ISJ 15: 66-82, 2018 ISSN 1824-307X

RESEARCH REPORT

Comparative proteomic analysis of the silkworm (*Bombyx mori* L.) silk gland reveals yield heterosis

Y Zhang, H Tang, Y Yang*, P Lü, Q Yao, K Chen*

Institute of Life Sciences, Jiangsu University, Zhenjiang 212013, China

Accepted February 26, 2018

Abstract

Heterosis is a prevalent phenomenon in nature and is widely found in animals and plants, but the molecular mechanism of heterosis is still unclear. As the model insect of Lepidoptera, *Bombyx mori* is an ideal material to carry out heterosis research. In current study, we employed proteomic and genetic cross approaches to globally identify differentially expressed proteins in parental silkworms JingSong and HaoYue and their F_1 hybrids. The results showed that there were significant differences between hybrids and their parents. In all, 28 differentially expressed proteins were successfully identified through MALDI-TOF-TOF mass spectrometry and database searches. Interestingly, three silk-related proteins were also obtained and identified, including fibroin L-chain, Nd-sD mutant fibroin light chain and fibrohexamerin, which were further confirmed by quantitative real-time polymerase chain reaction (qRT-PCR). Consistent trends were also found in other genes. Taken together, our work not only provides the theoretical basis for the study of molecular mechanism of heterosis, but also provides candidate proteins and genes for the improvement of yield of silkworm.

Key Words: Heterosis, *Bombyx mori*, Proteomics, qRT-PCR

Introduction

Heterosis was firstly proposed by Shull in 1914, referring to a hybrid offspring with better productivity, growth, development, and resistance compared with its parents (Shull, 1948). The theoretical hypothesis about heterosis mainly includes “dominance” model, “overdominance” model and epistasis theory. The dominant hypothesis emphasizes the interactions between dominant genes and recessive genes, suggesting that favorable dominant genes can cover up the effects of unfavorable recessive genes, and lead to the improvement of F_1 vigor. The overdominance model indicates that the heterozygous allele plays an important role in the formation of heterosis. The epistasis theory emphasizes nonadditive genetic effects based on the overdominance hypothesis and believes that non-alleles also play roles in heterosis. In recent years, many studies on heterosis have been carried out (Kawamura *et al*., 2016; Kwan *et al*., 2016; Li *et al*., 2016a; Li *et al*., 2016b; Wu *et al*., 2016). The results from transcriptome showed that overdominance, dominance, low dominance and additive effects coexisted in *Bombyx mori* heterosis, among which overdominance was the primary mode of gene action (Wang *et al*., 2015). Recent researches have also suggested that additional mechanisms may explain heterosis in *Arabidopsis* and maize, including expression of small RNA (sRNA) and DNA methylation, respectively (Hou *et al*., 2017; Lauria *et al*., 2017). To verify these mechanisms, transcriptional or gene expression profiles have been completed in maize, rice, pufferfish, and silkworm. Li *et al*. (2016a) used the transcriptomic method to analyze gene expression patterns and genetic variations between parents’ rubber trees and their productive hybrids, including single nucleotide polymorphisms (SNPs) and small insertions/deletions (InDels) (Li *et al*., 2016a). The results showed that the higher yield of F_1 hybrids was positively correlated with their higher

Corresponding author:
Yanhua Yang
Keping Chen
Institute of Life Sciences
Jiangsu University
301 Xuefu Road 212013, Zhenjiang, PR China
E-mail: yanhuayang@126.com

List of abbreviations:
MALDI-TOF MS, matrix-assisted laser desorption/ionization time of flight mass spectrometry; PMSF, phenylmethanesulfonyl fluoride; EGTA, ethylene glycol tetraacetic acid; DTT, 1,4-Dithio-DL-threitol; ACN, acetonitrile; SDS-PAGE, sodium dodecyl sulphate polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid

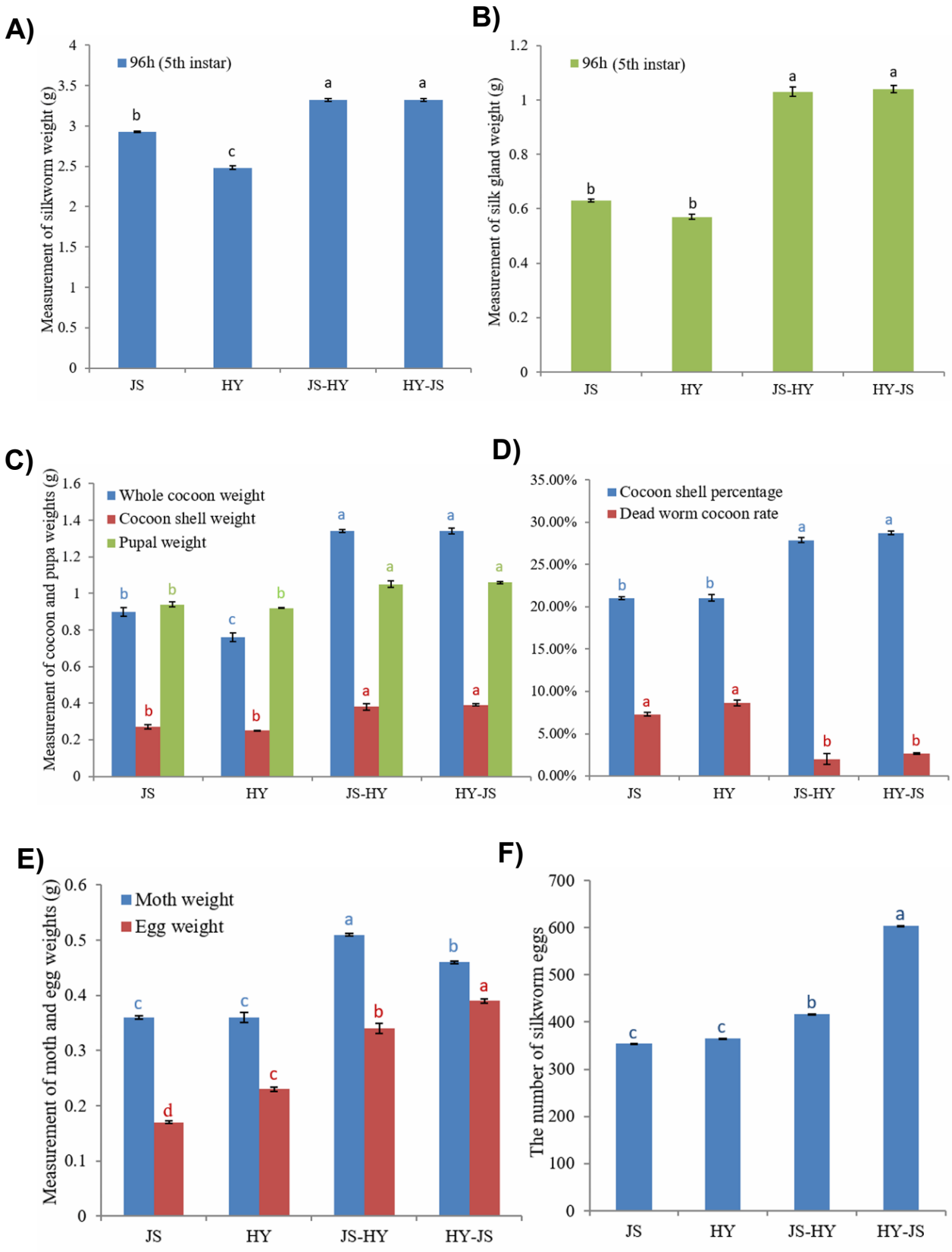


Fig.1 Results of character investigation from parental silkworms JS, HY and their F_1 hybrids. (A, B): Weight measurement of silkworm larva and silk gland at 96h of the fifth instar; (C): Measurement of whole cocoon weight, cocoon shell weight and pupal weight; (D): cocoon shell percentage and dead worm cocoon rate; (E): Weight measurement of moth and egg; (F): the number of silkworm eggs.

genomic heterozygosity, and the genome-wide genetic variations played a vital role in maintaining the heterosis of rubber tree. The research from maize illustrated that cell expansion played a major role in the heterosis of ABA (abscisic acid)-mediated maize seed germination. The up-regulation of inactivation gene *ZmABA8ox1b* also contributed to seed germination by promoting cell expansion (Li *et al*., 2016a). Zhu *et al*. suggested that cumulative small advantages of the single-locus effects and two-locus interactions could explain the genetic basis of seedling heterosis in F_1 hybrid (Zhu *et al*., 2016). Until now, however, there are no consistent conclusions in the investigation of molecular mechanism of heterosis.

