



Article

The Role of Chitooligosaccharidolytic β -N-acetylglucosaminidase in the Molting and Wing Development of the Silkworm *Bombyx mori*

Bili Zhang ¹, Chunlin Li ¹, Yue Luan ¹, Yaru Lu ¹, Hai Hu ¹, Yanyu Liu ¹, Kunpeng Lu ¹, Guizheng Zhang ², Fangyin Dai ^{1,*} and Xiaoling Tong ^{1,*}

¹ State Key Laboratory of Silkworm Genome Biology, Key Laboratory of Sericultural Biology and Genetic Breeding, Ministry of Agriculture and Rural Affairs, College of Biotechnology, Southwest University, Chongqing 400715, China; a18286002256@163.com (B.Z.); lclin13@163.com (C.L.); luanyue987@163.com (Y.L.); lulu510@email.swu.edu.cn (Y.L.); huhaiswu@163.com (H.H.); liuyanyu@email.swu.edu.cn (Y.L.); lukunpeng@swu.edu.cn (K.L.)

² Guangxi Academy of Sericultural Sciences, Nanning 530007, China; zhangdoudou1999@163.com

* Correspondence: fydai@swu.edu.cn (F.D.); xltong@swu.edu.cn (X.T.)

Supplementary Materials

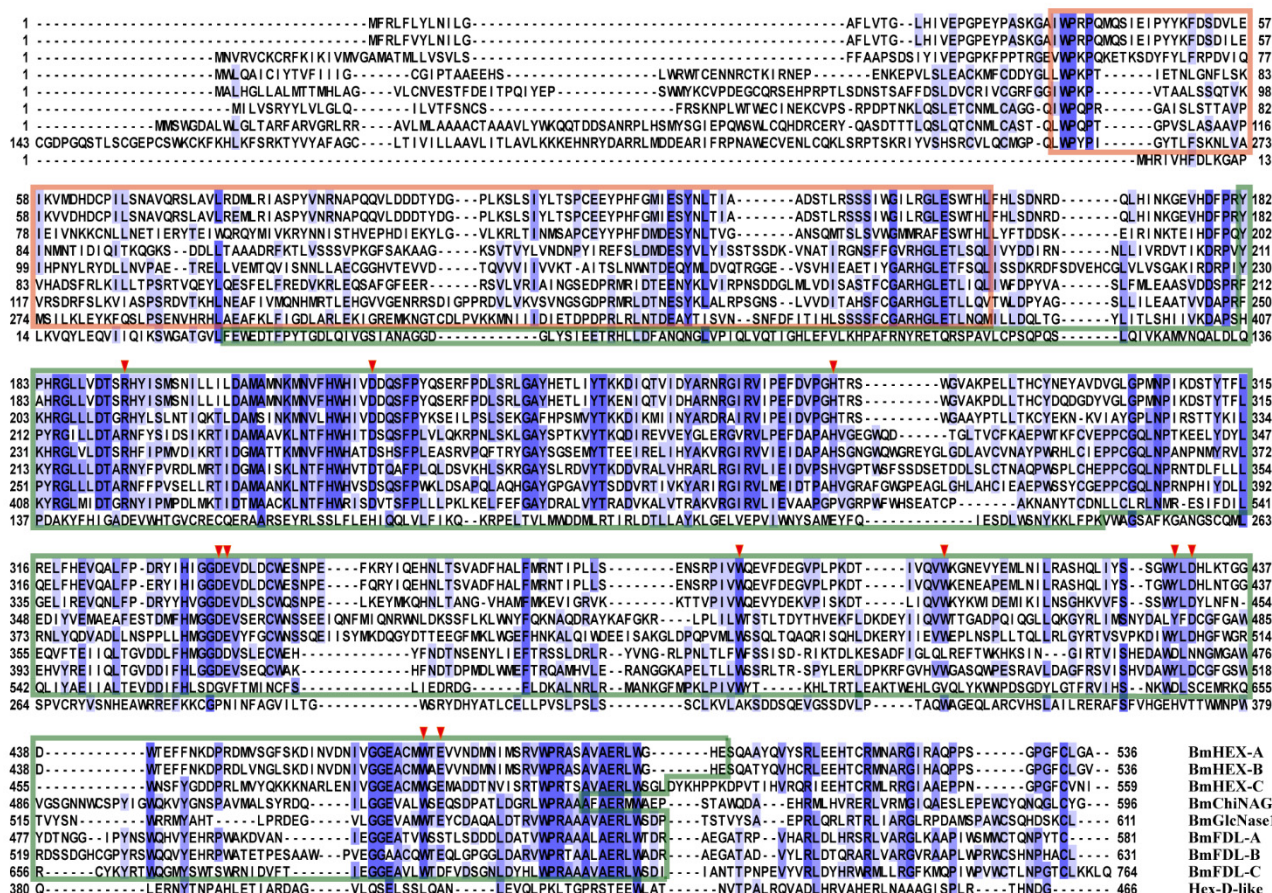


Figure S1. Multiple alignments of amino acid sequences of nine *HEX* genes in *B. mori*. The GH20 catalytic domain and the additional GH20b domain are marked with light pink and light green boxes respectively.

1 ATGTGGCTTCAAGCAATTGTATTACACCGTCTTCATAATTATAGGATGCGGATTCCGACGCGAGCCGAAGAGCACTATTGTGGAGA
 1 **M** W L Q A I C I Y T V F I I I G C G I P T A A E E H S L W R
 Signal peptide
 91 TGGACGTGCGAAAAAATAGATGCACGAAGATCAGGAACGAGCCGAGAATAAAGAGCCTGTCTCAGTCTGGAAGCATGCAAGATGTTCT
