Running title: SP1 increased hatchability during silkworm domestication

Artificial Selection on *Storage Protein 1* Contributes to Increase of Hatchability during Silkworm Domestication

Yanan Zhu[1,2,#], Lizhi Wang[1,#], Cencen Li[3], Yong Cui[1], Man Wang[1], Yongjian Lin[1], Ruoping Zhao[2], Wen Wang[2,4] , Hui Xiang[1,2]*

1. Guangzhou Key Laboratory of Insect Development Regulation and Application Research, Institute of Insect Science and Technology, School of Life Sciences, South China Normal University, Guangzhou, 510631, China.

2. State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China.

3. College of Life Sciences, Xinyang Normal University, Xinyang, 464000, China

4. Center for Ecological and Environmental Sciences, Key Laboratory for Space Bioscience & Biotechnology, Northwestern Poly-technical University, Xi'an, 710072, China.

[#] These authors are equally contributed.

*Correspondence to: xiang_shine@foxmail.com

Running title: SP1 increased hatchability during silkworm domestication

Abstract

Like other domesticates, efficient utilization of nitrogen resource is also important for the domestic insect, the silkworm. Deciphering how artificial selection act on silkworm genome for improved utilization of nitrogen resource and further human-favored domestication traits will provide unique cues from the insect scenario for understanding general rules of Darwin's evolutionary theory on domestication. Storage proteins (SP), which belong to a hemocyanin superfamily, basically serve as a source of amino acids and nitrogen during metamorphosis and reproduction in insects. Here through genomic search and further screening of artificial selection signature on silkworm SPs, we discovered a candidate domestication gene, i.e. the methionine-rich storage protein1 (*SP1)*, which is uniquely diverged from the others and showed increased expression in the ova of domestic silkworms. Knockout of *SP1* via CRISPR/Cas9 approach resulted in dramatic decrease in egg hatchability, without obvious impact on egg production, which was similar to the case in the wild silkworm compared with domestic one. Larval development or metamorphosis were not affected by *SP1* knockout. Comprehensive ova comparative transcriptomes indicated a general repression of gene expression, specifically vitellogenin, chorion proteins and structural component proteins in the extracellular matrix (ECM)-interaction pathway, as well as enzymes in folate biosynthesis, in both the mutant and the wild silkworm with the mutated allele, compared to the wild type domestic silkworm. Wild silkworms with the wild allele also showed generally down-regulated expression of genes enriched in structural constituent of ribosome and amide and peptide

39 biosynthesis. This study exemplified a novel case that artificial selection could
40 directly act on nitrogen resource protein to affect egg nutrient and eggshell formation,
41 and activate ribosome for improved biosynthesis and increased hatchability during
42 domestication. The findings shed new light on both understanding of artificial
43 selection and silkworm breeding from the angle of nitrogen and amino acid resource.
44

Author summary

Like other domesticates, nitrogen resource is also important for the domestic insect, the silkworm. Deciphering how artificial selection act on silkworm genome for improved utilization of nitrogen resource and further human-favored domestication traits, will provide unique cues from insect scenario, for understanding general rules of Darwin's evolutionary theory. However, mechanism of domestication in the silkworm is largely unknown to date. Here we focused on one important nitrogen resource, i.e, the storage proteins (SP). We discovered that the methionine-rich storage protein1 (*SP1)* which is divergent from the other SPs are the only target of the artificial selection. We proposed based on functional evidence together with the key findings of comprehensive comparative transcriptome, that artificial selection, on one hand favored higher expression of *SP1* in the domestic silkworm, which would subsequently up-regulate the genes or pathways vital for egg development and eggshell formation. On the other hand, artificial selection consistently favored activated ribosome activities and improved amide and peptide biosynthesis and in the ova, as it might act in the silk gland for increased silk-cocoon yield. We here exemplified a novel case that artificial selection could directly act on nitrogen resource protein for human desired domestication trait.

Introduction

The silkworm *B*ombyx *mori* is the only fully domestic insect species, which originated from its wild ancestor *B. mandarina* about 5000 years ago. During this process, the domestic silkworm evolved rapidly under human-preferred selection. Deciphering how artificial selection act on silkworm genome for human-favored domestication traits, will provide unique cues from insect scenario, for understanding general rules of Darwin's evolutionary theory. Recently, by population and evolutionary genomic analyses on domestic and wild silkworm individuals, we recently found that nitrogen and amino acid metabolism pathways and specifically genes in glutamate and aspartate metabolism, were under artificial selection and could affect the metamorphosis and cocoon yield [1]. These findings suggest that, like domestic plants and animals, domestic silkworms also tend to have efficient utilization of nitrogen resources to adapt to human-preference [1-3]. Besides the glutamate and aspartate metabolism which is an ammonia re-assimilated system[4], we further wonder whether other kind of nitrogen resources are also affected by artificial selection. If this is the case, how they contributed to silkworm phenotypic changes during domestication.

Insect storage proteins (SP) are another important source of amino acids and nitrogen, which belong to a special conserved arthropod hemocyanin superfamily [5]. Most insects have at least two main kinds of storage proteins, i.e, arylphorin and methionine-rich storage proteins and some species have other non-typical SPs [6]. *SP*s have been cloned or predicted in many insect species, including Lepidoptera moths

and butterflies [7-10]. Insect SPs are suspected to serves as a source of amino acids and nitrogen the for pupae and adults during metamorphosis and reproduction [11], however solid functional evidence on its biological significance is rather few [9, 12]. In plant, storage proteins are mainly reserved in seeds, together with other nutrient such as oil and starch, to supply energy for seed germination, growth [13, 14]. Especially in crops, seed SPs function in providing energy for humans and animals and they are of great interest and target for breeding and improvement [13-15].

In the domestic insect, the silkworm, previous studies preliminarily characterized three SPs mainly based on cues of gene or protein expression pattern [7, 16-18]. Especially the methionine-rich SP1 was implied to contribute to adult female characters [7] and related to synthesis of vitellogenin (Vg), the precursor of yolk protein [16]. SP2 coupled with SP3 form a heterohexamer and has the inhibitory effects on cell apoptosis [17, 18]. Whether SPs are also important in the silkworm domestication, as observed in domesticated plants, given the importance of nitrogen supply in silkworm domestication, pends deep exploration of their biological and evolutionary significances.

Development of genomics and genome-editing techniques provide tools for efficiently decipher the evolutionary and functional significances of interested genes [19, 20]. Here in this study, we conducted a genome-wild identification of silkworm SPs and taking advantages of the genomic data resource of a batch of representative domestic and wild silkworms [1], we conducted selection signature screening of these silkworm SPs followed by functional verification via CRISPR/Cas9 knockout system

and comprehensive comparative ova transcriptomes of the wild type and mutant silkworms as well as domestic and wild silkworms. The findings here suggested that artificial selection on *SP1* contribute to increased production during silkworm domestication, possibly by upregulation of vitellogenin and egg development and eggshell formation, to promote egg hatchability during domestication. These findings provide a novel case with functional evidence and figure out a frame of regulation on a silkworm domestication gene, which illuminated that artificial on nitrogen and amino acid supply will be also required for improved silkworm reproduction.

Results

SP1 *is diverged from the other* SPs *and is the only one targeted by artificial selection*

Totally, we identified 7 *SP*s in the silkworm genome by blast search. SP1 showed the dramatically highest methionine content (10.98%) (Table 1). Phylogenetic analysis showed that *SP1* was one distinct clade whereas the others were in another one, indicating an obvious divergence between *SP1* and the others (Fig 1A). *SP1* is located in Chromosome 23 and the others are clustered in Chromosome 3, suggesting possible tandem duplication events during evolution. Interestingly, by screening of artificial selection signatures on the genomic region bearing *SP1* and the other *SP*s respectively, we detected strong signatures in the *SP1* region (see material and methods) (Fig 1B and 1C). Furthermore, we detected strong differentiation in allelic frequency in the upstream of *SP1* (Fig 1D). Correspondingly, *SP1* were differentially

expressed in the ova between domestic and wild silkworms, with higher expression in the domestic one (Fig 1E). We also detected 11 SNPs that caused amino acid changes in the coding sequence of the gene (S1 Fig), despite that the biological significance of these SNPs needs further evaluation. These results suggest that artificial selection on *SP1* during silkworm domestication might affect the function of this gene in domestic silkworms, at least in terms of gene expression.

Knockout of* SP1 *by CRISPR/Cas9

To explore the functional impact of *BmSp1* in silkworm domestication, we firstly investigated the biological role of this gene in the silkworm through CRISPR/Cas9 knocking out system. For single guide RNA (sgRNA) design, we selected highly specific targets in the first exon close to the translation starting site, namely, S1 and S2 (Fig 2A). We choose another site S3 close to the end of the first exon, more than 60 bp downstream of S1 and S2 (Fig 2A and Table 2), to obtain a potentially large fragment deletion by injecting the pool of three gRNAs. After mutation screening of the injected eggs (G0 generation), the gRNAs targeting the above three sites successfully guided DNA editing and generated a variety of mutation types, including 4–9 bp deletions or small insertions followed by a large deletion (Fig 2B).

