


Article

Molecular Cloning and Expression Analysis of Three Suppressors of Cytokine Signaling Genes (SOCS5, SOCS6, SOCS7) in the Mealworm Beetle *Tenebrio molitor*

Bharat Bhusan Patnaik ^{1,2} , Bo Bae Kim ¹, Yong Hun Jo ^{1,*} and In Seok Bang ^{3,*}

¹ Division of Plant Biotechnology, Institute of Environmentally-Friendly Agriculture (IEFA), College of Agriculture and Life Sciences, Chonnam National University, Gwangju 61186, Korea; drbharatbhusan4@gmail.com (B.B.P.); kbb941013@gmail.com (B.B.K.)

² School of Biotech Sciences, Trident Academy of Creative Technology (TACT), Chandrasekharpur, Bhubaneswar, Odisha 751024, India

³ Department of Biological Science and the Research Institute for Basic Sciences, Hoseo University, Asan 31499, Korea

* Correspondence: yhun1228@jnu.ac.kr (Y.H.J.); isbang@hoseo.edu (I.S.B.)

Received: 24 January 2019; Accepted: 9 March 2019; Published: 16 March 2019



Abstract: Suppressors of cytokine signaling (SOCS) influence cytokine and growth factor signaling by negatively regulating the Janus kinase (JAK)-signal transducers and activators of transcription (STAT) pathway to maintain homeostasis during immune responses. However, functional characterization of SOCS family members in invertebrates is limited. Here, we identified and evaluated three SOCS genes (type I sub-family) in the mealworm beetle *Tenebrio molitor*. The full-length open reading frames (ORFs) of TmSOCS5, TmSOCS6, and TmSOCS7 comprised of 1389, 897, and 1458 nucleotides, encoding polypeptides of 462, 297, and 485 amino acids, respectively. The SH2 and SOCS box domains of the TmSOCS C-terminal region were highly conserved. Phylogenetic analysis revealed that these SOCS genes were clustered within the type I subfamily that exhibits the highest amino acid identity with *Tribolium castaneum* SOCS genes. Contrary to TmSOCS7 expression, the expression levels of TmSOCS5 and TmSOCS6 were lower in the larval, pupal, and adult stages. In larvae and adults, the expression levels of TmSOCS5 and TmSOCS6 were highest in the hemocytes and ovaries, respectively. SOCS transcripts were also highly upregulated in the hemocytes of *T. molitor* larvae within 3–6 h post-infection with the fungus *Candida albicans*. Collectively, these results provide valuable information regarding the involvement of TmSOCS type-I subfamily in the host immune response of insects.

Keywords: *Tenebrio molitor*; suppressor of cytokine signaling; insect immunity; gene expression

1. Introduction

Cytokines are secretory proteins that regulate inflammatory responses. Most cytokines promote gene expression through the Janus kinase (JAK)-signal transducers and activators of transcription (STAT) pathway. Cytokine signaling and the JAK-STAT pathway are known to play essential roles in metazoan development and homeostasis of immune responses [1,2]. A large number of JAKs and STATs (e.g., four JAKs and seven STATs in humans) exist, and there appears to be differential employment of specific JAK-STAT pathway components in response to signaling mediated by a plethora of cytokine molecules. Non-redundant JAK/STAT signaling occurs in mammals, with preferential usage of various JAKs and STATs in response to specific cytokines/growth factors. Additionally, multiple

regulators tightly control cytokine-mediated cellular effects by limiting inappropriate activity and the development of related diseases. Various proteins, including suppressor of cytokine signaling (SOCS) family members (a specific SH2-domain-containing tyrosine phosphatase [SHP] family); and protein inhibitors of STATs (PIAS), function as negative regulators of the JAK-STAT pathway, and help maintain homeostasis during defense reactions [3–5].

SOCS family proteins are prime regulators of the JAK-STAT pathway. They control inflammatory signaling by acting as pseudo-JAK substrates, blocking STAT signaling, and directing multiple pathway components towards ubiquitin-mediated proteasomal degradation [6]. Four members of the SOCS family, including SOCS1, SOCS2, SOCS3, and cytokine-inducible SH2-containing protein (CISH), were originally identified in mice [7]. These proteins negatively regulate cytokine induction, by either competing with STAT for binding to the cytoplasmic domains of the phosphorylated receptors or by inactivating the enzymatic activity of JAKs [8]. Subsequently, four additional members (SOCS4, SOCS5, SOCS6, and SOCS7) were identified in mice based on their conserved central SH2 and C-terminal SOCS-box domains. Evolutionary analysis of SOCS family proteins revealed a clear division, with CISH, SOCS1, SOCS2, and SOCS3 grouped to the type II subfamily, and the rest of the members grouped within the type I subfamily [9]. SOCS4 and SOCS5 are close homologs that are involved in a negative feedback loop for epidermal growth factor receptor (EGFR) signaling by directing the degradation of EGFR in a ligand- and E3 ubiquitin ligase c-Cbl-independent manner in cultured CHO cells [4]. SOCS6 associates with and inhibits the insulin receptor and is related to cytokine-mediated insulin resistance in SOCS6-overexpressing transgenic mice, while human SOCS7 interacts with STAT5 or STAT3, to prevent nuclear translocation and to attenuate prolactin, growth hormone, and leptin signaling [10,11]. Eight members of the SOCS family have been identified in vertebrates; however, the SOCS gene repertoire is greatly expanded in rainbow trout, with 26 expressed genes within the type I and type II subfamilies. This has been attributed to the expansion of type II SOCS genes from a single CISH/SOCS1-3 precursor via whole genome duplication events. [12].

Knowledge of SOCS gene families in invertebrates is limited to a few isolated reports from molluscan and arthropod species. In *Drosophila*, SOCS36E, which shares 64% identity with human SOCS5, was the first SOCS protein identified, and its expression during embryogenesis was studied [13]. Later, it was shown to function as a negative regulator of the JAK/STAT and EGFR pathways. In contrast, two identified genes, SOCS44A (34% and 33% identity with human SOCS6 and SOCS7, respectively) and SOCS16D (48% and 45% identity with human SOCS6 and SOCS7, respectively), exhibited limited involvement in the JAK/STAT cascade, although SOCS44A was considered to be a transcriptional target of STAT92E [14]. Additionally, in the pacific oyster (*Crassostrea gigas*), three SOCS genes (SOCS2, SOCS5, and SOCS7) have been identified as putative inducers of NF- κ B transcription. SOCS2 homologues exhibiting immune-related expression have also been identified in *Ruditapes philippinarum* [15], *Haliotis discus* [16], *Eriocheir sinensis* [17], *Procambarus clarkii* [18], and *Litopenaeus vannamei* [19]. Further, direct evidence suggests that SOCS6 plays a role in activation of the NF- κ B signaling pathway in *E. sinensis* [20]. The *Bombyx mori* SOCS2 homolog has been demonstrated to function as a negative regulator of the JAK/STAT and ecdysteroid signaling pathways [21], while the SOCS6 homolog regulates the EGFR pathway [22]. The type I SOCS genes, however, have not been previously reported in coleopteran insects. In the present study, we identified three type I SOCS genes (SOCS5, SOCS6, and SOCS7) in the coleopteran model insect *Tenebrio molitor* through bioinformatics analysis. We also cloned the SOCS homologs and characterized their evolutionary relationship through phylogenetic analysis. Finally, we studied their expression profiles in response to exposure to microorganisms, which provided insights into the putative immune functions of SOCS in insects.

