RESEARCH ARTICLE
doi: 10.2306/scienceasia1513-1874.2022.025

ScienceAsia 48 (2022): **171–180**

Transcriptome analysis of 20-hydroxyecdysone induced differentially expressed genes in the posterior silk gland of the silkworm, *Bombyx mori*

Yiling Zhang[a,b,†,*], **Mingzhen Ma**[a,†], **Haoran Huang**[a], **Yu Zhang**[a], **Guodong Zhao**[a,b,*]

[a] School of Biotechnology, Jiangsu University of Science and Technology, Jiangsu 212018 China
[b] Sericultural Research Institute, Chinese Academy of Agricultural Sciences, Jiangsu 212018 China

*Corresponding authors, e-mail: zhangyiling008@126.com, sdgdzhao@just.edu.cn
† These authors contributed equally to this work.

Received 15 Jul 2021
Accepted 29 Nov 2021

ABSTRACT: The silk gland is a larval-specific tissue of lepidopteran insects, including silkworm (*Bombyx mori*). Ecdysteroid is one of the two most important hormones that regulate larval molting and metamorphosis. However, the 20-hydroxyecdysone (20E)-induced gene expression profile in the posterior silk gland (PSG) of silkworm has not been clearly determined. In the present study, the silkworm larvae on the 2nd day of the fifth instar were treated with 20E. The analysis identified 2151, 4385, and 5632 differentially expressed genes (DEGs) at 4 h, 24 h, and 48 h, respectively, after 20E treatment; the downregulated transcripts were more than the upregulated transcripts at all time points. The possible functions of the DEGs were determined using the Gene Ontology (GO) classification system. The Kyoto Encyclopedia of Genes and Genomes (KEGG) database carried out the pathway enrichment analysis. Furthermore, few selected DEGs were verified by quantitative real-time PCR (qRT-PCR). Thus, the present study generated extensive transcriptome and DEG profiling data using the PSG of the 20E-fed silkworm, which provided comprehensive gene expression information at the transcriptional level. The study provides a comprehensive view of the molecular basis of gene regulation by 20E in silkworm PSG.

KEYWORDS: *Bombyx mori*, posterior silk gland, transcriptome, RNA-seq, 20E

INTRODUCTION

As the most important insect species, the silkworm *Bombyx mori* has been domesticated for silk production for more than 5000 years [1]. Silkworm pupae have recently been reported as one of the most suitable sources for preparation of pharmaceutical grade chitosan [2]. It is an economically important and well-studied lepidopteran model system for molecular genetics, biochemistry, developmental biology, and genomics research [3]. The functional adaptation of the *Bombyx* silk gland for silk protein synthesis and secretion makes it a suitable model for gene expression and regulation studies and biotechnological applications [4]. Due to the efficient protein synthesis production by the silk gland and the practicality of mass cultivation, the silkworm is considered a valuable candidate for producing recombinant proteins [5]. Therefore, understanding the molecular mechanism of silk gland development and regulation will be valuable for applied and basic research.

The silkworm silk gland, homologous to the *Drosophila* salivary gland, can be anatomically and physiologically divided into three distinct compartments: anterior silk gland (ASG), middle silk gland (MSG), and posterior silk gland (PSG) [6]. The PSG is the longest suborgan that is responsible for synthesizing the silk core protein fibroin, composed of heavy (H) and light (L) chains and P25, also named fibrohexamerin (fhx) [7, 8]. The fibroin secreted exclusively from the larval PSG moves to the MSG, where heterogeneous sericin proteins are enveloped and then toward the ASG, where the silk fiber is formed and spun.

Studies have proven that the function and development of silk glands are under the control of hormones [9]. Ecdysone is the principal steroid hormone in insects synthesized by the action of prothoracicotropic hormone (PTTH) on the prothoracic glands [10]. In silkworm, the inactive form of ecdysone is released into the hemolymph and converted to an active state, 20-hydroxyecdysone (20E), in the peripheral tissues [11, 12]. In *B. mori*, low ecdysteroid levels are necessary for the proper function and development of the silk gland [13]. On the other hand, high concentrations of ecdysteroids caused silk gland regression during the larval molts or degeneration during the pupal molts [14]. Typically, treatment with 20E can promote the maturation of silkworms and shorten the developmental duration. The hormonal studies on *B. mori* mainly intended to exploit silk commercially [15]. Meanwhile, microarray analysis has characterized the global gene expression profile of the silkworm PSG [16]. The proteomic profiles of PSGs from silkworms fed on different diets have also been compared to determine the effects of differential nutri-

ent conditions on silk protein synthesis [17]. However, the effect of 20E on the transcriptome in PSG has not been described.

The present study employed the Illumina Genome Analyzer platform to identify the differentially expressed genes (DEGs) and understand the transcriptome changes in the PSGs of 20E-fed silkworms. Furthermore, quantitative real-time PCR (qRT-PCR) was used to verify few selected DEGs. The study's findings provide the first comprehensive view of PSG development and gene regulation by 20E at the transcriptional level.

MATERIALS AND METHODS

Preparation of ecdysone solution

Ecdysone (20-hydroxyecdysone, 20E) (Sigma-Aldrich Trading Co, Ltd, Shanghai, China) solution was diluted to 2×10^{-3} μg/μl concentration [18]. Mulberry leaves were immersed in the solution for 5 s and allowed to dry naturally before being fed to the second day, fifth instar silkworms.

Insect rearing and PSG dissection

The domesticated silkworm strain Dazao was used in this study. These worms were reared at a standard temperature under a 12 h light/12 h dark cycle. Healthy and active fifth instar larvae were randomly selected for the experiments, with three replicates. The control group was fed with mulberry leaves soaked in double-distilled water, and the experimental groups were fed with mulberry leaves dipped in 20E solution for 1 h and subsequently with normal mulberry leaves. The experimental and control larvae were dissected at 4 h, 24 h, and 48 h after feeding 20E-coated leaves. The PSG of ten larvae were collected and rinsed three times in diethylpyrocarbonate (DEPC)-treated double-distilled water. All tissues were stored at −80 °C until further use.