By mutation screening of the exuviae of the fifth instar larva in the G0 cocoon, we successfully identified 26 mosaic mutant G0 moths. We then generated pairwise crosses of these G0 mutants with similar mutant genotypes for the G1 populations. After mutation screening of the G1 eggs, we selected two populations with large deletions for further feeding and mutation screening (see Material and methods).

Finally, in the G2 generation we obtained two types of homozygous mutants, i.e., MU1 and MU2 (Fig 2C). As to MU1, there was an 8 bp insertion followed by a 63 bp deletion in the *BmSP1* coding sequences. As to MU2, there was a 4 bp insertion followed by a 65 bp deletion. The mutations occurred at +29 bp of the first *SP1* exon in MU1 and MU2, respectively (Fig 2C), resulting in reading frame shift mutations and severe premature termination close to the translation starting site, with stop signals at +10 aa and +37 aa of the*SP1* protein (Fig 2D).

Both* SP1 *mutants and the wild silkworm had decreased hatching rate and Vg expression

We selected and maintained MU1 population for assay on phenotypes related to reproduction and metamorphosis, such as number of eggs, hatching rate, pupa weight, and cocoon weight. Compared with the wild-type, which showed hatching rates of about 90%, the hatching rates of *SP1* mutants were dramatically decreased, with a mean egg hatching rate of about 40% (Fig 2E), whereas the number of eggs produced was not obviously affected (Fig 2E) , neither did the whole pupa weight or cocoon shell weight (Fig 2E). Given that the data were obtained from large replicates (126 replicates for hatchability assay and 320 replicates for pupa and cocoon weight), the results are convincible. Loss-of-function mutation resulted in significant decreased expression of *SP1* and *Vg* in the ova, based on the RNA-seq data (Fig 2F).

Given that knockout of *SP1* caused reduced hatching rate (Fig 2E) and that expression of *SP1* in the ova of wild silkworm were significantly lower than that of domestic one, we further suspected that artificial selection on *SP1* might improve

silkworm hatching rate during domestication. Supporting this hypothesis, we found that hatching rates of the wild silkworm were generally lower compared with that of the domestic one and, similar to *SP1* mutant, no obvious differences was detected in egg production between the wild and the domestic silkworm (Fig 3A). The lower hatching rate of wild silkworm was also reported in other studies [21, 22]. Meantime, we also found that similar to SP1 mutants, expression of *Vg* in the ova of the wild silkworm was drastically lower compared with domestic one (Fig 3B). These results suggested that in the silkworm *SP1* may positively affect expression of *Vg* of ova and contribute to silkworm egg development. Promotion of *SP1* expression in the domestic silkworm thus results in the correspondingly up-regulation of *Vg*, which further contributes to increased hatchability during silkworm domestication.

Genes involved in egg development and eggshell formation were both repressed in SP1 mutant and the wild silkworm

In order to further exploring the regulation network and possible molecular mechanism of female specific *SP1* on egg hatchability, we generated comprehensively Ova comparative transcriptome analyses between the wild type and the mutant, as well as domestic and wild silkworm (*Bombyx mandarina*), with 4.87~9.15 Gb RNA-seq data for each sample (S1 Table). Totally, there were 561 genes identified as differential expressed genes (DEGs) in the SP1knockout mutants (MU1) compared to wild-type silkworm, with down-regulated genes (341) significantly more than up-regulated ones (220) ($p=0.0003$, Chi-squared test with Yates' continuity correction) (Fig. 4A and S2 Table). As expected, there are much more differential

196 expressed genes (2882) between the wild and domestic silkworm, since wild
197 silkworm are much more genetically and phenotypically different from the domestic
198 one, compared with the silkworm mutant. It is interesting that, down-regulated genes
199 (1761) were also significantly more than up-regulated (1121) ones (p=2.2e-16,
200 Chi-squared test with Yates' continuity correction) (Fig 3A and S3 Table). The results
201 suggested that transcriptome repression in ova might be an output of SP1 depletion, in
202 both the SP1-KO mutant (Fig 2F and Fig 4A) and in the wild silkworm (Fig 1E and
203 Fig 4A).

204 We identified 302 common genes shared in the two sets of DEGs. KEGG
205 enrichment analysis indicated that these common differential expressed genes were
206 significantly enriched in pathways related to cell proliferation, such as ECM-receptor
207 interaction and folate biosynthesis. GO enrichment analysis indicated that they were
208 enriched in reproduction related biological processes such as ovarian follicle cell
209 development, chorion-containing eggshell formation (Fig 4B) and these genes was
210 also enriched in the molecular function of structural constituent of chorion (Fig 4B).
211 Factually they are all annotated as chorion proteins, including 4 chorion class CB
212 protein M5H4-like genes (*BGIBMGA003248, BGIBMGA009720, BGIBMGA009719,*
213 *BGIBMGA009715*), a chorion class B protein PC10 (BGIBMGA009721) and a
214 Chorion 1 domain containing gene (*BGIBMGA005877*). All these chorion like genes
215 are down regulated except *BGIBMGA009720* (Fig 4B, Table S2, S3). Genes in
216 ECM-receptor interaction pathway includes collagens and integrins (S2 Fig) and those
217 in folate biosynthesis includes folylpolyglutamate synthase, involved in

7,8-Dihydrofolate (DHF) and 5,6,7,8-Tetrahydrofolate (THF) which are substrates for subsequent one carbon pool mediated by folate (S3 Fig). Extend to all the enriched genes, it is notable that they were mostly down-regulated, in both the SP1 mutant and the wild silkworm (Fig 4B). This results suggested that the common influence of repression of *SP1* in both the SP1 mutant and the wild silkworm at transcriptome level is on genes or pathways involved in reproduction, such as ovarian follicle cell development or proliferation, and eggshell formation.

We further generated enrichment analyses on DEGs in the two sets of comparison independently and observed consistent pattern (Table 3 and 4). Functional enrichment analysis of DGEs between wild-type silkworm and *SP1* mutant silkworm reveals significantly enriched the KEGG pathway "ECM-receptor interaction" and the GO terms, including ovarian follicle cell development and eggshell formation process, functioning as structural constituent of chorion (Table 3) and the related genes are mostly down-regulated. Consistently, These GO items were also in the top rank with the lowest p value when analyzing the DEGs between wild and domestic silkworm (Table 4). The involved genes in these KEGG or GO terms were nearly all down-regulated in the wild silkworm (Table 4). These results further supported that in the wild silkworm, repression of *SP1* may result in suppressed expression of ovarian follicle cell development and eggshell formation process.

Ribosome proteins and genes in amide and peptide biosynthetic processes were also repressed in the wild silkworm.

DEGs between wild and domestic silkworms were significantly in enriched in

function of structural constituent of ribosome. The related genes are mostly ribosome proteins (S4 Fig). We also noted that the related biological progresses, such as amide and peptide biosynthetic processes was also in the top rank with the lowest p value (Table 4). The related genes were also mostly down-regulated in the wild silkworm. However, during domestication, there might be other factors that contributes to improved hatchability, that is, improved amide and peptide biosynthesis and activated ribosome activities in the ovarian.

Discussion

Nitrogen resource is very important for silkworm domestication. The domestic silkworm tend to have efficient utilization of nitrogen resources for yielding protein outputs to adapt to human-preference. Here in this study, we discovered that artificial selection could directly act on nitrogen resource gene, i.e, storage protein1 (*SP1*), for improved silkworm hatchability. SPs are also of target loci of breeding in crops [23]. However, in the crops, human could directly benefit from the nutrient of these improved SPs [15] whereas in the silkworm, the benefits of SPs are increased silkworm reproductive capacity.

Among all the SPs identified, SP1 is quite diverged and somewhat unique from the others, in terms of both genomic location and phylogenetic position. Similar pattern was also observed in other Lepidoptera species, such as tobacco hornworm, *Manduca sexta* [24], suggesting that *SP1* might evolve dependently while the other type of *SPs* might have experienced duplication during Lepidoptera evolution. Methionine-rich SP1 seems to be of special interest, since methionine is reported to be

an important amino acid for the tradeoff between growth and reproduction [25]. In *Drosophila*, dietary methionine restriction extends lifespan [25], while in grasshoppers, a reduced reproduction-induced increase in expression methionine-rich protein occurred during life extension [26]. Similarly, in the beet armyworm, silencing of Sp1 by RNA interference (RNAi) decreases larval survival, which indicate the role of the methionine-rich SP in growth and metamorphosis[12]. We therefore added an new evidence that different to grasshopper [26] and the beet armyworm[12], but similar to *Drosophila*[25], silkworm methionine-rich SP1 functions in reproduction process but not obviously affect growth.