As an important economic insect and model insect of Lepidoptera, the *Bombyx mori* (L.) is an ideal material for molecular biological research because of its some advantages, such as clear genome and genetic background, high breeding coefficient, and easy control of growth environment and so on. The silk gland of *B. mori*, including three functionally distinct regions: anterior silk gland (ASG), middle silk gland (MSG), and posterior silk gland (PSG), is a special organ for synthesis and secretion of the silk fibroin and the sericin. As the matter of fact, the silk gland of silkworm is not developed before the 4th instar, which grows rapidly at the 5th instar and will become the largest organ of silkworm at the late 5th instar (Lv, 2008). The silk from silkworm is a very good and extensively utilized textile material, whose yield and quality mainly depends on silk fibroin and sericin synthesized and secreted by the silk gland. Therefore, in recent years, researchers have paid more and more attentions on the silk gland of silkworm (Dhawan *et al*., 2003; Ma *et al*., 2011; Miller *et al*., 2015; Chen *et al*., 2017; Xue *et al.*, 2017). Xue *et al*. used DGE-seq (digital gene expression sequencing) to analyze gene expression profile of silk gland and examine the effect of TiO_2 nanoparticles on the silk protein synthesis of silkworm (Xue *et al*., 2017). The results showed that the addition of TiO_2 could promote the synthesis of silk protein. Chen *et al*. found that *BmDredd*, a homologue of the starting apoptotic protease, played a key role in the apoptosis of silk gland (Chen *et al*., 2017). Ma *et al*. investigated the effect of Ras1 oncogene overexpression on the production of silk protein. The result showed that the activation of Ras1 could increase the size of silk gland cells, promote protein synthesis of silk gland, activate mRNA translational ribosome biogenesis, and eventually result in the increase of silk yield (Ma *et al*., 2011). However, these studies mainly focused on the genome, transcriptome and proteome of silkworm silk glands, the heterosis research at the protein level are still few. Currently, the proteomic methods, including two-dimensional electrophoresis and mass spectrometry, have successfully been employed to identify the specific proteins from different tissues and organs of silkworm (Moghaddam *et al*., 2008; Hou *et al*., 2010; Li *et al*., 2010; Liu *et al*., 2010; Qin *et al*., 2012; Zhong *et al*., 2013; Feng *et al*., 2014; Kannan *et al*., 2016; Li *et al*., 2016a). The results from proteome provided some helpful evidences and clues for understanding the developmental and physiological processes of silkworm. In this study, we employed proteomic and genetic cross approaches to globally identify differentially expressed proteins of silk glands between parental silkworms JingSong, HaoYue and their F_1 hybrids. Our objectives were 1) to detect, identify and classify the different proteins between parental silkworms and their F_1 hybrids, 2) to find out the heterosis-related proteins and genes and attempt to decipher the molecular mechanism of silkworm heterosis.

Materials and methods

Experimental materials

Parental silkworms JingSong (JS, Chinese line) and HaoYue (HY, Japanese line), and their direct cross (JS-HY, D-F_1), reciprocal cross (HY-JS, R-F_1) hybrids were used for this experiment, which were fed in the Institute of Life Sciences, Jiangsu University, Zhenjiang, China. The larvae of 1st~3rd instar were fed with the chopped tender leaves and 4th~5th instar larvae were fed with the matured leaves. The larvae were dissected to obtain silk glands at 96h of 5th instar, immediately frozen in liquid nitrogen and stored at −70 °C refrigerator for later use (at least 50 silkworms of each sample). The experiments were repeated three times.

Protein sample preparation

The silk glands were quickly grounded into fine powder in liquid nitrogen and homogenized on ice for 5 min in pre-cooled extraction buffer (20 mM Tris-HCl pH 7.5, 250 mM sucrose, 10 mM EGTA, 1 mM PMSF, 1 mM DTT, and 1% Triton (v/v) X-100). Then the proteins were extracted by phenol method, as described by Cilia *et al* (Cilia *et al*., 2009). Finally, the protein samples were vacuum-dried. The dry powder was dissolved with lysis buffer (7 M urea, 2 M thiourea, 4% *w/v* CHAPS, 1% *w/v* DTT and 2% *v/v* Biolyte, pH 3-10), and overnight at 4 °C. The mixtures were centrifuged at 12,000g for 20 min, 4 °C. The protein concentration was measured by Bradford method (Bradford, 1976).

Two-dimensional electrophoresis (2-DE)

The isoelectric focusing (IEF) was carried out through ETTAN IPGphor 3 electrophoresis system and 17 cm immobilized pH gradient (IPG) dry gel strips with pH 4-7 liner range (Bio-Rad, USA) was used. About 2 mg protein samples were loaded (passive rehydration, room temperature, 11~12 h). Then IEF was performed at 20 °C with a voltage gradient of 100 V for 1 h, 300 V for 1 h, 500 V for 1 h, 1000 V for 1 h (linear), next 10000 V, linear for 5 h, and then remained 10000 V until a total voltage of 40000 Vh. Subsequently, the gel strips were equilibrated for 15 min in equilibration buffer (50 mM Tris-HCl pH6.8, 2.5% *w/v* SDS, 30% *v/v* glycerol and 1% *w/v* DTT) and then equilibrated for 15 min again with the same equilibration buffer without DTT. The second dimension SDS-PAGE was performed with a Laemmli buffer system (12% resolving gels) (Laemmli *et al*., 1970). The electrophoresis was divided into two steps, that is, 5 w/gel, 50min, then 15 w/gel, 4~5h. At last, the gels were dyed with 0.116% Coomassie Brilliant blue (CBB) R-250 and scanned with a high precision scanner (ImageScanner III, GE Healthcare Life Sciences) at a resolution of 300 dpi.

Image analysis

Image analysis was carried out using ImageMaster 7.0 software (GE Healthcare Life Sciences). The optimized parameters were set according to the gel background. The corresponding spots in different gels was normalized and calculated. The protein spots with ratios ≥ 1.5 or ratios ≤ 0.67 and ANOVA ($P < 0.05$) were considered to be different proteins.

In-gel digestion and MALDI-TOF-TOF analysis

The protein spots were excised from the stained gels, washed three times with milli-Q water, destained by sonication in 50 mM NH_4HCO_3, 50% ACN, and 100% CAN (50 μL/each), respectively, then dried in room temperature. The dried protein spots were incubated with 10~15 μL trypsin solution (10 μg/mL) at 37 °C for about 12 h. Peptides were collected and 50% ACN containing 0.1% TFA 20 μL were added to each spot to incubate at 60 °C for 10 min, then treated by sonication for 1 min, and dried in centrifugal concentrator. The dried sample was dissolved using 3 μL of α-cyano-4-hydroxycinnamic acid (CHCA) matrix. Finally, the mixtures were analyzed through MALDI TOF/TOF Mass Spectrometer (MALDI-TOF/TOF™ 5800 System, AB SCIEX).

Protein identification

MALDI-TOF/TOF data were analyzed using MASCOT (Matrix Science, London, UK, http://www.matrixscience.com/). Using the NCBI and SwissProt databases, the primary and secondary mass spectra of the corresponding protein spots were obtained. The parameters were set as follows: Metazoa (Animals) was selected as the taxonomic category, trypsin was selected as enzyme, carbamidomethyl was chosen as fixed modification, Deamdated (NQ) and Oxidation (HW) were selected as variable modification, the peptide tolerance is 100 ppm, and MS/MS tolerance is 0.6 Da. Totally, 28 proteins with significant differences were successfully identified.

Gene Ontology (GO) and KEGG analysis

According to the annotations of the identified proteins from UniProt knowledgebase (http://www.expasy.org/sprot/) (Ye *et al*., 1970), the corresponding Gene Ontology (GO) IDs of these proteins were obtained. Based on Gene Ontology Database, the GO classification of these proteins was conducted with WEGO (http://wego.genomics.org.cn/). The KEGG results of the corresponding proteins were obtained from the NCBI database. The KEGG maps were summarized according to the KEGG network (http://www.genome.jp/kegg/pathway.html) to obtain the relevant pathways of the different proteins.

RNA extraction and quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR)

The silk glands were thawed, ground into powder, weighed (0.1 g) and transferred into the centrifuge tube. Using Trizol reagent (Invitrogen, USA), total RNA was extracted. The extracted RNA was assayed with 1% agarose. At last, 1 μg RNA was used for the first strand synthesis. Reverse transcription was performed using a reverse transcription kit. The qRT-PCR was carried out in a total volume of 20 μL containing 2 μL of cDNA (200 ng), 10 μL of SYBR Green Master Mix (Vazyme Biotech Co., Ltd, China), 0.4 μL of 50×ROX Reference Dye I, 0.4 μL of primers (10 μM) and 7.2 μL of H_2O. Amplification was performed using an ABI7300 PCR thermocycler (Applied Biosystems, USA). The α-tubulin gene (NCBI accession No. NM_001043419.1) of silkworm was selected as a reference and the experiments were repeated three times.

Results

The economic characters of JS, HY and their F_1 hybrids

In our study, the weights of larva and silk gland, whole cocoon weight, cocoon shell weight, pupal weight, cocoon shell percentage, dead worm cocoon rate, moth weight, egg weight and the number of eggs were investigated. As shown in Figure 1, all traits of F_1 generation were significantly superior to their parents, but the differences of most characters between D-F_1 and R-F_1 generations, including the weights of larva and silk gland, whole cocoon weight, cocoon shell weight, pupal weight, cocoon shell percentage and dead worm cocoon rate were not significant. As the matter of fact, the investigated traits, especially silk gland weight and whole cocoon weight, showed obvious over-dominant effect in F_1 hybrid (Table S1). Although dead worm cocoon rate of F_1 generation appeared negative over-parent heterosis, the lower dead worm cocoon rate is actually advantageous for silkworm. Therefore, our data illustrated that the heterosis of silkworm is over-dominant. Unfortunately, as a result of the restriction of the feeding condition, the whole cocoon weight in our experiment was lower than other research laboratories. However, in current study, the experiments were biologically repeated three times and the data of the triplicates were very similar to each other. Therefore, the conclusions from these results will still remain interesting and reliable.