RNA extraction

Total RNA was isolated from the 20E-treated PSG and the control PSG using TRIzol (Invitrogen, Camarillo, CA) according to the manufacturer's instructions to perform Illumina sequencing. Further, the RNA samples were pretreated with RNase-free DNase I (TaKaRa, Beijing, China) for 30 min at 37 °C to remove any residual DNA. RNA quality was verified using agarose gel electrophoresis and NanoDrop ND-1000 spectrophotometer (NanoDrop, Wilmington, DE).

Library preparation and Illumina sequencing

The RNA-seq libraries were prepared using the TruSeq™ DNA Sample Preparation Kit (Illumina, San Diego, USA), following the manufacturer's instructions. Approximately 20 μg of total RNA obtained from the 20E-treated PSG was used to isolate mRNA with Sera Mag Magnetic Oligo (dT) Beads. The purified mRNA was fragmented into small pieces (100–400 bp) using divalent cations for 5 min at 94 °C to avoid priming bias. The double-stranded cDNA was first synthesized using the SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen, Camarillo, CA) with random hexamer (N6) primers. The libraries were sequenced using Illumina's HighSeq 2000 platform according to the manufacturer's protocol.

Data analysis

Preprocessing was carried out following a stringent filtering process. First, the raw reads were filtered to remove low-quality sequences, and the generated clean reads were aligned to the deposited reference gene sequences generated from the SilkDB database (v2.0; ftp://silkdb.org/pub/current/Gene/) using TopHat software (v2.0.6) [19]. The DEGs were identified at a p-value of 0.05 [20]. For genes expressed only in one sample, the DEGs were identified at a fold-change value of 1 (FC-test).

Functional annotation

The possible functions of all DEGs were annotated using the GO (Gene Ontology) classification system (http://www.geneontology.org/). GO term with a corrected $p \leqslant 0.05$ was designated as a significantly enriched term. WEGO (Web Gene Ontology Annotation Plot) software was used to visualize, compare, and plot the GO annotation results [21]. The KEGG (Kyoto Encyclopedia of Genes and Genomes) database (http://www.genome.jp/kegg/) was used to analyze the pathway enrichment [22]. The DEGs were functionally annotated following BLAST comparisons with the manually curated KEGG GENES database. Pathways with a corrected $p \leqslant 0.05$ were designated as significantly enriched pathways.

Quantitative real-time PCR

Quantitative real-time PCR was carried out on a real-time reverse transcription-PCR system (ABI 7300, Applied Biosystems, USA) to validate the RNA-seq data. Primers used to determine the relative expression level of the genes were designed based on the cDNA sequences using the Primer premier software (v5.0) (Table S1). The housekeeping gene of *B. mori* actin3 was used as an internal control to normalize the gene relative expression levels. A total of 20 μl of the real-time PCR solution contained 0.2 μg of cDNA, 5 pmol of each primer, and 10 μl of SYBR Green PCR Master Mix. PCR was run in triplicates with two biological replicates; the thermal cycling program was 95 °C for 10 min followed by 40 cycles at 95 °C for 30 s, 60 °C for 1 min. The relative expression level of these genes was determined according to the $2^{-\Delta\Delta CT}$ method [23].

RESULTS

Global statistics of RNA-seq data

To better understand the 20E-induced gene regulation in *Bombyx* PSG, four expression libraries were constructed from the 20E-treated and control PSGs of silkworms, and the transcriptomes were compared. The RNA-seq raw data were deposited to NCBI Sequence Read Archive (SRA) with the accession numbers SRR15522085, SRR15522084, SRR15522083, and SRR15522082 (https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA756109). The total length of the reads was about 21 Gb, representing about a 45-fold coverage of the *B. mori* genome and more than 637-fold coverage of the annotated transcriptome. After removing the low-quality reads, a total of 221 733 704 paired-end reads with an average length of 97.75 bp were obtained and used to map the *Bombyx* genome using TopHat software [24]. Approximately 81.87% of the reads for NC (negative control), and 79.55% , 79.46%, and 79.96% after 4, 24, and 48 h 20E treatment, respectively, were found mapped to the silkworm genome (Table S2).

Further, RPKM (Reads Per Kilobase of transcript per Million mapped reads) was used to estimate the gene expression abundance. Of the predicted 13 340 protein-coding silkworm genes that were built by merging the different gene datasets using GLEAN in the SilkDB [25], 10 330, 10 235, 10 206, and 10 172 genes were found expressed in NC, 4 h, 24 h, and 48 h, respectively, which covered 77.44%, 76.72%, 76.51%, and 76.25% of the predicted genes, respectively (Table S3).

Identification of DEGs after 20E treatment

The DEGs identified at the early feeding stage may provide an important clue to the *Bombyx* response against 20E. First, the expression abundance of tag-mapped genes was normalized by counting the number of transcripts per million (TPM) clean tags, and then, $p \leqslant 0.05$ and $\log_2$ Ratio$\geqslant$1 were used as a threshold to distinguish the DEGs during feeding. Analysis showed differences in DEGs identified at 4 h, 24 h, and 48 h (Fig. 1). A total of 2151, 4385, and 5632 genes were differentially expressed at 4 h, 24 h, and 48 h after 20E treatment.

Venn analysis showed that the total number of upregulated and downregulated genes were 345 and 660, respectively (Fig. 2). Meanwhile, 378, 386, and 429 genes were uniquely upregulated, while 149, 592, and 1683 genes were downregulated at 4 h, 24 h, and 48 h, respectively (Fig. 2).

Functional annotation of DEGs

The GO analysis was performed to classify the functions of DEGs revealed by transcriptomic analysis. A total of 202 upregulated and 202 downregulated genes were annotated in the GO functional groups for 4 h.