Given that in those cocoon-producing silk moths, another nitrogen utilization system such as gluminate/glutamine cycle system were reported to be vital in metamorphosis silk-cocoon production [1, 4, 27], we suspect that strategy of nitrogen resource allocation via storage proteins may be diverged or modified during Lepidoptera insect evolution. Here in the silkworm, function of SP1 limited to influencing egg hatching rate. Artificial selection acted only on *SP1*, again suggesting functional importance of SP1 rather than the other SPs, for human-preferred domestication traits, i.e, increased hatchability.

Ova comparative transcrtiome analyses further illustrated a frame of regulation network of SP1 on hatchability. Vitellogenin(Vg), chorion proteins, structural component proteins in the extracellular matrix (ECM)-interaction pathway such as collagen and integrins, and synthetase in folate biosynthesis are all generally repressed in both the *SP1* mutant and the wild silkworm. Thus, artificial selection

acting on *SP1* for increased hatchability, possibly thorough direct or indirect influence on those genes, pathway or biological processes. Vg is the main nutrition for silkworm egg formation and embryonic development. It appears and accumulates at the stage when SP1 rapidly declined and disappeared in the fat body shortly before adult emergence [16, 28]. SP1 may supplies amino acids for synthesis of Vg, as previously reported in *Plutella xylostella* (Yaginuma & Ushizima, 1991). Chorion proteins are the major component of the silkworm eggshell that have the essential function of protecting the embryo from external agents during development while allowing gas exchange for respiration. Eggshell (chorion) is constructed by the ovarian follicle cells. The follicle cell epithelium surrounds the developing oocyte and in the absence of cell division synthesizes a multilayer ECM [29]. Eggshell ECM were usually linked by integrins, a family of transmembrane receptor proteins to the cytoskeleton of oocyte. Via a series of signal transduction, ECM-integrins functions in oocyte movement, differentiation, and proliferation [29]. Integrins was reported to function in formation of actin arrays in the egg cortex [30] and it is also involved in tracheole morphogenesis which is for respiration [31]. Folate is one of B-vitamin cofactors with and functions in transferring various single-carbon units. The key folylpolyglutamate synthase is involved in production of 7,8-Dihydrofolate (DHF) and 5,6,7,8-Tetrahydrofolate (THF) which are substrates for subsequent one carbon pool mediated by folate. In human, folate is used as a supplement by women during pregnancy to prevent neural tube defects (NTD) in the baby, and low levels in early pregnancy are believed to be the cause of more than half of babies born

306 with neural tube defects, indicating its important role for fetal development [32]. In
307 insects, folate also plays important roles in egg development. It could promote
308 biosynthesis of nucleic acids in the ovaries, and evoke mitoses in cells of the
309 collicular epithelium [33-35].

310 Notably, increased hatchability during domestication may not be solely attributed
311 to increased expression of *SP1* and the associated downstream genes, given that
312 artificial selection acts on hundreds of gene loci in silkworm genome [1, 36] and that
313 ova comparative transcriptome between wild and domestic silkworm identified much
314 more genes than that between SP1 mutant and wild type silkworm. We factually
315 observed significantly enriched pathway and structural constituent of ribosome, the
316 protein translation machinery and involved biological processes in amide and peptide
317 biosynthesis, are generally lower expressed in the wild silkworm compared with
318 domestic silkworm (Table 4). These results again supported the importance of
319 nitrogen and amino acid in silkworm domestication, not only for silkworm protein
320 output [1], but also for productivity.

321 Similar to other domesticates, hatchability of silkworm eggs directly determine
322 quantity of offspring, and thus it is an important productivity trait human favorably
323 selected during domestication. Based on the results and the discussion above, we
324 proposed that artificial selection, on one hand favored higher expression of *SP1* in the
325 domestic silkworm, which would subsequently up-regulate the genes or pathways
326 vital for egg development and eggshell formation. On the other hand, artificial
327 selection consistently favored activated ribosome activities and improved amide and

peptide biosynthesis and in the ova, as it might act in the silk gland for increased silk-cocoon yield [1]. As an output, domestic silkworm demonstrates improved increase egg hatchability compared with it wild ancestor.

Materials and Methods

Silkworm strains

A multivoltine silkworm strain, Nistari, was used in all experiments. Larvae were reared on fresh mulberry leaves under standard conditions at 25°C. The wild silkworms were collected in Zhejiang province, China and maintained as laboratory population in our lab.

Silkworm genomic data resource

The silkworm *Bombyx mori* reference genome and related data used for searching and screening of artificial selection signature on silkworm SPs is our updated version (https://doi.org/10.5061/dryad.fn82qp6); those for RNA-seq data analyses were obtained from Ensemble database (http://metazoa.ensembl.org/Bombyx_mori/Info/Index).

Genomic search of B.mori SPs and Phylogenetic analysis of SPs

The reference sequences of *B. mori* storage proteins (SP1, SP2, SSP2 and SP3) were retrieved from the NCBI GenBank. These sequences were used as quires searching for homologs in the *B. mori* genome by tblastn with e-value $<10^{-7}$. Other insect homologs of the silkworm SPs were searched in GenBank (https://blast.ncbi.nlm.nih.gov/) by BLASTP with an e-value $<10^{-7}$. We selected sequences from several representative Lepidoptera species and *Drosophila melanogaster* as candidate proteins for further

analyses. The sequences of the SP1 homologs were aligned using MEGA 6.0 software [37]. A gene tree was constructed using The Bayes tree was constructed by MrBayes-3.1.2 with GTR + gamma substitution model. The gene-ration number was set as 1000000 and the first 25% was set as burn-in. Other parameters were set as default.

Molecular selection of Sp1 in domesticated and wild silkworm populations

Based on the available whole genomic Single nuclear polymorphic data (SNP) of domesticated and wild silkworm populations [1], (https://doi.org/10.5061/dryad.fn82qp6), we screened the selection signatures of the silkworm *SPs*, according to Xiang et al's pipeline [1]. Allelic frequency and SNP annotation was calculated by in-house Perl scripts.

Design of sgRNA target and in vitro synthesis of sgRNA and Cas9 mRNA

The 20 bp sgRNA targets immediately upstream of PAM were designed by the online platform CRISPRdirect (http://criSpr.dbcls.jp/)[38]. The sgRNA DNA template was synthesized by PCR, with Q5® High-Fidelity DNA Polymerase (NEB, USA). The PCR conditions were 98°C for 2 min, 35 cycles of 94°C for 10 s, 60°C for 30 s, and 72°C for 30 min, followed by a final extension period of 72°C for 7 min. The sgRNA were synthesized based on the DNA template *in vitro* using a MAXIscript® T7 kit (Ambion, Austin, TX, USA) according to the manufacturer's instructions. The Cas9 construct was a kind gift provided by the Shanghai Institute of Plant Physiology and Ecology (Shanghai, China). The Cas9 vector was pre-linearized with the NotI-HF® restriction enzyme (NEB, USA). The Cas9 mRNA was synthesized *in vitro* with a

372 mMESSAGE mMACHINE® T7 kit (Ambion, Austin, TX, USA) according to the
373 manufacturer's instructions. All related primers are shown in Table 2.

374 **Microinjection of Cas9/gRNAs**

375 Fertilized eggs were collected within 1 h after oviposition and microinjection was
376 within 4 h. The Cas9-coding mRNA (500 ng/μL) and total gRNAs (500 ng/μL) were
377 mixed and injected into the preblastoderm Nistari embryos (about 8 nl/egg) using a
378 micro-injector (FemtoJet®, Germany), according to standard protocols (Tamura,
379 2007). The injected eggs were then incubated at 25°C for 9–10 d until hatching.

380 **Cas9/gRNAs-mediated mutation screening and analyses of germline mutation**
381 **frequency**

382 To calculate the efficiency of Cas9/gRNAs-mediated gene mutation in the injected
383 generation (G0), we collected ~10% of the eggs (64 out of 600) 5 d after injection to
384 extract genomic DNA for PCR, with primers Sp1-F and Sp1-R (Table 2). The
385 amplified fragments were cloned into a pMD^{TM}19-T simple vector (Takara, Japan)
386 and sequenced to determine mutation type and mutagenic efficiency.

387 When the injected G0 silkworms pupated, we collected silkworm exuviae from
388 fifth instar larvae in each cocoon. Individual DNA was then extracted Genomic DNA
389 was extracted using a TIANamp Blood DNA Kit (Tian gen Biotech, Beijing)
390 according to the manufacturer's instructions. Individual mutation screening was
391 generated with PCR with 94°C for 2 min, 35 cycles of 94°C for 30 s, 57°C for 30 s,
392 and 72°C for 45 s, followed by a final extension period of 72°C for 5 min. PCR
393 products were cloned to pMD^{TM}19-T simple vector (Takara, Japan) and sequenced.