 31 W T C E N N R C T K I R N E P E N K E P V L S L E A C K M F
 181 TGTGACGATTACGGTCTCTCTGGCCGAAACACGATCGAGACAAATCTGGGTAACCTCTCTCCAAGATTATATGAACACCATCGAC
 61 C D D Y G L L W P K P T I E T N L G N F L S K I N M N T I D
 271 ATTCAGATCACCAAGCAGGAAAGAGCGACACCTCTCACAGCGGCTGCCGACAGATTAAGACACTGGTTCCAGTCCGTGCCAAAA
 91 **I** Q I T K Q G K S D D L L T A A A D R F K T L V S S S V P K
 Additional GH20b catalytic domain
 361 GGCTTCTCGGCGAAGGCTGCAGGAAATCAGTCACAGTTTATTTAGTCAATGACAATCCTTATATCCGAGAATTCTCCCTGGACATGGAT
 121 G F S A K A A G K S V T V Y L V N D N P Y I R E F S L D M D
 451 GAGAGCTACGTCTCTACATTCATCAACGTCGAGCGACAAAGTGAACGCTACCATACGTGGAACTCGTTCTTCGGGGTCGTCACGGT
 151 E S Y V L Y I S S T S S D K V N A T I R G N S F F G V R H G
 541 CTGGAGACGCTCTCTCAGCTAATAGTGTACGACGACATCAGGAATAATCTGTTGATTGCTCGTGACGTCACAATTAAGGACAGCGCTGC
 181 L E T L S Q L I V Y D D I R N N L L I V R D V T I K D R P V
 631 TATCCTTACCGAGGTATTTGTCTGGACACCGCGAGGAACCTTCTACTCCATAGATTCAATCAAAAGAACTATAGACCAATGGCCGACGTG
 211 Y P Y R G I L L D T A R N F Y S I D S I K R T I D A M A A V
 721 AAATTGAACACATTCCTGACCATACGACAGCAGAGTTCCCGTTGGTCTTACAGAAAAGACCGAACTGAGTAAGCTCGGGCA
 241 K L N T F H W H I T D S Q S F P L V L Q K R P N L S K L G A
 811 TACAGTCCCAACGAGGTATACAAAGCAAGATATTCGTGAAGTTGTGCAATATGGCTAGAGAGGGAGTTCGTGTTTACCGGAATT
 271 Y S P T K V Y T K Q D I R E V V E Y G L E R G V R V L P E F
 901 GATGCGCCGCTCATGTGGAGAGGGATGCGAGACACCGACTACCGTGTGCTTTAAGCGGAGCGGTGGACGAAATCTGCTGGAG
 301 D A P A H V G E G W Q D T G L T V C F K A E P W T K F C V E