2. Materials and Methods

2.1. Insect Rearing

Tenebrio molitor were reared in the laboratory on an artificial diet (1.1 g of sorbic acid, 1.1 mL of propionic acid, 20 g of bean powder, 10 g of brewer's yeast powder and 200 g of wheat bran in 4400 mL of DW; autoclaved at 121 °C for 15 min) in the dark at 26 ± 1 °C, under $60\% \pm 5\%$ relative humidity.

2.2. Microorganisms

The Gram-negative bacterium *Escherichia coli* (strain K12), Gram-positive bacterium *Staphylococcus aureus* (strain RN4220), and the fungus *Candida albicans* were used for the immune challenge experiments. *E. coli* and *S. aureus* were cultured overnight in Luria Bertani (LB) broth at 37 °C, and *C. albicans* was cultured in Sabouraud Dextrose broth. The microorganisms were harvested, washed twice in phosphate-buffered saline (PBS; pH 7.0), and centrifuged at 3500 rpm for 10 min. The samples were then suspended in PBS, and the optical density was measured at 600 nm (OD_{600}). *E. coli* and *S. aureus* were diluted to 10^6 cells/ μ L, and *C. albicans* was diluted to 5×10^4 cells/ μ L for the immune challenge studies.

2.3. Identification and In Silico Characterization of TmSOCS Genes

The sequences of TmSOCS5, TmSOCS6, and TmSOCS7 were retrieved from the *T. molitor* RNAseq (unpublished) and expressed sequence tag (EST) databases by using local-tblastn analysis [23]. For the retrieval, the amino acid sequences of *T. castaneum* SOCS5 (XP_015833441.1), SOCS6 (XP_008190646.1), and SOCS7 (XP_008190646.1) were used as queries. Conserved domains were identified using InterProScan (<http://www.ebi.ac.uk/interpro/search/sequence-search>) and blastx (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). A domain-specific multiple alignment that included representative SOCSs from other insects retrieved from GenBank was generated using Clustal X2 [24]. The percent identity and phylogenetic analyses were performed using Clustal X2 and MEGA7 [25], respectively. The evolutionary relationships were inferred by the neighbor-joining method [26]. The bootstrap consensus tree was inferred from 1000 replicates, and the evolutionary distances were computed using the Poisson correction method. Human SOCS1, which belongs to the type II subfamily, was used as an outgroup.

2.4. Cloning the TmSOCS ORF

Primers were designed to amplify the TmSOCS ORFs based on the identified sequences, and the ORFs were amplified by polymerase chain reaction (PCR) using AccuPower[®] PyroHotStart Taq PCR PreMix (Bioneer, Korea). The primers are listed in Table 1. Briefly, 1 μ g of RNA was used as the template to synthesize cDNA using Oligo(dT) primers. Then, the generated cDNA was diluted 20 times, and 1 μ L was used for PCR amplification under conditions that included an initial denaturation step at 95 °C for 5 min, followed by 35 cycles of 95 °C for 30 s, 53 °C for 30 s, and 72 °C for 2 min. The PCR products were purified by the AccuPrep[®] PCR Purification Kit (Bioneer, Korea) and immediately ligated into a T-Blunt vector (Solgent, Korea). The PCR product-vector ligation was transformed into *E. coli* DH5 α competent cells according to the manufacturer's instructions. After validation by colony PCR, the plasmid DNA was extracted from the cells using the AccuPrep[®] Nano-Plus Plasmid Extraction Kit (Bioneer, Korea), and the cloned ORF was sequenced by using the M13 forward and reverse primers (Table 1).

Table 1. Primers used in this study.

Primer Name	Sequences (5' → 3')
TmSOCS5-cloning_Fw	CCCCCTGAGATGTGATTTCC
TmSOCS5-cloning_Rv	AACACCGCACATGAAAACAA
TmSOCS5-qPCR_Fw	CGCGCCCAAAGACAAGAAAATC
TmSOCS5-qPCR_Rv	TTTGGTGGGCCTTCTTGTTG
TmSOCS5-qPCR_Fw2	TCACGTTTTCCCTGCAACAC
TmSOCS5-qPCR_Rv2	TGCAGGAGGTTGATGTTGTC
TmSOCS6-cloning_Fw	AGTGTCCGTTGTGCGTGGT
TmSOCS6-cloning_Rv	GCGCGATTACTAAAAGTACGG
TmSOCS6-qPCR_Fw	TAAAGAGAAGCCTGCAGGACAG
TmSOCS6-qPCR_Rv	TCCGAGTGCTCCAAAACCTTC
TmSOCS7-cloning_Fw	CAGTGTCTCACGATACGCTTTC
TmSOCS7-cloning_Rv	AGTCTCAGGATTGTCTGGGATT
TmSOCS7-qPCR_Fw	ATTGAAGCGCGACAGTACAG
TmSOCS7-qPCR_Rv	AAGTCATGTGGATGGGTTCCC
TmL27a_qPCR_Fw	TCATCCTGAAGGCAAAGCTCCAGT
TmL27a_qPCR_Rev	AGGTTGGTTAGGCAGGCACCTTTA
M13_Fwd	GTAAAACGACGGCCAGT
M13_Rev	CAGGAAACAGCTATGAC

2.5. Developmental and Tissue-Specific Expression of TmSOCS Transcripts

Total RNA was isolated from eggs, early larvae (12th–15th instar; length, ~2.4 cm), late larvae (19th–20th instar; length ~3 cm), pre-pupae, 1–7-day-old pupae, and 1–5-day-old adult insects to monitor the expression of TmSOCS5, TmSOCS6, and TmSOCS7 during development. The pupal (PP to P7) and the adult stages (A1 to A5) of *T. molitor* considered for development expression were not sexed before sampling. For the tissue-specific expression analysis, total RNA was isolated from the hemocytes, gut, Malpighian tubules, fat body and integument of late-instar larvae. Total RNA was also isolated from the ovaries and testes of 5-day-old adults. The LogSpin RNA isolation method [27], with minor modifications, was used to isolate total RNA from the tissue and whole-body samples. Briefly, the samples were homogenized in 1 mL of guanidine thiocyanate RNA lysis buffer (20 mM EDTA, 20 mM MES buffer, 3 M guanidine thiocyanate, 200 mM sodium chloride, 0.005% Tween-80, and 1% isoamyl alcohol in DW, pH 5.5), and centrifuged at $21,000 \times g$ for 5 min at 4 °C. After a 1 min incubation in absolute ethanol, the samples were transferred to silica spin columns, and centrifuged at $21,000 \times g$ for 30 s at 4 °C to remove debris. After DNase treatment and two washes with 3 M sodium acetate and 80% ethanol, the total RNA was eluted with DNase and RNase-free water. The cDNAs were synthesized from 2 µg of total RNA using AccuPower® RT PreMix (Bioneer, Korea) and Oligo(dT)_{12–18} primers on a MyGenie96 Thermal Block (Bioneer, Korea). Then, the developmental and tissue-specific expression of the TmSOCS genes was analyzed by quantitative real-time reverse transcriptase polymerase chain reaction (qRT-PCR). The qRT-PCR assay was performed in a 20-µL reaction containing AccuPower® 2X GreenStar qPCR Master Mix (Bioneer) and primers for TmSOCS5, TmSOCS6, and TmSOCS7 (Table 1). The cycling parameters included an initial denaturation step at 95 °C for 5 min, followed by 40 cycles of 95 °C for 30 s, 60 °C for 30 s and 72 °C for 40 s. The qRT-PCR assays were performed on an AriaMx real-time PCR system (Agilent Technologies, Santa Clara, CA, USA) and the results analyzed using AriaMx real-time PCR software. The *T. molitor* Ribosomal protein L27a gene (*TmRpL27a*) was used for normalization, and the results were calculated by the $\Delta\Delta C_t$ method [28]. The reactions were performed in triplicate, and the results represent mean \pm S.E. of three biological replications.

