Journal of Insect Science, (2018) 18(1): 6; 1–11
doi: 10.1093/jisesa/iex107
Research

Protein Profile Analysis of *Ericerus pela* (Hemiptera: Coccoidea) Egg

Yanhong Hu,[1,2,3] Xiaoming Chen,[1,3,4] and Pu Yang[3]

[1]College of Forest Resources and Environment, Nanjing Forestry University, Nanjing 210037, China, [2]College of Forest, Henan Science and Technology University, Luoyang, Henan Province, 371003, China, [3]The Research Institute of Resource Insect, Chinese Academy of Forestry, Kunming, Yunnan Province, 650224, China, and [4]Corresponding author, e-mail: cafcxm@139.com

Subject Editor: John Ewer

Received 17 August 2017; Editorial decision 27 November 2017

Abstract

The transformation from embryo to first instar nymph is an essential process in the insect life cycle. In order to characterize protein expression in the *Ericerus pela* Chavannes (Hemiptera: Coccoidea) egg, high-throughput proteomics and bioinformatics methods were used. A total of 678 peptides were identified and assigned to 358 protein groups. The proteins exhibited a wide range of molecular weight (3.50–495.12 kDa) and isoelectric points (3.50–13.1). Gene Ontology annotation showed that the majority of proteins were associated with cellular processes, metabolic processes, and response to stimulus processes. The predominant molecular functions of *E. pela* egg proteins included binding, catalytic activity, transporter activity, and structural molecule activity. Kyoto Encyclopedia of Genes and Genomes annotations identified 137 pathways, and most proteins were assigned to metabolism events, including many enzymes participating in energy metabolism, protein folding, sorting, and degradation. The processes and functions of the identified proteins were closely related to the physiological status of egg and embryo development. We conclude that some identified proteins are related to important egg biological characteristics, and regard them as the target proteins for future study.

Key words: *Ericerus pela* Chavannes, egg, protein profile

Ericerus pela Chavannes (Hemiptera: Coccoidea) is one of the oldest economic insect. It has been used successfully in commercial wax production, and reared by humans for more than a thousand years. White wax is secreted only by the male *E. pela*, and has an increasingly wide utilization; examples include candle production, printing, medicine, food, cosmetic industries, and precision machinery (Chen and Feng 2009).

The oviposition of *E. pela* occurs about in March and April each year, and incubation happens in April and May. Eggs are laid in the capsule of the female adult every day during the oviposition period; the maximum and minimum amount of eggs laid per adult are 18,047 and 12,000, respectively (Wu and Zhong 1983). When the oviposition period is over, the ventral side of the female adult body is very close to its dorsal side, and the adult will soon die. The eggs, with wax powder on their surface, will incubate in the closed capsule for at least 29 d (Wu and Zhong 1983). After the newly hatched nymphs crawl out of the capsule, the male nymphs live in the shadow of the host plant in a gregarious manner, and secret an amount of white wax to cover themselves until eclosion, while the female nymphs scatter on the host plant and do not secrete the white wax (Chen 2011).

Research on *E. pela* eggs at molecule level has not been reported up to date. There are only proteomic analyses of the male adult cuticle and the male pupal stage (Yang et al. 2011, Yang and Chen 2014), and transcriptome analyses of adults and pupae (Yang et al. 2012, 2015; Yu et al. 2016). The results of these aforementioned studies, coupled with in-depth study of biology and ecology of *E. pela* will lay the foundation for further proteomic study of *E. pela* eggs and understanding the roles of proteins at this stage. Proteomic analyses of other insect eggs or embryos have been well documented (Amenya et al. 2010; Li et al. 2010, 2011; Müller et al. 2010; Gala et al. 2013), and can provide beneficial references for the present study.

The present study was performed to explore the protein expression profile of *E. pela* eggs, generate hypotheses about the connections between certain proteins and biological characteristics of *E. pela* eggs, and provide basic information for studying target proteins in future.

Materials and Methods

Insect Culture

E. pela were cultured on the branches of *Ligustrum lucidum*, in the experimental field of Research Institute of Resources Insects of the Chinese Academy of Forestry in Kunming (longitude: 102°42′E;

latitude: 25°02′N). When the first nymphs crawled out of their capsules in March, the eggs in these capsules were collected in tubes and stored at -80℃.

Protein Extraction

Before protein extraction, the eggs collected from different mother scale insects were mixed and randomly divided into three groups, and washed in phosphate buffer solution (pH 7.2) three times. Total protein of each group was extracted on ice in 1 ml lysis buffer (containing 10% glycerin, 2.5% SDS, 5% β-mercaptoethanol, and 62.5 mM Tris–HCl, pH 6.8) per mg of sample homogenate. The homogenate extraction was kept for 10 min at room temperature and then subjected to five rounds of sonication treatment in an ice-bath, each time for 20 s with a 20 s interval. After centrifugation at 20,000 *g*, the supernatant was aliquoted and stored at –20℃. Protein concentration was determined with a Bradford protein assay kit (Beyotime, Shanghai, China), BSA (Sigma, USA) was taken as standard.

SDS–PAGE Separation and In-gel Digestion

Before being subjected to SDS–PAGE, the extracted total protein was dissolved in SDS–PAGE loading buffer, boiled for 5 min, centrifuged at 20,000 *g* for 10 min, loaded on a 5% stacking gel and 12% separating gel, and was run at 15 mA for 30 min and then 30 mA for 2 hr in a mini-vertical electrophoresis system (Bio-Rad, USA). After electrophoresis, the gel was stained overnight in a solution of 0.1% (w/v) Coomassie Brilliant Blue G-250 (Sangon Biotech, Shanghai, China), 30% (v/v) methanol, and 10% (v/v) glacial acetic acid. After decolorization, the gel was analyzed for bands along with a molecular weight marker.

Thereafter, the gel containing all bands was cut into 1 mm^3 particles for in-gel digestion: gel particles were washed three times in deionized water and subsequently dehydrated with 100% acetonitrile (ACN) for 10 min. The particles were incubated with 100 mM DTT for 30 min at 56℃. The resulting free thiol (–SH) groups were subsequently alkylated by incubating the samples with 200 mM iodoacetamide for 20 min in the dark. Gels were washed with 25 mM ammonium bicarbonate and dehydrated with 100% ACN sequentially. Thereafter, 10 ng/μl trypsin (Promega, USA) was added and incubated for 20 hr at 37℃ for protein digestion. Supernatants were transferred to fresh tubes for mass spectrometric analysis.

Protein Identification Using LC–MS/MS

The resuspended extracts were separated and identified using HPLC (Easy nLC system, ThermoQuest, San Jose, CA) coupled with Q-Exactive mass spectrometer (thermo Fisher, San Jose, CA). One microliter of sample was loaded on a trapping column (Thermo scientific EASY column (2 cm × 100 μm 5 μm-C_{18})) each time. After flow-splitting, peptides were transferred to the analytical column (Thermo scientific EASY column (75 μm × 100 mm 3 μm-C_{18})) for separation equilibrated with buffer A (0.1% methanoic acid in water) and buffer B (84% ACN, 0.1% methanoic acid in water), a 280 min linear gradient was set: buffer B started from 0 to 60% at a flow rate of 250 nl/min, came to 100% subsequently, and then maintained constantly at this flow rate.