Homozygous mutant screening and mutation effect on inferred protein

Mosaic mutant moths were obtained from the above mutation screening of exuviae DNA from fifth instar larvae. Moths with the same mutation site were pairwise crossed with each other to acquire G1 offspring. About 7 d after the G1 eggs were laid, we collected ~30 eggs of each offspring population from one parental pair and pooled them to extract genomic DNA for mutation screening by PCR. The amplified fragments were cloned into a pMD^{TM}19-T simple vector (Takara, Japan) and sequenced to determine the exact mutation type. Two G1 offspring populations with large deletions in *BmSp1* were selected for further breeding. At the pupa stage, 20 randomly selected individuals within each population were subjected to mutation screening of exuviae DNA. Homozygous mutant moths with the same identified mutant genotype were crossed for G2 offspring. Mutation effects on proteins were evaluated using MEGA 6/0 software[37] by codon alignment of the wild type and the mutant.

Phenotypic assay

On the fourth day of pupation (P4), the silkworms had successfully pupated. We weighed every 10 individuals as a group, recorded the whole cocoon weight, pupa weight, and cocoon shell weight, respectively, and calculated the ratio of pupa weight to whole cocoon weight. In total, 32 replicates for the mutants and 11 replicates for the wild-type silkworms were set, respectively. Offspring of the homozygous mutants and wild-type silkworms were incubated at 25°C for 9–10 d until hatching. Number of Egg produced and egg hatching rate were determined. Eighty-three, forty-three, and

fourteen replicates were set for the two mutant (SP1-MU1, SP1-MU2) and wild-type populations.

Number of eggs produced and the hatching rate of the wild silkworm were also recorded. Ten replicated was set repeated two times.

Ova dissection, total RNA isolation and RNA-seq

Ova from virgin moth of the domestic wild type silkworm, SP1 mutant and the wild type were dissected and used to extract total RNA with three replicates. Total RNA were isolated using TRIzol (Invitrogen). For each sample, RNA were sent to Novogene Bioinformatics Institute (Beijing, China) for cDNA library construction and RNA-seq. Six cDNA libraries were sequenced by Illumina Hiseq 2500 (Illumina, San Diego, CA, USA) with 125 bp paired-end reads according to the manufacturer's instructions.

Analyses on RNA-seq data

Raw data were filtered with the following criteria: (1) reads with ≥ 10% unidentified nucleotides (N); (2) reads with > 10 nt aligned to the adapter, allowing ≤ 10% mismatches; and (3) reads with > 50% bases having phred quality < 5. The clean data were mapped to the *Bombyx mori* reference genome using Tophat with 2 nt fault tolerance and analyzed using Cufflinks [39]. The expression value of each gene was calculated and normalized using fragments per kilobase of exon per million reads mapped (FPKM) [39]. In order to identify differentially expressed genes, Cuffdiff was used to perform pairwise comparisons between wild-typed and SP1 mutant samples, as well as the wild and domestic silkworm, respectively, with corrected *P*-value of

438 0.05 <5 and Log2-fold change>1.

439 KEGG and GO enrichment analyses of differentially expressed genes were

440 performed with an online platform (http://www.omicshare.com/tools/), using all the

441 annotated genes in *Bombyx mori* as background.

442

443 **Data availability.** RNA-seq data were deposited in the NCBI Short Read Archive

444 database under the accession SAMN09700389-SAMN09700397.

445

446 **Acknowledgements**

447 We thank Anli Chen, Muwang Li and Lang You for their kind helps in silkworm

448 maintaining, and Anjiang Tan for his kindly providing the Cas9 vector. We thank

449 Shuai Zhan for his help in genomic data analyses and mining. This work was

450 supported by a National Basic Research Program of China (No. 2013CB835204) and

451 a grant of National Natural Science Foundation of China (No. 31371286).

452 **References**

453

454 1. Xiang H, Liu X, Li M, Zhu Y, Wang L, Cui Y, et al. The evolutionary road from wild moth to
455 domestic silkworm. Nat Ecol Evol. 2018. doi: 10.1038/s41559-018-0593-4. PubMed PMID: 29967484.
456 2. Hufford MB, Xu X, van Heerwaarden J, Pyhajarvi T, Chia JM, Cartwright RA, et al. Comparative
457 population genomics of maize domestication and improvement. Nat Genet. 2012;44(7):808-11. doi:
458 10.1038/ng.2309. PubMed PMID: 22660546; PubMed Central PMCID: PMCPMC5531767.
459 3. Wiener P, Wilkinson S. Deciphering the genetic basis of animal domestication. Proc Biol Sci.
460 2011;278(1722):3161-70. doi: 10.1098/rspb.2011.1376. PubMed PMID: 21885467; PubMed Central
461 PMCID: PMCPMC3169034.
462 4. Hirayama C, Nakamura M. Regulation of glutamine metabolism during the development of
463 Bombyx mori larvae. Biochim Biophys Acta. 2002;1571(2):131-7. PubMed PMID: 12049793.
464 5. Xie W, Luan YX. Evolutionary implications of dipluran hexamerins. Insect biochemistry and
465 molecular biology. 2014;46:17-24. doi: 10.1016/j.ibmb.2014.01.003. PubMed PMID: 24462816.
466 6. Haunerland NH. Insect storage proteins: gene families and receptors. Insect biochemistry and

467 molecular biology. 1996;26:11.
468 7. Sakurai H, Fujii T, Izumi S, Tomino S. Structure and expression of gene coding for sex-specific
469 storage protein of Bombyx mori. The Journal of biological chemistry. 1988;263(16):7876-80. PubMed
470 PMID: 2836399.
471 8. Gkouvitsas T, Kourti A. Juvenile hormone induces the expression of the SnoSP2 gene encoding a
472 methionine-rich hexamerin in Sesamia nonagrioides (Lepidoptera). Comparative biochemistry and
473 physiology Part B, Biochemistry & molecular biology. 2009;153(2):206-15. doi:
474 10.1016/j.cbpb.2009.02.017. PubMed PMID: 19275945.
475 9. Sonoda S, Fukumoto K, Izumi Y, Ashfaq M, Yoshida H, Tsumuki H. Methionine-rich storage
476 protein gene in the rice stem borer, Chilo suppressalis, is expressed during diapause in response to
477 cold acclimation. Insect molecular biology. 2006;15(6):853-9. doi: 10.1111/j.1365-2583.2006.00670.x.
478 PubMed PMID: 17201777.
479 10. Zhu YC, Muthukrishnan S, Kramer KJ. cDNA sequences and mRNA levels of two hexamerin
480 storage proteins PinSP1 and PinSP2 from the Indianmeal moth, Plodia interpunctella. Insect
481 biochemistry and molecular biology. 2002;32(5):525-36. PubMed PMID: 11891129.
482 11. Telang A, Buck NA, Wheeler DE. Response of storage protein levels to variation in dietary protein
483 levels. Journal of insect physiology. 2002;48(11):1021-9. PubMed PMID: 12770025.
484 12. Tang B, Wang S, Zhang F. Two storage hexamerins from the beet armyworm Spodoptera exigua:
485 cloning, characterization and the effect of gene silencing on survival. BMC Mol Biol. 2010;11:65. doi:
486 10.1186/1471-2199-11-65. PubMed PMID: 20807423; PubMed Central PMCID: PMCPMC2939506.
487 13. Gacek K, Bartkowiak-Broda I, Batley J. Genetic and Molecular Regulation of Seed Storage
488 Proteins (SSPs) to Improve Protein Nutritional Value of Oilseed Rape (Brassica napus L.) Seeds.
489 Frontiers in plant science. 2018;9:890. doi: 10.3389/fpls.2018.00890. PubMed PMID: 30013586;
490 PubMed Central PMCID: PMCPMC6036235.
491 14. Chen P, Shen Z, Ming L, Li Y, Dan W, Lou G, et al. Genetic Basis of Variation in Rice Seed Storage
492 Protein (Albumin, Globulin, Prolamin, and Glutelin) Content Revealed by Genome-Wide Association
493 Analysis. Frontiers in plant science. 2018;9:612. doi: 10.3389/fpls.2018.00612. PubMed PMID:
494 29868069; PubMed Central PMCID: PMC5954490.
495 15. Wu Y, Yuan L, Guo X, Holding DR, Messing J. Mutation in the seed storage protein kafirin creates
496 a high-value food trait in sorghum. Nature communications. 2013;4:2217. doi: 10.1038/ncomms3217.
497 PubMed PMID: 23948869.
498 16. Ogawa K, Tojo S. Quantitative Changes of Storage Proteins and Vitellogenin during the
499 Pupal-Adult Development in the Silkworm, Bombyx Mori (Lepidoptera, Bombycidae). Applied
500 Entomology and Zoolology. 1981;16:9.
501 17. Yu W, Wang M, Zhang H, Quan Y, Zhang Y. Expression and Functional Analysis of Storage Protein
502 2 in the Silkworm, Bombyx mori. International journal of genomics. 2013;2013:145450. doi:
503 10.1155/2013/145450. PubMed PMID: 23671839; PubMed Central PMCID: PMC3647547.
504 18. Chen J, Shu T, Chen J, Ye M, Lv Z, Nie Z, et al. Purification and Initial Functions of Sex-Specific
505 Storage Protein 2 in Bombyx mori. The protein journal. 2015;34(4):256-66. doi:
506 10.1007/s10930-015-9619-9. PubMed PMID: 26231571.
507 19. Xia Q, Li S, Feng Q. Advances in silkworm studies accelerated by the genome sequencing of
508 Bombyx mori. Annu Rev Entomol. 2014;59:513-36. doi: 10.1146/annurev-ento-011613-161940.
509 PubMed PMID: 24160415.
510 20. Daimon T, Kiuchi T, Takasu Y. Recent progress in genome engineering techniques in the