2.6. Immune Challenge Studies

Healthy *T. molitor* larvae were administered with a 1 μ L suspension containing either 10^6 cells of *E. coli* or *S. aureus* or 5×10^4 cells of *C. albicans* by intra-abdominal injection. The preparation methods for microorganisms and pretreatment methods were performed according to one of our previous studies [29]. The dose of microorganisms in the healthy larvae has been validated to be sufficient for immune challenge studies and the mortality of the larvae was limited to 10% or less in most cases. A similar volume of phosphate buffered saline (PBS) was injected into a separate group of larvae as a wounded control. Due to their pivotal roles in the humoral and cell-mediated immune responses, hemocytes, fat body and gut tissues were collected at 3, 6, 12, and 24 h post-injection to analyze the expression of TmSOCS5, TmSOCS6, and TmSOCS7 after microorganism challenge. Total RNA isolation, cDNA synthesis, and qRT-PCR were performed as described above. To determine the effects of microbial challenge on the expression of TmSOCS transcripts, the fold-change at each time point was determined by comparison to that of the PBS-injected control. All data are presented as mean \pm standard error (SE). One-way analysis of variance (ANOVA) and Tukey's multiple range tests were used to evaluate the differences between groups ($p < 0.05$). All statistical tests were performed using the Statistical Analysis Software (SAS) suite (SAS Institute Inc., Cary, NC, USA).

3. Results and Discussion

3.1. Identification of TmSOCS Homologs and Molecular Characterization

A local tblastn search was conducted using TcSOCS5, TcSOCS6, and TcSOCS7 to retrieve SOCS gene family sequences from the *Tenebrio* RNAseq and EST libraries, and the retrieved gene sequences were named TmSOCS5, TmSOCS6, and TmSOCS7, respectively. The putative ORF sequences for the TmSOCS genes were identified using the gene-finding program FGENESH (<http://www.softberry.com/berry.phtml?topic=fgenes&group=programs&subgroup=gfind>). Next, primers were designed to amplify and clone the ORF sequences of TmSOCS5, TmSOCS6, and TmSOCS7. The TmSOCS constructs (T-blunt vector + full-length TmSOCS ORF insert) were sequenced to validate the TmSOCS genes. The sequences of TmSOCS5, TmSOCS6, and TmSOCS7 (nucleotide and translated protein sequences) were submitted to GenBank (NCBI) under the accession numbers MK292064, MK292065, and MK292066, respectively. The genomic architecture (exons and introns) of the *T. molitor* SOCS family genes was difficult to determine due to lack of a fosmid DNA library.

The TmSOCS5 ORF is 1389 bp and encodes a polypeptide of 462 amino acids (Figure 1). The computed molecular mass and isoelectric point of TmSOCS5 protein are 52.8 kDa and 6.23, respectively. These values are similar to those of the SOCS5 homolog found in *C. gigas*, which has a molecular mass of 51.5 kDa and an isoelectric point of 6.40 [30]. TmSOCS5 also contains an SH2 domain (residues 301–379) that includes a phosphotyrosine hydrophobic binding pocket and a C-terminal SOCS box domain (residues 381–437) containing the putative elongin B/C interaction residues. The TmSOCS6 ORF is 897 bp in length and encodes a polypeptide of 297 amino acids (Figure 2). Its calculated molecular mass and isoelectric point is 34.5 kDa and 9.40, respectively. The full-length TmSOCS6 sequence is shorter in comparison to the SOCS6 homologue found in *E. chinensis* [20], although typical SH2 and SOCS box domains are present. The TmSOCS7 ORF is 1458 bp in length and encodes a polypeptide of 485 amino acid residues (Figure 3). Its predicted molecular mass is 56 kDa, and its theoretical isoelectric point is 9.41. Similar to other SOCS family members, the phosphotyrosine, hydrophobic residues, and putative elongin B/C interaction sites are found in the conserved SH2 and SOCS box domains. This study and other closely-related studies have shown that the sequences in the amino-terminal region are diverse, while the central SH2 and carboxyl-terminal SOCS box domains are well conserved [31,32]. The SOCS box amino acid consensus sequence has also been found in protein families containing WD-40 repeats (IPR017986), ankyrin repeats (IPR020683), and SPRY domains (IPR003877) [8]. The elongin B/C region within the SOCS box domain may be involved

in the negative regulation of signaling pathways by directing the target proteins for ubiquitination and proteasomal degradation [33,34].