The mass spectra of the peptides were recorded on a Q-Exactive mass spectrometer. The positive ions were adopted as the mode of Scanning MS spectra, the MS analysis was performed with one full MS scan (*m/z* 300–1800) with the resolution (*R* = 70,000) at *m/z* 200 and dynamic exclusion (40.0 s), followed by MS/MS scans on the 10 most intense ions from the MS spectrum. Collision-induced dissociation was conducted with normalized collision energy of 35% and voltage of 27 eV.

The raw data from LC–MS/MS was analyzed using maxquant 1.3, the parameters were set as: the maximum number of missed cleavages a peptide for 2, the trypsin digestion, carbamidomethyl (C) for fixed modification, oxidation (M), and acetyl (N-term) for variable modifications, proteins false discovery rate ≤0.01, and as well as peptides, specifying the string for reverse and contaminant hits. In this study, an in-house database was used for proteomic data analysis, constructed by combining the coding sequences from *E. pela* Illumina transcriptome sequencing databases (Yang et al. 2012) with insect sequence data downloaded from the National Center for Biotechnology Information (NCBI) Nr (nonredundant) database and the SwissProt protein database.

Bioinformatic Analysis

Functional analysis of the identified proteins was performed using UniProt Knowledgebase (Swiss-Prot + TrEMBL) (http://www.uniprot.org), and proteins were grouped on the basis of their biological process and molecular function of Gene Ontology terms (Gala et al. 2013). The GO annotation terms were obtained from Web Gene Ontology Annotation Plotting (http://wego.genomics.org.cn/). All the identified proteins were searched against the Kyoto Encyclopedia of Genes and Genomes (KEGG) (http:// www.genome.jp/kegg) to identify the correlated pathways. The protein enzyme commission numbers were obtained based on the best matches (*E*-value ≤e-15).

Results

Molecular Weight and Isoelectric Point of the Identified Proteins

In this study, about 678 unique peptides were identified, ascribed to 358 protein groups. These identified proteins exhibited a broad range of theoretical molecular weight (MW); one major group comprised proteins with MW between 10 and 30 kDa, and remarkably, there were eight proteins with MW exceeding 300 kDa. With regard to theoretical isoelectric point (*pI*), about 70% of the total proteins had *pI* in the range of 4–8, the *pI* of three proteins was <4, and the *pI* of 15 proteins was >11 (Fig. 1 and Table 1).

Categorization of the Identified Proteins

The identified proteins were categorized based on the biological process and molecular function as predicted from associated GO terms (Table 1). Most of the proteins were predicted to act on material conversion and transportation, and information modulation, and the proteins associated with metabolism of carbohydrates and energy, amino acid metabolism, nucleotide metabolism, transcription and translation, and other categories are shown in Table 1.

We found that a large number of proteins associated transformation of matter and energy were expressed during egg development. Many identified enzymes were related to the glycolytic pathway, pentose phosphate pathway and tricarboxylic acid cycle (Supp Table 1 [online only]). Pyruvate kinase (NP_001036906.1), malate dehydrogenase (XP_001659012.1), phosphoglyceromutase (NP_001037540.1), aldehyde dehydrogenase 2 family (mitochondrial) (NP_001087022.1), ATP synthase (NP_001040233.1), GI20614 (XP_002005703.1), and other enzymes were identified, and assigned to different carbohydrate metabolism pathways and energy metabolism pathways, which suggested the carbohydrate and energy metabolism was vigorously performing in this stage.

RNA-binding protein 8A (XP_001849141.1), ribosomal protein L28 (NP_001155658.1), and other 20 proteins took part

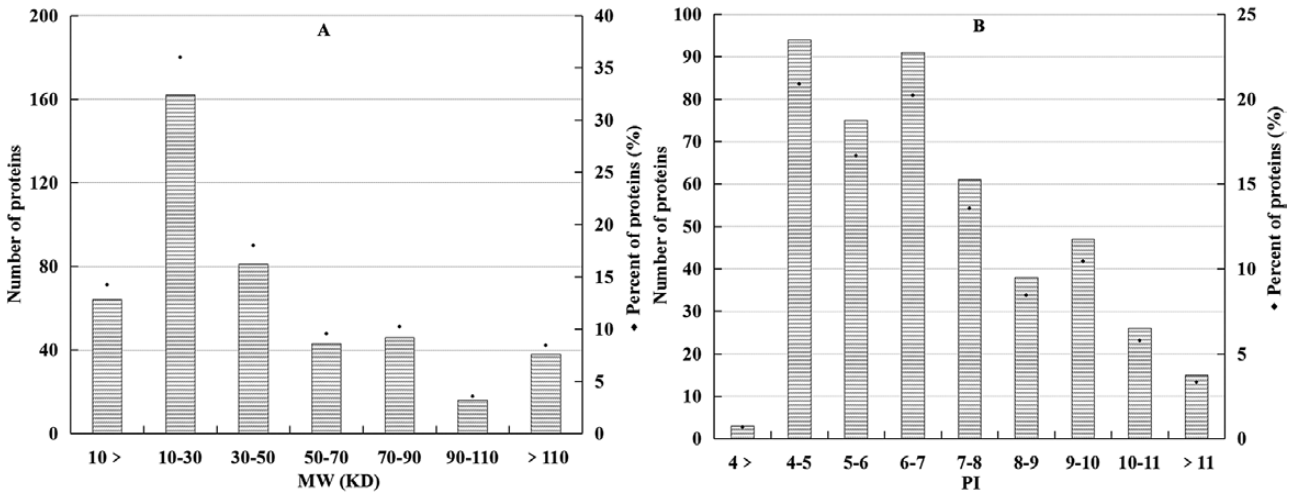

Fig. 1. Theoretical molecular weight (MW) and isoelectric point (pI) distribution of the identified proteins. The bar represents the number of proteins and the solid square represent the percentage of per group of proteins to the total identified proteins. (A) distribution of MW, (B) distribution of pI.

in spliceosome pathway (ID: ko03040), ribosome pathway (ID: ko03010), and other seven pathways (Supp Table 2 [online only]), which suggested a number of transcription and translation programs were now in process, which is consistent with an embryo in a vigorously developing stage.

In the identified proteins, tubulin, actin, profilin, myosin, microtubule-associated proteins, other cytoskeleton proteins and some chaperonin proteins, such as t-complex protein 1, were identified. These proteins involved in the phagosome pathway (ID: ko04145), tight junction pathway (ID: ko04530), and gap junction pathway (ID: ko04540). The organization of the cytoskeleton plays important roles in cell morphogenesis, and the identification of these proteins might be due to the various development programs acting during embryogenesis (Supp Table 3 [online only]).

Another group of identified proteins were HSPs, among them, HSP 90 and HSP 70 were predominant (Supp Table 3 [online only]). Many members of this group might perform chaperone function by stabilizing new proteins to ensure correct folding or by helping to refold proteins that were damaged by the cell stress.

In addition, some proteins were protective for the embryo development. In this study, various proteasomes and ubiquitin proteins were identified (Supp Table 3 [online only]). The identified proteasomes were involved in the proteasome pathway (ID: ko03050) and antigen processing and presentation pathway (ID: ko04612). These might be necessary to regulate specific proteins and remove protein misfolding during the progress of *E. pela* embryo development. Moreover, some proteins associated with antioxidant system were identified (Supp Table 3 [online only]), they might play important roles in removing harmful metabolites conducted in the process of embryo development.