511 silkworm, Bombyx mori. Dev Growth Differ. 2014;56(1):14-25. doi: 10.1111/dgd.12096. PubMed
512 PMID: 24175911.
513 21. Chu JM, Yang R, Xia D, Gu W. The influence of light hours for copulation and oviposition of wild
514 silkworms. Sci Sericulture. 1990;04:2.
515 22. Zhao QL, Li L, Wei ZJ, Zhang ZF. The selections affect analysis of silkworm in laying eggs. Bullet
516 Sericulture 2001;3:3.
517 23. Chao H, Wang H, Wang X, Guo L, Gu J, Zhao W, et al. Genetic dissection of seed oil and protein
518 content and identification of networks associated with oil content in Brassica napus. Scientific reports.
519 2017;7:46295. doi: 10.1038/srep46295. PubMed PMID: 28393910; PubMed Central PMCID:
520 PMCPMC5385559.
521 24. Burmester T. Expression and evolution of hexamerins from the tobacco hornworm, Manduca
522 sexta, and other Lepidoptera. Insect biochemistry and molecular biology. 2015;62:226-34. doi:
523 10.1016/j.ibmb.2014.11.009. PubMed PMID: 25497415.
524 25. Lee BC, Kaya A, Ma S, Kim G, Gerashchenko MV, Yim SH, et al. Methionine restriction extends
525 lifespan of Drosophila melanogaster under conditions of low amino-acid status. Nature
526 communications. 2014;5:3592. doi: 10.1038/ncomms4592. PubMed PMID: 24710037; PubMed
527 Central PMCID: PMC4350766.
528 26. Tetlak AG, Burnett JB, Hahn DA, Hatle JD. Vitellogenin-RNAi and ovariectomy each increase
529 lifespan, increase protein storage, and decrease feeding, but are not additive in grasshoppers.
530 Biogerontology. 2015;16(6):761-74. doi: 10.1007/s10522-015-9599-3. PubMed PMID: 26298568;
531 PubMed Central PMCID: PMC4605866.
532 27. Osanai M, Okudaira M, Naito J, Demura M, Asakura T. Biosynthesis of L-alanine, a major amino
533 acid of fibroin in Samia cynthia ricini. Insect biochemistry and molecular biology. 2000;30(3):225-32.
534 PubMed PMID: 10732990.
535 28. Tojo S, Nagata M, Kobayashi M. Storage Proteins in the Silkworm, Bombyx Mori. . Insect
536 Biochem. 1980;10:5.
537 29. Khokhar A, Chen N, Yuan JP, Li Y, Landis GN, Beaulieu G, et al. Conditional switches for
538 extracellular matrix patterning in Drosophila melanogaster. Genetics. 2008;178(3):1283-93. doi:
539 10.1534/genetics.106.065912. PubMed PMID: 18245854; PubMed Central PMCID: PMCPMC2278091.
540 30. Chan D, Thomas CJ, Taylor VJ, Burke RD. Integrins on eggs: focal adhesion kinase is activated at
541 fertilization, forms a complex with integrins, and is necessary for cortex formation and cell cycle
542 initiation. Mol Biol Cell. 2013;24(21):3472-81. doi: 10.1091/mbc.E13-03-0148. PubMed PMID:
543 23985318; PubMed Central PMCID: PMCPMC3814141.
544 31. Levi BP, Ghabrial AS, Krasnow MA. Drosophila talin and integrin genes are required for
545 maintenance of tracheal terminal branches and luminal organization. Development.
546 2006;133(12):2383-93. doi: 10.1242/dev.02404. PubMed PMID: 16720877.
547 32. Jin J. Folic Acid Supplementation for Prevention of Neural Tube Defects. JAMA. 2017;317(2):222.
548 doi: 10.1001/jama.2016.19767. PubMed PMID: 28097357.
549 33. Grzelakowska B, Zielinska ZM. Folate Derivatives in the Metabolism of Insects. Ii. Biosynthesis of
550 Nucleic Acids in the Polytrophic Ovaries of the Diapausing Larvae of Acantholyda Nemoralis Thoms. As
551 Promoted by Folic Acid and Its 4-Aminoanalogue. Journal of insect physiology. 1965;11:431-42.
552 PubMed PMID: 14288372.
553 34. Blatch SA, Stabler SP, Harrison JF. The effects of folate intake on DNA and single-carbon pathway
554 metabolism in the fruit fly Drosophila melanogaster compared to mammals. Comparative

555 biochemistry and physiology Part B, Biochemistry & molecular biology. 2015;189:34-9. doi:
556 10.1016/j.cbpb.2015.07.007. PubMed PMID: 26219578.
557 35. Zielinska ZM, Grzelakowska B. Folate Derivatives in the Metabolism of Insects. I. Mitoses in Cells
558 of the Follicular Epithelium as Evoked in Acantholyda Nemoralis Thoms. By Folate and Its
559 4-Aminoanalogue. Journal of insect physiology. 1965;11:405-11. PubMed PMID: 14288370.
560 36. Xia Q, Guo Y, Zhang Z, Li D, Xuan Z, Li Z, et al. Complete Resequencing of 40 Genomes Reveals
561 Domestication Events and Genes in Silkworm (Bombyx). Science. 2009. PubMed PMID: 19713493.
562 37. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics
563 Analysis version 6.0. Molecular biology and evolution. 2013;30(12):2725-9. doi:
564 10.1093/molbev/mst197. PubMed PMID: 24132122; PubMed Central PMCID: PMCPMC3840312.
565 38. Naito Y, Hino K, Bono H, Ui-Tei K. CRISPRdirect: software for designing CRISPR/Cas guide RNA
566 with reduced off-target sites. Bioinformatics. 2015;31(7):1120-3. doi: 10.1093/bioinformatics/btu743.
567 PubMed PMID: 25414360; PubMed Central PMCID: PMCPMC4382898.
568 39. Ghosh S, Chan CK. Analysis of RNA-Seq Data Using TopHat and Cufflinks. Methods Mol Biol.
569 2016;1374:339-61. doi: 10.1007/978-1-4939-3167-5_18. PubMed PMID: 26519415.

570

571

572

573 **Figure Legends**

574 **Figure 1. Molecular evolution of silkworm SPs. (A)** Phylogenetic tree of insect
575 Storage proteins based on Bayesian inference analyses. Full-length amino acid
576 sequences were aligned to generate a phylogenetic tree. Bayesian posterior probability
577 was shown for each node. Cf, *Choristoneura fumiferana*; Bm, *Bombyx mori*; Hc,
578 *Hyalophora cecropia*; Sr, *Samia ricini*; Ms, *Manduca sexta*; Ob, *Operophtera*
579 *brumata*; Hyc, *Hyphantria cunea*; Sl, *Spodoptera litura*; Sn, *Sesamia nonagrioides*;
580 Cs, *Chilo suppressalis*; Pi, *Plodia interpunctella*; Px, P*lutella xylostella*; Ad,
581 *Anopheles darlingi*; Ca, *Corethrella appendiculata*; Dm, *Drosophila melanogaster*;
582 Pm, *Perla marginata*; Td, *Thermobia domestica*. Accession number for each protein is
583 indicated ahead of abbrevation of each species. (**B-C)** . Selection signatures of the
584 silkworm SPs. Signature index— Population divergence coefficient (Fst) between the
585 Chinese local trimoulting(CHN_L_M3) domestic silkworm group and the wild
586 silkworms and nucleic acid diversity (π) in the silkworm populaton is shown along the
587 genomic regions covering the *SP* genes. Dashed lines represent the top 1% Fst cutoff.
588 The red square represents *SP1* region which is located in Chromosome 23 and the
589 blue squares represents the other *SPs* which are clustered in Chromosome 3. **(D)**
590 Plotting of frequency of reference genotype for each SNP position in the upstream 2
591 kb region of *SP1* , indicating many mutant alleles with fairly high allelic frequency in
592 the wild silkworm population. **(E)** Expression level indicated as FPKM of *SP1* in the
593 ova of domestic and wild silkworms. The three data pot indicate the highest value
594 with 95% confidence, the average and the lowest value with 95% confidence,

595 respectively. ***, FDR corrected p <0.001. Dome, the domestic silkworm. Wild, the
596 wild silkworm.