ATG GAC ACG AGC AGC AAC TCG AGC ATC GAG AAG TGC AGC TGC GAC AAG TCG GAG CCG TCG	60
M D T S S N S S I E K C S C D K S E P S	20
CAG GAG AGC GAC AGC AGC AAC TCG GAG AGC TCC GCC GCG CCC AAA GAC AAG AAA ATC AAA	120
Q E S D S S N S E S S A A P K D K K I K	40
ATC AGC CTG ACG TCG CTG GGC CTG AAC CTG CGG AGG GGC AGG TCG AAG CGC AAC AAG AAG	180
I S L T S L G L N L R R G R S K R N K K	60
GCC CAC CAA ACG ACG CCG TCC AGC TCT TTC GGT GCC AGC ACC AGC AAA GAG ATC ACC AAG	240
A H Q T T P S S S F G A S T S K E I T K	80
AAG TCG CCG AAA TGG GGC ATC AAG TTC AAC TAC ACC AAA AAA GAA CCG AAA GTG GCC TTC	300
K S P K W G I K F N Y T K K E P K V A F	100
GCC GAG GAG AAC ACG AAC TGT TGC AAG TGC ACG TGC TAC AGG AGG ACG GAG GAG CAC GAG	360
A E E N T N C C K C T C Y R R T E E H E	120
GCG GGG CCG GGT ACG GTG TTT TCG GCT GAA GAA GGC GAA AGG GTT AAC GAT GTT CCT GTA	420
A G P G T V F S A E E G E R V N D V P V	140
GAA GCC GCC AAC GAG GGG GAC TTT GCC GAG AGT AGA GAG GCG GAA CAG CAC CAG AAC GAC	480
E A A N E G D F A E S R E A E Q H Q N D	160
GAA GAC GAC GGG GAA GGG GTC AGA GGG ATA TAT GGC ATG TCC ATG TCT AGG AAT TTG ACG	540
E D D G E G V R G I Y G M S M S R N L T	180
TGT TTG TGC ATC GGT GCG TGG GAC ATG CAC TGG GTG CGC ACC GTT CAC GAA GAC TGC GAT	600
C L C I G A W D M H W V R T V H E D C D	200
GAA GCG GCC AGA ATA GCC AGG GCG CGC GAA ATC GAG CAG GGC GTC GAA GCG CCC GCG AAC	660
E A A R I A R A R E I E Q G V E A P A N	220
TTC CGA CCC GTC AGG CGA GTG CAA CTC TTG TGT TCC GAC ATG GAC TCC GAC GGC GGT CAA	720
F R P V R R V Q L L C S D M D S D G G Q	240
GTG CCG AGA AGG TTG CAG CTG GTA TGC CCT TCG GAC CTG ACC GTC GAC TCC CTG CGG GCG	780
V P R R L Q L V C P S D L T V D S L R A	260
CTC TTC CAG AAC CAC GTG ACT CTG CGA TCC TGT CCC TTG GAC ACC ATC GCC GGC TGT CAC	840
L F Q N H V T L R S C P L D T I A G C H	280
ACT CAG GTG GAC TTC ATA CAC TGT CTC GTA CCT GAT CTC TTG GAC ATC ACC AAG TGC AGC	900
T Q V D F I H C L V P D L L D I T K C S	300
TTC TAC TGG GGT AAG ATG GAC CGC TAC GAG GCC GAG AGG CTG CTG GAC GGC AAG CCG GAA	960
F Y W G K M D R Y E A E R L L D G K P E	320
GGC ACG TTC CTC CTG CGG GAC TCG GCC CAA GAG GAG TTC CTC TTC AGC GTG TCG TCG CGC	1020
G T F L L R D S A Q E E L F S V S F R	340
AAG TAC AGC CGC TCC CTG CAC GCC AGG ATA GAA CAG TGG AAT CAC AAG TTC AGT TTC GAT	1080
K Y S R S L H A R I E Q W N H K F S F D	360
TCT CAC GAT CCC GGC GTC TAC ACG TCG GAC ACA GTT TGT GGC CTT ATC GAA CAT TAC AAA	1140
S H D P G V Y T S D T V C G L I E H Y K	380
GAT CCT AGT AGT TGC ATG TTT TTC GAG CCC ATG CTC ACC TGG CCC CTG CAC AGG AAC TTC	1200
D P S S C M F F E P M L T W P L H R N F	400
ACG TTT TCC CTG CAA CAC CTG TGC AGG GCC GTG ATC GTC AAC CGT TTG TCC TAC GAC AAC	1260
T F S L Q H L C R A V I V N R L S Y D N	420
ATC AAC CTC CTG CAG TTG CCC AAG ACG CTG AAG AGT TAC CTC AAG GAG TAC CAC TAC AGG	1320
I N L L Q L P K T L K S Y L K E Y H Y R	440
CAG AAA GTG AGA GTG GAG AGG TTC GAC GAC GAC ATG CAG TGG CTG GAT TTG AGG AAG GTA	1380
Q K V R V E R F D D D M Q W L D L R K V	460
TCT TTG TAA	1389
S L *	462

Figure 1. The nucleotide and deduced amino acid sequence of *Tenebrio molitor* SOCS5 (TmSOCS5). The Src Homology 2 (SH2) and suppressor of cytokine signaling (SOCS) box domains are shown in black and blue boxes, respectively. The polypeptide binding sites, including the phosphotyrosine binding pocket, hydrophobic binding pocket, and putative elongin B/C interaction residues, are shown in red, blue, and green text, respectively.

ATG GAT CAT CCG GAG AGC AAC AAA AAC AAA GCG AAA AAC TGG CTG TAT CGC TTC TTG AAG	60
M D H P E S N K N K A K N W L Y R F L K	20
ATA AAG AGA AGC CTG CAG GAC AGG TCG AGC ATC TCC GAG GTA GAC CCC AGC ATC TAC AGG	120
I K R S L Q D R S S I S E V D P S I Y R	40
AGG GAA GTT TTG GAG CAC TCG GAG AGT CGA TCC CAC CGC CAC AGC TTC CGC AGG AGT TTC	180
R E V L E H S E S R S H R H S F R R S F	60
ACC CTG AAG CGG TGG CAC ACC AAG GTA AAA AGC TGC ATG TTC CCG CAA GAG GCG CCC ACC	240
T L K R W H T K V K S C M F P Q E A P T	80
CCG AAG CGT CCG GTC ACC ATA TCC GAG GCC CCT TGC ACG TCC GGC CTG CCC GAC TTG CCA	300
P K R P V T I S E A P C T S G L P D L P	100
CCG CCT CCC GTC CCT GAG CGC AAC AGC CAG TGG ACG GAG CGC CGC CAA GCG CCG TCG TCC	360
P P P V P E R N S Q W T E R R Q A P S S	120
CGC GAG TCC ACC CCC ACG GAG ACG CCC AAC CAT CGA ATA CCG ACC GAC GAG ATC GTG AGC	420
R E S T P T E T P N H R I P T D E I V S	140
CTT TCG AAT TGC TAC TGG TAC TGG GGG CCG ATG GCC ACC GCC GAC GCC GAG GAG CGC CTC	480
L S N C Y W Y W G P M A T A D A E E R L	160
CAG TTC AAG CCC GAC GGC ACC TTC CTG GTC CGG GAC TCG TCC TCC ACC TCG TAC TTG TTC	540
Q F K P D G T F L V R D S S S T S Y L F	180
AGC ATC AGC TTC CGC AGC GTG GGC AAG ACG ATG CAC GCC AGG ATC GAA TAC AGC CGC GGC	600
S I S F R S V G K T M H A R I E Y S R G	200
AGG TAC AAC CTG TGC GGG ACC TAC TCC GAG GGC TTC TCC ACC GTC ACG GAG CTG ATC CAG	660
R Y N L C G T Y S E G F S T V T E L I Q	220
GAC GCG ATG AAG ACC TCC GAG AAC GGC GTC TAC TGC TAC TCG AGG CGG GGC GAA CAG ACG	720
D A M K T S E N G V Y C Y S R R G E Q T	240
TGC GTG GAG TTC CCG GTC AGG CTC ACC AAA CCC ATA TCC AGG TAC ACC GAA GTG CGC TCC	780
C V E F P V R L T K P I S R Y T E V R S	260
CTC AAA TAC CTG TGC AAG TTC GTC ATA AGA CAG TGC ACC AAC ATT AAC GAC ATA CCG AAA	840
L K Y L C K F V I R Q C T N I N D I P K	280
TTG CCT CTG CCC ACC GTG CTT CAC GGT TAT CTC CTG GAG AAG CCG TAC TTT TAG TAA	897
L P L P T V L H G Y L L E K P Y F * *	297

Figure 2. The nucleotide and deduced amino acid sequence of *T. molitor* SOCS6 (TmSOCS6). The SH2 and SOCS box domains are shown in black and blue boxes, respectively. The phosphotyrosine binding pocket, hydrophobic binding pocket, and putative elongin B/C interaction residues are highlighted in red, blue, and green text, respectively.