In the protein profile of the *E. pela* egg, some kinds of proteins were likely to be related to the specific differentiation and morphogenesis programs for various tissues and organs, as well as early nymphal morphogenesis (Supp Table 4 [online only]). These proteins included chitinase, cuticular protein analogous to peritrophins 3-B precursor, prophenoloxidase, similar to *n*-synaptobrevin CG17248-PA, muscular protein 20, transformer-2 sex-determining protein, and other proteins, which were likely to be the important proteins for cuticular, nerve tissue, and reproductive organ formation of the new nymph. From these results, it was presumed that the embryo development was in its later stage. The result was consistent with what we would expect at a stage close to egg hatching.

Biological Process and Molecular Functions of the Identified Proteins

Based on the biological process and molecular function according to the GO terms, in total, 213 proteins were found to be involved in 22 categories of biological processes (Fig. 2). Most proteins were related to cellular process (84.04%) and metabolic process (73.71%), which was consistent with the active cell division and vigorous metabolism during the course of embryo development. Proteins related to the GO term 'response to stimulus' showed high representation (25.82%), and proteins involved in 'cell proliferation', 'multi-organism process', and 'growth' made the lowest representation (0.47%) in our protein profile. Molecular function terms associated with *E. pela* egg proteins revealed that most of the proteins were involved in binding (49.09%), followed by catalytic activity, transporter activity, and structural molecule activity. Proteins related to other functions were represented as small groups (Fig. 3). In order to reveal the enzyme classes in *E. pela* egg, proteins with catalytic feature were further classified (Fig. 4). The enzyme distribution illustrated that hydrolases accounted for the largest proportion (49.28%), followed by oxidoreductases and transferases.

KEGG Pathway Analysis

When searched against KEGG reference pathway database, 129 proteins were assigned to 137 KEGG pathways, which were ascribed to five categories: organismal systems, metabolism, genetic information processing, environmental information processing, and cellular processes (Fig. 5). In the metabolism term, there were 56 pathways identified. In particular, 20, 24, and 12 proteins were, respectively, found in connection with 14 carbohydrate metabolism pathways, five energy metabolism pathways, and nine amino acid metabolism pathways. There were 18 identified biological pathways in genetic information processing, folding sorting and degradation were so complex and active that six pathways were associated with 32 proteins in this processing. Under the cellular process category, 18 proteins were involved in four pathways and linked with transport and catabolism. Environmental information processing included signal transduction (eight pathways) and signaling molecules and interaction term (three pathways), the proteins involved in the former were more numerous than the proteins involved in the latter. In the organismal systems term category, 10

Table 1. Categorization of the identified proteins of *E. pela* egg based on the GO analysis

Protein description	ACC	Mol. weight [kDa]/*pI*	Biological_process	Molecular_function
Carbohydrate and energy metabolism				
GF23287	XP_001964635.1	56/10	Acetyl-CoA biosynthetic process from pyruvate	Pyruvate dehydrogenase (acetyl-transferring) activity
AGAP011066-PA	XP_309579.4	3.5/8.1		Oxidoreductase activity
Isocitrate dehydrogenase [NADP]	XP_001971666.1	54/6.7	Isocitrate metabolic process	NAD binding
MGC80785 protein	NP_001087022.1	6.5/7.7		Oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor
Putative uncharacterized protein	XP_967960.2	56/7.2		Oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor
GJ19670	XP_002058888.1	80/7.6	Carbohydrate metabolic process	Hydrolase activity, hydrolyzing *O*-glycosyl compounds
AAEL004297-PA	XP_001648848.1	123/7.2	Cellular carbohydrate metabolic process	ATP binding
AGAP009039-PA	XP_319791.4	10/7.9	carbohydrate metabolic process	Carbohydrate binding
Glucose-6-phosphate isomerase	XP_002005703.1	62/7	Gluconeogenesis	Glucose-6-phosphate isomerase activity
phosphoglyceromutase	NP_001037540.1	28/6.8	Glycolysis	Phosphoglycerate mutase activity
GJ15342	XP_002059015.1	37/9.2		ATP citrate synthase activity
GI22035	XP_002001326.1	6.5/7.5		Catalytic activity
pyruvate kinase	NP_001036906.1	31/5.2	Glycolysis	Pyruvate kinase activity
GJ17558	XP_002052468.1	28/8.8	Porphyrin-containing compound biosynthetic process	Coproporphyrinogen oxidase activity
GE24063	XP_002098232.1	8.1/7.2		Holocytochrome-*c* synthase activity
GH22974	XP_001995110.1	73/7.3	Tricarboxylic acid cycle	Oxidoreductase activity, acting on the CH-CH group of donors
AAEL008167-PB	XP_001658987.1	52/8.6	Fumarate metabolic process	Fumarate hydratase activity
ATP synthase subunit alpha	NP_001040233.1	13/9.6	ATP hydrolysis coupled proton transport	ATP binding
Oligomycin sensitivity-conferring protein	XP_968733.1	7.2/10	ATP synthesis coupled proton transport	Proton-transporting ATP synthase activity, rotational mechanism
vacuolar ATP synthase catalytic subunit A	NP_001091829.1	45/5	ATP hydrolysis coupled proton transport	ATP binding
V-type proton ATPase subunit E	P31402.1	7.4/9.5	ATP hydrolysis coupled proton transport	Proton-transporting ATPase activity, rotational mechanism
GJ16665	XP_002051636.1	93/5.9	ATP hydrolysis coupled proton transport	Hydrogen ion transmembrane transporter activity
electron-transfer-flavoprotein beta polypeptide	NP_001040123.1	18/9		Electron carrier activity
Acyl carrier protein	XP_311483.3	7.6/4.7	Fatty acid biosynthetic process	
malate dehydrogenase, putative	XP_002432539.1	36/9.4	Malate metabolic process	l-Malate dehydrogenase activity
malate dehydrogenase	XP_001659012.1	14/9.3	Malate metabolic process	l-Malate dehydrogenase activity
luciferase-like protein	BAI66602.1	60/8.8	Bioluminescence	catalytic activity
Amino acid metabolism				
peroxidase	XP_001867956.1	9.6/6.6	Response to oxidative stress	Peroxidase activity
GE13192	XP_002090578.1	12/6.5	Response to oxidative stress	Peroxidase activity
AGAP000901-PA	XP_316880.5	61/8.3	Biosynthetic process	Catalytic activity
phosphoserine aminotransferase	XP_001849403.1	30/8.6	l-Serine biosynthetic process	*O*-Phospho-l-serine:2-oxoglutarate aminotransferase activity
GF18621	XP_001955137.1	62/7.4	Proteolysis	Aminopeptidase activity
GI23169	XP_001999260.1	5.3/6.6	Proteolysis	Aminopeptidase activity
prophenoloxidase	XP_001661190.1	6.4/6.7	Oxidation–reduction process	Oxidoreductase activity
Lipid metabolism				
AGAP011066-PA	XP_309579.4	3.5/8.1		Oxidoreductase activity
aldehyde dehydrogenase (NAD+)	NP_001087022.1	6.5/7.3		Oxidoreductase activity
fatty acid synthase *S*-acetyltransferase	XP_001845135.1	249/6	Biosynthetic process	Hydrolase activity, acting on ester bonds