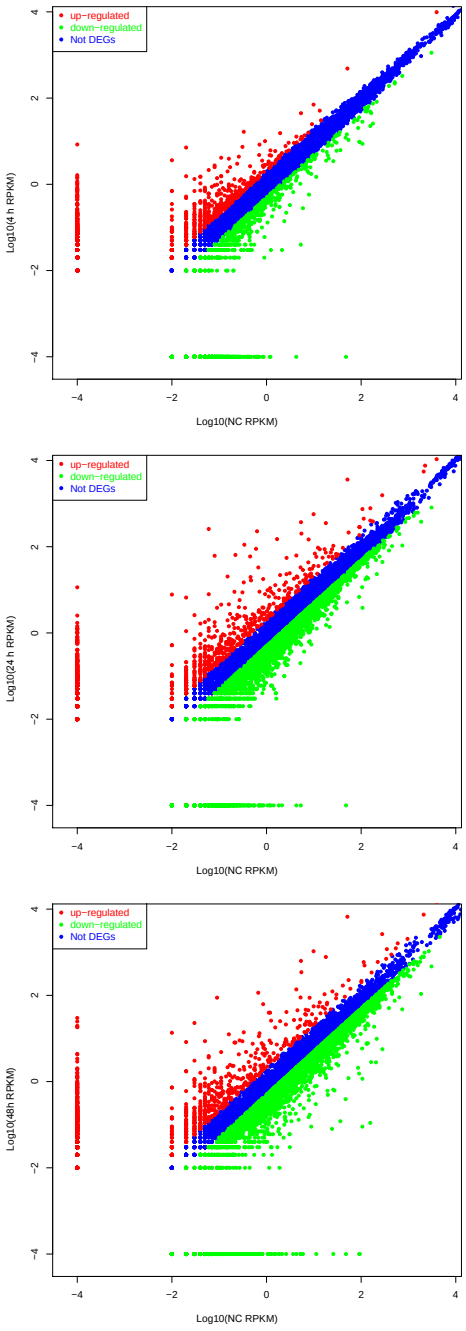

Fig. 1 Differential expressed genes of the 4 h, 24 h and 48 h 20E-treated PSGs. All reads were aligned to predicted genes and are shown as log10 values derived from cDNAs of the PSGs. (A) 4 h; (B) 24 h; (C) 48 h. The up-regulated genes and down-regulated genes are marked with red and green spots, respectively.

Meanwhile, the numbers of upregulated and downregulated genes with GO annotations were 230 and 187, respectively, for 24 h and 230 and 207, respectively,

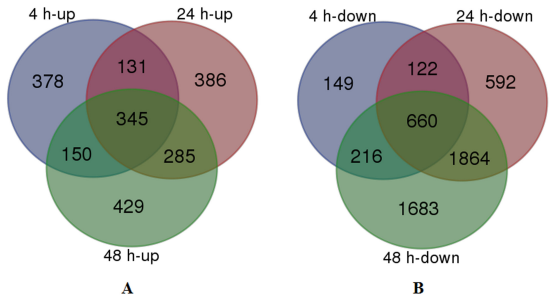

Fig. 2 Number of common differential expressed genes between different time points after 20E treatment. (A) up-regulated; (B) down-regulated.

Table 1 Summary of GO and KEGG enrichment analysis of the 4 h, 24 h and 48 h 20E-treated PSGs.

			GO		KEGG	
		DEG	$p \leqslant 0.05$	$p \leqslant 0.01$	$p \leqslant 0.05$	$p \leqslant 0.01$
4 h	all	2151	265	151	72	48
	up	1004	202	119	65	37
	down	1147	202	124	52	24
24 h	all	4385	253	145	88	59
	up	1147	230	108	62	38
	down	3238	187	128	75	47
48 h	all	5632	223	108	78	37
	up	1209	230	108	76	50
	down	4423	207	77	59	31

for 48 h (Table 1). The DEGs were termed by GO ontology into three categories, namely cellular component, molecular function, and biological process. In each of the three main categories, "cell and cell parts", "binding", "catalytic activity", "cellular process", and "metabolic process" terms were dominant. Interestingly, few genes related to "chemoattractant" were upregulated in 20E-treated PSG, while few related to "virion and virion part", "protein tag", and "viral reproduction" were downregulated (Fig. 3).

KEGG analysis of DEGs

Furthermore, the DEGs were mapped to the reference canonical pathways using the KEGG database to identify the biological pathways active in the 20-treated PSG [21]. KEGG enrichment analysis revealed that 65 upregulated and 52 downregulated pathways were predicted for 4 h. Meanwhile, the numbers of upregulated and downregulated genes annotated were 62 and 75, respectively, for 24 h and 76 and 59, respectively, for 48 h (Table 1). The ten most important pathways of 4 h, 24 h, and 48 h are displayed in Fig. 4. KEGG pathway analysis revealed that the upregulated genes at 4 h, 24 h, and 48 h were mainly enriched in the "lysosome", "pentose and glucuronate interconversions", and "purine metabolism" pathways, respectively. Conversely, the downregulated genes at 4 h, 24 h, and 48 h were mainly enriched in the "biosynthesis of unsaturated fatty acids", "pathways in cancer", and "splicesome", respectively.

DEGs and pathways identified at 4 h

At 4 h after feeding 20E-treated mulberry leaves, 2151 DEGs were identified, including 1004 upregulated and 1147 downregulated genes. A total of 117 KEGG pathways were identified at this time point, of which 65 were upregulated, such as "lysosome", "phagosome", "glycerolipid metabolism", "neuroactive ligand-receptor interaction", "amino acids (glycine, serine, and threonine) metabolism", "focal adhesion", "gap junction", and "pentose and glucuronate interconversions"; 15, 14, and 11 upregulated DEGs were involved in "lysosome", "phagosome" and "glycerolipid metabolism", respectively. Meanwhile, 52 pathways were downregulated, including "biosynthesis of unsaturated fatty acids", "Huntington's disease", "peroxisome", "neuroactive ligand-receptor interaction", "pancreatic secretion", "starch and sucrose metabolism", "fatty acid biosynthesis", and "biotin metabolism"; 17, 16, and 11 downregulated DEGs were involved in "biosynthesis of unsaturated fatty acids", "Huntington's disease", and "peroxisome", respectively.

DEGs and pathways identified at 24 h

Furthermore, 4385 DEGs, including 1147 upregulated and 3238 downregulated ones, were identified at 24 h after 20E treatment. At this time point, 137 pathways were identified, and 62 of them, including "pentose and glucuronate interconversions", "glycine, serine and threonine metabolism", "starch and sucrose metabolism", "metabolism of xenobiotics by cytochrome P450", "phagosome", "lysosome", "drug metabolism-other enzymes", "glycerolipid metabolism", "drug metabolism-cytochrome P450", and "peroxisome" were upregulated; 15, 15, and 14 upregulated DEGs were enriched in "pentose and glucuronate interconversions", "glycine, serine and threonine metabolism", and "starch and sucrose metabolism", respectively. Meanwhile, 75 identified pathways were downregulated, including "pathways in cancer", "endocytosis", "MAPK signaling pathway", "insulin signaling pathway", "chemokine signaling pathway", "influenza A", "regulation of actin cytoskeleton", "tight junction", "vascular smooth muscle contraction", and "ErbB signaling pathway"; 39, 34, and 33 downregulated DEGs were involved in "pathways in cancer", "endocytosis" and "MAPK signaling pathway", respectively.