ATG TAT GCC GTT CCG GTT GAT GTT ATT AAA CCA CCC TTG AAG CCC AAG AGG AAT CAA CAA	60
M Y A V P V D V I K P P L K P K R N Q Q	20
AAT AAG AAA CGA AGA AGA AAC ACG TCG TCA GGA TGT AAG GAA ATT GAA GCG CGA CAG TAC	120
N K K R R R N T S S G C K E I E A R Q Y	40
AGG GTT CTT TCA AGC AAG AAA AAT ACC CTC AGA TCT GAG AGG TAT CAT TGG TGT TTA CAT	180
R V L S S K K N T L R S E R Y H W C L H	60
CAA GAA AAC AAA CGA CAC AGT GTC GCC GGT ACA TCT GCT CCG GTA GAA GGG GAA CCC ATC	240
Q E N K R H S V A G T S A P V E G E P I	80
CAC ATG ACT TTA CAA GAA GTC CGC CAT TAT CTT CAA ACG CTA TAT TCA AGT TCA TCC GAT	300
H M T L Q E V R H Y L Q T L Y S S S S D	100
TCA TCC GAT CAC AAA GAG AGA AGC TAC AGA CCA AAA CCT CCG TTA ATC GTC GCA ACA GAC	360
S S D H K E R S Y R P K P P L I V A T D	120
AAC AAG TGT ACA ACA GAA ACA AAC AAC ACT CAT GTA AAT AGA GAT GTT AAT TGC GTA TCA	420
N K C T T E T N N T H V N R D V N C V S	140
AAC GCA AAC GGA AGA ACG AAA AAG AGT ACA TTC CTT ATT AAC ATT AAG AAC AAA AAA GTG	480
N A N G R T K K S T F L I N I K N K K V	160
AAA GAC TCC TGT GAT AAT GTC GGT AAA CAG ATA AAT ATT TCG CAA ACG AAT AAG CCA GAC	540
K D S C D N V G K Q I N I S Q T N K P D	180
AAA CGC AAA AAG ACA CCA GCG AAG TTG TTT TCG TTT AAG CAG ACC TTA TGT AAT ATG TTC	600
K R K K T P A K L F S F K Q T L C N M F	200
AGA TTC AAG CGG TTT TTG TCA CCT GAA CAC GTA AAA TGT GAT GTT AGC GAC GTA CAA GAA	660
R F K R F L S P E H V K C D V S D V Q E	220
AAA ACC AAT TAC ATT AGC AAT GCC AGT TTG GTG GAG GAA GCG CAC AAC AAC ATA TCC AGC	720
K T N Y I S N A S L V E E A H N N I S S	240
AGG GCG TTG CCT CCT TTG CCT TTG AAA GAA GAG GAG GAA GTC TCA GAG GAA CCT ATC TTG	780
R A L P P L P L K E E E E V S E E P I L	260
GAT TTT GCC ACC AGC ATC CAA AGA GTC AAA GAC TAT GGC TGG TAC TGG GGT CCA TTA CCG	840
D F A T S I Q R V K D Y G W Y W G P L P	280
AGC GAA GTC GCC GAA AAA ATC TTG TCC AAC GAA CCA GAC GGA TCA TTT ATT GTA CGT GAT	900
S E V A E K I L S N E P D G S F I V R D	300
AGT AGC GAC GAT CAC TAT ATT TTT TCA TTA ACT TTT AAG CTG AAC AAC TGC GTT AGA CAC	960
S S D D H Y I F S L T F K L N N C V R H	320
GTC AGA ATT GAT CAT GAT CAG GGT AAC TTT AGT TTC GGT AGT TGT ACA AAG TTC AAG TCG	1020
V R I D H D Q G N F S F G S C T K F K S	340
CAA ACA ATC GTA GAG TTT ATA GAG AAT GCG GTG GAA CAT TCT CGT AGC GGT AGG TAT TTG	1080
Q T I V E F I E N A V E H S R S G R Y L	360
TTC TTT TTG CAC CGG AGG CCC GTG ATT GGG CCG GTG AGA GTG CAG TTA TTA CAT CCC GTG	1140
F F L H R R P V I G P V R V Q L L H P V	380
TCC AGG TTT AAG CAG GTG CAG AGC TTA CAG CAT ATC TGC AGA TTT GCC ATA CAC AAA GTG	1200
S R F K Q V Q S L Q H I C R F A I H K V	400
GTG CGA AGG GAT TTA ATT CCC AGT CTT CCA TTA CCA AGA AGA ATG ATC GAT TAT TTG AAT	1260
V R R D L I P S L P L P R R M I D Y L N	420
ACG CCA CAT TAT TAT TCA GAG CAT CTG ATC GAC GTA GAA GAT ACA AGT AAT GAA CAA GTC	1320
T P H Y Y S E H L I D V E D T S N E Q V	440
CCA CAA GTC CCT CCT AGA AAT CCC ACC CCA GCA ATT CGT CGT GAT TAT CAA GAA GAG GAA	1380
P Q V P P R N P T P A I R R D Y Q E E E	460
GTC CAA GCA CTC GTC GTT CCC AAC GTA CCT TTA AAC AAT TAT GTT GTT TTA AAT AAC ACA	1440
V Q A L V V P N V P L N N Y V V L N N T	480
AAT CCC AGA CAA TCC TGA	1458
N P R Q S *	485

Figure 3. The nucleotide and deduced amino acid sequence of *T. molitor* SOCS7 (TmSOCS7). The conserved SH2 and SOCS box domains are shown in black and blue boxes, respectively. The phosphotyrosine binding pocket, hydrophobic binding pocket, and putative elongin B/C interaction residues are highlighted in red, blue, and green text, respectively.