Table 1. Continued

Protein description	ACC	Mol. weight [kDa]/*pI*	Biological_process	Molecular_function
Putative uncharacterized protein	XP_971757.1	32/9.2		Catalytic activity
GK10733	XP_002061054.1	11/9.6		Fatty-acyl-CoA reductase (alcohol-forming) activity
Nucleotide metabolism				
MGC130953 protein	NP_001090100.1	1/6.50	Purine nucleotide biosynthetic process	IMP cyclohydrolase activity
GE23527	XP_002099157.1	9.2/8	Purine nucleotide biosynthetic process	IMP cyclohydrolase activity
gualynate kinase-1	ACD69431.1	1/6.58	Purine nucleotide metabolic process	Guanylate kinase activity
Nudix (Nucleoside diphosphate linked moiety X)-type motif 2	NP_001002323.1	7.3/5.9		Bis(5′-nucleosyl)-tetraphosphatase activity
GMP synthase, putative	XP_002427615.1	76/7.1	GMP biosynthetic process	GMP synthase (glutamine-hydrolyzing) activity
GK17059	XP_002061681.1	17/4.6	Transcription, DNA-dependent	
Ribonucleoside-diphosphate reductase	XP_001660977.1	30/7	DNA replication	Ribonucleoside-diphosphate reductase activity, thioredoxin disulfide as acceptor
AAEL003193-PB	XP_001656515.1	42/6.6	Phosphate-containing compound metabolic process	Inorganic diphosphatase activity
Inosine-5'-monophosphate dehydrogenase	XP_309514.2	14/8.3	GMP biosynthetic process	IMP dehydrogenase activity
GD20335	XP_002103291.1	12/5.9		Flavin adenine dinucleotide binding
transformer-2 sex-determining protein, putative	XP_002432020.1	15.6/10.5		Nucleic acid binding,nucleotide binding
Transcription and translation				
heterogeneous nuclear ribonucleoprotein A1	NP_001093319.1	22/7.1		Nucleic acid binding
Putative uncharacterized protein GLEAN_07585	XP_973561.1	31/10		Nucleic acid binding
AAEL007239-PA	XP_001658243.1	44/8.3		Nucleic acid binding
GK13948	XP_002073099.1	43/12		Nucleotide binding
GL19239	XP_002014541.1	86/7.9	Threonyl-tRNA aminoacylation	Threonine-tRNA ligase activity
leucyl-tRNA synthetase, putative	XP_002422927.1	97/8.3	Leucyl-tRNA aminoacylation	LEUCINE-tRNA ligase activity
GL25810	XP_002018701.1	12/5.8	Alanyl-tRNA aminoacylation	Alanine-tRNA ligase activity
60S ribosomal protein L10A, putative	XP_002426587.1	25/11	Translation	RNA binding
ACYPI006342 protein	NP_001155658.1	16/12	Translation	Structural constituent of ribosome
GH23036	XP_001995231.1	19/11	Translation	RNA binding
Putative uncharacterized protein	XP_967571.1	24/9.8	Translation	RNA binding
eukaryotic translation initiation factor 5	XP_001842254.1	1/4.68	RNA metabolic process	Translation initiation factor activity
Lysine-tRNA ligase	NP_572573.1	38/6.5	Lysyl-tRNA aminoacylation	Lysine-tRNA ligase activity
GTP-binding nuclear protein RAN1, putative	XP_002423913.1	50/9.2	Nucleocytoplasmic transport	GTP binding
RNA-binding protein 8A	XP_001849141.1	19/4.8	RNA processing	RNA binding
Folding, sorting, and degradation				
proteasome subunit alpha type, putative	XP_002422679.1	4.7/5.8	Ubiquitin-dependent protein catabolic process	Threonine-type endopeptidase activity
Proteasome subunit alpha type	NP_001040387.1	5.9/5.1	Ubiquitin-dependent protein catabolic process	Threonine-type endopeptidase activity
Proteasome subunit beta type	XP_317882.3	25/6.9	Proteolysis involved in cellular protein catabolic process	Threonine-type endopeptidase activity
GK22666	XP_002072105.1	16/4.8	Regulation of catalytic activity	Enzyme regulator activity
heat shock 70 kDa protein cognate 4	XP_001850527.1	71/5.2	Response to stress	ATP binding
HSP 70 B2	XP_001861436.1	70/5.5	Response to stress	ATP binding

Table 1. Continued

Protein description	ACC	Mol. weight [kDa]/*pI*	Biological_process	Molecular_function
60 kDa HSP, mitochondrial	XP_001850501.1	60/5.2	Protein refolding	ATP binding
GG25088	XP_001968832.1	61/6.1	Protein refolding	ATP binding
HSP 90 kDa alpha (cytosolic), class B member 1	NP_001025655.1	77/4.7	Protein refolding;response to stress	
HSP 83	XP_001865484.1	82/4.6	Response to stress	ATP binding
HSP 90 protein, putative	XP_002432348.1	83/4.7	Protein refolding; response to stress	ATP binding
disulfide isomerase	XP_001866126.1	11/6.2	Glycerol ether metabolic process	Isomerase activity
GG17350	XP_001980787.1	47/4.2	Protein folding	Calcium ion binding
Prkcsh-prov protein	NP_001087124.1	5.5/4.1	*N*-glycan processing	Calcium ion binding
AGAP001424-PA	XP_321706.5	91/4.6	Protein refolding	ATP binding
AAEL012827-PA	XP_001662951.1	27/4.5	Protein refolding	ATP binding
78 kDa glucose-regulated protein	XP_001845218.1	72/4.8		ATP binding
GH11975	XP_001991597.1	72/5		ATP binding
Putative uncharacterized protein	XP_971446.1	45/7.6	Response to heat	ATP binding
Thioredoxin domain-containing protein 1	ACO12744.1	26/4.6	Cell redox homeostasis	
GJ22764	XP_002054516.1	55/9.1		Ubiquitin-protein ligase activity
Transitional endoplasmic reticulum ATPase TER94	XP_966692.1	89/5.1		Nucleoside-triphosphatase activity
Transport and catabolism				
Dnase2-prov protein	NP_001086671.1	9.7/6.8	DNA metabolic process	Deoxyribonuclease II activity
GL12416	XP_002019473.1	136/7.4	ATP catabolic process	ATPase activity, coupled to transmembrane movement of substances
GG20906	XP_001974813.1	53/4.8	Microtubule-based process	GTP binding
Rab-protein 5	XP_001813105.1	24/8.6	Protein transport	GTP binding
GK13103	XP_002073518.1	21/4.6	Sphingolipid metabolic process	
RAC GTPase, putative	XP_002429222.1	21/6.8	Small GTPase mediated signal transduction	GTP binding
Development and organism system				
GD25430	XP_002082004.1	13/6.5	Oxidation–reduction process	Oxidoreductase activity
AAEL012062-PC	XP_001662217.1	45/5.2	ATP biosynthetic process	Monovalent inorganic cation transmembrane transporter activity
clathrin light chain	XP_001868264.1	6.7/4.2	Intracellular protein transport	Structural molecule activity
GL25029	XP_002021066.1	10/9.5		Fatty-acyl-CoA binding
GF20350	XP_001963345.1	10/8.9		Zinc ion binding
AAEL003413-PA	XP_001656777.1	28/6.9		Serine-type endopeptidase inhibitor activity
AGAP007452-PA	XP_001687921.1	299/7.1	Regulation of Rho protein signal transduction	Rho guanyl-nucleotide exchange factor activity
lumbrokinase-3(1)	XP_001844812.1	5.4/6.7	Proteolysis	Kinase activity
GK12466	XP_002072489.1	9/4.7	Proteolysis	Serine-type endopeptidase activity
leukocyte elastase inhibitor	NP_001089382.1	17/6		Serine-type endopeptidase inhibitor activity
Alpha-2-antiplasmin	NP_777095.1	13/5.5	Acute-phase response	serine-type endopeptidase inhibitor activity
Signaling				
AGAP007523-PB	XP_308355.3	231/5.2		Motor activity
GA13959	XP_001360276.2	5.4/5.2		Motor activity
AAEL004141-PA	XP_001648499.1	11/6.5	Transport	
GTP-binding protein alpha subunit, gna	XP_001858618.1	41/5.1	G-protein coupled receptor signaling pathway	Signal transducer activity
GK23973	XP_002064648.1	4/7.8		Voltage-gated anion channel activity
GK17256	XP_002061919.1	59/4.8	Regulation of cell adhesion	
Peptidyl-prolyl cis–trans isomerase	XP_002134398.1	25/10	Protein peptidyl-prolyl isomerization	Peptidyl-prolyl cis-trans isomerase activity