DEGs and pathways identified at 48 h

At 48 h after 20E treatment, 5632 DEGs were identified; 1209 were upregulated, and 4423 were down-

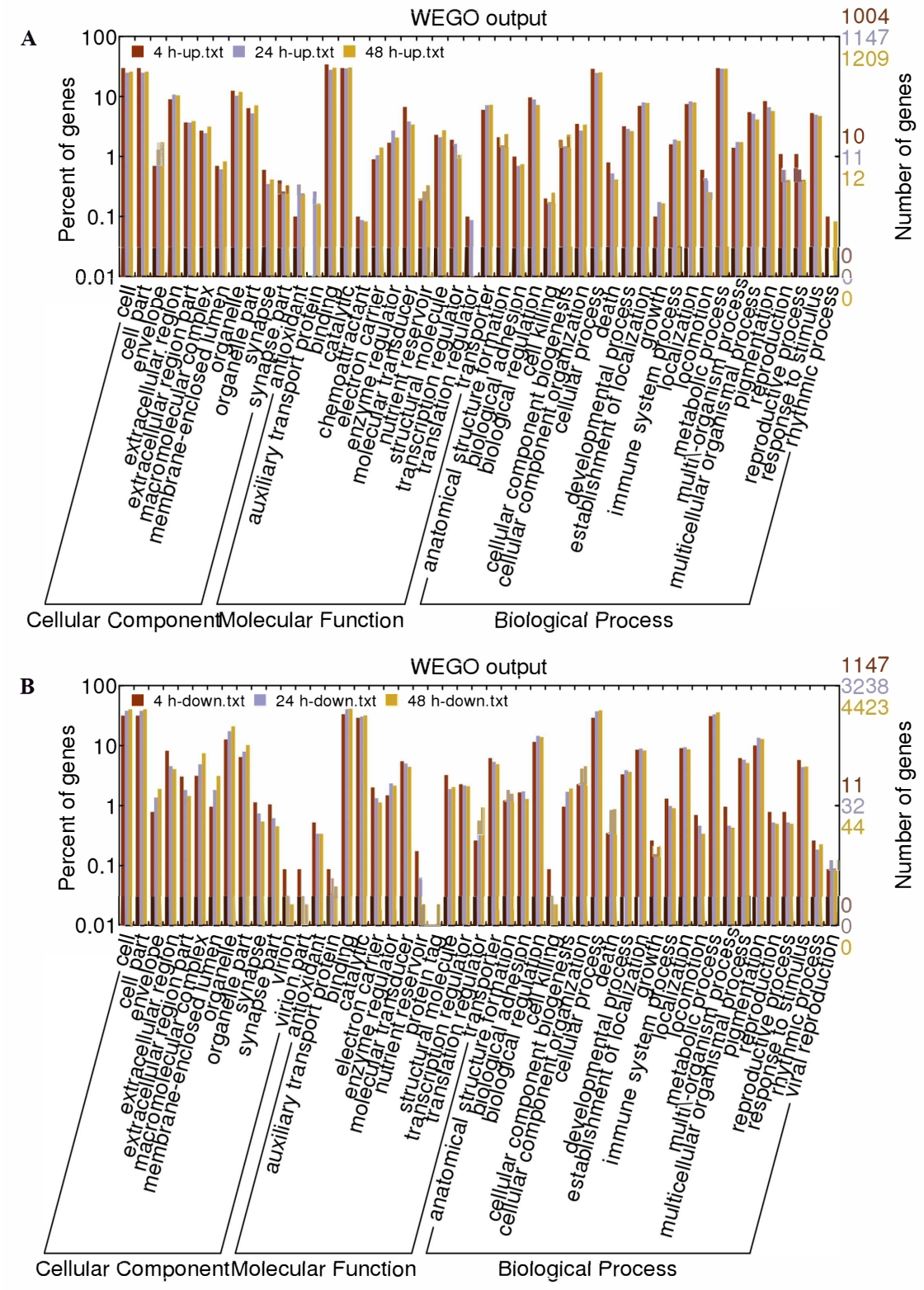

Fig. 3 Histogram presentation of Gene Ontology classification of DEGs (A) up-regulated genes; (B) down-regulated genes. The results are summarized in three main categories: cellular component, molecular function and biological process. The left y-axis indicates the percentage of a specific category of genes. The right y-axis indicates the number of genes in a category.