 991 CCGCCTTGTGGTCAACTGAACCCGACTAAAGAGGAACCTTACGACTACTTGAAGATATTTACGTGAATGGCTGAGGCGTTTGAGAGC
 331 P P C G Q L N P T K E E L Y D Y L E D I Y V E M A E A F E S
 GH20 catalytic domain
 1081 ACCGACATGTTCCACATGGGAGGAGACGAGGTACGCAACGCTGTTGGAACCTCTCAGAGGAGATCCAGAACTTTATGATTGAGAACCGA
 361 T D M F H M G G D E V S E R C W N S S E E I Q N F M I Q N R
 1171 TGGAAATTTGACAAGAGCAGTTTCTGAAGCTTTGGAACCTACTCCAGAAGAAGCTCAAGACAGAGCGTATAAGGCCCTTCGGTAAACGA
 391 W N L D K S S F L K L W N Y F Q K N A Q D R A Y K A F G K R
 1261 CTGCTCTGATTCTATGACACGACATGACCGACTACACTACGTAAGAAATCTTGGACAAAGACGAATACATCATAAGGTCTGG
 421 L P L I L W T S T L T D Y T H V E K F L D K D E Y I I Q V W
 1351 ACCACTGGAGCCGACCCACAAATCAAGGTTTACTCCAGAAAGGATATCGCTGATCATGTCAAATATGACGCTCTACTTTGACTGT
 451 T T G A D P Q I Q G L L Q K G Y R L I M S N Y D A L Y F D C
 1441 GGATTCGGGCATGGTGGGTCTGTAATAATTGGTGTACCGTACATCGGTGGCAGAAAGTGTACGGTAACAGTCCAGCGGTGATG
 481 G F G A W V G S G N N W C S P Y I G W Q K V Y G N S P A V M
 1531 GCGCTCTGTACCGAGATCAGATCTTAGTGGTGAAGTAGCGTGTGGTCCGAGCAGTCGACCTGCCACGCTGGACGGCCGACTGTGG
 511 A L S Y R D Q I L G G E V A L W S E Q S D P A T L D G R L W
 1621 CCCAGAGCGGCTTCGCCGAGCGCATGTGGGCGGAACCTTCACCGCGTGGCAGGACCGGACCGGATGCTCCATGTCAGAGAA
 541 P R A A A F A E R M W A E P S T A W Q D A E H R M L H V R E
 1711 OGTTTGGTAAGAATGGGAATCAAGCTGAATCGCTTGAGCCGGAGTGGTGTATCAGAAATCAAGGACTTTGCTATGGTTAG
 571 R L V R M G I Q A E S L E P E W C Y Q N Q G L C Y G *

