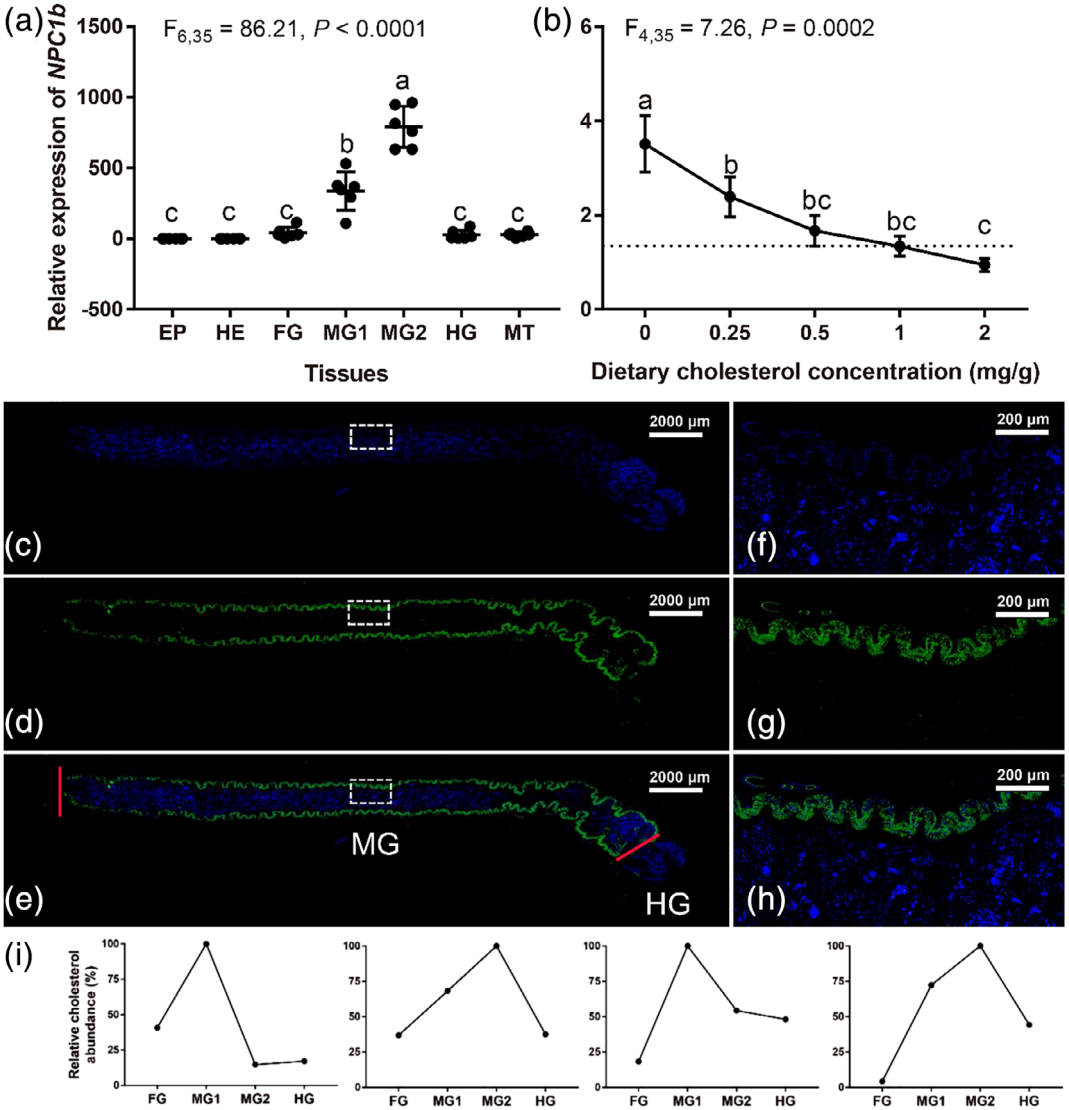

Figure 3. Expression of *Helicoverpa armigera NPC1b* and localization of cholesterol enrichment in gut tissue. Different letters in (A) and (B) indicate significant differences. (A) Relative expression of *NPC1b* (mean ± standard deviation) in different tissues of the 3rd instar larvae. Expression was normalized to epidermis, and six biological replicates were used. EP, epidermis; HE, head; FG, foregut; MG1, the first half of midgut; MG2, the second half of midgut; HG, hindgut; MT, Malpighian tubule. (B) Relative expression of *NPC1b* (mean ± standard error) in the midgut of the 3rd instar larvae fed with different concentrations of cholesterol. Expression and statistical analyses were performed same as (A) except eight biological replicates were used here. (C) DAPI-stained whole mount gut tissue. (D) DIG-dUTP-labeled *NPC1b* RNA probes (green) hybridized to gut tissue. (E) Overlaid image of (C) and (D). Green signal was mainly detected in the midgut. (F), (G), and (H) are magnified display of the white rectangle areas in (C), (D), and (E), respectively. (I) Relative cholesterol abundance in different gut sections of the 5th instar larvae. Each small panel illustrated the relative cholesterol content in the four sections of a gut from one same individual insect, and four insects were used. Cholesterol abundance in each section was calculated relative to the one with highest cholesterol content (set as 100%).

molecule (1.89 Å) and forms one hydrogen bond with the hydroxyl of cholesterol, is conserved between human and *H. armigera* orthologs (Fig. S8). Other residues were predicted to interact with cholesterol within the range of 3.05 Å to 5.34 Å, and are mostly hydrophobic except for Lys76 and Asn106, consistent