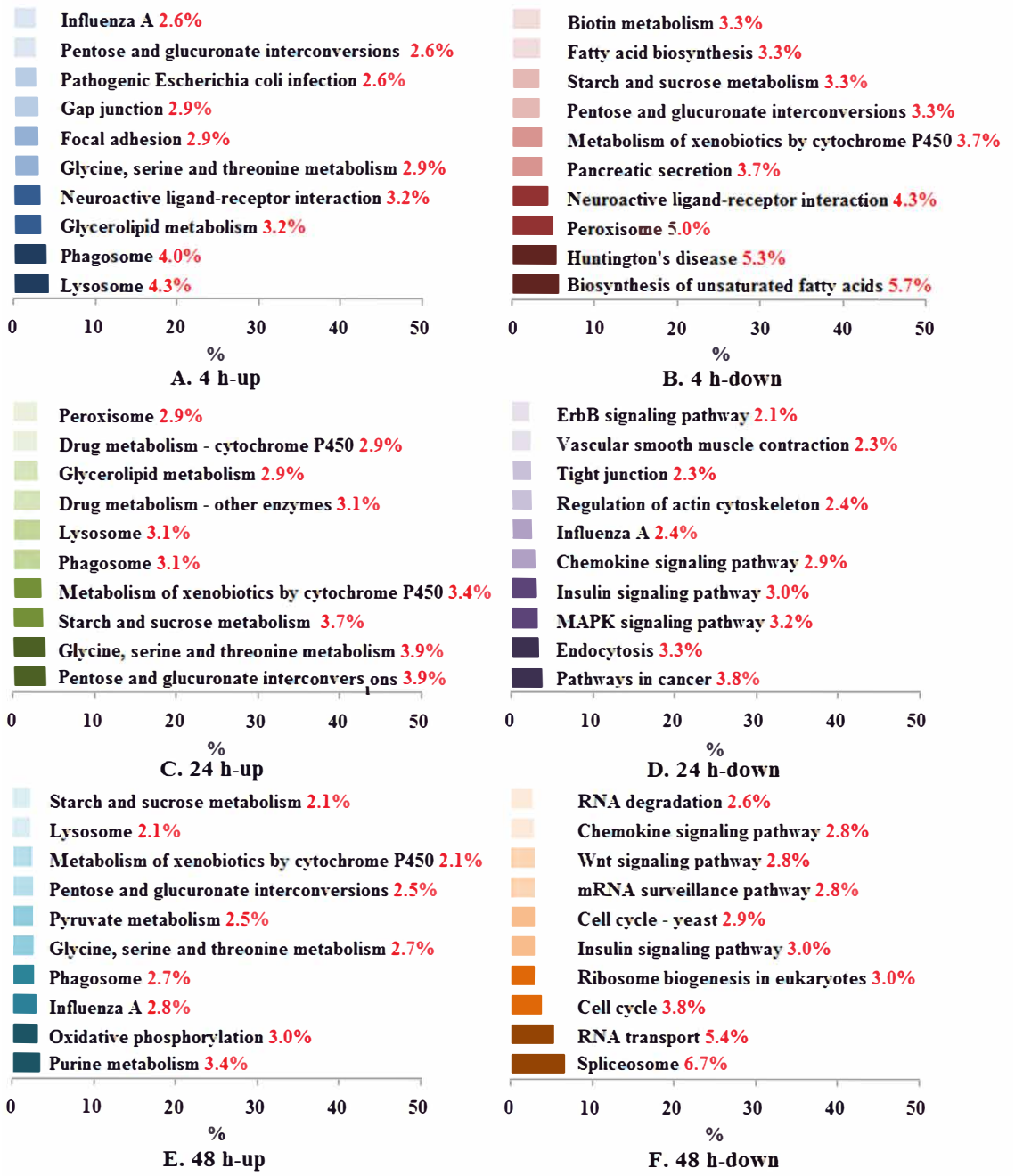

Fig. 4 KEGG analysis for differential expressed genes. p-value < 0.05 and FDR < 0.05 were used as a threshold to select significant KEGG pathways. A,B: KEGG analysis for up-/down-regulated DEGs of 4 h; C,D: KEGG analysis for up-/down-regulated DEGs of 24 h; E,F: KEGG analysis for up-/down-regulated DEGs of 48 h.

regulated. Further, 135 pathways were identified at 48 h, of which 76, including "purine metabolism", "oxidative phosphorylation", "influenza A", "phagosome", "glycine, serine and threonine metabolism", "pyruvate metabolism", "pentose and glucuronate interconversions", "metabolism of xenobiotics by cytochrome P450", "lysosome", and "starch and sucrose metabolism" were upregulated; 19, 17, and 16 upregulated DEGs were enriched in "purine metabolism", "oxidative phosphorylation" and "influenza A", respectively. Besides, 59 identified pathways were downregulated at 48 h, including "spliceosome", "RNA

transport", "cell cycle", "ribosome biogenesis in eukaryotes", "insulin signaling pathway", "cell cycle-yeast", "mRNA surveillance pathway", "wnt signaling pathway", "chemokine signaling pathway", and "RNA degradation"; 83, 67, and 48 downregulated DEGs were involved in "splicesome", "RNA transport", and "cell cycle", respectively.

Validation of DEGs by quantitative real-time PCR

To validate the DEGs identified using Solexa Sequencing, a time-course analysis of the gene expression was performed by qRT-PCR. Twelve random DEGs were selected for the qRT-PCR, and their transcription levels are shown in Fig. 5. The qRT-PCR data of these genes were consistent with the sequencing result. Both analyses revealed that the genes encoding lysozyme, sex-specific storage-protein 2, and serine protease inhibitor 14 were upregulated at 4 h, 24 h, and 48 h after 20E treatment. Likewise, the genes encoding cyclin-dependent kinase 12, receptor-type guanylyl cyclase, cuticular protein RR-2 motif 127, and cytochrome P450 18a1 were significantly suppressed at all time points. Meanwhile, the expression levels of the genes encoding glycolipid transfer protein, DNA polymerase accessory subunit, ethanolamine kinase-like, and translation machinery-associated protein 16 were upregulated in the PSG at 4 h and 24 h and downregulated at 48 h.

DISCUSSION

The present study systematically analyzed the gene expression profile and regulation in 20E-treated silkworm PSG for the first time. The genome-wide transcriptome profiles provided deep insights into the biological pathways and molecular mechanisms [26, 27]. This study employed RNA-seq, a cost-effective and robust technology, to analyze the transcriptome changes in silkworm PSG after 20E treatment. GO and KEGG analyses were used to classify the functions of DEGs and identify the 20E regulatory pathway.

The PSG analysis revealed that many 20E-induced genes were found mainly distributed in "purine metabolism", "oxidative phosphorylation", "starch and sucrose metabolism", "glycerolipid metabolism", and "pentose and glucuronate interconversions" pathways. Many genes involved in energy metabolism, including genes in the "starch and sucrose metabolism", "glycerolipid metabolism", "oxidative phosphorylation", and "pentose and glucuronate interconversions" pathways, were upregulated by 20E, implying the importance of these genes to meet the energy requirement during the physiological and developmental processes. The top KEGG pathway enriched mainly by the DEGs at 48 h was "purine metabolism", which probably provides the components of the nucleotides, DNA, and RNA and the energy currency of the cell, ATP [28]. Typically, hormones, including ecdysones, are required for the formation of the *Bombyx* cuticle. The expression of ecdysis-related genes suggests that the PSG surface layer is renewed along with the shedding of epidermal cells, which are under the hormones' control. Meanwhile, the cuticular proteins are rich in glycine, alanine, and serine, the primary fibroin synthesis materials. Many genes upregulated after 20E treatment were enriched in the "glycine, serine and threonine metabolism" pathway, implying the importance of these genes in providing amino acids to meet the physiological and developmental requirements. However, whether the shed cuticle is recycled as an amino acid source for fibroin synthesis remains unclear.