Figure S2. The nucleotide and deduced amino acid sequence encoded by *BmChiNAG*. The numbers along the left margin refer to the position of nucleotide and the deduced amino acids. The start codon is indicated in bold and the stop codon in bold with an asterisk. The predicted signal peptide, GH20 catalytic domain, and an additional GH20b domain are underlined. The conserved HMGG-DEV××CW motif is shaded and the catalytically active sites are boxed.

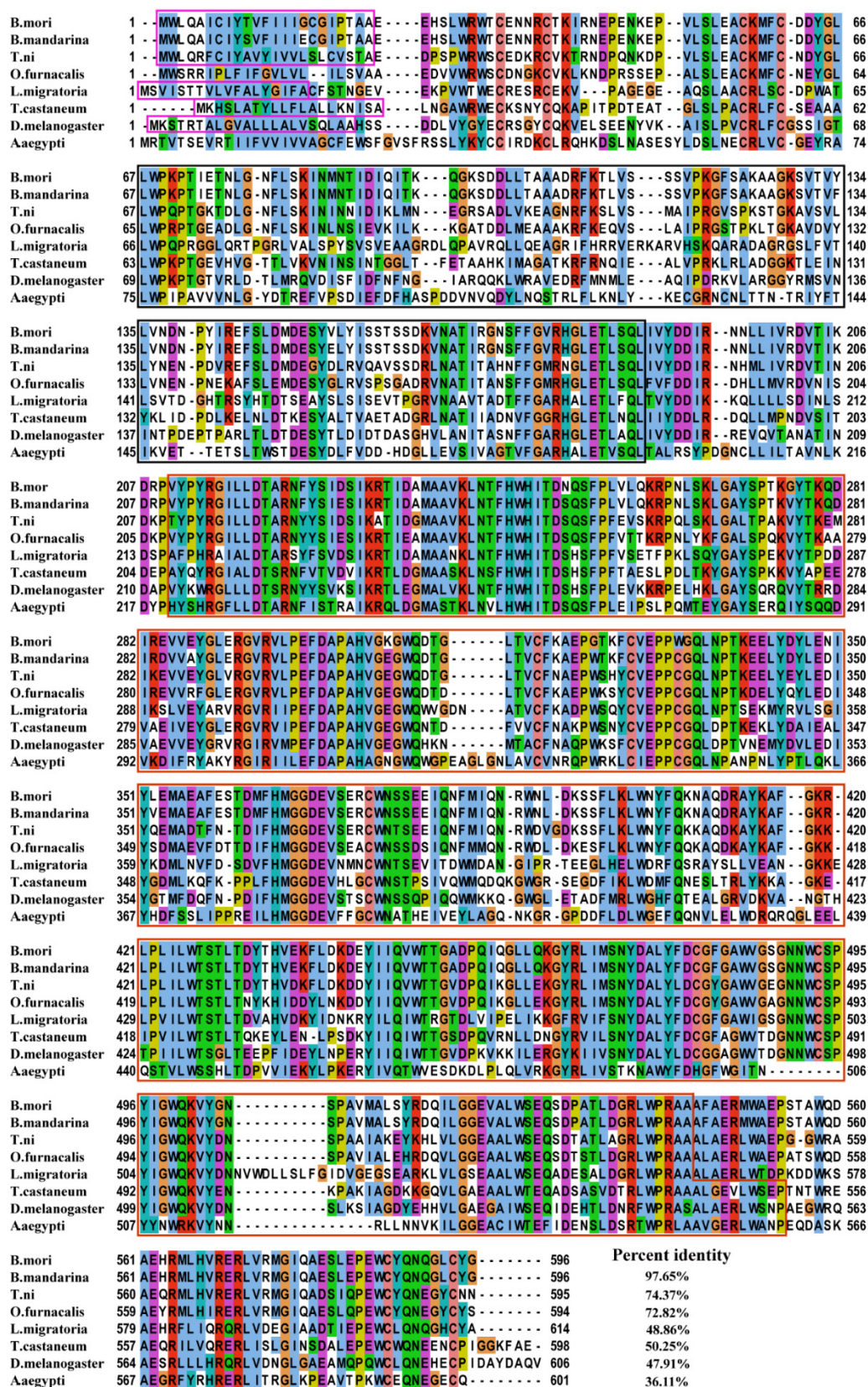


Figure S3. Multiple alignments of amino acid sequences of BmChiNAG with the homologous sequences of other species. The predicted signal peptide is marked with a purple box. The GH20 catalytic domain and the additional GH20b domain are marked with black and red boxes respectively. The similarity between BmChiNAG and other NAGs is listed at the end.