with the hydrophobicity of cholesterol. This is similar to the hydrophobic characteristics of the binding pocket in human NPC1_NTD and NPC1L1_NTD. Interestingly, the varied amino acids at these sites tend to show similar polarity or hydrophobicity, suggesting these key residues have been subject to purifying selection. For example, Lys76 in *H. armigera* can be replaced by Ser or Arg in other lepidopteran species, and the three amino acids have similar polarity. The replacements of Ile102 by Val, Met164 by Leu or Ile, Leu188 by Ile or Met and Ala189 by Val also occur, and these are hydrophobic ones.

Impact on bees is a major concern of developing new pest control agents.[4] We found that the cholesterol-binding domain, NTD, in bees was very divergent from that in caterpillars, and two groups of proteins only shared 23–30% identity (Table S5). Importantly, none of the predicted binding sites in *H. armigera* is identical to the aligned amino acid residues in bees (Fig. 6).

4 DISCUSSION

Since the discovery of CRISPR/Cas (clustered regularly interspaced short palindromic repeats/CRISPR-associated protein) technology,

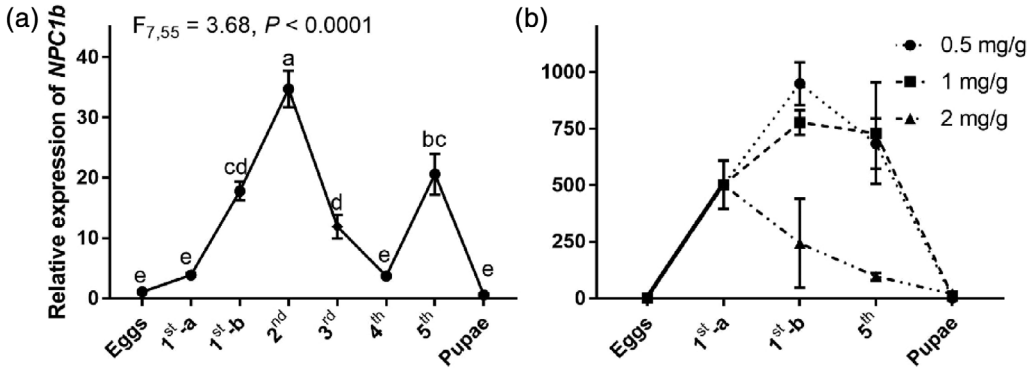

Figure 4. Relative expression of *NPC1b* (mean ± standard error) at different *Helicoverpa armigera* life stages. (A) Relative expression in eight life stages, each including eight biological replicates, of insects that were fed with the diet containing 1 mg/g of cholesterol. (B) Relative expression in five life stages, each including four biological replicates, of insects that were fed with the diet containing 0.5, 1, and 2 mg/g of cholesterol, respectively. On *x*-axis, 1st through 5th indicate different larval instars; 1st-a and 1st-b indicate 1st instar larvae unfed and fed, respectively.