The detection and identification of differentially expressed proteins

In current study, the genetic cross and proteomic approaches (2-DE and MS) were employed to investigate the molecular mechanism of heterosis at protein level. In order to improve the compatibility of MS identification, we employed CBB staining for visualization of protein spots, and triplicates were performed for each sample. As shown in Figure 2, the gels with high resolution and sensitivity were obtained, which provided enough information for our study. The gels were analyzed using ImageMaster 7.0 software (GE Healthcare Life Sciences). Totally, 768 ± 12, 755 ± 16, 759 ± 13, 762 ± 22 protein spots were detected in JS, HY, JS-HY, and HY-JS, respectively, in CBB R-250 stained gels. Totally, 28 protein spots with significant differences were successfully identified, and 14, 21, 16 and 5 differentially expressed proteins were ultimately determined between JS/HY, JS/JS-HY, JS/HY-JS, and JS-HY/HY-JS, respectively (Supplementary files, Figure S1 and Table S2). These

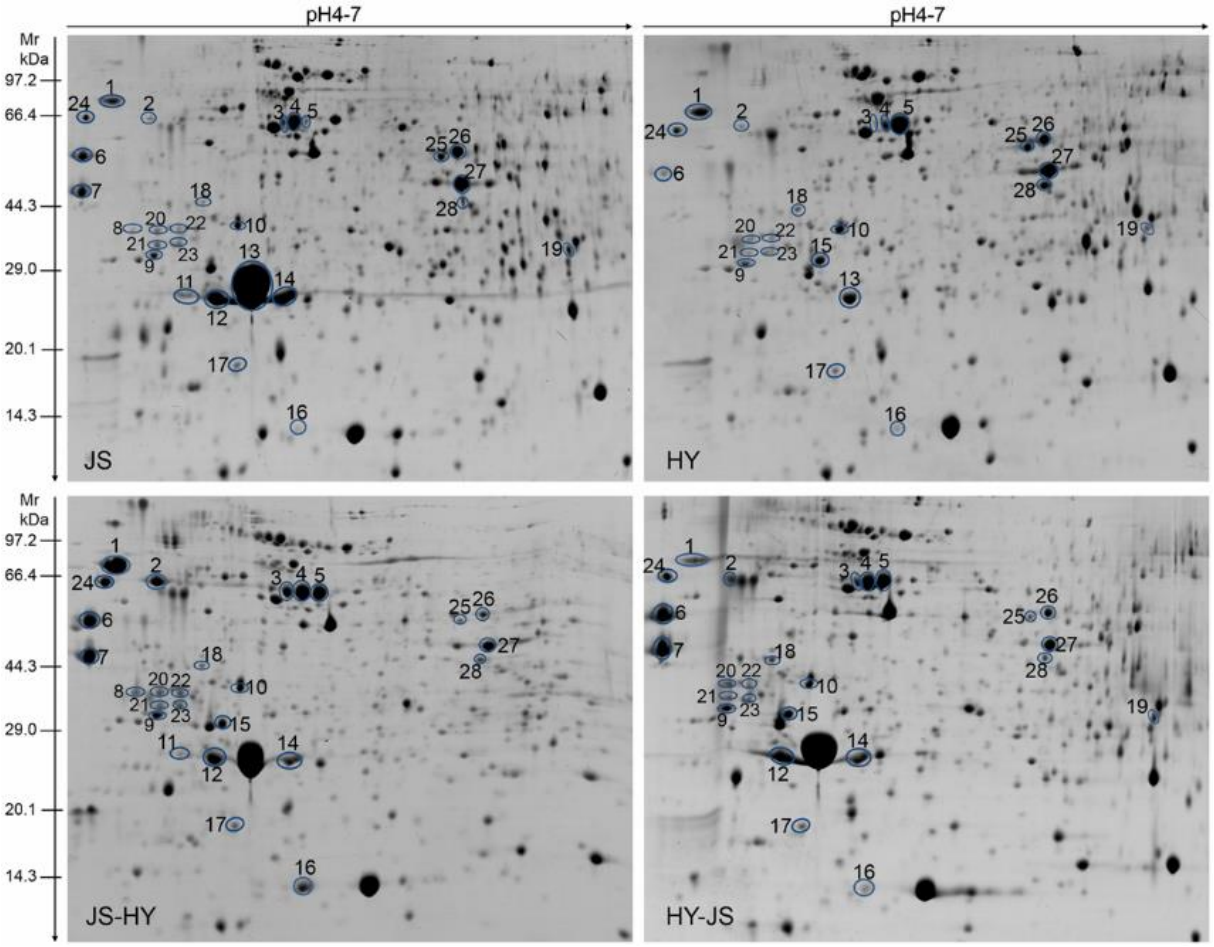


Fig. 2 The 2-DE profiles from silkworm silk gland of parental silkworms JS, HY and their F_1 hybrids. Note: The differentially expressed protein spots are indicated by circles and labeled with Arabic numerals. The numbers shown on the left indicate the protein markers in kDa.

differentially expressed proteins were numbered uniformly (Table 1) and labeled in Figure 2. As shown in Table 1, some different spots were the same protein, e.g., spots 3, 4, and 5 corresponded to protein disulfide-isomerase like protein ERp57 precursor, and spots 20, 21 and 22 were fibrohexamerin-like. Therefore, these 28 differentially expressed proteins can be classified into 22 different proteins, involving in genetic information processing, metabolism, cellular process, environmental information processing related proteins and organismal system (Table 1, Figure 3). As shown in Figure 3, most of the identified proteins participated in genetic information processing (33%), and 28% proteins belonged to the metabolism group, which accounted for 61% of the identified proteins. Interestingly, three silk-related proteins were also identified, that is, Nd-sD mutant fibroin light chain (Figure 2, spot 11), fibroin L-chain (Figure 2, spot 13) and fibrohexamerin (Fhx) (Figure 2, spots 20, 21 and 22). Because the sequence similarity of fibroin light chain precursor (spots 12 and 14, 261 amino acids) with fibroin L-chain (spot 13, 261 amino acids) reached 99.24%, protein spots 12, 13 and 14 were the same protein. The fibroin L-chain, encoded by the *Fib-L* gene on the 14th chromosome, its expression level in JS was 21.6 times higher than that of HY (Figure 2, spot 13). However, the difference between JS and F_1 hybrids was not significant.

GO analysis

These 22 different proteins were used for GO analysis, which can be divided into cell components, molecular functions and biological processes three categories, mainly involving in catalytic, cellular process, metabolic process, binding and extracellular region (Figure 4).

qRT-PCR validation

In current study, the silk gland from three biological samples were used for qRT-PCR to further validate the proteomic results. Eventually, 10 different proteins (spot nos. 1, 5, 11, 12, 13, 16, 18, 22, 26, 27) were selected for qRT-PCR. The gene sequences of these proteins were obtained from the silkworm genome sequences. The primer sequences used for qRT-PCR were listed in Table S3. As shown in Figure 5, the qRT-PCR expression results were basically consistent with the proteomic results, which further confirmed the reliability of the results.

Table 1 The identification of 28 differentially expressed proteins through MALDI-TOF/TOF

Protein spot number	Protein ID	Protein Name	Sequence coverage	MW (kDa)/pI[a]	Mascot score	Amino acid	Function [b]
1	gi|112983032	calreticulin	45%	46.1/4.49	636	398	calcium ion binding; carbohydrate binding; protein folding
2	gi|827542145	zonadhesin-like isoform X2	18%	73.9/4.87	204	621	sperm-specific transmembrane protein
3	gi|112983366	protein disulfide-isomerase like protein ERp57 precursor	26%	55.4/5.31	793	491	protein disulfide isomerase activity, cell redox homeostasis
4	gi|112983366	protein disulfide-isomerase like protein ERp57 precursor	30%	55.4/5.31	875	491	protein disulfide isomerase activity, cell redox homeostasis
5	gi|112983366	protein disulfide-isomerase like protein ERp57 precursor	41%	55.4/5.31	885	491	protein disulfide isomerase activity, cell redox homeostasis
6	gi|827563326	serine protease inhibitor 18 isoform X1	50%	44.2/4.44	596	392	extracellular space
7	gi|226342906	serine protease inhibitor 22 precursor	38%	44.1/4.41	562	392	extracellular space
8	gi|512894896	uncharacterized protein LOC101739721	46%	32.9/4.56	708	289	unknown function
9	gi|148298752	14-3-3 epsilon protein	62%	29.8/4.66	632	262	monooxygenase activity
10	gi|512892462	inorganic pyrophosphatase	62%	32.2/4.96	503	288	Oxidative phosphorylation
11	gi|452392	Nd-sD mutant fibroin light chain	24%	30.9/7.66	127	276	extracellular region; integral component of membrane
12	gi|112984494	fibroin light chain precursor	51%	27.9/5.23	405	262	extracellular region
13	gi|304443558	fibroin L-chain	62%	27.8/5.06	447	262	extracellular region
14	gi|112984494	fibroin light chain precursor	88%	27.9/5.23	513	262	extracellular region
15	gi|512891724	phosphatidylethanolamine-binding protein homolog F40A3.3-like	62%	20.2/5.02	729	185	construction and remodeling of biofilm; signal transduction, nervous system growth and development
16	gi|112983600	cellular retinoic acid binding protein	59%	15.0/5.66	471	132	lipid binding; transporter activity
17	gi|827559362	cuticular protein RR-2 motif 90 isoform X1	19%	22.9/5.94	87	408	structural constituent of cuticle
18	gi|404276811	farnesoic acid O-methyltransferase	41%	33.0/4.88	446	312	physiological processes
19	gi|512917993	uncharacterized protein LOC101743840 isoform X2	68%	26.4/6.31	873	238	unknown function
20	gi|512891160	fibrohexamerin-like	38%	27.8/4.81	491	241	maintain the integrity of silk fibroin complexes
21	gi|512891160	fibrohexamerin-like	31%	27.8/4.81	200	241	maintain the integrity of silk fibroin complexes
22	gi|512891160	fibrohexamerin-like	63%	27.8/4.81	510	241	maintain the integrity of silk fibroin complexes
23	gi|827548944	zinc finger protein 569-like	22%	23.9/8.41	59	204	cell differentiation; embryonic development
24	gi|512918988	lisH domain-containing protein C1711.05-like, partial	54%	39.1/9.61	36	635	protein dimerization
25	gi|112983898	elongation factor 1 gamma	46%	48.6/5.83	632	423	glycolytic process; magnesium ion binding; phosphopyruvate hydratase activity
26	gi|148298800	enolase	65%	47.2/5.62	951	433	magnesium ion binding; phosphopyruvate hydratase activity; glycolytic process
27	gi|827541166	arginine kinase	42%	40.3/5.87	640	355	ATP binding; kinase activity
28	gi|827562394	aldo-keto reductase AKR2E4-like isoform X2	37%	35.8/5.55	430	310	oxidoreductase activity