597 **Figure 2. Cas9/sgRNA mediated gene editing of *SP1* in the silkworm. (A)**
598 Schematic diagram of sgRNA targeting sites. The five boxes indicate the five exons
599 of *BmSp1*, and the black line with blocks represents the gene locus (blocks, exons;
600 lines between blocks, introns). All sgRNA targeting sites are located on the first exon
601 (S1, S2, and S3) and the PAM sequence is labeled in red. Sp1-F and Sp1-R were the
602 primers used for mutation screening. **(B)** Various types of mutations in G0 injected
603 embryos. Deletions are indicated by hyphens and insertions are shown in blue
604 lowercase letters. The PAM sequence is in red. **(C)** Two types of mutations identified
605 from homozygous mutant silkworms. W, wild type. MU1, *SP1* mutant type 1. There
606 was an 8bp insertion followed by a 63 bp deletion in the *SP1* coding sequences; MU2,
607 *SP1* mutant type 2. there was a 4 bp insertion followed by a 65 bp deletion. **(D)**
608 Comparison of inferred amino acid sequences of homozygous mutant silkworms and
609 the wild type. The missing amino acids are replaced with dashes. Premature stop
610 codons are shown in red asterisk. **(E)** Phenotype assay of the mutants. Hatching rate
611 of *SP1* mutants (MU1) decreased to about 51%. In total, 83 replicates of hatchings for
612 MU1, 43 replicates for MU2 and 14 replicates for wild-type were used for statistical
613 analysis. Number of egg produced, whole cocoon weight, pupa weight and cocoon
614 layer weight, and pupa weight/whole cocoon weight ratio of mutants and wild-type
615 silkworms, indicating no significant differences. In total, 32 replicates for mutants and
616 11 replicates for wild-type were used. Error bar: SD. **(F)** Expression level indicated as

617 FPKM of *SP1* and *Vg* in the ova of domestic and wild silkworms. The three data pots
618 indicate the highest value with 95% confidence, the average and the lowest value with
619 95% confidence, respectively. *, ** and *** represented significant differences at the
620 0.05, 0.01, 0.001 level (t-test).

621 **Figure 3. The artificial selection on *SP1* might improve silkworm hatching rate**
622 **during domestication. (A)** The egg production and hatching rates in the domestic
623 group and wild silkworms. Error bar: SD. Wild_1 and Wild_2 indicated the results
624 from two repeat times **(B)** Expression level indicated as FPKM of *Vg* in the ova of
625 domestic and wild silkworms. The three data pots indicate the highest value with 95%
626 confidence, the average and the lowest value with 95% confidence, respectively. *, **
627 and *** represented significant differences at the 0.05, 0.01, 0.001 level.

628 **Figure 4. Differential expressed genes (DEGs) and functional enrichment**
629 **analysis of the common genes from DEGs of two comparisons, i.e., from mutant**
630 **VS the wild type, and from the wild and the dometic silworm. (A)** Venn diagram
631 of the DEGs between the wild type and the mutant, as well as domestic and wild
632 silkworm (*Bombyx mandarina*). The blue venn diagram represents the total
633 differentially expressed gene, The green venn diagram and the red venn diagram
634 represent down-regulated differentially expressed and up-regulated differentially
635 expressed, respectively. **(B)** Functional enrichment analysis of common genes shared
636 in the two sets of DEGs. Corrected p-value: p-value in hypergeometric test after FDR
637 correction. All non-redundant terms that had Corrected p-value < 0.05 is shown.

638

639 **Supporting Information Caption**

640 **S1 Fig. SNPs between wild and domestic silkworm groups that cause non-**
641 **synonymous mutations of SP1. (A).** Amino acid alignment of SP1 protein of the
642 domestic silkworm (*B.mori*) and the deduced SP1 protein sequence of the wild
643 silkworm (*B. mandarina*) by inferred SNP data. (**B**). Allelic frequencies of the 11
644 non- synonymous mutations in the wild (the left column) and domestic silkworm (the
645 right column). Blue, reference allele; Black, alternative allele. The 11 mutations of
646 amino acid were labelled with numerals in red.

647 **S2 Fig. Scheme of ECM-receptor interaction pathway.** The shared DEGs of ova in
648 two sets of comparison (SP1 mutant vs wild type and wild vs domestic silkworm) are
649 indicated by red frames.

650 **S3 Fig. Scheme of folate biosynthesis pathway.** The shared DEGs of ova in two sets
651 of comparison (SP1 mutant vs wild type and wild vs domestic silkworm) are indicated
652 by red frames.

653 **S4 Fig. Scheme of Ribosome pathway.** The DEGs of ova between the wild and
654 domestic silkworm are indicated by red frames.

655 **S1 Table. Summary of RNA-seq data.**

656 **S2 Table. Information of the differentially expressed genes between SP1 mutant**
657 **(MU1) and the wild type (WT) domestic silkworm.**

658 **S3 Table. Information of the differentially expressed genes between the wild and**
659 **the domestic (Dome) silkworms.**

Table 1. Basic information and methionine composition of silkworm SPs.

	Protein length (KD)	Protein weight (No. amino acid)	No. Met	Proportion of Met (%)
SP1	87.25	747	82	10.98
SSP2	83.45	703	23	3.27
SP2	83.46	704	27	3.84
SP3	82.85	696	22	3.16
Bmor_9524	87.39	743	49	6.59
Bmor_9525	99.02	866	18	2.08
Bmor_9527	77.61	651	20	3.07
Bmor_9880	47.42	441	3	0.73

660

Table 2 Primers used in this study

Primer	Name	Primer
Sp1gRNA1-f	gaaattaatacgactcactataTCTAGTACTACTGGCCTGCTgttttagagctagaaatagc	Preparation of sgRNA templates
Sp1gRNA2-f	gaaattaatacgactcactataACTACTGGCCTGCTTGGCCGgttttagagctagaaatagc	Preparation of sgRNA templates
Sp1gRNA3-f	gaaattaatacgactcactataGGTCTTCACCAAGGAACCAAgttttagagctagaaatagc	Preparation of sgRNA templates
Sp1gRNA-r	AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC	Preparation of sgRNA templates
Sp1-F	CGGAAATATGGCCTATAATGCT	Identification of somatic mutations
Sp1-R	TGTCTACACGGATCATCACC	Identification of somatic mutations

Note: Sp1gRNA1-f , Sp1gRNA2-f , Sp1gRNA3-f , Sp1gRNA-r were used for amplification of DNA templates to generate gRNAs targeting S1, S2, S3 sites; Sp1-F and Sp1-R for detection of targeting sites.

Table 3. Functional Enrichment of differential expressed genes between the loss-of-funtion *SP1* mutants and the wild-type domestic silkworms.

ID	Description	Corrected p value	Enrichment fold	No. DEGs	No. down regulated genes
ko04512	ECM-receptor interaction	0.07278	1.78230	7	5
Molecular Function					
GO:0005213	Structural constituent of chorion	0.00008	16.35621	6	5
GO:0004497	Monooxygenase activity	0.00204	3.73348	14	4
GO:0016491	Oxidoreductase activity	0.00295	1.92136	39	21
GO:0005506	Iron ion binding	0.01655	2.91085	14	6
GO:0004517	Nitric-oxide synthase activity	0.04483	24.53431	2	0
GO:0046906	Tetrapyrrole binding	0.04483	2.99199	10	4
GO:0005215	Transporter activity	0.04483	1.75245	30	23
Biological Process					
GO:0051704	Multi-organism process	0.00000	10.6227	13	7
GO:0007292	Female gamete generation	0.00001	15.2532	6	5
GO:0030703	Eggshell formation	0.00001	15.2532	6	5
GO:0030707	Ovarian follicle cell development	0.00001	15.2532	6	5
GO:0002376	Immune system process	0.00005	7.9582	8	2
GO:0009607	Response to biotic stimulus	0.00008	8.8977	7	1
GO:0042742	Defense response to bacterium	0.00027	9.1519	6	0
GO:0006810	Transport	0.00070	1.7538	44	32
GO:0019731	Antibacterial humoral response	0.00208	8.1713	5	0