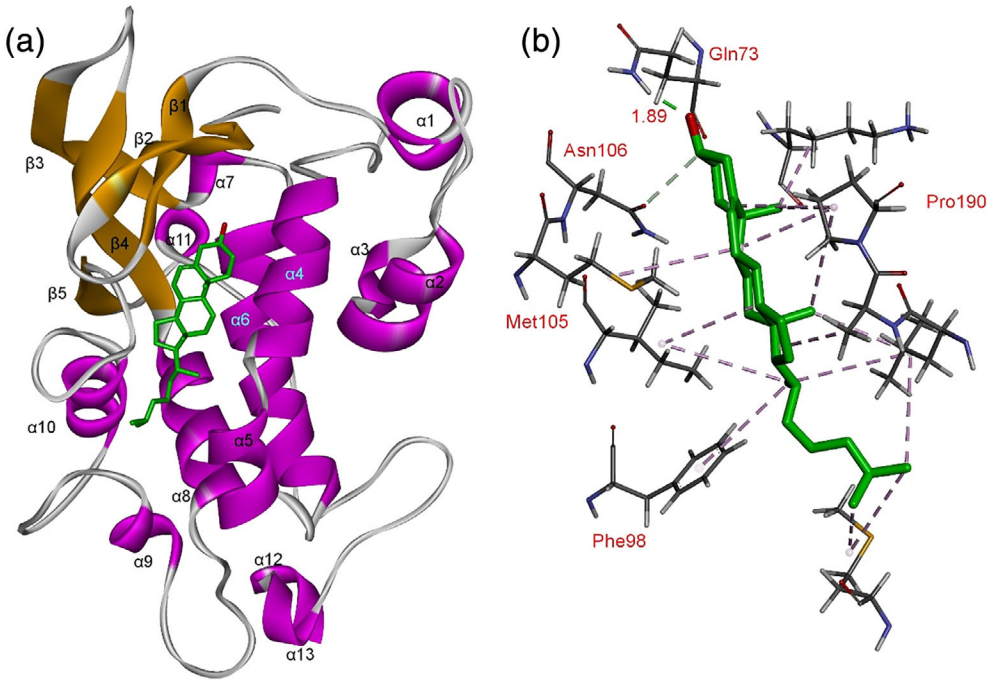

Figure 5. Predicted three-dimensioanl (3D)-structure of the *Helicoverpa armigera* NPC1b_NTD domain and its interaction with cholesterol. Cholesterol molecule (green) and its terminal three-oxygen atom (red) are indicated. (A) Ribbon diagram of NPC1b_NTD binding cholesterol. The secondary structures of NPC1b_NTD, including α-helices (purple), β-sheets (orange), and loops (gray), are indicated. (B) NPC1b_NTD amino acid residues interacting with cholesterol in the binding pocket. The five labeled residues are conserved among caterpillar species.

its application was initially limited to the model species but has now been largely expanded to non-model organisms.[40–43] In Lepidoptera, this technology has been successfully used to edit the genomes of several important pest species, for example, *H. armigera*, *Plutella xylostella* and *Spodoptera exigua*, to identify novel target genes for pest control and understand molecular mechanisms of pesticide resistance.[1,2,44–46] Moreover, it was demonstrated to be a powerful tool in the analysis of editing effects on target genes that would otherwise be lethal in homozygous mutants.[47] In this study, we used it generating a *NPC1b* mutant line of *H. armigera*, to demonstrate the essential role that *NPC1b* plays in dietary sterol uptake in insects.