Note:
[a] MW: molecular weight; pI: isoelectric point
[b] KEGG results

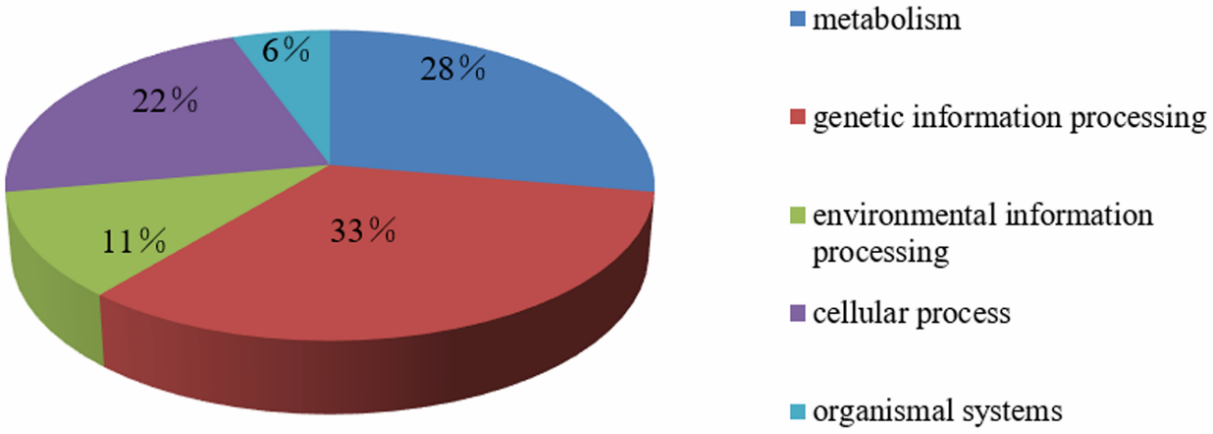


Fig. 3 The classification of differentially expressed proteins by KEGG

Discussion

The silkworm is not only a domestic insect for silk production, but also is a model system for research on developmental as well as physiological processes of Lepidoptera. In current study, we selected the silk gland of silkworm varieties JS and HY and their F_1 hybrids as research materials to investigate the molecular mechanism of heterosis. The hybrid JS-HY is the most popular silkworm variety for spring rearing in China because of its high yield, high quality, especially its excellent silk production. The silk fibroin, produced by PSG, is comprised of a heavy chain (fibroin H-chain, Fib-H), a light chain (fibroin L-chain, Fib-L), and a glycoprotein, P25 (Inoue *et al.*, 2000). The sericins, produced by MSG, are a protein family and can ensure the cohesion of the cocoon by sticking the silk threads together. According to our proteomic results, 28 significantly differentially expressed proteins were successfully identified, involving in 22 different proteins. GO analysis showed that these proteins were mainly involved in catalytic, cellular process, metabolic process, binding and extracellular region. KEGG analysis indicated that most of the identified proteins were involved in metabolic and genetic information processing, which accounted for 61% of the identified proteins, suggesting that metabolic and

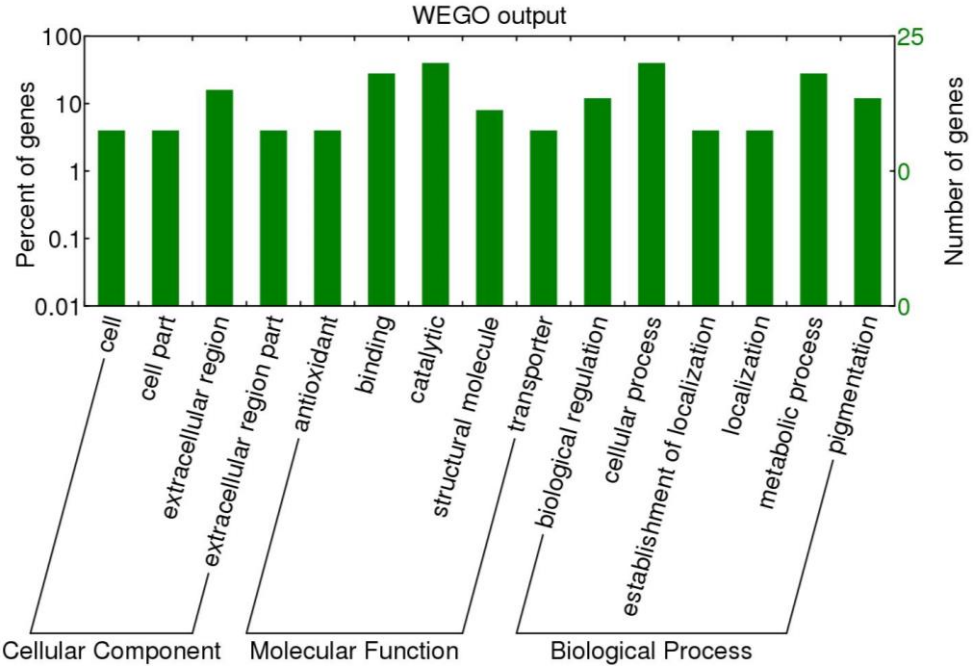
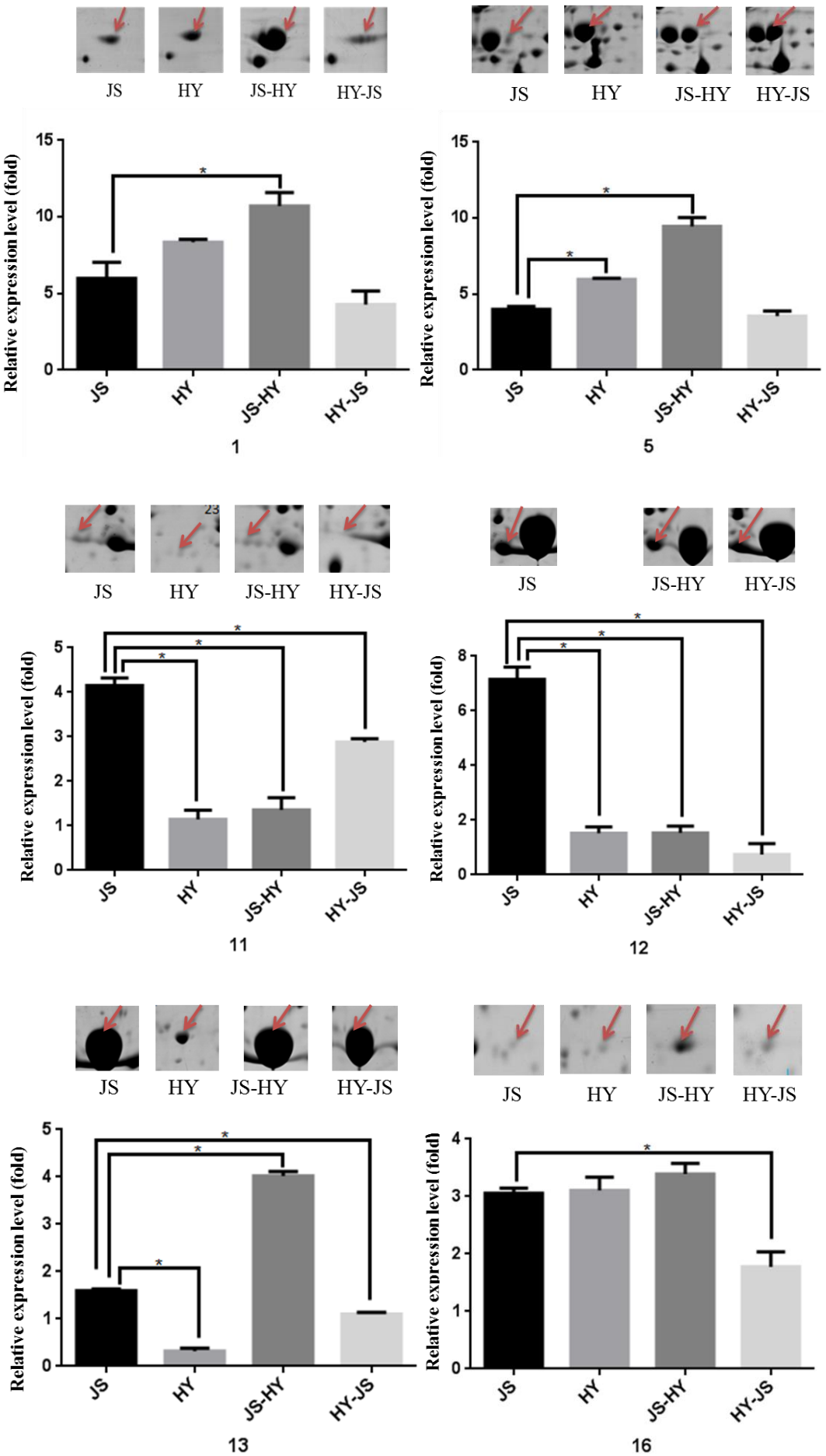


Figure 4 GO analysis of 28 differentially expressed proteins.