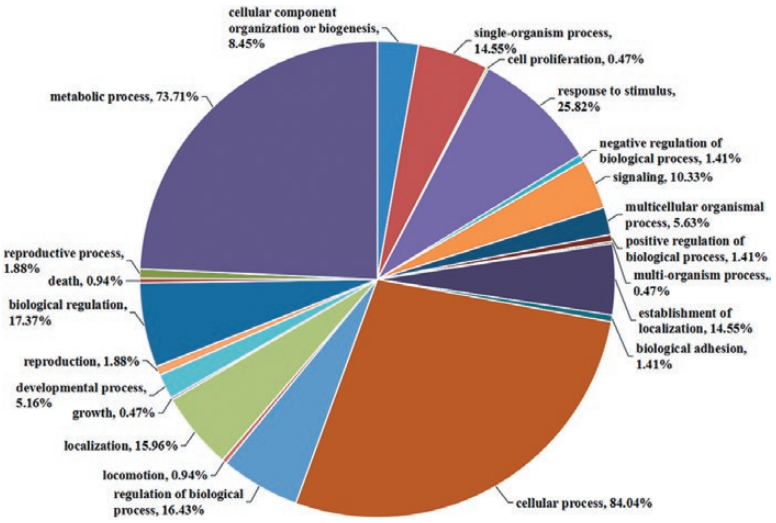

Fig. 2. Biological process components by proportion according to GO classification.

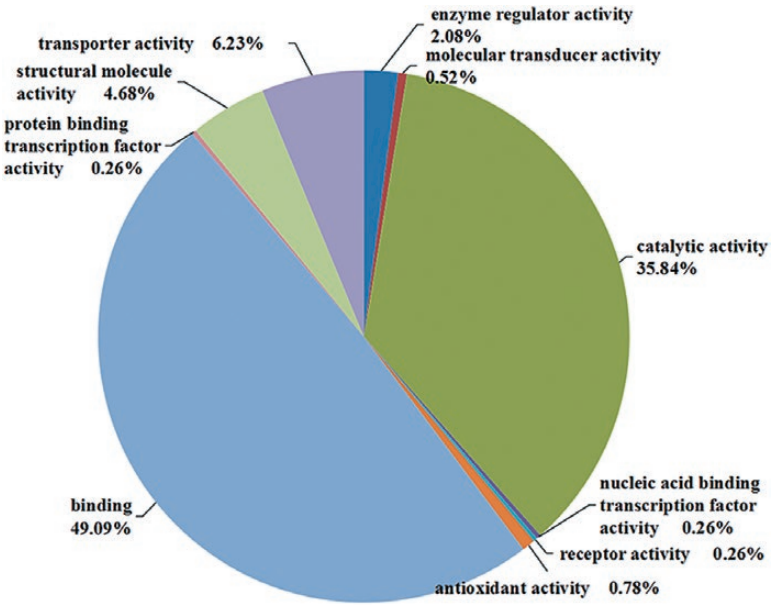

Fig. 3. Molecular function components by proportion according to GO classification.

pathways were included in the immune system, and the proteins related to immune system were the most numerous, which would be consistent with the embryo requiring active protection from pathogens at this stage.

Similarity Distribution of the Identified Proteins

The identified nonredundant proteins were analyzed for their similarity distribution in the database. A majority of the annotated proteins shared similarity with proteins from arthropods (Fig. 6). *E. pela* shared maximum similarity with *Drosophila* (26.19%), followed by different mosquito fauna, beetles, *Pediculus humanus corporis*, and *Maconellicoccus hirsutus* Green (Hemiptera: Pseudococcidae), etc. Out of the identified nonredundant proteins, only about 6.09% proteins exhibited similarity with the known proteins of scale insects.

Target Proteins Selection

One aim of our study was to identify for future study proteins that may be closely connected with biological and ecological characteristics of *E. pela*. These proteins were selected through comparison and analysis using the known functional information of the most similar protein in another insect as a justification, combining this information with the biological and ecological characteristics of *E. pela* and the known information on the molecular level about *E. pela*. The protein named GK10733 (XP_002061054.1) has fatty-acyl-CoA reductase (alcohol-forming) activity and is related to cutin, suberine, and wax biosynthesis pathway according to GO and KEGG analysis (Table 1 and Supp Table 4 [online only]). We hypothesized that fatty-acyl-CoA reductase is relevant to wax secretion on the surface of *E. pela* eggs. A number of HSPs with different molecular weight were identified (Table 1 and Supp Table 3 [online only]), and we

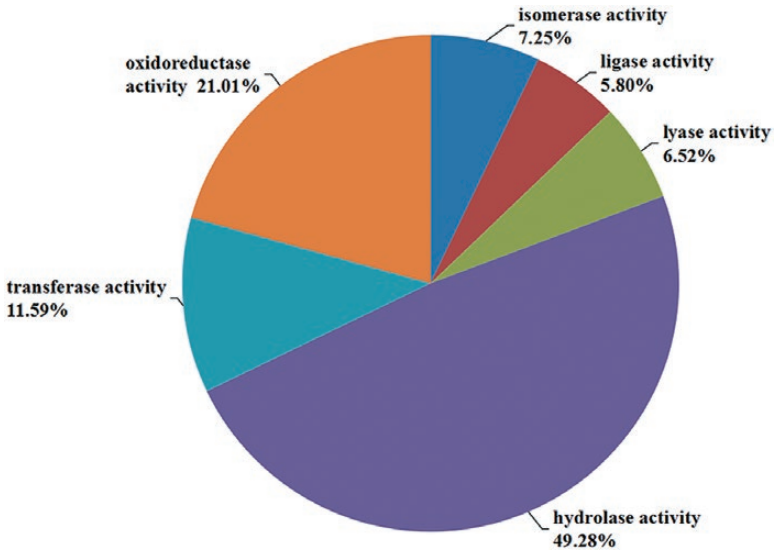

Fig. 4. Enzyme classes of nonredundant proteins according to GO classification.