Our results showed that *NPC1b* expression was dramatically enriched in the midgut rather than in the FG and the HG, and this segment-specific expression along the alimentary tract seems common among animals. In *D. melanogaster*, *NPC1b* expression was also observed to be confined in the midgut epithelium.[25] In mammals, *NPC1L1*, the *NPC1b* ortholog, is highly expressed in the small intestine, a tissue that functions equivalently to insect midgut, and very lowly in the colon,[48] which is regulated by the DNA methylation status of the promoter.[49] In coincidence with the tissue-specific expression of *NPC1b*, we observed that the enrichment of cholesterol mainly occurred in the midgut of *H. armigera* (Fig. 3(I)), and previous radioactive tracer feeding

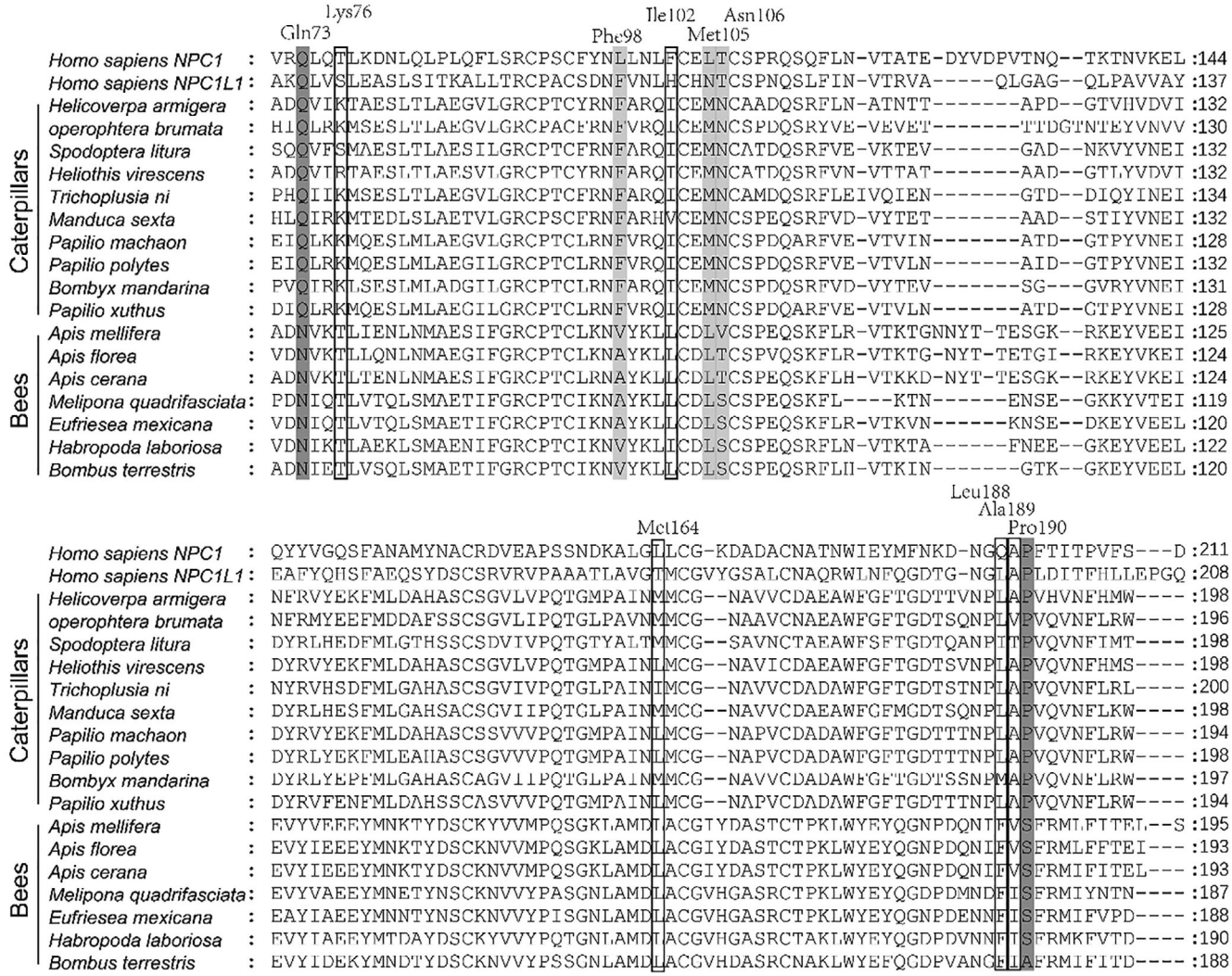

Figure 6. Amino acid sequence alignments of the NPC1b_NTD domain derived from human and insect species. Critical cholesterol-binding amino acid residues were highlighted in dark gray, light gray, or boxed, depending on the degree of conservation shared by the species analyzed. Note that dark gray-highlighted residues share conservation between human and lepidopteran insects, light gray-highlighted residues share conservation among lepidopteran insects, and boxed residues do not share conservation among lepidopteran insects. None of these residues share conservation between caterpillars and bees.