The multiple sequence alignment (MSA) and percent identity matrix of the C-terminal region of TmSOCS5 and included insect orthologs are shown in Figure 4. The full-length amino acid sequences of the *T. molitor* SOCS proteins were not included in the phylogenetic analyses due to high level of sequence divergence within the N-terminus. As indicated in previous studies, the N-terminal region of SOCS proteins is largely disordered, and a region of ~70 residues is conserved across different species and between SOCS4 and SOCS5 [35,36]. Therefore, alignment of the conserved SH2 and SOCS box domains was deemed appropriate for understanding the sequence diversity and evolutionary position of *T. molitor* SOCS type I subfamily proteins. The SH2 domain contains numerous identical residues across species, whereas the SOCS box domain possesses lower, but still high levels of similarity (Figure 4A). TmSOCS5 is most similar to TcSOCS5 (97% identity), followed by OtSOCS5 from the dung beetle, *Onthophagus taurus* (88% identity) (Figure 4B). TmSOCS6 exhibited less similarity in the SH2 and SOCS box domains, but it was most similar to TcSOCS6 (91% identity; Figure 5). This protein exhibited significantly less similarity to other insect orthologs, with percent identities of 35%–55%. BmSOCS6 however, exhibited sequence similarity to lepidopteran SOCS6 family members, and *E. sinensis* SOCS6 possessed high levels of sequence similarity to other molluscan SOCS family members [20,21]. In contrast, high levels of sequence identity and similarity were observed in the SH2 and SOCS box domains of the SOCS7 homologs (Figure 6A). Interestingly, the SH2 and SOCS box domains of TmSOCS7 possessed 100% identity to TcSOCS7, and 74%–86% identity to other insect species (Figure 6B). Additionally, TmSOCS5, TmSOCS6, and TmSOCS7 were closely related to TcSOCS5, TcSOCS6, and TcSOCS7, respectively.

3.2. Molecular Evolutionary Relationships of the TmSOCS Proteins

To explore the phylogenetic relationships of the *T. molitor* type I SOCS proteins, we constructed a phylogenetic tree using the amino acid sequences of insect SOCS proteins (Figure 7). Human SOCS1, which belongs to the type II subfamily, was used as an outgroup. The phylogram indicated three conspicuous clusters of type II SOCS proteins comprising of the SOCS5, SOCS6, and SOCS7 proteins. The SOCS5 and SOCS7 proteins formed more compact clusters than those of SOCS6 cluster. As expected, TmSOCS5, TmSOCS6, and TmSOCS7 were closely related to TcSOCS5, TcSOCS6, and TcSOCS7, respectively. Additionally, the SOCS5 protein from another coleopteran species, OtSOCS5 was found to be closely related to TmSOCS5 and TcSOCS5. The TmSOCS6 cluster consisted of two sub-clusters. We described earlier, through MSA and percent identity analysis, that the SH2 and SOCS box domains of TmSOCS6 proteins are comparatively less well conserved than those of the TmSOCS5 and TmSOCS7 proteins. Also, the longer N-terminal region of the insect SOCS6 proteins, similar to vertebrate SOCS6 proteins, is thought to be involved in nuclear localization of this protein [37]. These results are consistent with similar report examining *E. sinensis* SOCS6 [20]. The observed close relationship of *Bombyx mori* SOCS6 (BmSOCS6) and *Danaus plexippus* SOCS6 (DppSOCS6) proteins is also consistent and is consistent with previous findings [22].

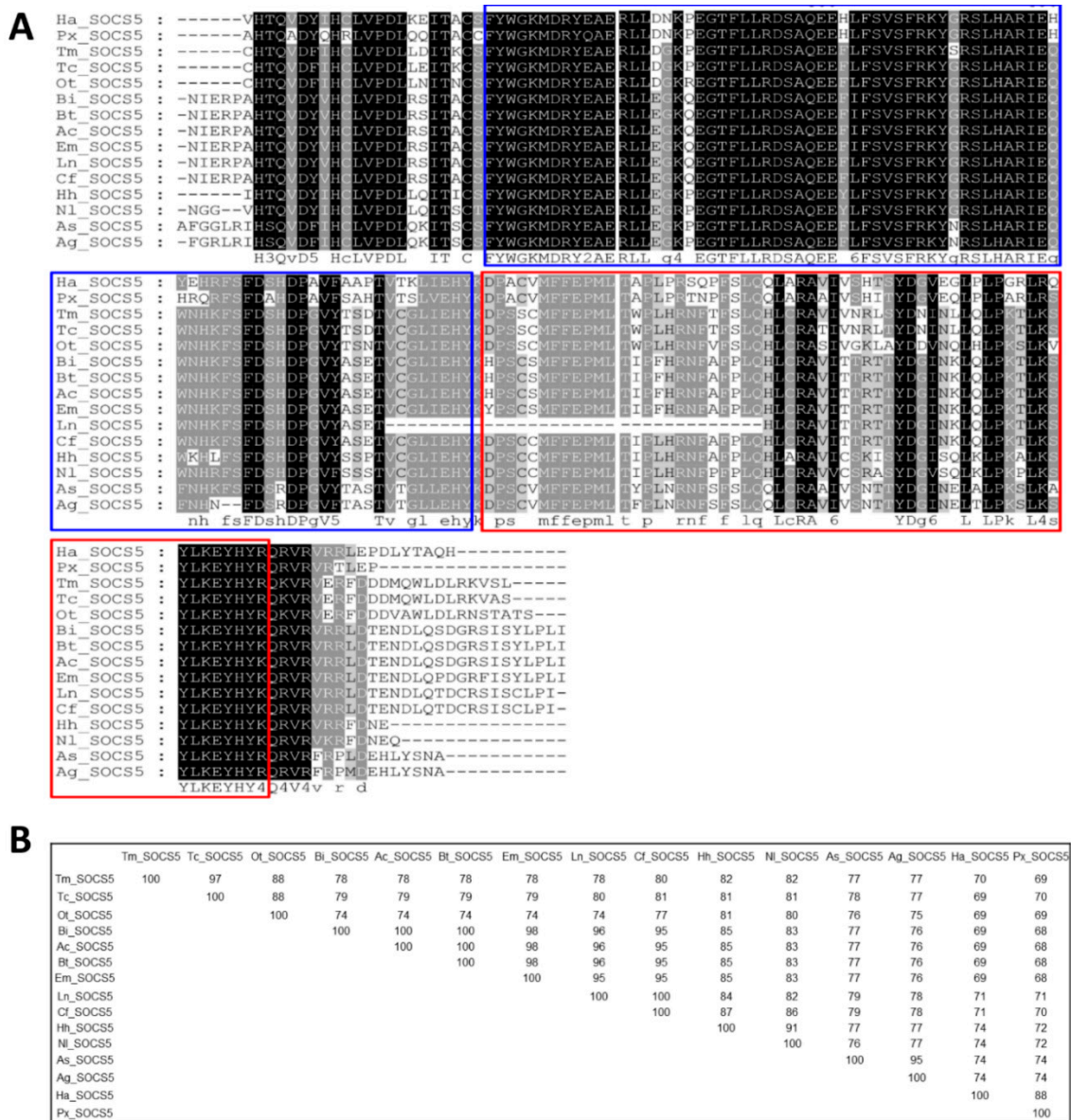


Figure 4. Alignment of TmSOCS5 to other SOCS5 proteins and percent identities. (A) Multiple sequence alignment of the SOCS5 C-terminal region of insect orthologs. The SH2 and SOCS box domains are shown in blue and red boxes, respectively. Identical residues in all sequences are shaded black, and well-conserved sequences are shaded grey. Deletions are indicated by dashes; (B) Percent identity matrix of TmSOCS5 and representative SOCS5 proteins. The analysis was performed by clustalX2 using representative amino acid sequences from *Tribolium castaneum* (XP_015833443.1), *Plutella xylostella* (XP_011566498.1), *Helicoverpa armigera* (XP_021196679.1), *Onthophagus taurus* (XP_022915891.1), *Bombus impatiens* (XP_012245047.1), *Bombus terrestris* (XP_012163125.1), *Apis cerana* (AEY61566.1), *Eufriesea mexicana* (OAD52276.1), *Lasius niger* (KMQ91264.1), *Camponotus floridanus* (EFN62367.1), *Halyomorpha halys* (XP_014279177.1), *Nilaparvata lugens* (XP_022206249.1), *Anopheles sinensis* (KFB35251.1), and *Anopheles gambiae* (ABV01933.1).