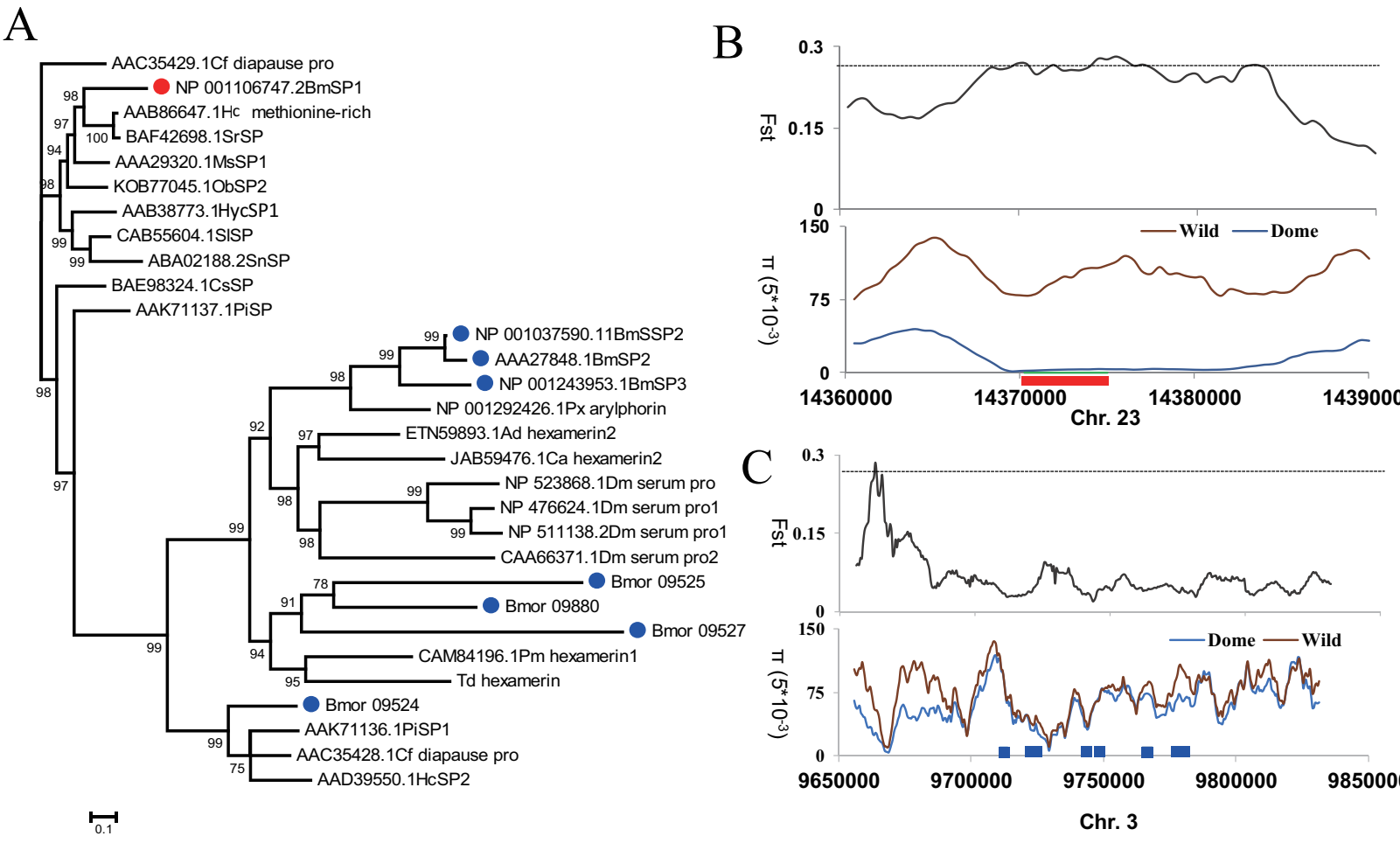
GO:0044707	Single-multicellular organism process	0.00715	3.5503	9	7
GO:0006809	Nitric oxide biosynthetic process	0.01570	22.8797	2	0
GO:0000041	Transition metal ion transport	0.03106	8.5799	3	1
GO:0051188	Cofactor biosynthetic process	0.04184	5.3835	4	3
GO:0018149	Peptide cross-linking	0.04248	15.2532	2	0
GO:0072593	Reactive oxygen species metabolic process	0.04273	7.6266	3	0

676

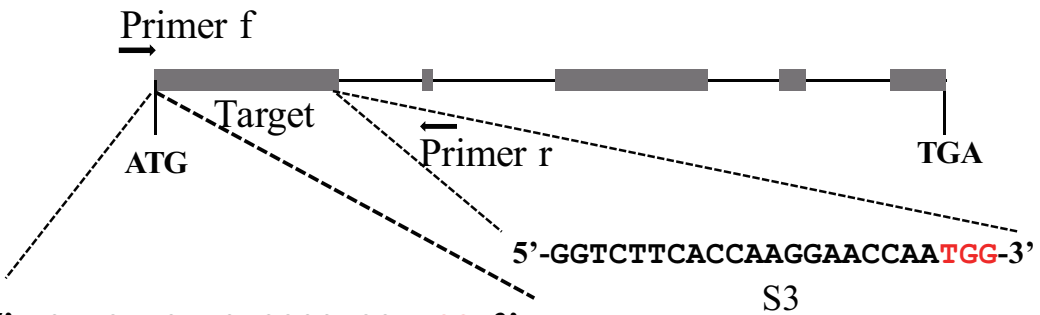
Table 4. Functional enrichment of differential expressed genes between the wild and domestic silkworms.

ID	Description	Corrected p value	Enrichment fold	No. DEGs	No. down regulated genes
ko03010	Ribosome	0.00001	2.08066	50	49
Molecular Function					
GO:0003735	Structural constituent of ribosome	0.00003	2.12234	44	42
Biological Process					
GO:0007292	Female gamete generation	0.12304	2.99379	6	5
GO:0007304	Chorion-containing eggshell formation	0.12304	2.99379	6	5
GO:0046700	Heterocycle catabolic process	0.12304	2.56610	8	5
GO:0009653	Anatomical structure morphogenesis	0.12304	2.99379	6	5
GO:0019953	Sexual reproduction	0.12304	2.99379	6	5
GO:0030707	Ovarian follicle cell development	0.12304	2.99379	6	5
GO:1901361	Organic cyclic compound catabolic process	0.12304	2.69441	9	6
GO:1901564	Organonitrogen compound metabolic process	0.12304	1.29411	117	98
GO:0043603	Cellular amide metabolic process	0.12304	1.50499	62	57
GO:0043604	Amide biosynthetic process	0.12304	1.48862	60	55
GO:0043043	Peptide biosynthetic process	0.12304	1.42885	56	52

677


A
AAC35429.1Cf diapause pro
NP 001106747.2BmSP1
AAB86647.1Hc methionine-rich
BAF42698.1SrSP
AAA29320.1MsSP1
KOB77045.1ObSP2
AAB38773.1HycSP1
CAB55604.1SlSP
ABA02188.2SnSP
BAE98324.1CsSP
AAK71137.1PiSP
NP 001037590.11BmSSP2
AAA27848.1BmSP2
NP 001243953.1BmSP3
NP 001292426.1Px arylphorin
ETN59893.1Ad hexamerin2
JAB59476.1Ca hexamerin2
NP 523868.1Dm serum pro
NP 476624.1Dm serum pro1
NP 511138.2Dm serum pro1
CAA66371.1Dm serum pro2
Bmor 09525
Bmor 09880
Bmor 09527
CAM84196.1Pm hexamerin1
Td hexamerin
Bmor 09524
AAK71136.1PiSP1
AAC35428.1Cf diapause pro
AAD39550.1HcSP2
0.1
B
Fst
π (5*10-3)
Wild
Dome
14360000
14370000
14380000
Chr. 23
C
Fst
π (5*10-3)
Dome
Wild
9650000
9700000
9750000
9800000
Chr. 3

A



S1 5'-TCTAGTACTACTGGCCTGCTTGG-3'

S2 5'-ACTACTGGCCTGCTTGGCCGCGG-3'

B

```
ATGAGGGTTCTAGTACTACTGGCCTGCTTGGCCGCGG(49 bp)GGAACCAATGG(W)
ATGAGGGTTCTAGTACTAC---------TGGCCGCGG(49 bp)GGAACCAATGG(1)
ATGAGGGTTCTAGTACTACTGGCCTGCTTGGCCGCGG(49 bp)G----CAATGG(2)
ATGAGGGTTCTAGTACTACTGGCCTGCTTGGCCGCGG(49 bp)--------TGG(3)
ATGAGGGTTCTAGTACTACTGGCC---------GCGG(49 bp)GGAACCAATGG(4)
ATGAGGGTTCTAGTACTACTGGCCTGCT-G-------/-----/-----CAATGG(5)
ATGAGGGTTCTAGTACTACTGGCCTGCT-GAGTGCTG/-----/-----CAATGG(6)
```