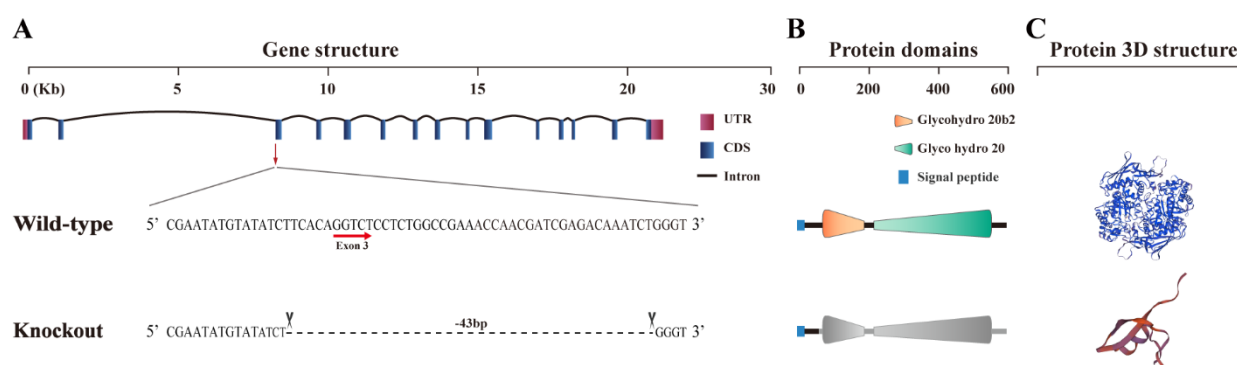


Figure S4. The gene structure of *BmChiNAG* and protein domains structure in wild-type and knockout line. (A) The gene structure of *BmChiNAG*. (B) The protein domains structure of *BmChiNAG*. (C) The protein 3D structure of *BmChiNAG*.

Table S1. The GenBank accession numbers used in the phylogenetic analysis of GH20 genes in *Bombyx mori*.

Gene name	GenBank accession numbers
<i>BmHeX-A</i>	NP_001093291.1
<i>BmHeX-B</i>	NP_001037096.1
<i>BmGlcNase1</i>	XP_012550370.1
<i>BmChiNAG</i>	AAC60521.1
<i>BmHexD-like</i>	XP_004926682.1
<i>BmFDL-A</i>	XP_004922444.1
<i>BmFDL-B</i>	NP_001165928.1
<i>BmHeX-C</i>	XP_004924240.1
<i>BmFDL-C</i>	XP_021202023.1

Table S2. The GenBank accession numbers and the protein sequences of hexosaminidase used in the phylogenetic tree.

Gene name	Species	GenBank accession numbers
<i>MsCHINAG</i>	<i>Manduca sexta</i>	XP_030029670.1
<i>DpCHINAG</i>	<i>Danaus plexippus</i>	XP_032514907.1
<i>SiCHINAG</i>	<i>Spodoptera litura</i>	XP_022814621.1
<i>OfCHINAG</i>	<i>Ostrinia furnacalis</i>	Q06GJ0.1
<i>HaCHINAG</i>	<i>Helicoverpa armigera</i>	XP_021201252.1
<i>DmHEX1</i>	<i>Drosophila melanogaster</i>	NP_728974.2
<i>AaCHINAG</i>	<i>Aedes aegypti</i>	XP_001649548.1
<i>HiCHINAG</i>	<i>Hermetia illucens</i>	XP_037908553.1
<i>DaCHINAG</i>	<i>Diachasma alloeum</i>	XP_015118507.1
<i>AcCHINAG</i>	<i>Apis cerana</i>	XP_016905893.1
<i>TcNAG</i>	<i>Tribolium castaneum</i>	NP_001092318.1
<i>LsNAG</i>	<i>Lasioderma serricorne</i>	AYK27452.1
<i>NiCHINAG</i>	<i>Nilaparvata lugens</i>	XP_039295029.1
<i>ApCHINAG2</i>	<i>Acyrtosiphon pisum</i>	XP_001947177.2
<i>AaNAG2</i>	<i>Aedes aegypti</i>	EAT40440
<i>AmNAG2</i>	<i>Anopheles merus</i>	XP_041788645.1
<i>DmNAG2</i>	<i>Drosophila melanogaster</i>	NM_080342
<i>LsNAG2</i>	<i>Lasioderma serricorne</i>	MN310607
<i>PhNAG2</i>	<i>Pediculus humanus corporis</i>	XP_002425212
<i>TcNAG2</i>	<i>Tribolium castaneum</i>	EF592537.1
<i>AgFDL</i>	<i>Anopheles gambiae</i>	XP_308677
<i>MSFDL</i>	<i>Manduca sexta</i>	XP_037295010.1
<i>DmFDL</i>	<i>Drosophila melanogaster</i>	NP_725178
<i>OfFDL</i>	<i>Ostrinia furnacalis</i>	ADF97235.1
<i>TcFDL</i>	<i>Tribolium castaneum</i>	NM_001098826
<i>DpFDL</i>	<i>Danaus plexippus</i>	XP_032519501.1
<i>AaHEX</i>	<i>Aedes aegypti</i>	EAT43655