studies on *Manduca sexta* larvae also indicated that dietary cholesterol reached internal organs of the body mainly through the midgut.[50]

Not only the spacious coincidence, but also the temporal linkage of *NPC1b* expression and cholesterol utilization in insects, suggest that *H. armigera NPC1b* functions in cholesterol uptake. As an essential nutrient, cholesterol must be absorbed through insects' feeding, and we observed that *H. armigera NPC1b* was only highly expressed in the feeding larval stages, but much low in the non-feeding stages including eggs and pupae. The disruption of *NPC1b* had no effect on insect embryogenesis (Fig. S4), indicating that *NPC1b* has no critical function during embryonic development. More importantly, we revealed that the loss-of-function mutation of *H. armigera NPC1b* led to the aberrant growth and death of insects in their early larval stage, similar to the cholesterol-deficient symptoms which were previously reported in insects when dietary sterol supply was limited[12,33] or the uptake process was interrupted.[14,16,18] Meanwhile, as observed in Drosophila,[25] we also demonstrated that the cholesterol content in *H. armigera NPC1b* mutant was significantly reduced, strongly supporting that *NPC1b* is indispensable for cholesterol uptake in insects.

Given this correlation between *NPC1b* and cholesterol, whether and how insects regulate the expression of *NPC1b* to facilitate or hinder the acquirement of cholesterol from food becomes an interesting subject. Compared to the unfed, the fed 1st instar caterpillars supplied with cholesterol-containing diet increased their *NPC1b* expression, suggesting that diet or cholesterol availability might serve as an induction signal for *NPC1b* expression. *NPC1b* expression in the wild-type insects feeding on a constant concentration of cholesterol (1 mg/g) varied among larval instars, with the level boosting in the 1st and 2nd instar, then dropping in the 3rd and 4th instar, and boosting again in the 5th instar (Fig. 4(A)), suggesting that the regulation of *NPC1b* expression may not, or not simply, be determined by dietary cholesterol concentration. The high expression of *NPC1b* in the first two instars implied that the young larvae may be in high demand of cholesterol intake. The subsequent high cholesterol absorbing rate can

possibly lead to a peak cholesterol storage in the 2nd instar, which can serve as a signal to reduce cholesterol intake through down-regulating *NPC1b* expression in the 3rd and 4th instar. Consequently, the decreasing expression of *NPC1b* in these 'middle-aged' larvae resulted into low level of cholesterol storage in insects, which, however, served as a boosting signal to upregulate *NPC1b* expression in the 5th instar. We also observed that *NPC1b* expression reduced significantly when insects were given twice the amount of cholesterol (2 mg/g), likely because excessive supply led to sufficient cholesterol storage, which guided the down-regulation of *NPC1b* expression even in the 1st instar (Fig. 4(B)). The varying *NPC1b* expression among different instars and among different dietary cholesterol concentrations indicated that cholesterol storing capability could be low in *H. armigera* larvae, similar as demonstrated in *M. sexta*.[48] Actually, the insects cannot grow much with the significantly reduced food intake if they were transferred onto the sterol-free diet from the diet containing 1 mg/g cholesterol (Fig. 1).

In summary, we provided a proof-of-principle study here to demonstrate that *NPC1b* can be developed as a novel target for the management of *H. armigera* and possibly other lepidopteran pests. This conjecture is based on (i) that *NPC1b* is an essential gene for insect survival and growth; (ii) that *NPC1b* expression and its function being confined in the midgut makes this target more accessible to inhibitors and other potential controlling agents because insect midgut, unlike the FG and the HG, is not protected by the cuticle;[51] (iii) that *NPC1b* is mainly expressed in the actively-feeding larval stage, in which the controlling agents can be easily delivered to the target; and (iv) that *H. armigera NPC1b*-targeting insecticides may have little off-target effect on bees and sterol-autotrophic animals. This study laid a foundation to further develop *NPC1b*-targeting pest management strategies on *H. armigera*.