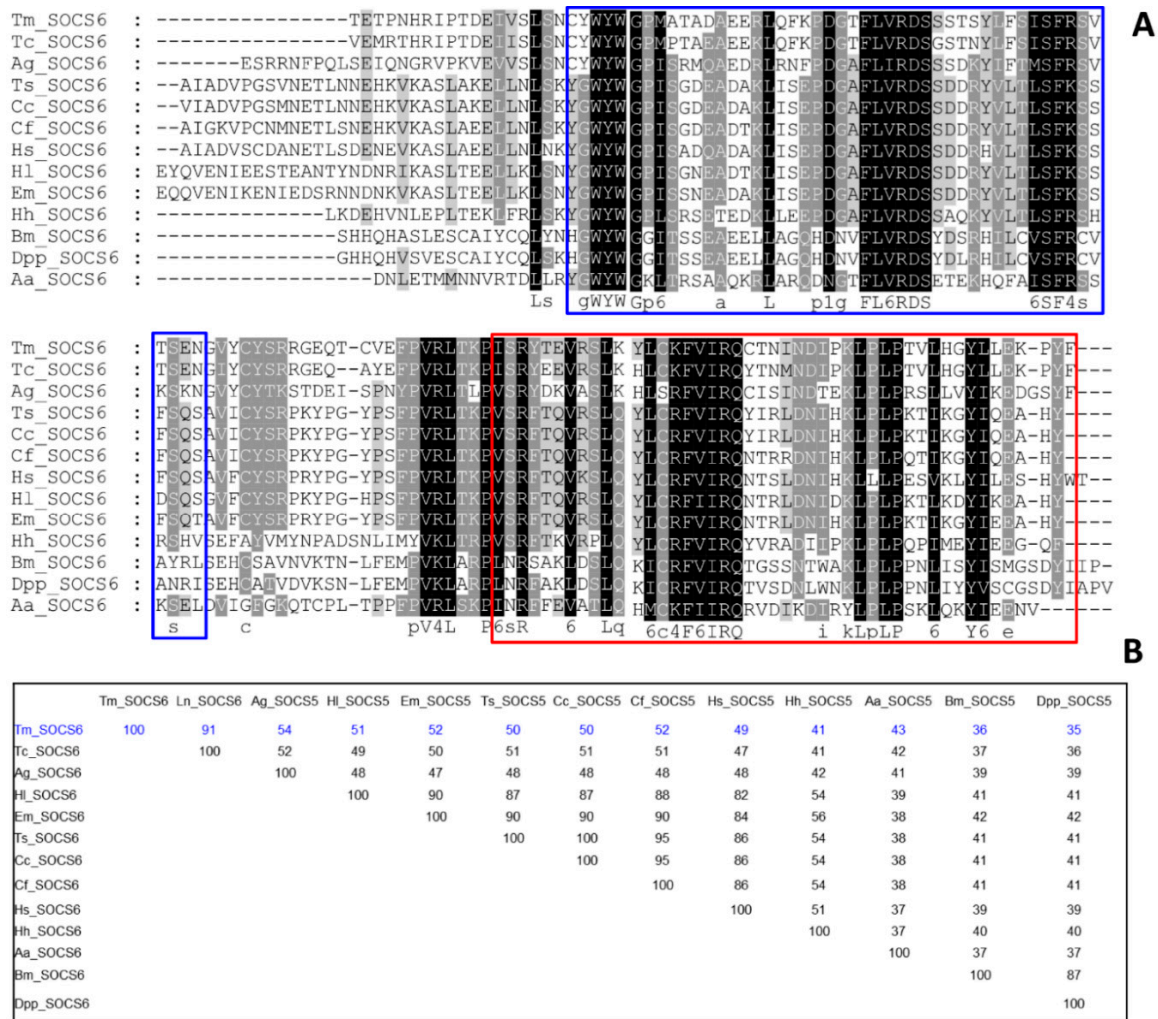


Figure 5. Alignment of TmSOCS6 to other SOCS6 proteins and percent identity: (A) Multiple sequence alignment of insect SOCS6 to insect proteins. Only the C-terminal region was conserved. SH2 and SOCS box domains are shown in blue and red boxes, respectively. Identical residues in all sequences are shaded black, and similar sequences are shaded grey. Deletions are indicated by dashes; (B) Percent identity matrix of SOCS6 members from representative insect species. Analysis was performed by clustalX2 using representative amino acid sequences from *Tribolium castaneum* (XP_008190646.1), *Anopheles gambiae* (JAB61954.1), *Trachymyrmex septentrionalis* (KYN39913.1), *Ceratitis capitata* (KYN07884.1), *Camponotus floridanus* (EFN74396.1), *Harpegnathos saltator* (EFN81362.1), *Habropoda laboriosa* (KOC63154.1), *Eufrisea mexicana* (OAD52403), *Halymorpha halys* (XP_014279265.1), *Bombyx mori* (NP_001185652.1), *Danaus plexippus plexippus* (OWR49085.1), *Aedes aegypti* (XP_001660156.3).