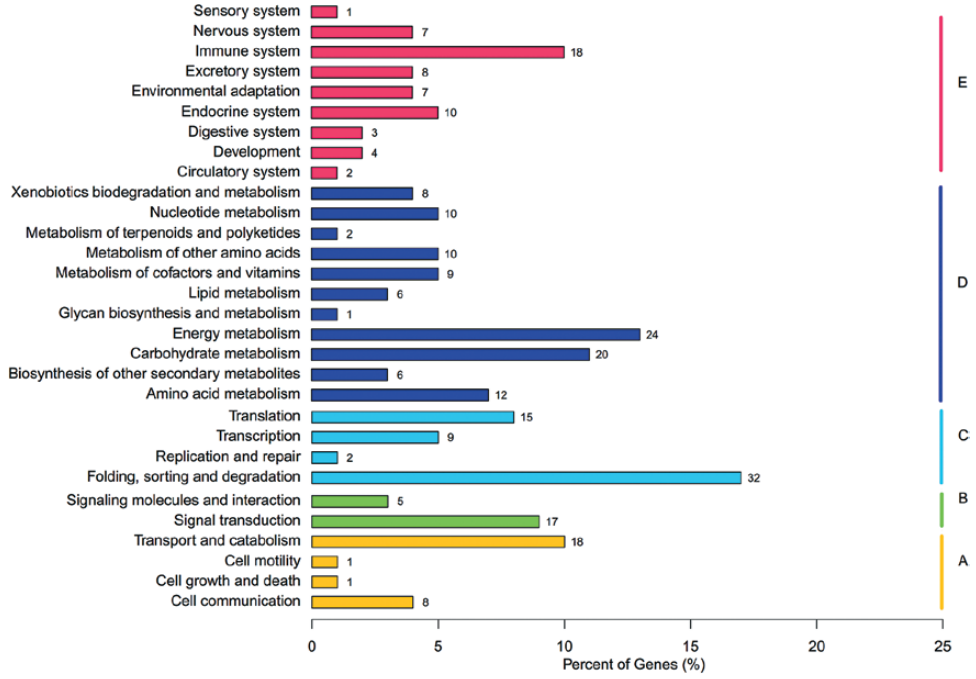

Fig. 5. Classification of pathways according to the definition in KEGG. The pathways were clustered into cellular processes (A), environmental information processing (B), genetic information processing (C), metabolism (D), and organismal systems (E).

have previously studied some HSP genes of *E. pela* (Liu et al. 2013) in an earlier study. We hypothesized that these HSPs are very likely tied to stress resistance to the environment. We analyzed the possible relationship between the genes identified in this study and the typical ecological and biological characteristics of the *E. pela* egg in part of our discussion, and will use this as a basis for further investigation.

Discussion

E. pela is one of the most economically valuable insects, belonging to the family Coccidae. There exists minimal research about *E. pela* embryo development. The protein component of the *E. pela* egg during embryo development stage has not been reported until this study. The egg is light yellow in color, and has features commonly found in an insect egg. *E. pela* embryonic development begins with cleavage by karyokinesis, goes through the formation of blastoderm and germ band, formation and disappearing of the amnion and serosa, differentiation of the germinal layer, germ band sectionalization, the formation of appendage, as well as the formation of the alimentary canal, nerve tissue, dorsal blood vessel, and generative cells (Zhao and Wu 1990). The external body is well developed on the 15th day after oviposition, and the nymph crawls out of the chorion on the 18th day after oviposition. However, the duration of the egg stage varies with temperature and other factors (Zhao and Wu 1990). During the egg development process, protein expression is active, and biosynthesis and catabolism programs are performed.

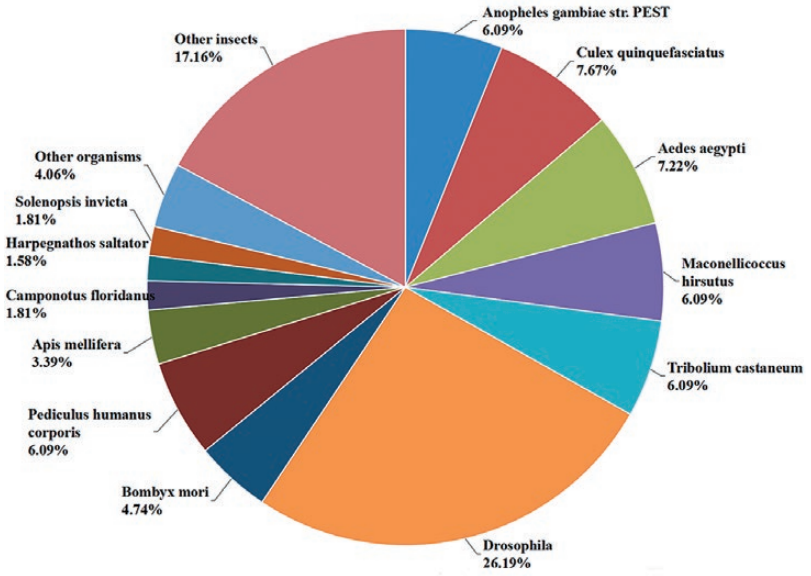

Fig. 6. Pie diagram showing similarity distribution of nonredundant proteins assigned by GO analysis (represented as the percentage of total similarity proteins).

Embryo development is a sequential and complex process controlled by genes. Some proteins are constitutively expressed throughout the development process—these are indispensable for egg development. The existence of proteins expressed in specific stages of the egg suggests that the different developmental stages need specific protein(s) to proceed correctly (Fang and Li 2010).

In this study, we obtained the protein profile of the *E. pela* egg at stages close to hatching. A large number of identified proteins were related to metabolism and organismal systems pathways, and the results were in accordance with the physiological development features of *E. pela* egg. The identified enzymes in the *E. Pela* egg shared maximum similarity with proteins in *Drosophila*, and the fact that *Drosophila* proteins are generally very well characterized was helpful to predict the function of *E. pela* proteins. On the basis of KEGG pathway analysis and GO analysis, we discuss the possible relationship among some identified proteins and the biological and ecological characteristics of the *E. pela* egg.

FAR and Secreting Wax Behavior

According to the wax ester biosynthesis pathway in organism (Cheng and Russell 2004, Doan et al. 2009, Liénard et al. 2010, Teerawanichpan and Qiu 2010, Teerawanichpan et al. 2010), fatty-acyl-CoA reductase (FAR) and wax synthase are the key enzymes. In this study, some detected proteins were predicted to be involved in white wax synthesis. Among them, a protein named GK10733 had fatty-acyl-CoA reductase (alcohol-forming) activity, and was found to be related to the cutin, suberin, and wax biosynthesis pathways according to KEGG analysis (pathway ID: ko00073). The mRNA level of *E. pela* FAR gene in nymphs has previously been analyzed using qRT-PCR, and *E. pela* FAR was assumed the key enzyme to white wax biosynthesis (Yang et al. 2012). In each capsule, all eggs, without exception, are covered with some wax powder in natural conditions. Before this study, some researchers postulated that the wax adhering to the surface of eggs was secreted by the mother (Wu and Zhong 1983), but there was not enough evidence to support this hypothesis. In the present study, we predict that FAR is likely involved in the wax formation on the surface of *E. pela* eggs, though the quantity of wax is so small that this has not generally been a priority for study by researchers. On the other hand, we also cannot rule out that the expression of FAR protein in the egg stage is possibly in preparation for the secreting wax behavior of *E. pela* nymphs.