ACKNOWLEDGEMENTS

The authors thank Dr Jiyuan Liu and Dr Pengjun Shi for helpful discussion on the manuscript. The authors are grateful to Dr Yayun Zuo and Mr Jing Wang for technical assistance in producing CRISPR/Cas9-edited mutants. This work was supported by funds awarded to X.J. by the National Natural Science Foundation of China (31672369), the Introduction of Talent Research Start-up Fund of Shaanxi Province (A279021711) and Northwest A&F University (Z111021503).

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

1 Wang H, Shi Y, Wang L, Liu S, Wu S, Yang Y *et al.*, CYP6AE gene cluster knockout in *Helicoverpa armigera* reveals role in detoxification of phytochemicals and insecticides. *Nat Commun* **9**:4820 (2018).
2 Khan SA, Reichelt M and Heckel DG, Functional analysis of the ABCs of eye color in *Helicoverpa armigera* with CRISPR/Cas9-induced mutations. *Sci Rep* **7**:40025 (2017).
3 Murúa MG, Scalora FS, Navarro FR, Cazado LE, Casmuz A, Villagrán ME *et al.*, First record of *Helicoverpa armigera* (Lepidoptera: Noctuidae) in Argentina. *Fla. Entomol* **97**:854–856 (2014).
4 Siviter H, Brown MJF and Leadbeater E, Sulfoxaflor exposure reduces bumblebee reproductive success. *Nature* **561**:109–112 (2018).
5 Mironidis GK, Despina K, Maria B, Evangelia M, Savopoulou-Soultani M and John V, Resurgence of the cotton bollworm *Helicoverpa armigera* in northern Greece associated with insecticide resistance. *Insect Sci* **20**:505–512 (2013).
6 Joußen N, Agnolet S, Lorenz S, Schöne SE, Ellinger R, Schneider B *et al.*, Resistance of Australian *Helicoverpa armigera* to fenvalerate is due to the chimeric P450 enzyme CYP337B3. *Proc Natl Acad Sci* **109**: 15206–15211 (2012).
7 Behmer ST and Nes WD, Insect sterol nutrition and physiology: a global overview, in *Advances in Insect Physiology*, Vol. **31**. Academic Press, London, pp. 1–72 (2003).
8 Jing X and Behmer ST, Insect sterol nutrition: physiological mechanisms, ecology, and applications. *Annu Rev Entomol* **65**:251–271 (2020).
9 Zhang T, Yuan D, Xie J, Lei Y, Li J, Fang G *et al.*, Evolution of the cholesterol biosynthesis pathway in animals. *Mol Biol Evol* **36**:2548–2556 (2019).
10 Jing X, Grebenok RJ and Behmer ST, Plant sterols and host plant suitability for generalist and specialist caterpillars. *J Insect Physiol* **58**: 235–244 (2012).
11 Jing X, Grebenok RJ and Behmer ST, Sterol/steroid metabolism and absorption in a generalist and specialist caterpillar: effects of dietary sterol/steroid structure, mixture and ratio. *Insect Biochem Mol Biol* **43**:580–587 (2013).
12 Jing X, Grebenok RJ and Behmer ST, Diet micronutrient balance matters: how the ratio of dietary sterols/steroids affects development, growth and reproduction in two lepidopteran insects. *J Insect Physiol* **67**:85–96 (2014).