JS
HY
JS-HY
HY-JS
Relative expression level (fold)
1
5
11
12
13
16

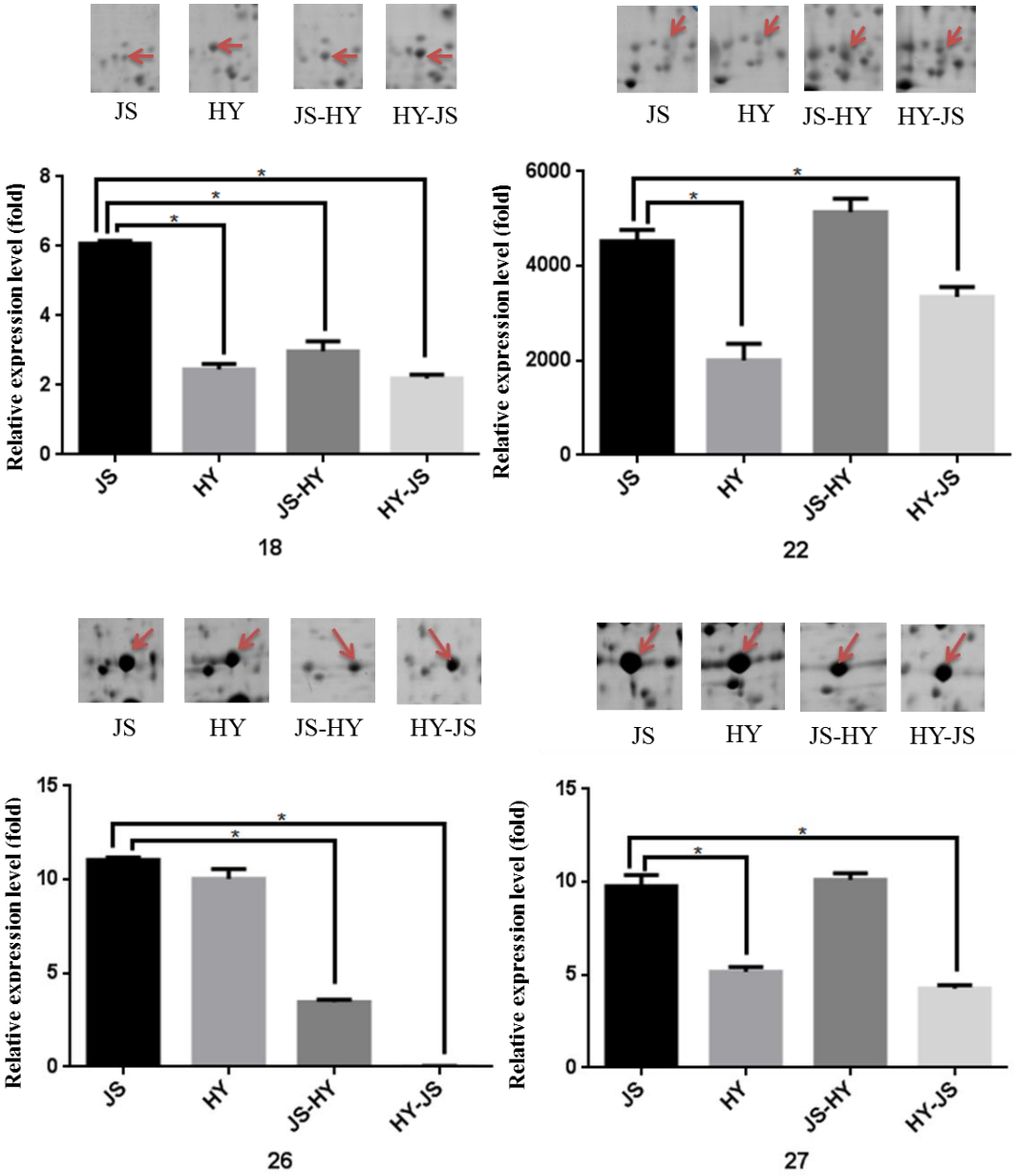


Fig. 5 Validation of 10 differentially expressed genes using qRT-PCR. X-axis represents different samples, Y-axis represents the relative expression level of mRNA

genetic information processing related proteins may play an important role in the heterosis of F_1 hybrids. Studies have shown that the epigenetic modification of genes associated with circadian rhythms changed the expression of downstream genes in *Arabidopsis*. They speculated that hybrids obtained advantages of circadian-mediated physiological and metabolic pathways and therefore promoted the improvement of vigor and the increase of biomass, consistent with our findings (Ni *et al*., 2009).

In current study, fibroin L-chain, Nd-sD mutant fibroin light chain, and Fhx were detected to be differentially expressed between parental silkworms and their F_1 hybrids. The *Fib-L* gene locates on the 14th chromosome, contains 7 exons and 6 introns, and encodes a total of 262 amino acids, of which the first 18 amino acids are signal peptides (Kikuchi *et al*., 1992). The H-chain and L-chain of fibroins are linked by a single disulfide bond between Cys-172 of L-chain and Cys-c20 (the twentieth residue from the C terminus) of H-chain. The H-L linkage is essential for the effective secretion of large amounts of fibroin. Compared with H-chain, the amino acids of L-chain are highly conserved, even a very little variation also affect the function of silk, indicating the functional importance of *Fib-L* gene in the performance of silk fibroin. Studies have illustrated that the silkworm with *Nd-s* or *Nd-sD* mutation of the *Fib-L* gene could not form the disulfide linkage with H-chain, leading to less than 1% fibroin of the normal level to be secreted (Tanaka *et al*., 1999). Previous research has shown that the cocoon of transgenic silkworm with antimicrobial peptide *cecropin* gene possessed antibacterial properties, and its silk performance can

be improved through the modification of *Fib-L* gene (Li *et al*., 2015). Because of the specific function of *Fib-L* gene in silk fibroin synthesis, it well explained the advantages of F_1 generation in silk gland weight, whole cocoon weight, and cocoon shell weight. The *Nd-s* gene is a dominant gene that controls the formation of silk cocoons (Takemura *et al*., 2016). *Nd-sD* is the allele gene of *Nd-s*, and only one base change exists in the coding region of *Nd-s* and *Nd-sD* genes. The mutation of *Nd-sD* gene occurs on the third intron of *Fib-L* gene, caused by a deletion from the third exon, which leads to the recombination of sequences in the third intron, and therefore affects the synthesis and secretion of silk fibroin (Takeiet *et al*., 1987; Mori *et al*., 1995). In fact, *Nd-sD* homozygous mutants can form immature PSGs, that's because the lack of *Fib-L* in the endoplasmic reticulum (ER) inhibits the development of PSG cells (Takei *et al*., 1987). Eventually, the *Nd-sD* mutant only secretes less than 1% fibroin of the normal level, and produces very thin and naked-pupa cocoon (Barbosa *et al*., 2008). Fhx, a class of glycoproteins, is highly expressed in PSGs but is inhibited in MSG. Studies have shown that Fhx plays roles in maintaining the integrity of silk fibroin complexes and promoting the expression of secretory cell genes (Julien *et al*., 2002; Barbosa *et al*., 2009). The expression levels of Fhx in F_1 hybrids were significantly higher than that of in the parents, consistent with the measurement result of the silk gland weight, indicating that Fhx indeed played important roles in promoting the synthesis of silk fibroin and maintaining the integrity of silk fibroin.

Meanwhile, we also identified two molecular chaperone-related proteins, including calreticulin (CRT, spot 1) and ERp57 protein (spots 3, 4 and 5). As shown in Figure 2, the total expression level of ERp57 are significantly up-regulated in F_1 hybrids. The similar results were also obtained from silk glands at 24h, 48h of 5th instar silkworms, verifying the reliability of our work (data not shown here). Calreticulin (CRT) is a unique ER protein and affects many cellular functions inside and outside the ER (Michalak *et al*., 1999). As a multifunctional lectin-like molecular chaperone, CRT is involved in the synthesis of various molecules, including ion channels, surface receptors, integrins and transporters (Martin *et al*., 2006). The ERp57 is a member of the protein disulfide isomerase (PDI) family. As a soluble protein of ER, ERp57 can form discrete complexes with the ER lectins, CRT and calnexin (CNX). These complexes can interact specifically with newly synthesized glycoproteins and modulate glycoprotein folding within the ER lumen. Additionally, ERp57 also acts as a thiol-disulfide oxidoreductase for proteins carrying N-linked glycans (Frickel *et al*., 2002). The thiol-disulfide oxidoreductase ERp57 can efficiently catalyze disulfide reduction, disulfide isomerization, and dithiol oxidation in substrate proteins (Frickel *et al*., 2004). Thereby, it can be speculated that the up-regulation of ERp57 in F_1 hybrids ensures the correct folding and prevents the degradation of proteins, and thus promotes rapid growth of F_1 hybrids (Figure 6).