Many genes downregulated after 20E treatment were mainly distributed in the "spliceosome", "RNA transport", "cell cycle", and "ribosome biogenesis in eukaryotes" pathways. The spliceosome is complex molecular machinery involved in the removal of introns from mRNA precursors [29]. Meanwhile, the transport of RNA from the nucleus to the cytoplasm is fundamental for gene expression and regulation. Most eukaryotic RNAs are produced in the nucleus by RNA polymerase I, II, or III. The RNA molecules undergo various post-transcriptional processing events, and many are localized in their functional sites in the cell [30]. In the present study, the ribosome biogenesis for mRNA translation was downregulated in the 20E-treated PSG, which indicates that 20E transcriptionally inactivates cell cycle genes for increasing DNA content and ribosome biogenesis for increasing protein content [31].

Cytochrome P450 is a complex and ubiquitous superfamily of heme-containing enzymes that participate in the metabolism of both exogenous and endogenous substrates [32]. A total of 84 P450 genes were identified, of which 78 were functional in the silkworm. Most P450s of the silkworm, tandemly arranged on chromosomes, were upregulated in the fat body after exposure to insecticides [33]. In this study, P450 was upregulated at 24 h and 48 h but downregulated at 4 h after feeding 20E, which indicates the response of P450 genes to 20E.

Furthermore, studies have demonstrated the hormonal regulation of silk gland function and development [9]. Comparative transcriptome analysis of the seven segments of the silk gland of the 5th instar larvae revealed that the genes highly expressed in the PSG were mainly involved in the ribosome, proteasome, citrate cycle, and glycolysis/gluconeogenesis pathways, consistent with the present study [34]. Besides, 20E regulated the transcription of the L-chain and P25 genes of fibroin, and the effect was dose-dependent [35]. The present study revealed that the transcription levels of L-chain and P25 were significantly upregulated by 20E at a concentration of 2×10^{-3} μg/μl. Several protein factors binding the tran-

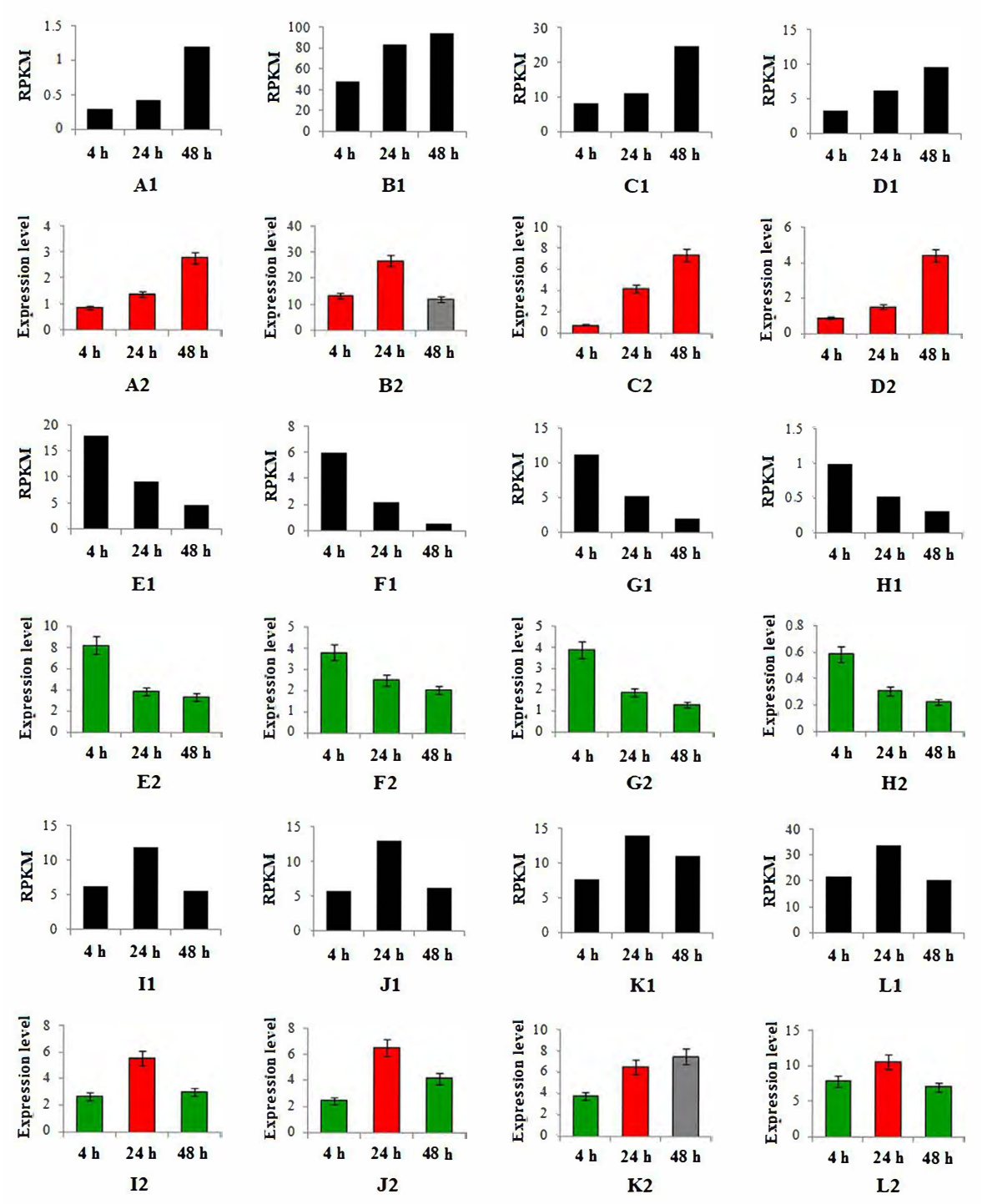

Fig. 5 Verification of transcriptomics results of DEGs by qPCR. The transcriptional results of DEGs (RPKM) are marked with black. The qPCR data of DEGs up-regulated and down-regulated by 20E are marked with red and green. The mismatched DEGs are marked with gray. (A1,A2) lysozyme; (B1,B2) glutathione S-transferase omega 2; (C1,C2) sex-specific storage-protein 2; (D1,D2) serine protease inhibitor 14; (E1,E2) cyclin dependent kinase 12; (F1,F2) receptor type guanylyl cyclase; (G1,G2) cuticular protein RR-2 motif 127; (H1,H2) cytochrome P450 18a1; (I1,I2) glycolipid transfer protein; (J1,J2) DNA polymerase accessory subunit; (K1,K2) ethanolamine kinase-like; (L1,L2) translation machinery-associated protein 16. Asterisks denote significant differences between treatment group and negative control (NC) group, as determined using pairwise *t*-tests (* $p \leqslant 0.05$, ** $p \leqslant 0.01$).