13 Ma H, Ma Y, Liu X, Dyer DH, Xu P, Liu K *et al.*, NMR structure and function of *Helicoverpa armigera* sterol carrier protein-2, an important insecticidal target from the cotton bollworm. *Sci Rep* **5**:18186 (2015).
14 Du X, Ma H, Zhang X, Liu K, Peng J, Lan Q *et al.*, Characterization of the sterol carrier protein-x/sterol carrier protein-2 gene in the cotton bollworm, *Helicoverpa armigera*. *J Insect Physiol* **58**:1413–1423 (2012).
15 Larson RT, Lorch JM, Pridgeon JW, Becnel JJ, Clark GG and Lan Q, The biological activity of alpha-Mangostin, a larvicidal botanic mosquito sterol carrier protein-2 inhibitor. *J Med Entomol* **47**:249–257 (2010).
16 Kim M-S and Lan Q, Sterol carrier protein-x gene and effects of sterol carrier protein-2 inhibitors on lipid uptake in *Manduca sexta*. *BMC Physiol* **10**:9 (2010).
17 Guo X-R, Zheng S-C, Liu L and Feng Q-L, The sterol carrier protein 2/-3-oxoacyl-CoA thiolase (SCPx) is involved in cholesterol uptake in the midgut of *Spodoptera litura*: gene cloning, expression, localization and functional analyses. *BMC Mol Biol* **10**:102 (2009).
18 Kim M-S, Wessely V and Lan Q, Identification of mosquito sterol carrier protein-2 inhibitors. *J Lipid Res* **46**:650–657 (2005).
19 Behmer ST, Grebenok RJ and Douglas AE, Plant sterols and host plant suitability for a phloem-feeding insect. *Funct Ecol* **25**:484–491 (2011).
20 Krebs KC and Lan Q, Isolation and expression of a sterol carrier protein-2 gene from the yellow fever mosquito, *Aedes aegypti*. *Insect Mol Biol* **12**:51–60 (2003).
21 Lan Q and Massey RJ, Subcellular localization of the mosquito sterol carrier protein-2 and sterol carrier protein-x. *J Lipid Res* **45**: 1468–1474 (2004).
22 Zhang L, Li D, Xu R, Zheng S, He H, Wan J *et al.*, Structural and functional analyses of a sterol carrier protein in *Spodoptera litura*. *PLoS One* **9**:e81542 (2014).
23 Carstea ED, Morris JA, Coleman KG, Loftus SK, Zhang D, Cummings C *et al.*, Niemann-Pick C1 disease gene: homology to mediators of cholesterol homeostasis. *Science* **277**:228–231 (1997).
24 Altmann SW, Davis HR Jr, Zhu LJ, Yao X, Hoos LM, Tetzloff G *et al.*, Niemann–Pick C1 Like 1 protein is critical for intestinal cholesterol absorption. *Science* **303**:1201–1204 (2004).
25 Voght SP, Fluegel ML, Andrews LA and Pallanck LJ, *Drosophila* NPC1b promotes an early step in sterol absorption from the midgut epithelium. *Cell Metab* **5**:195–205 (2007).
26 Zhang J-H, Ge L, Qi W, Zhang L, Miao H-H, Li B-L *et al.*, The N-terminal domain of NPC1L1 protein binds cholesterol and plays essential roles in cholesterol uptake. *J Biol Chem* **286**: 25088–25097 (2011).
27 Zheng J-C, Sun S-L, Yue X-R, Liu T-X and Jing X, Phylogeny and evolution of the cholesterol transporter NPC1 in insects. *J Insect Physiol* **107**:157–166 (2018).