<i>AgHEX</i>	<i>Anopheles gambiae</i>	XP_319210
<i>AmHEX</i>	<i>Apis mellifera caucasica</i>	KAG6801275.1
<i>NvHEX</i>	<i>Nasonia vitripennis</i>	XP_031781641.1
<i>PhHex</i>	<i>Pediculus humanus corporis</i>	XP_002428512.1
<i>SfHEX</i>	<i>Spodoptera frugiperda</i>	ABA27427
<i>TcHEX</i>	<i>Tribolium castaneum</i>	XM_970563
<i>HsHEXA</i>	<i>Homo sapiens</i>	NM_000520
<i>HsHEXB</i>	<i>Homo sapiens</i>	NM_000521.4
<i>RnHEXB</i>	<i>Rattus norvegicus</i>	NP_001011946.1
<i>RnHEXA</i>	<i>Rattus norvegicus</i>	NP_001004443.1
<i>BmHEX-A</i>	<i>Bombyx mori</i>	NP_001093291.1
<i>BmHEX-B</i>	<i>Bombyx mori</i>	NP_001037096.1
<i>BmGlcNaseI</i>	<i>Bombyx mori</i>	XP_012550370.1
<i>BmChiNAG</i>	<i>Bombyx mori</i>	AAC60521.1
<i>BmHEX-D-like</i>	<i>Bombyx mori</i>	XP_004926682.1
<i>BmFDL-A</i>	<i>Bombyx mori</i>	XP_004922444.1
<i>BmFDL-B</i>	<i>Bombyx mori</i>	NP_001165928.1
<i>BmHEX-C</i>	<i>Bombyx mori</i>	XP_004924240.1
<i>BmFDL-C</i>	<i>Bombyx mori</i>	XP_021202023.1
<i>TnHEX-D-like</i>	<i>Trichoplusia ni</i>	XP_026729051.1
<i>OfHEX-D-like</i>	<i>Ostrinia furnacalis</i>	XP_028173019.1
<i>PaHEX-D-like</i>	<i>Pararge aegeria</i>	XP_039754573.1
<i>DpHEX-D-like</i>	<i>Danaus plexippus plexippus</i>	OWR42230.1
<i>BaHEX-D-like</i>	<i>Bicyclus anynana</i>	XP_023935131.1
<i>MhHEX-D-like</i>	<i>Maniola hyperantus</i>	XP_034830853.1
<i>VtHEX-D-like</i>	<i>Vanessa tameamea</i>	XP_026498959.1

Table S3. The species and GenBank accession numbers used in the multiple alignments.

Species	GenBank accession numbers
<i>Bombyx mandarina</i>	XP_028037958.1
<i>Trichoplusia ni</i>	XP_026732584
<i>Ostrinia furnacalis</i>	Q06GJ0.1
<i>Locusta migratoria</i>	AFZ76982.1
<i>Tribolium castaneum</i>	NP_001092318.1
<i>Drosophila melanogaster</i>	NP_728974.2
<i>Aedes aegypti</i>	XP_021693811.1
<i>Bombyx mori</i>	AAC60521.1

Table S4. Primers were used to clone the open reading frame (ORF) of *BmChiNAG*.

Name of primer	Forward primer	Reverse primer
BmChiNAG ORF	ATGTGGCTTCAAGCAATTGTA	CTAACCATAGCAAAGTCCTTGA

Table S5. Primers used for qRT-PCR.

Name of primer	Forward primer	Reverse primer
BmChiNAG	GCAGGAAAATCAGTCACAGTTT	GTCGTCGTACACTATTAGCTG A
PPO1	ACGTGCTGGGCAACCTTA	ACGGACGACACCCCTGATG
Ddc	CTTGGACTGCGGTGATGG	TAGCCGTGCCCTGGATTA
Th	CCCAGCAAGATGAAACGA	AAGGGCGAGTCTCAAGGT
At	CTCGGCGGCTTTACATTA	TGAGCGTCTGCTGGTTCT
eif4A	TTCGTACTGGCTCTTCTCGT	CAAAGTTGATAGCAATTCCCT