scriptional regulatory elements of silk genes have been identified and characterized in 20E-treated silkworms, including Fkh, FMBP1, POUM1, BMFA, SGF-1, and SGFB [36, 37]. The differences in gene response to 20E indicate their different roles in the transcriptional regulation of silk genes.

CONCLUSION

In summary, the global transcriptome profiling of the 20E-treated PSG provided a comprehensive view of the molecular basis of PSG development and gene regulation by 20E at the transcriptional level for the first time in this study. At 4 h, 24 h, and 48 h after 20E treatment, 2151, 4385, and 5632 genes were differentially expressed. More downregulated transcripts than upregulated transcripts were identified at all time points. Further, GO annotation and KEGG analysis were performed to classify the functions of DEGs and identify the pathways involved in the 20E regulation. Furthermore, few selected DEGs were verified by qRT-PCR.

Appendix A: Supplementary data

Supplementary data associated with this article can be found at http://dx.doi.org/10.2306/scienceasia1513-1874.2022.025.

Acknowledgements: This work was supported by the Natural Science Foundation of Jiangsu Province (BK20201229), Youth Science and Technology Innovation Project of Jiangsu University of Science and Technology and Open Project of Key Laboratory of Silkworm and Mulberry Genetic Improvement (KL201901).

REFERENCES

1. Xia QY, Wang J, Zhou ZY, Li RQ, Fan W, Cheng DJ, Cheng TC, Qin JJ, et al (2009) The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. *Insect Biochem Mol Biol* **38**, 1036–1045.
2. Zhu L, Fan ZQ, Shi XQ, Wang N, Bo YY, Guo HE (2020) A novel silkworm pupae carboxymethyl chitosan inhibits mouse L929 fibroblast proliferation. *ScienceAsia* **46**, 30–36.
3. Goldsmith MR, Shimada T, Abe H (2005) The genetics and genomics of the silkworm, *Bombyx mori*. *Annu Rev Entomol* **50**, 71–100.
4. Tomita M, Munetsuna H, Sato T, Adachi T, Hino R (2003) Transgenic silkworms produce recombinant human type III procollagen in cocoons. *Nat Biotechnol* **21**, 52–56.
5. Masashi I, Shingo O, Atsushi T, Shinichi N, Yuhki K, Yoshitaka M, Jun H, Masahiro T (2009) Production of a recombinant mouse monoclonal antibody in transgenic silkworm cocoons. *FEBS J* **276**, 5806–5820.
6. Matsunami K, Kokubo H, Ohno K, Xu PX, Ueno K, Suzuki Y (1999) Embryonic silk gland development in *Bombyx*: molecular cloning and expression of the *Bombyx* trachealess gene. *Dev Genes Evol* **209**, 507–514.
7. Suzuki Y, Brown DD (1972) Isolation and Identification of the messenger RNA for silk fibroin from *Bombyx mori*. *J Mol Biol* **63**, 409–429.
8. Couble P, Moine A, Garel A, Prudhomme JC (1983) Developmental variations of a nonfibroin mRNA of *Bombyx mori* silk gland, encoding for a low-molecular-weight silk protein. *Dev Biol* **97**, 398–407.
9. Sehnal F, Akai H (1990) Insect silk glands: their types, development and function, and effects of environmental factors and morphogenetic hormones on them. *Int J Insect Morphol Embryol* **19**, 79–132.
10. Gilbert LI, Combest WL, Smith WA, Meller VH, Rountree DB (1988) Neuropeptides, second messengers and insect molting. *BioEssays* **8**, 153–157.
11. Gilbert LI, Bollenbacher WE, Granger NA (1980) Insect endocrinology: regulation of endocrine glands, hormone titer, and hormone metabolism. *Annu Rev Physiol* **42**, 493–510.
12. Riddiford LM, Cherbas P, Truman JW (2000) Ecdysone receptors and their biological actions. *Vitam Horm* **60**, 1–73.
13. Dai YJ (1985) Moulting hormone and silk protein synthesis. Effect of ecdysone on nucleic acid synthesis in the posterior portion of silk gland in *Bombyx mori*. *Acta Entomol Sin* **28**, 8–14.
14. Shigematsu H, Moriyama H (1970) Effect of ecdysterone on fibroin synthesis in the posterior division of the silk gland of the silkworm, *Bombyx mori*. *J Insect Physiol* **16**, 2015–2022.
15. Miao YG, Shi LG, Nair KS (2010) Ecdysteroid as a mediator in the regulation of silk protein synthesis and its influence on silkworm (*Bom., Lepidoptera*) genome. *J Appl Entomol* **128**, 348–353.
16. Xia QY, Cheng DJ, Duan J, Wang GH, Cheng TC, Zha XF, Liu C, Zhao P, et al (2007) Microarray-based gene expression profiles in multiple tissues of the domesticated silkworm, *Bombyx mori*. *Genome Biol* **3**, 19–31.
17. Zhou ZH, Yang HJ, Chen M, Lou CF, Zhang YZ, Chen KP, Wang Y, Yu ML, et al (2008) Comparative proteomic analysis between the domesticated silkworm (*Bombyx mori*) reared on fresh mulberry leaves and on artificial diet. *J Proteome Res* **7**, 5103–5111.
18. Lu AC, Wei ZG, Li B, Shen WD (2009) Induced expression of cytochrome P450 CYP305 B1 V1 gene in different tissues of wild mulberry silkworm (*Bombyx mandarina*). *Agric Sci Technol* **10**, 54–58.
19. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Baren MJV, Salzberg SL, Wold BJ, et al (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**, 511–515.
20. Wang L (2010) DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26**, 136–138.
21. Ye J, Fang L, Zheng HK, Zhang Y, Chen J, Zhang ZJ, Wang J, Li ST, et al (2006) WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res* **34**, 293–297.
22. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **27**, 29–34.
23. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**, 402–408.
24. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111.