28 Ritter KS and Nes WR, The effects of cholesterol on the development of *Heliothis zea*. *J Insect Physiol* **27**:175–181 (1981).
29 Jing X, *Effects of Sterol Structure on Insect Herbivore Physiology, Biochemistry and Molecular Biology, Entomology*. Texas A&M University, College Station, TX (2011).
30 Chung SH, Jing X, Luo Y and Douglas AE, Targeting symbiosis-related insect genes by RNAi in the pea aphid-*Buchnera* symbiosis. *Insect Biochem Mol Biol* **95**:55–63 (2018).
31 Zhang L and Reed RD, A practical guide to CRISPR/Cas9 genome editing in Lepidoptera, ed. by Sekimura T and Nijhout H. *Diversity and Evolution of Butterfly Wing Patterns*. Springer, Singapore, pp. 155–172 (2017).
32 Akcakaya P, Bobbin ML, Guo JA, Malagon-Lopez J, Clement K, Garcia SP *et al*., In vivo CRISPR editing with no detectable genome-wide off-target mutations. *Nature* **561**:416–419 (2018).
33 Jing X, Vogel H, Grebenok RJ, Zhu-Salzman K and Behmer ST, Dietary sterols/steroids and the generalist caterpillar *Helicoverpa zea*: physiology, biochemistry and midgut gene expression. *Insect Biochem Mol Biol* **42**:835–845 (2012).
34 Zhang S, An S, Li Z, Wu F, Yang Q, Liu Y *et al*., Identification and validation of reference genes for normalization of gene expression analysis using qRT-PCR in *Helicoverpa armigera* (Lepidoptera: Noctuidae). *Gene* **555**:393–402 (2015).
35 Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS *et al*., AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* **30**:2785–2791 (2009).
36 Waldbauer GP, The consumption and utilization of food by insects, in *Advances in Insect Physiology*, ed. by JWL B, Treherne JE and Wigglesworth VB. Academic Press, London, pp. 229–288 (1968).
37 Kwon HJ, Abi-Mosleh L, Wang ML, Deisenhofer J, Goldstein JL, Brown MS *et al*., Structure of N-terminal domain of NPC1 reveals distinct subdomains for binding and transfer of cholesterol. *Cell* **137**: 1213–1224 (2009).
38 Kwon HJ, Palnitkar M and Deisenhofer J, The structure of the NPC1L1 N-terminal domain in a closed conformation. *PLoS One* **6**:e18722 (2011).
39 Yang A-S and Honig B, An integrated approach to the analysis and modeling of protein sequences and structures. I. Protein structural alignment and a quantitative measure for protein structural distance. *J Mol Biol* **301**:665–678 (2000).
40 Cong L, Ran FA, Cox D, Lin S, Barretto RPJ, Habib N *et al*., Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**: 819–823 (2013).
41 Li X, Fan D, Zhang W, Liu G, Zhang L, Zhao L *et al*., Outbred genome sequencing and CRISPR/Cas9 gene editing in butterflies. *Nat Commun* **6**:8212 (2015).
42 Wang Y, Li Z, Xu J, Zeng B, Ling L, You L *et al*., The CRISPR/Cas system mediates efficient genome engineering in *Bombyx mori*. *Cell Res* **23**:1414–1416 (2013).
43 Gaj T, Gersbach CA and Barbas CF, ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* **31**:397–405 (2013).
44 Wang J, Wang H, Liu S, Liu L, Tay WT, Walsh TK *et al*., CRISPR/Cas9 mediated genome editing of *Helicoverpa armigera* with mutations of an ABC transporter gene HaABCA2 confers resistance to *Bacillus thuringiensis* Cry2A toxins. *Insect Biochem Mol Biol* **87**:147–153 (2017).
45 Huang Y, Chen Y, Zeng B, Wang Y, James AA, Gurr GM *et al*., CRISPR/-Cas9 mediated knockout of the abdominal-A homeotic gene in the global pest, diamondback moth (*Plutella xylostella*). *Insect Biochem Mol Biol* **75**:98–106 (2016).
46 Zuo Y, Wang H, Xu Y, Huang J, Wu S, Wu Y *et al*., CRISPR/Cas9 mediated G4946E substitution in the ryanodine receptor of *Spodoptera exigua* confers high levels of resistance to diamide insecticides. *Insect Biochem Mol Biol* **89**:79–85 (2017).
47 Zhang L and Reed RD, Genome editing in butterflies reveals that *spalt* promotes and *Distal-less* represses eyespot colour patterns. *Nat Commun* **7**:11769 (2016).
48 Davies JP, Scott C, Oishi K, Liapis A and Ioannou YA, Inactivation of NPC1L1 causes multiple lipid transport defects and protects against diet-induced hypercholesterolemia. *J Biol Chem* **280**:12710–12720 (2005).
49 Malhotra P, Soni V, Kumar A, Anbazhagan AN, Dudeja A, Saksena S *et al*., Epigenetic modulation of intestinal cholesterol transporter Niemann–Pick C1-like 1 (NPC1L1) gene expression by DNA methylation. *J Biol Chem* **289**:23132–23140 (2014).
50 Jouni ZE, Zamora J and Wells MA, Absorption and tissue distribution of cholesterol in *Manduca sexta*. *Arch Insect Biochem Physiol* **49**: 167–175 (2002).
51 Napoleão TH, Albuquerque LP, Santos ND, Nova IC, Lima TA, Paiva PM *et al*., Insect midgut structures and molecules as targets of plant-derived protease inhibitors and lectins. *Pest Manag Sci* **75**: 1212–1222 (2019).