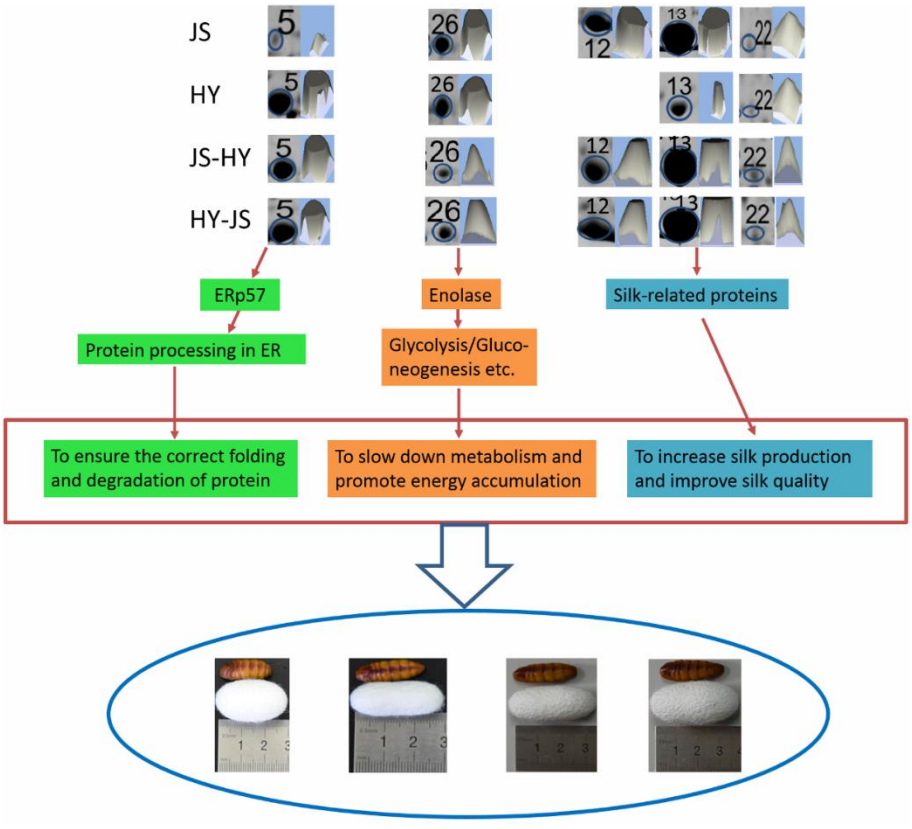


Fig. 6 The possible explanation of silkworm heterosis

At the same time, the metabolism, and/or energy-related proteins, 14-3-3 protein (spot 9) and enolase (spot 26) were also identified. The 14-3-3 protein is not only an important signal transduction protein, but also a conserved regulatory molecule. The main feature of 14-3-3 protein is to combine various functional signal proteins, including kinases, phosphatases, and transmembrane receptors. Consequently, 14-3-3 protein plays vital roles in a variety of important regulatory processes, such as mitotic signal transduction, apoptotic cell death, and cell cycle control (Fu *et al*., 2000). Enolase is a key enzyme in the glycolysis that can provide large amounts of energy for organisms. Studies have shown that the enolase of insect differs from mammalian enolase, which has relatively low conservation and shows species specificity in function (Kikuchi *et al*., 2017). In addition, enolase is involved in RNA degradation, glycolysis/gluconeogenesis, biosynthesis of amino acids, carbon metabolism and other pathways. As shown in Figure 2, the expression of enolase was down-regulated in hybrids, which was confirmed by the result of qRT-PCR, suggesting that the down-regulation of enolase might lead to the decline of metabolic rate in F_1 hybrids and therefore cause the accumulation of large amounts of energy. Finally, the growth and developmental time of F_1 generation was prolonged. As a result, body weight, cocoon shell weight, and other economic traits were improved accordingly (Figure 6). These observations, in conjunction with the qRT-PCR and proteomic results, suggesting that these different proteins may play important roles the formation of silkworm heterosis. These findings described here not only provide candidate proteins and genes for the improvement of yield in silkworm but also demonstrate that comparative proteomic approach can be feasible in investigating molecular mechanism of heterosis at the protein level. It should be noted, however, that the power to detect heterosis using the above methods may be not enough. Further investigations using other methods on the molecular mechanisms of heterosis should be carried out in the future.

Conclusion

Based on proteomic and genetic cross approaches, 28 differentially expressed proteins were successfully identified, including three silk-related proteins, Nd-sD mutant fibroin light chain, fibroin L-chain and Fhx, which were further confirmed by qRT-PCR. In addition, ERp57 and enolase were found to be involved in the heterosis of silkworm. Taken together, our work not only provides novel insights and references for the understanding of molecular mechanism of silkworm heterosis, but also provides candidate proteins and genes for the improvement of yield in silkworm.

References

Barbosa J, Bravo J, Takeda K, Zanatta D, Silva J, Balani V, *et al*. Intrinsic bent DNA colocalizes with the sequence involved in the Nd-sD mutation in the *Bombyx mori* fibroin light chain gene. BMB Reports 41: 394-399, 2008.

Barbosa JF, Bravo JP, Zanatta DB, Silva JL, Fernandez MA. Allelic variability in the third intron of the fibroin light chain gene in *Bombyx mori* (Lepidoptera: Bombycidae). Genet. Mol. Res. 8: 197-206, 2009.

Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72: 248-254, 1976.

Chen R, Jiao P, Liu Z, Lu Y, Xin H, Zhang D, *et al*. Role of *BmDredd* during apoptosis of silk gland in silkworm, *Bombyx mori*. PloS One 12: e0169404, 2017.

Cilia M, Fish T, Yang X, Mclaughlin M, Thannhauser TW, Gray S. A comparison of protein extraction methods suitable for gel-based proteomic studies of aphid proteins. J. Biomolecular Techniques 20: 201-215, 2009.

Dhawan S, Gopinathan KP. Cell cycle events during the development of the silk glands in the mulberry silkworm *Bombyx mori*. Dev. Genes Evol. 213: 435-444, 2003.

Feng F, Chen L, Lian C, Xia H, Yang Z, Yao Q, *et al*. Comparative proteomic analysis reveals the suppressive effects of dietary high glucose on the midgut growth of silkworm. J. Proteomics 108: 124-132, 2014.

Frickel EM, Frei P, Bouvier M, Stafford WF, Helenius A, Glockshuber R, *et al*. ERp57 is a multifunctional thiol-disulfide oxidoreductase. J. Biol. Chem. 279: 18277-18287, 2004.

Frickel EM, Riek R, Jelesarov I, Helenius A, Wuthrich K, Ellgaard L. TROSY-NMR reveals interaction between ERp57 and the tip of the calreticulin P-domain. P. Natl. Acad. Sci. USA 99: 1954-1959, 2002.

Fu H, Subramanian RR, Masters SC. 14-3-3 proteins: structure, function, and regulation. Annu. Rev. Pharmacol. 40: 617-647, 2000.

Hou Y, Ma W. Small RNA and mRNA Profiling of *Arabidopsis* in Response to Phytophthora, Infection and PAMP Treatment. Plant Pattern Recognition Receptors: Springer New York, 2017.

Hou Y, Zou Y, Wang F, Gong J, Zhong X, Xia Q, *et al*. Comparative analysis of proteome maps of silkworm hemolymph during different developmental stages. Proteome Sci. 8: 45-54, 2010.

Inoue S, Tanaka K, Arisaka F, Kimura S, Ohtomo K, Mizuno S. Silk fibroin of *Bombyx mori* is secreted, assembling a high molecular mass elementary unit consisting of H-chain, L-chain, and P25, with a 6:6:1 molar ratio. J. Biol. Chem. 275: 40517-40528, 2000.

Julien E, Brdeaux MC, Garel A, Couble P. Fork head alternative binding drives stage-specific gene expression in the silk gland of *Bombyx mori*. Insect Biochem. Molec. 32: 377-387, 2002.

Kannan M, Suryaaathmanathan V, Saravanakumar M, Jaleel A, Romanelli D, Tettamanti G, *et al*. Proteomic analysis of the silkworm midgut during larval-pupal transition. Invert. Surviv. J. 13: 191-204, 2016.

Kawamura K, Kawanabe T, Shimizu M, Nagano AJ, Saeki N, Okazaki K, *et al*. Genetic distance of inbred lines of Chinese cabbage and its

relationship to heterosis. Plant Gene 5: 1-7, 2016.

Kikuchi A, Nakazato T, Ito K, Nojima Y, Yokoyama T, Iwabuchi K, *et al*. Identification of functional enolase genes of the silkworm *Bombyx mori*, from public databases with a combination of dry and wet bench processes. BMC Genomics 18: 83-94, 2017.

Kikuchi Y, Mori K, Suzuki S, Yamaguchi K, Mizuno S. Structure of the *Bombyx mori* fibroin light-chain-encoding gene: upstream sequence elements common to the light and heavy chain. Gene 110: 151-158, 1992.

Kwan KD, Dominica R, Song Q, Taylor SH, Juenger TE, Harmon FG, *et al*. Temporal shift of circadian-mediated gene expression and carbon fixation contributes to biomass heterosis in maize hybrids. PloS Genet. 12: e1006197, 2016.

Laemmli UK. Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature 227: 680-685, 1970.

Lauria M, Echegoyennava RA, Rodríguezríos D, Zaina S, Lund G. Inter-individual variation in DNA methylation is largely restricted to tissue-specific differentially methylated regions in maize. BMC Plant Biol. 17: 52-63, 2017.

Li B, Shi Y, Gong J, Li J, Liu A, Shang H, *et al*. Genetic effects and heterosis of yield and yield component traits based on *Gossypium* barbadense chromosome segment substitution lines in two *Gossypium hirsutum* backgrounds. PloS One 11: e0157978, 2016.

Li D, Zeng R, Yan L, Zhao M, Chao J, Yu L, *et al*. Gene expression analysis and SNP/InDel discovery to investigate yield heterosis of two rubber tree F_1 hybrids. Sci. Rep. 6: 24984, 2016.

Li J, Moghaddam SHH. Shotgun proteomic analysis on the embryos of silkworm *Bombyx mori* at the end of organogenesis. Insect Biochem. Molec. 40: 293-302, 2010.

Li JY, Ye LP, Che JQ, Song J, You ZY, Wang SH, *et al*. Proteome identification of the silkworm middle silk gland. Data in Brief 6: 903-907, 2016.