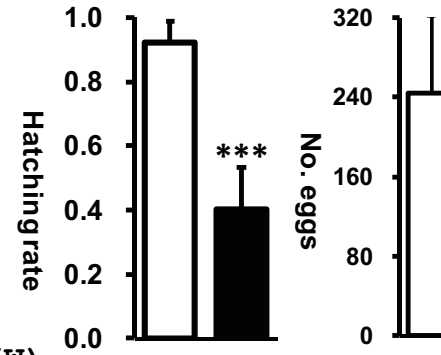
C

```
ATGAGGGTTCTAGTACTACTGGCCTGCTTGGCCGCGG(54 bp)CAATGG(W)
ATGAGGGTTCTAGTACTACTGGCCTGCT-G-------/-----/CAATGG(MU1)
ATGAGGGTTCTAGTACTACTGGCCTGCT-GAGTGCTG/-----/CAATGG(MU2)
```

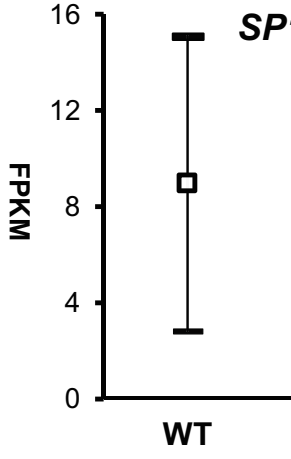
D

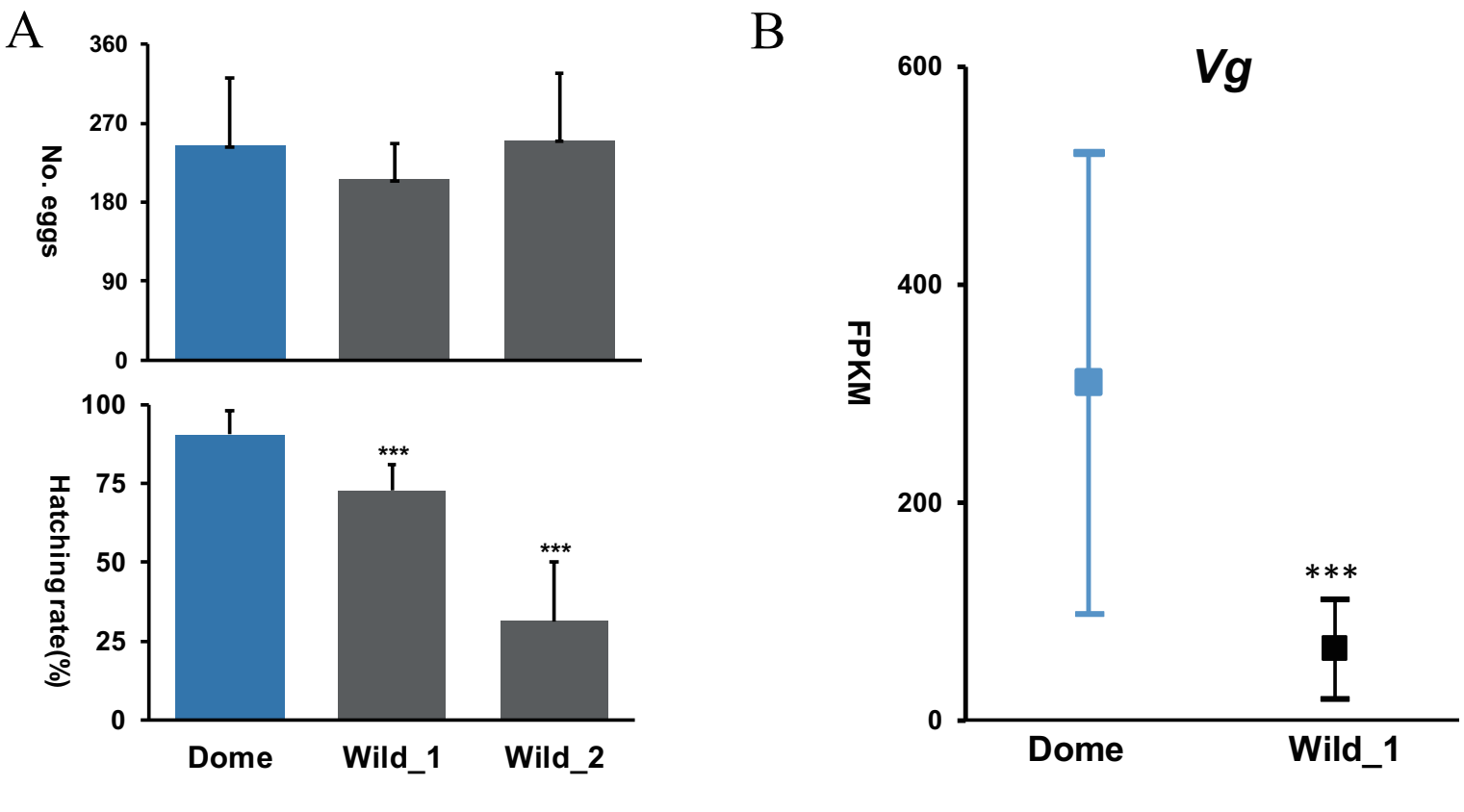
```
MRVLVLLACLAAASATAISGGYGTMVFTKEPMVNLDMK(W)
MRVLVLLAC*VLQWVSSSIILYNLHINLASLGYFDG*A(MU1)
MRVLVLLA--WVQWVSSSIILYNLHINLASLGYFDG*A(MU2)
```

E

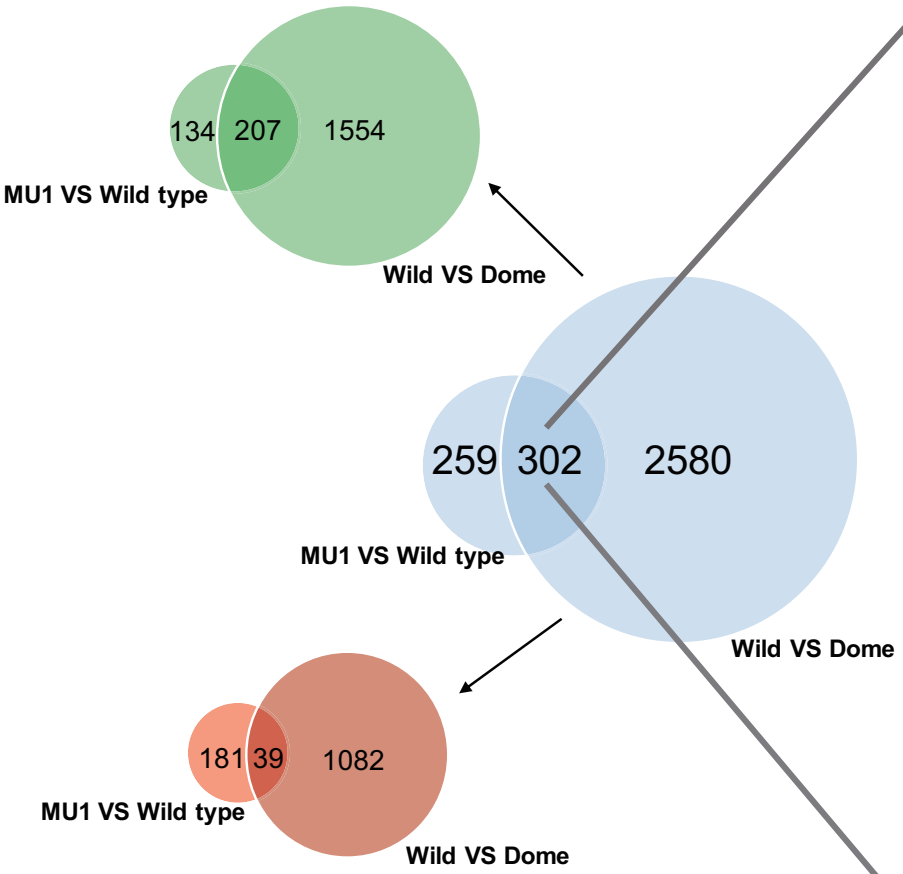


F




A
360
270
180
90
0
No. eggs
100
75
50
25
0
Hatching rate(%)
***
***
Dome
Wild_1
Wild_2
B
Vg
600
400
200
0
FPKM
***
Dome
Wild_1

A



B

	Corrected p value
Significant enriched pathway	
ECM-receptor interaction (ko04512)	0.0015
Folate biosynthesis (ko00790)	0.0330
Significant enriched GO term	
Female gamete generation (0007292) BP	0.0006
Chorion-containing eggshell formation (0007304) BP	0.0006
Ovarian follicle cell development (0030707) BP	0.0006
Multicellular organism reproduction (0032504) BP	0.0006
Single-multicellular organism process (0044707) BP	0.0057
Multi-organism process (0051704) BP	0.0057
Single-organism developmental process (0044767) BP	0.0098
Coenzyme biosynthetic process (0009108) BP	0.0148
Transport (0006810) BP	0.0307
Structural constituent of chorion (0005213) MF	0.0046