HSPs in *E. pela* Egg

HSPs are known to play a vital role in both normal cellular homeostasis and stress response, and are involved in many biological functions such as cellular communication, immune response, protein transport, cell cycle regulation, apoptosis, gametogenesis, and aging (Sarkar et al. 2011). A report documented that Hsp70 and small HSPs are probably the major players in midgut metamorphosis in *Spodoptera litura* (Gu et al. 2012); this viewpoint provides valuable insight into the roles of the HSP superfamily in insect metamorphosis. Furthermore, HSPs are documented widely as defensive response proteins to stress factors including heat shock, cold shock, and other abiotic stresses and biotic stresses in insects (Zhao and Jones 2012). The functions of various HSP often overlap but can be different between different proteins (Zhang and Denlinger 2010, Benoit et al. 2011, Michaud et al. 2011, Xu et al. 2011). In the present study, heat shock 70 kDa protein cognate 4 and HSP 70 B2 were identified, and these were primarily involved in the spliceosome pathway, the protein processing in endoplasmic reticulum pathway, the MAPK signaling pathway, the endocytosis pathway and antigen processing and presentation pathway on the basis of KEGG analysis, which showed these proteins were associated with multiple biological processes. Moreover, some higher molecular weight proteins, HSP 90 kDa alpha (cytosolic) and HSP 83, were identified for the plant–pathogen interaction pathway, the progesterone-mediated oocyte maturation pathway, the antigen processing and presentation, the NOD-like receptor signaling pathway, the PI3K-Akt signaling pathway, and the protein processing in endoplasmic reticulum pathway. In addition, other HSPs were identified, including HSP cognate 5, HSP beta-6-like isoform 1, HSP 68a, 60 kDa HSP (mitochondrial), and a few putative small HSPs. We hypothesized that the identified HSPs likely exhibit very important role in aiding organogenesis by folding newly synthesized proteins, binding other non-native proteins, and assisting proteins in the correct folding and functional actualization.

Furthermore, synthesis of the relevant literature about *E. pela*, particularly as regards its ecological strategy, some HSPs likely

primarily function in response to very high heat and humidity stress. We reasoned that, first, thousands of eggs are laid in relatively closed ootheca, and they need to respond to intense heat, and likely hypoxia. Secondly, temperature and humidity is very high in the source region of white wax during the period of *E. pela* hatching, thus the egg likely responds to this abiotic stress by using the regulatory mechanism of HSPs. The host plant is often infested by pathogens because of the scale insect colonization, and there are many pollutants on the surface of egg capsule, but surprisingly impaired eggs were not found in our investigation. For this reason, we hypothesized that there is some inhibitory mechanism, potentially HSPs, in effect to protect the eggs from these harms. Therefore, considerable further work is needed to fully understand these mechanisms, and the HSPs will be target proteins in our future work.

Amount of Metabolic Energy Required for Development

A very radical morphological transformation is exhibited from egg to nymphal stage, and large amount of carbohydrate metabolism and energy production is needed to undergo extensive organogenesis during this process. In this study, according to GO and KEGG classification, some important proteins were implicated in carbohydrate metabolism and energy production. In particular, the citrate cycle (TCA cycle), glycolysis/gluconeogenesis, pyruvate metabolism, propanoate metabolism, and pentose and glucuronate interconversions were assigned 41 proteins (32%), and 24 proteins (19%) involved in energy metabolism. This suggested that, similar to other insects (Zhong et al. 2005, Li et al. 2009), large amounts of metabolic energy produced by all types of metabolism are required for *E. pela* embryo development.

Cytoskeletal Proteins Being Essential for Metamorphosis

Cytoskeletal proteins have a number of essential cellular functions including maintaining the stability of cell shape and structure, and play important roles in intracellular transport and cellular division (Wulfkuhle et al. 1998). One report showed that controlled actin assembly is crucial to a wide variety of cellular processes (Quinlan 2013), and polymerization of actin filaments against cellular membranes provides necessary force for a number of cellular processes leading to protein recruitment (Saarikangas et al. 2010, Lucas et al. 2013). Tubulins are the major constituents of microtubules, and have a range of post-translational modifications, potentially regulating the microtubule cytoskeleton (Janke and Kneussel 2010). The Tcp-1 complex belongs to Type II Chaperonin; it is a multi-subunit molecular machine that assists in the folding of 10% of newly translated cytosolic proteins in eukaryotes (Coghlin et al. 2006, Posokhova et al. 2011). In this study, the majority of identified cytoskeletal proteins were belong to tubulin, actin, and myosin proteins. These proteins are associated with ultrastructure, cell division, and cellular morphology. We hypothesized these proteins are likely essential for *E. pela* embryo.

Protective Proteins for Embryo Development

Throughout developmental, various quantities of metabolites can be beneficial or harmful to cells and tissues. For instance, excess reactive oxygen species (ROS) can induce oxidative modification of biological micromolecules, and inhibit protein function by protein oxidation, lipid peroxidation, DNA base modifications, and strand break (Circu and Aw 2010). Aerobic organisms have developed complicated antioxidant mechanisms to protect themselves against toxic ROS, superoxide dismutases, catalases, peroxidases, thioredoxin, glutathione peroxidase, and other enzymes (Wang et al. 2008). Superoxide dismutase (XP_002048532.1), peroxidase (XP_001867956.1), manganese superoxide dismutase (AEL79287.1), glutathione *S*-transferase theta (ACB36909.1), aldo-keto reductase (XP_001844819.1), and other proteins were found to be expressed in the *E. pela* eggs stage. These proteins might participate in the protective pathways in order to provide essential protection from harmful metabolites during embryo development of *E. pela*.

Conclusions

E. pela is a model for scale insects. This study provided the first proteomic analysis in the eggs of *E. pela* near hatching, which provided a basis to elucidate the mechanism underlying embryogenesis, and illuminated candidate proteins for deeper research. Some identified proteins might be directly correlated to the biological characteristics of the eggs at the stage at which they were examined. Further research is needed to verify the functions of these important proteins.

Supplementary Data

Supplementary data are available at *Journal of Insect Science* online.

Acknowledgments

This work is supported by Research Funds for the Central Non-profit Research Institution of CAF (CAFYBB2017ZB005); Forestry Industry Research Special Funds for Public Welfare Projects (201204602) and the National High Technology Research and Development Program of China (2014AA021801).

References Cited

Amenya, D. A., W. Chou, J. Y. Li, G. Y. Yan, P. D. Gershon, A. A. James, and O. Marinotti. 2010. Proteomics reveals novel components of the *Anopheles gambiae* eggshell. J. Insect Physiol. 56: 1414–1419.

Benoit, J. B., G. L. Martinez, K. R. Patrick, Z. P. Phillips, T. B. Krause, and D. L. Denlinger. 2011. Drinking a hot blood meal elicits a protective heat shock response in mosquitoes. Proc. Natl Acad. Sci. USA. 108: 8026–8029.

Chen, X. M. 2011. Natural population ecology of *Ericerus pela*. Science Press, Beijing, China.

Chen, X. M., and Y. Feng. 2009. An introduction to resource entomology. Science Press, Beijing, China.

Cheng, J. B., and D. W. Russell. 2004. Mammalian wax biosynthesis. I. Identification of two fatty acyl-coenzyme A reductases with different substrate specificities and tissue distributions. J. Biol. Chem. 279: 37789–37797.

Circu, M. L., and T. Y. Aw. 2010. Reactive oxygen species, cellular redox systems, and apoptosis. Free Radical Biol. Med. 48: 749–762.

Coghlin, C., B. Carpenter, S. R. Dundas, L. C. Lawrie, C. Telfer, and G. I. Murray. 2006. Characterization and over-expression of chaperonin t-complex proteins in colorectal cancer. J. Pathol. 210: 351–357.

Doan, T. T. P., A. S. Carlsson, M. Hamberg, L. Bülow, S. Stymne, and P. Olsson. 2009. Functional expression of five *Arabidopsis* fatty acyl-CoA reductase genes in *Escherichia coli*. J. Plant Physiol. 166: 787–796.

Fang, Y., and J. K. Li. 2010. Analysis of developmental proteome at egg stage of drone honeybees (*A. m. ligustica*). Sci. Agric. Sinc. 3: 392–400.