Li R, Li W, Min T, Qin Q, Jing L, Jing C, *et al*. Homoeologue expression insights into the basis of growth heterosis at the intersection of ploidy and hybridity in Cyprinidae. Sci. Rep. 6: 27040, 2016.

Li Y, Cheng W, Liu X, Jian S, Li H, Sui Z, *et al*. Up-regulating the abscisic acid inactivation gene *ZmABA8ox1b* contributes to seed germination heterosis by promoting cell expansion. J. Exp. Bot. 67: 2889-2900, 2016.

Li Z, Jiang Y, Cao G, Li J, Xue R, Gong C. Construction of transgenic silkworm spinning antibacterial silk with fluorescence. Mol. Biol. Rep. 42: 19-25, 2015.

Liu XY, Yao Q, Wang Y, Chen KP. Proteomic analysis of nucleopolyhedrovirus infection resistance in the silkworm, *Bombyx mori* (Lepidoptera: Bombycidae). J. Invertebrate Pathol. 105: 84-90, 2010.

Lv, HS. Principles of Insect Immunology. Shanghai Scientific and Technical Publishers. 2008.

Ma L, Xu H, Zhu J, Ma S, Liu Y, Jiang RJ, *et al*. Ras1CA overexpression in the posterior silk gland improves silk yield. Cell Res. 21: 934-943, 2011.

Martin V, Groenendyk J, Steiner SS, Guo L, Dabrowska M, Parker JM, *et al*. Identification by mutational analysis of amino acid residues essential in the chaperone function of calreticulin. J. Biol. Chem. 281: 2338-2346, 2006.

Michalak M, Corbett EF, Mesaeli N, Nakamura K, Opas M. Calreticulin: one protein, one gene, many functions. Biochem. J. 344: 281-292, 1999.

Miller M, Song Q, Shi X, Juenger TE, Chen ZJ. Natural variation in timing of stress-responsive gene expression predicts heterosis in intraspecific hybrids of *Arabidopsis*. Nat. Commun. 6: 7453-7465, 2015.

Moghaddam SHH, Du X, Li J, Cao J, Zhong B, Chen YY. Proteome analysis on differentially expressed proteins of the fat body of two silkworm breeds, *Bombyx mori,* exposed to heat shock exposure. Biotechnol. Bioproc. E. 13: 624-631, 2008.

Mori K, Tanaka K, Kikuchi Y, Waga M, Waga S, Mizuno S. Production of a chimeric fibroin light-chain polypeptide in a fibroin secretion-deficient naked pupa mutant of the silkworm *Bombyx mori*. J. Mol. Biol. 251: 217-228, 1995.

Ni Z, Kim ED, Ha M, Lackey E, Liu J, Zhang Y, *et al*. Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. Nature 457: 327-333, 2009.

Qin L, Xia H, Shi H, Zhou Y, Chen L, Yao Q, *et al*. Comparative proteomic analysis reveals that caspase-1 and serine protease may be involved in silkworm resistance to *Bombyx mori* nuclear polyhedrosis virus. J. Proteomics 75: 3630-3638, 2012.

Shull GH. What is “heterosis”?. Genetics 33: 439-446, 1948.

Takei F, Shimura K. Further evidence for importance of the subunit combination of silk fibroin in its efficient secretion from the posterior silk gland cells. J. Cell Biol. 105: 175-180, 1987.

Takemura Y, Ohnuma A, Mochida Y, Mastumoto M. A simple method for homozygotic selection of *Nd* genes using artificial diet rearing in the silkworm, *Bombyx mori*. J. Insect Biotechno. Sericol. 85: 1-6, 2016.

Tanaka K, Kajiyama N, Ishikura K, Waga S, Kikuchi A, Ohtomo K, *et al*. Determination of the site of disulfide linkage between heavy and light chains of silk fibroin produced by *Bombyx mori*. B. B. A. 1432: 92-103, 1999.

Wang H, Fang Y, Wang L, Zhu W, Ji H, Wang H, *et al*. Heterosis and differential gene expression in hybrids and parents in *Bombyx mori* by digital gene expression profiling. Sci. Rep. 5: 8750, 2015.

Wu X, Li R, Li Q, Bao H, Wu C. Comparative transcriptome analysis among parental inbred and crosses reveals the role of dominance gene expression in heterosis in *Drosophila melanogaster*. Sci. Rep. 6: 21124, 2016.

Xue B, Li F, Hu J, Tian J, Li J, Cheng X, *et al*. Differentially expressed genes in the silk gland of silkworm (*Bombyx mori*) treated with TiO_2 NPs. Gene. 611: 21-26, 2017.

Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, *et al*. WEGO: a web tool for plotting GO

annotations. Nucleic Acids Res. 34: W293-297, 2006.
Zhong XW, Zou Y, Liu SP, Yi Q, Hu CM, Wang C, *et al*. Proteomic-based insight into malpighian tubules of silkworm *Bombyx mori*. PloS One 8: e75731, 2013.
Zhu D, Zhou G, Xu C, Zhang Q. Genetic components of heterosis for seedling traits in an elite rice hybrid analyzed using an immortalized F_2 population. J. Genet. Genomics 43: 87-97, 2016.

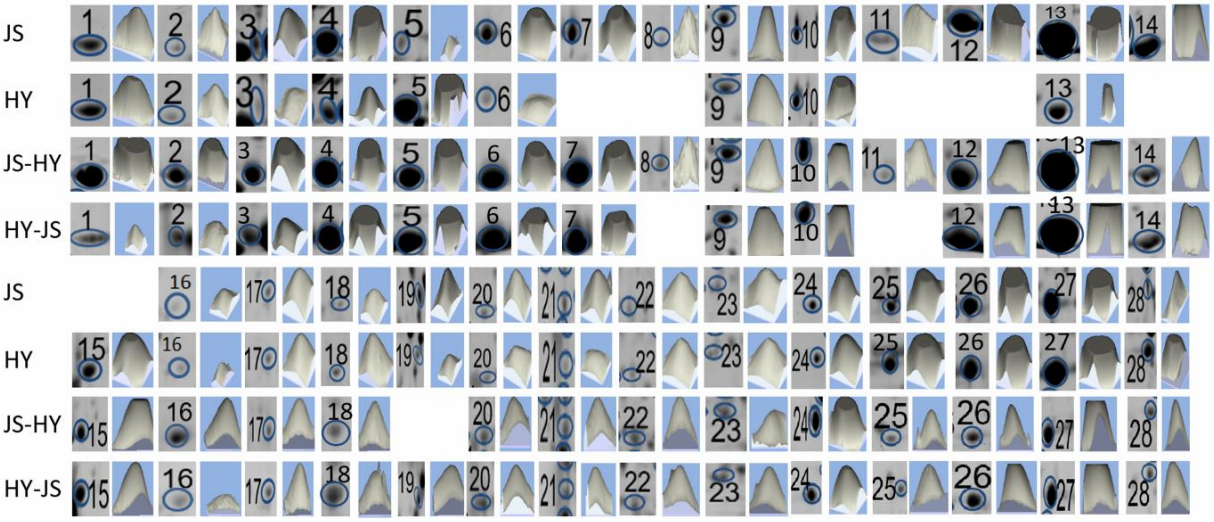


Fig. S1 Spot volume analyses of differentially expressed proteins among JS, HY and their F_1 hybrids. The gels are shown on the left panel, and the corresponding spot volume analyses performed using ImageMaster are shown on the right panel.

Table S1 Results of character investigation and evaluation of heterosis

Characters	F_1 [a]	HP [b]	LP [c]	MR [d]	The evaluation of heterosis [e]
Silkworm weight (g)	3.32	2.92	2.48	2.7	over-parent heterosis
Silk gland weight (g)	1.35	0.62	0.53	0.58	over-parent heterosis
Whole cocoon weight (g)	1.34	0.90	0.76	0.83	over-parent heterosis
Cocoon shell weight (g)	0.38	0.27	0.25	0.26	over-parent heterosis
Pupal weight (g)	1.06	0.94	0.92	0.93	over-parent heterosis
Cocoon shell percentage (%)	28.3	21.05	20.98	21.02	over-parent heterosis
Dead worm cocoon rate (%)	2.33	8.66	7.3	7.98	negative over-parent heterosis
Moth weight (g)	0.49	0.36	0.36	0.36	over-parent heterosis
Egg weight (g) (500 eggs/group)	0.37	0.23	0.17	0.2	over-parent heterosis
The number of eggs	510	365	354	360	over-parent heterosis

Note:
[a] F_1: the average value of the F_1 hybrids;
[b] HP: the average value of the parent with a high trait value;
[c] LP: the average value of the parent with a low trait value;
[d] MR: the average value of two parents;
[e] F_1 > HP, over-parent heterosis; MR < F_1 < HP, mid-parent heterosis; LP < F_1 < MR, negative mid-parent heterosis; F_1 < LP, negative over-parent heterosis.