25. Duan J, Li RQ, Cheng DJ, Fan W, Zha XF, Cheng TC, Wu YQ, Wang J, et al (2010) SilkDB v2.0: a platform for silkworm (*Bombyx mori*) genome biology. *Nucleic Acids Res* **38**, 453–456.
26. Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10**, 57–63.
27. Marguerat S, Bähler J (2010) RNA-seq: from technology to biology. *Cell Mol Life Sci* **67**, 569–579.
28. Di VF (2012) Purines, purinergic receptors, and cancer. *Cancer Res* **72**, 5441–5447.
29. Bonnal S, Vigevani L, Valcárcel J (2012) The spliceosome as a target of novel antitumour drugs. *Nat Rev Drug Discovery* **11**, 847–859.
30. Lécuyer E, Yoshida H, Parthasarathy N, Alm C, Babak T, Cerovina T, Hughes TR, Tomancak P, et al (2007) Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131**, 174–187.
31. Ma L, Xu H, Zhu JQ, Ma SY, Liu Y, Jiang RJ, Xia QY, Li S (2011) Ras1CA overexpression in the posterior silk gland improves silk yield. *Cell Res* **21**, 934–943.
32. Feyereisen R (2005) Insect Cytochrome P450. *Compr Mol Insect Sci* **5**, 1–77.
33. Yamamoto K, Ichinose H, Aso Y, Fujii H (2010) Expression analysis of cytochrome P450s in the silkmoth, *Bombyx mori*. *Pestic Biochem Physiol* **97**, 1–6.
34. Shi R, Ma SY, He T, Peng J, Zhang T, Chen XX, Wang XG, Chang JS, et al (2019) Deep insight into the transcriptome of the single silk gland of *Bombyx mori*. *Int J Mol Sci* **20**, ID 2491.
35. Chaitanya RK, Duttagupta A (2010) Chaitanya RK, Dutta-Gupta A. Light chain fibroin and P25 genes of Corcyra cephalonica: Molecular cloning, characterization, tissue-specific expression, synchronous developmental and 20hydroxyecdysone regulation during the last instar larval development. *Gen Comp Endocrinol* **167**, 113–121.
36. Durand B, Drevet J, Couble P (1992) P25 gene regulation in *Bombyx mori* silk gland: two promoter-binding factors have distinct tissue and developmental specificities. *Mol Cell Biol* **12**, 5768–5777.
37. Horard B, Julien E, Nony P, Garel A, Couble P (1997) Differential binding of the *Bombyx* silk gland-specific factor SGFB to its target DNA sequence drives posterior-cell-restricted expression. *Mol Cell Biol* **17**, 1572–1579.

Appendix A. Supplementary data

Table S1 Primers used in this study.

Target gene	Accession number	Primer Sequence (5′–3′)
Lysozyme	NM_001043983.1	AGGAACTGGGTATGTCTGG
		TCGCTGCCTTAGTAATGTC
glutathione S-transferase omega 2 (GSTo2)	NM_001043941.1	TCTCCGACACTGTCAATGAGGA
		TCAGCGTATCCAGGTTCACTTC
sex-specific storage-protein 2 (Sp2)	NM_001044125.1	ATGACAAGATGAGGGATG
		GGATAACAGCGATGTAGAA
serine protease inhibitor 14 (serpin-14)	NM_001146234.1	CAACATTACTTATTTCCCTC
		CGTCAACCTTGTCGCTTC
cyclin dependent kinase 12 (CDK12)	NM_001258363.1	AAGGCTCATTCTATTTGGTC
		TTCCGATGACAGTACGCTA
receptor type guanylyl cyclase (Gcy)	NM_001043405.1	TGCTGTGCTGACCGTGAC
		CAGGAGGACCGCATCGTA
cuticular protein RR-2 motif 127 (CPR127)	NM_001173172.1	GGCGTCGCTGACCCTAAT
		CAGCATTGAAGCCGTGGA
cytochrome P450 18a1 (Cyp18a1)	NM_001083609.2	TTGGAAATGGCTGAAGGTG
		GCCGACATGACGAAGATGAG
glycolipid transfer protein (LOC692943)	NM_001046783.1	TAATAATGGTAAAGCTCCTG
		GACCCAACTGTGATACTTCT
DNA polymerase accessory subunit (LOC692827)	NM_001046672.1	TAGGAAATACTGGTGGAG
		ACTGAGGCGATAATGGTT
ethanolamine kinase-like (LOC101745087)	XM_004927829.1	GAGTAGCTTTGGGTCAGT
		TGGTCCCTTCATCTTTGT
translation machinery-associated protein 16	XM_004928943.1	AACTAGAACAGATCGCATTA
		GTTCATTTGTTGTGGGTC

Table S2 Raw data preprocessed results.

Sample	Raw data		Valid data			Valid ratio
	Read	Base	Read	Base	Average length	(reads)
NC	63101320	6373233320	60161304	5917618167	98.36	95.34%
4 h	56179596	5674139196	53347028	5203708238	97.54	94.96%
24 h	58905136	5949418736	56129632	5481791612	97.66	95.29%
48 h	54413422	5495755622	52095740	5072287797	97.36	95.74%
All	232599474	23492546874	221733704	21675405814	97.75	95.33%

Table S3 Evaluation of valid reads mapped to reference genome of each sample.

	Valid data	Map data	Data%	All gene	Expression gene	Gene%
NC	60161304	49253916	81.87%	13340	10330	77.44%
4 h	53347028	42437547	79.55%	13340	10235	76.72%
24 h	56129632	44600496	79.46%	13340	10206	76.51%
48 h	52095740	41654489	79.96%	13340	10172	76.25%
all	221733704	177946448	80.25%	13340	11412	85.55%