Gala, A., Y. Fang, D. Woltedji, L. Zhang, B. Han, M. Feng, and J. Li. 2013. Changes of proteome and phosphoproteome trigger embryo-larva transition of honeybee worker (*Apis mellifera* Ligustica). J. Proteomics 78: 428–446.

Gu, J., L. X. Huang, Y. Shen, L. H. Huang, and Q. L. Feng. 2012. Hsp70 and small Hsps are the major heat shock protein members involved in midgut metamorphosis in the common cutworm, *Spodoptera litura*. Insect Mol. Biol. 21: 535–543.

Janke, C., and M. Kneussel. 2010. Tubulin post-translational modifications: encoding functions on the neuronal microtubule cytoskeleton. Trends Neurosci. 33: 362–372.

Li, J. Y., X. Chen, S. H. Hosseini Moghaddam, M. Chen, H. Wei, and B. X. Zhong. 2009. Shotgun proteomics approach to characterizing the embryonic proteome of the silkworm, *Bombyx mori*, at labrum appearance stage. Insect Mol. Biol. 18: 649–660.

Li, J.Y., S. H. Hosseini Moghaddam, J. E. Chen, M. Chen, and B. X. Zhong. 2010. Shotgun proteomic analysis on the embryos of silkworm *Bombyx mori* at the end of organogenesis. Insect Biochem. Mol. Biol. 40: 293–302.

Li, J. K., Y. Fang, L. Zhang, and D. Begna. 2011. Honeybee (*Apis mellifera ligustica*) drone embryo proteomes. J. Insect Physiol. 57: 372–384.

Liénard, M. A., A. K. Hagström, J. M. Lassance, and C. Löfstedt. 2010. Evolution of multicomponent pheromone signals in small ermine moths involves a single fatty-acyl reductase gene. Proc. Natl Acad. Sci. USA 107: 10955–10960.

Liu, W. W., P. Yang, and X. M. Chen. 2013. Expression analysis of heat shock protein genes in *Ericerus pela* under cold stress. Forest Res. 26: 681–685.

Lucas, E. P., I. Khanal, P. Gaspar, G. C. Fletcher, C. Polesello, N. Tapon, and B. J. Thompson. 2013. The Hippo pathway polarizes the actin cytoskeleton during collective migration of *Drosophila* border cells. J. Cell Biol. 201: 875–885.

Michaud, M. R., N. M. Teets, J. T. Peyton, B. M. Blobner, and D. L. Denlinger. 2011. Heat shock response to hypoxia and its attenuation during recovery in the flesh fly, *Sarcophaga crassipalpis*. J. Insect Physiol. 57: 203–210.

Müller, H., D. Schmidt, S. Steinbrink, E. Mirgorodskaya, V. Lehmann, K. Habermann, F. Dreher, N. Gustavsson, T. Kessler, H. Lehrach, et al. 2010. Proteomic and functional analysis of the mitotic *Drosophila* centrosome. EMBO. J. 29: 3344–3357.

Posokhova, E., H. M. Song, M. Belcastro, L. A. Higgins, L. R. Bigley, N. A. Michaud, K. A. Martemyanov, and M. Sokolov. 2011. Disruption of the chaperonin containing TCP-1 function affects protein networks essential for rod outer segment morphogenesis and survival. Mol. Cell. Proteomics 1: 1–12.

Quinlan, M. E. 2013. Direct interaction between two actin nucleators is required in *Drosophila* oogenesis. Development. 140: 4417–4425.

Saarikangas, J., H. X. Zhao, and P. Lappalainen. 2010. Regulation of the actin cytoskeleton-plasma membrane interplay by phosphoinositides. Physiol. Rev. 90: 259–289.

Sarkar, S., M. D. Singh, R. Yadav, K. P. Arunkumar, and G.W. Pittman. 2011. Heat shock proteins: molecules with assorted functions. Front. Biol. 6: 312–327.

Teerawanichpan, P., and X. Qiu. 2010. Fatty acyl-CoA reductase and wax synthase from *Euglena gracilis* in the biosynthesis of medium chain wax esters. Lipids. 45: 263–273.

Teerawanichpan, P., A. J. Robertson, and X. Qiu. 2010. A fatty acyl-CoA reductase highly expressed in the head of honey bee (*Apis mellifera*) involves biosynthesis of a wide range of aliphatic fatty alcohols. Insect Biochem. Mol. Biol. 40: 641–649.

Wang, Q., K. Chen, Q. Yao, Y. Zhao, Y. J. Li, H. X. Shen, and R. H. Mu. 2008. Identification and characterization of a novel 1-Cys peroxiredoxin from silkworm, *Bombyx mori*. Comp. Biochem. Physiol. 149: 176–182.

Wu, C. B., and Y. H. Zhong. 1983. Study on the bionomics of the white wax scale *Ericerus pela* Chavannes part I. J. Sichuan Univ. 1: 91–99.

Wulfkuhle, J. D., N. S. Petersen, and J. J. Otto. 1998. Changes in the F-actin cytoskeleton during neurosensory bristle development in *Drosophila*: the role of singed and forked proteins. Cell Motil. Cytoskeleton 40: 119–132.

Xu, Q., Q. Zou, H. Z. Zheng, F. Zhang, B. Tang, and S. G. Wang. 2011. Three heat shock proteins from *Spodoptera exigua*: gene cloning, characterization and comparative stress response during heat and cold shocks. Comp. Biochem. Physiol. B. 159: 92–102.

Yang, P., and X. M. Chen. 2014. Protein profiles of Chinese white wax scale, *Ericerus pela*, at the male pupal stage by high-throughput proteomics. Arch. Insect Biochem. Physiol. 87: 214–233.

Yang, P., J. Y. Zhu, M. Li, J. M. Li, and X. M. Chen. 2011. Soluble proteome analysis of male *Ericerus pela* Chavannes cuticle at the stage of the second instar larva. Afr. J. Microbiol. Res. 5: 1108–1118.

Yang, P., J. Y. Zhu, Z. J. Gong, D. L. Xu, X. M. Chen, W. W. Liu, X. D. Lin, and Y. F. Li. 2012. Transcriptome analysis of the Chinese white wax scale *Ericerus pela* with focus on genes involved in wax biosynthesis. PLoS One. 7: e35719.

Yang, P., X. M. Chen, W. W. Liu, Y. Feng, and T. Sun. 2015. Transcriptome analysis of sexually dimorphic Chinese white wax scale insects reveals key differences in developmental programs and transcription factor expression. Sci. Rep. 5: 8141.

Yu, S.H., Q. Qi, T. Sun, X. Q. Wang, P. Yang, and Y. Feng. 2016. Transcriptome analysis of male white wax scale pupae. Forest Res. 29: 413–417.

Zhang, Q. R., and D. L. Denlinger. 2010. Molecular characterization of heat shock protein 90, 70 and 70 cognate cDNAs and their expression patterns during thermal stress and pupal diapause in the corn earworm. J. Insect Physiol. 56: 138–150.

Zhao, L., and W. A. Jones. 2012. Expression of heat shock protein genes in insect stress responses. Invertebrate Surviv. J. 9: 93–101.

Zhao, X. P., and C. B. Wu. 1990. The embryonic development of *Ericerus pela* Chavannes. J. Sichuan Univ. 2: 222–231.

Zhong, B. X., J. K. Li, J. R. Lin, J. S. Liang, S. K. Su, H. S. Xu, H. Y. Yan, P. B. Zhang, and H. Fujii. 2005. Possible effect of 30 K proteins in embryonic development of silkworm *Bombyx mori*. Acta Biochem. Biophys. Sin. 37: 355–361.