Table S2 Detailed information of the differentially expressed proteins from silk glands at 96h of the fifth instar

Material and expression level	Protein spot number	Protein ID	Protein name	SC [a]	MW (kDa)/pI[b]	Mascot Score	AA [c]	Function [d]	Fold change [e]
JS and HY	3	gi|112983366	protein disulfide-isomerase like protein ERp57 precursor	26%	55.4/5.31	793	491	protein disulfide isomerase activity, cell redox homeostasis	3.7
Up-regulated proteins in JS	4	gi|112983366	protein disulfide-isomerase like protein ERp57 precursor	30%	55.4/5.31	875	491	protein disulfide isomerase activity, cell redox homeostasis	5.9
	6	gi|827563326	serine protease inhibitor 18 isoform X1	50%	44.2/4.44	596	392	extracellular space	2.9
	13	gi|304443558	fibroin L-chain	62%	27.8/5.06	447	262	extracellular region	21.6
	19	gi|512917993	uncharacterized protein LOC101743840 isoform X2	68%	26.4/6.31	873	238		1.6
	21	gi|512891160	fibrohexamerin-like	31%	27.8/4.81	200	241	maintain the integrity of silk fibroin complexes	2.1
Down-regulated proteins in JS	5	gi|112983366	protein disulfide-isomerase like protein ERp57 precursor	41%	55.4/5.31	885	491	protein disulfide isomerase activity, cell redox homeostasis	20.2
	18	gi|404276811	farnesoic acid O-methyltransferase	41%	33.0/4.88	446	312	physiological processes	2.0
	28	gi|827562394	aldo-keto reductase AKR2E4-like isoform X2	37%	35.8/5.55	430	310	oxidoreductase activity	3.6
Unique proteins in JS	7	gi|226342906	serine protease inhibitor 22 precursor	38%	44.1/4.41	562	392	extracellular space	
	8	gi|512894896	uncharacterized protein LOC101739721	46%	32.9/4.56	708	289		
	11	gi|452392	Nd-sD mutant fibroin light chain	24%	30.9/7.66	127	276	extracellular region, integral component of membrane	
	12	gi|112984494	fibroin light chain precursor	51%	27.9/5.23	405	262	extracellular region	
Unique proteins in HY	15	gi|512891724	phosphatidylethanolamine-binding protein homolog F40A3.3-like	62%	20.2/5.02	729	185	construction and remodeling of biofilm, signal transduction, nervous system growth and development	
JS and JS-HY	1	gi|112983032	calreticulin	45%	46.1/4.49	636	398	calcium ion binding, carbohydrate binding, protein folding	5.6
Up-regulated proteins in JS-HY	2	gi|827542145	zonadhesin-like isoform X2	18%	73.9/4.87	204	621	sperm-specific transmembrane protein	3.8
	3	gi|112983366	protein disulfide-isomerase like protein ERp57 precursor	26%	55.4/5.31	793	491	protein disulfide isomerase activity, cell redox homeostasis	3.2
	5	gi|112983366	protein disulfide-isomerase like protein ERp57 precursor	41%	55.4/5.31	885	491	protein disulfide isomerase activity, cell redox homeostasis	13.1
	6	gi|827563326	serine protease inhibitor 18 isoform X1	50%	44.2/4.44	596	392	extracellular space	3.2
	7	gi|226342906	serine protease inhibitor 22 precursor	38%	44.1/4.41	562	392	extracellular space	3.4
	8	gi|512894896	uncharacterized protein LOC101739721	46%	32.9/4.56	708	289		2.4
	9	gi|148298752	14-3-3 epsilon protein	62%	29.8/4.66	632	262	monooxygenase activity	1.8
	10	gi|512892462	inorganic pyrophosphatase	62%	32.2/4.96	503	288	oxidative phosphorylation	3.6

	16	gi|112983600	cellular retinoic acid binding protein	59%	15.0/5.66	471	132	lipid binding, transporter activity	4.2
	18	gi|404276811	farnesoic acid O-methyltransferase	41%	33.0/4.88	446	312	physiological processes	2.8
	20	gi|512891160	fibrohexamerin-like	38%	27.8/4.81	491	241	maintain the integrity of silk fibroin complexes	3.2
	21	gi|512891160	fibrohexamerin-like	31%	27.8/4.81	200	241	maintain the integrity of silk fibroin complexes	2.8
	22	gi|512891160	fibrohexamerin-like	63%	27.8/4.81	510	241	maintain the integrity of silk fibroin complexes	1.8
	23	gi|827548944	zinc finger protein 569-like	22%	23.9/8.41	59	204	cell differentiation, embryonic development	2.4
	24	gi|512918988	lisH domain-containing protein C1711.05-like, partial	54%	39.1/9.61	36	635	protein dimerization	2.4
Down-regulated proteins in JS-HY	11	gi|452392	Nd-sD mutant fibroin light chain	24%	30.9/7.66	127	276	extracellular region, integral component of membrane	2.2
	25	gi|112983898	elongation factor 1 gamma	46%	48.6/5.83	632	423	glycolytic process, magnesium ion binding, phosphopyruvate hydratase activity	3.0
	26	gi|148298800	enolase	65%	47.2/5.62	951	433	magnesium ion binding, phosphopyruvate hydratase activity, glycolytic process	3.1
	27	gi|827541166	arginine kinase	42%	40.3/5.87	640	355	ATP binding, kinase activity	2.5
Unique proteins in JS-HY	15	gi|512891724	phosphatidylethanolamine-binding protein homolog F40A3.3-like	62%	20.2/5.02	729	185	construction and remodeling of biofilm, signal transduction, nervous system growth and development	
JS and HY-JS	2	gi|827542145	zonadhesin-like isoform X2	18%	73.9/4.87	204	621	sperm-specific transmembrane protein	1.8
Up-regulated proteins in HY-JS	5	gi|112983366	protein disulfide-isomerase like protein ERp57 precursor	41%	55.4/5.31	885	491	protein disulfide isomerase activity, cell redox homeostasis	11.9
	6	gi|827563326	serine protease inhibitor 18 isoform X1	50%	44.2/4.44	596	392	extracellular space	3.0
	7	gi|226342906	serine protease inhibitor 22 precursor	38%	44.1/4.41	562	392	extracellular space	2.9
	10	gi|512892462	inorganic pyrophosphatase	62%	32.2/4.96	503	288	oxidative phosphorylation	3.8
	18	gi|404276811	farnesoic acid O-methyltransferase	41%	33.0/4.88	446	312	physiological processes	3.2
	20	gi|512891160	fibrohexamerin-like	38%	27.8/4.81	491	241	maintain the integrity of silk fibroin complexes	3.0
	21	gi|512891160	fibrohexamerin-like	31%	27.8/4.81	200	241	maintain the integrity of silk fibroin complexes	3.2
	22	gi|512891160	fibrohexamerin-like	63%	27.8/4.81	510	241	maintain the integrity of silk fibroin complexes	2.0
	23	gi|827548944	zinc finger protein 569-like	22%	23.9/8.41	59	204	cell differentiation, embryonic development	2.1
	24	gi|512918988	lisH domain-containing protein C1711.05-like, partial	54%	39.1/9.61	36	635	protein dimerization	2.4
Down-regulated proteins in HY-JS	25	gi|112983898	elongation factor 1 gamma	46%	48.6/5.83	632	423	glycolytic process, magnesium ion binding, phosphopyruvate hydratase activity	2.8
	26	gi|148298800	enolase	65%	47.2/5.62	951	433	magnesium ion	2.6

								binding, phosphopyruvate hydratase activity, glycolytic process	
	27	gi|827541166	arginine kinase	42%	40.3/5.87	640	355	ATP binding, kinase activity	2.1
Unique proteins in HY-JS	15	gi|512891724	phosphatidylethanolamine-binding protein homolog F40A3.3-like	62%	20.2/5.02	729	185	construction and remodeling of biofilm, signal transduction, nervous system growth and development	
Unique proteins in JS	8	gi|512894896	uncharacterized protein LOC101739721	46%	32.9/4.56	708	289		
JS-HY and HY-JS	1[f]	gi|112983032	calreticulin	45%	46.1/4.49	636	398	calcium ion binding, carbohydrate binding, protein folding	6.2
	16[f]	gi|112983600	cellular retinoic acid binding protein	59%	15.0/5.66	471	132	lipid binding, transporter activity	3.2
Unique proteins in JS-HY	8	gi|512894896	uncharacterized protein LOC101739721	46%	32.9/4.56	708	289		
	11	gi|452392	Nd-sD mutant fibroin light chain	24%	30.9/7.66	127	276	extracellular region, integral component of membrane	
Unique proteins in HY-JS	19	gi|512917993	uncharacterized protein LOC101743840 isoform X2	68%	26.4/6.31	873	238		

Note:
[a] SC: sequence coverage
[b] MW: molecular weight; pI: isoelectric point
[c] AA: amino acid
[d] KEGG results
[e] $P < 0.05$
[f] Up-regulated proteins in JS-HY

Table S3 Primer sequences used for quantitative real-time PCR

Spot No.[a]	Protein name	Forward primer (5'-3')	Reverse primer (5'-3')
1	calreticulin	GCGACCCAGAGGATGACAAA	AGACCTTGAGGTATCCGCCT
5	protein disulfide-isomerase like protein ERp57 precursor	CGGCTTCCCCACAATCTTCT	TGGCGTTACCCTTTCTGTCC
11	Nd-sD mutant fibroin light chain	TCTCACGTGCATGGGACTAC	ACCCATACTGTTAGCGGCTG
12	fibroin light chain precursor	CTCACGTCGATGGGACTACG	CACCGCTGATGCTTGACTTG
13	fibroin L-chain	CGATACTCTGTCGGACCAGC	AAGTGAGCGGTGATGTAGGC
16	cellular retinoic acid binding protein	AAGAGTTCGAAGAGGACCGC	AGACTCTGGTGCAAGTCACG
18	farnesoic acid O-methyltransferase	AAATCGAAGTGCCACCGACT	CAGTCCTGTCCTCCCACAAC
22	fibrohexamerin-like	GGCGGATGAGGTAGACTGTG	GTACTGGCCCCAGCGTTATT
26	enolase	ATGATTGGTCTGCATGGGCA	CTCTGTTACGCTGCCGATCT
27	arginine kinase	TTGGAGGCTGGTTTCAGCAA	AGGAGGGTGGATCCGAATGA

Note:
[a] The spot number of identified proteins (see Table 1)