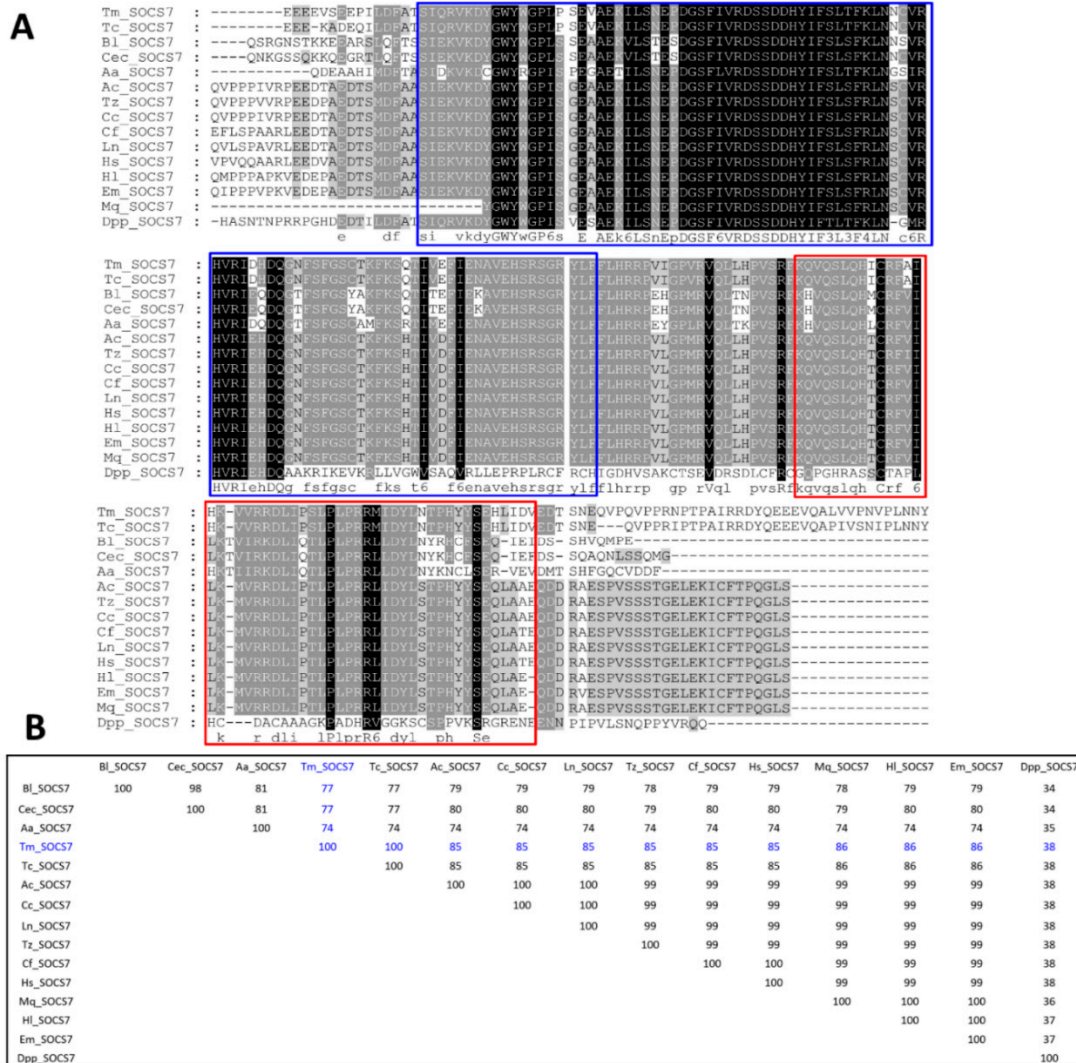


Figure 6. Alignment and identity of TmSOCS7 to other SOCS7 proteins: (A) Multiple alignment of the SOCS7 C-terminal sequences of insect orthologs. The SH2 and SOCS box domains are shown in blue and red boxes, respectively. Identical residues in all sequences are shaded black, and similar sequences are shaded grey. Deletions are indicated by dashes; (B) Percent identity matrix of SOCS7 proteins using the amino acid sequences from representative insect species. ClustalX2 was used to generate the identity matrix. The sequences used in the analysis were from *Tribolium castaneum* (XP_008190646.1), *Bactrocera latifrons* (JAI49983.1), *Ceratitis capitata* (JAC02138.1), *Aedes aegypti* (XP_021707758.1), *Acromyrmex echinaiator* (EGI60822.1), *Trachymyrmex zeteki* (KYQ56054.1), *Cyphomyrmex costatus* (KYM98051.1), *Camponotus floridanus* (EFN69786.1), *Lasius niger* KMQ86622), *Harpegnathos saltator* (EFN83798.1), *Habropoda laboriosa* (KOC61557.1), *Eufrisea mexicana* (OAD58426.1), *Melipona quadrifasciata* (KOX74779.1), and *Danaus plexippus plexippus* (OWR46815.1).

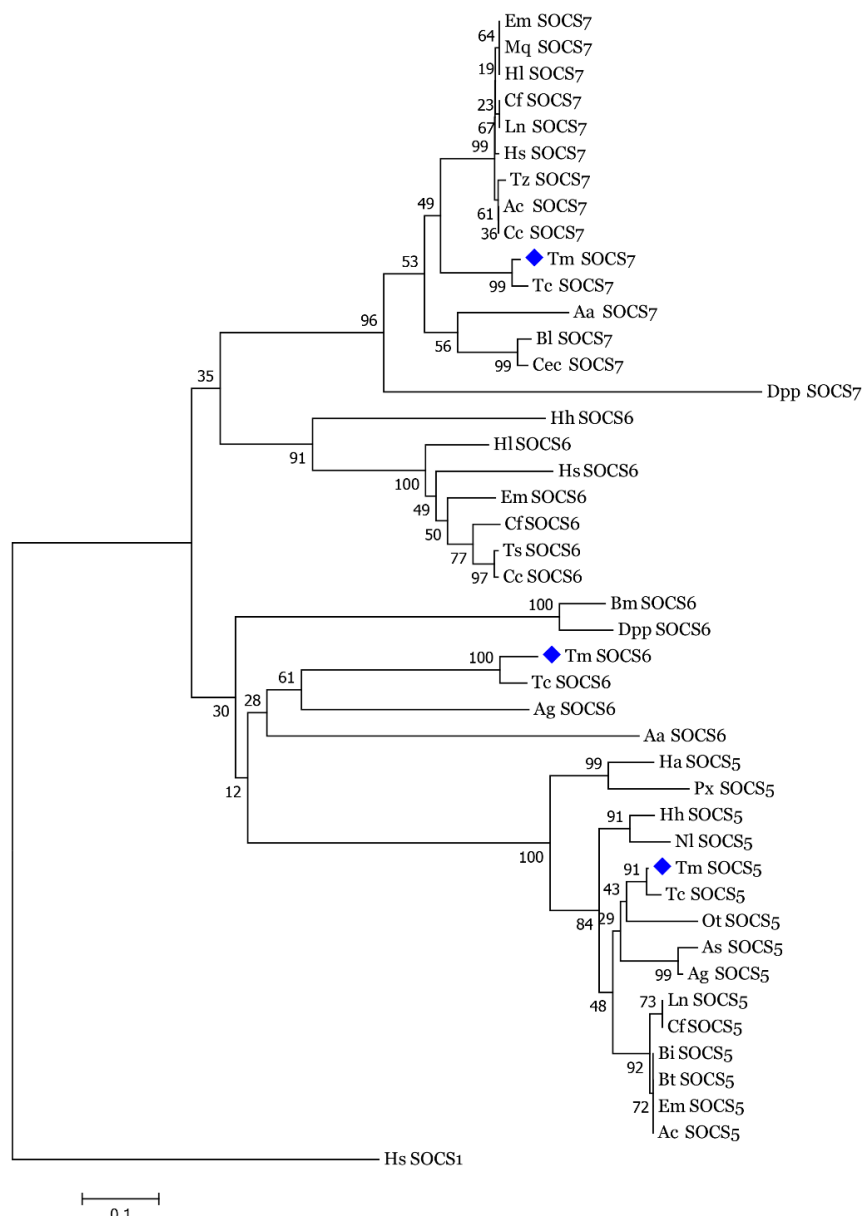


Figure 7. Phylogenetic analysis of TmSOCS5, TmSOCS6, and TmSOCS7 and insect SOCS5, SOCS6, and SOCS7 proteins. The tree was constructed in MEGA7 using the neighbor-joining method. Numbers at the nodes are the bootstrap support of 1000 replicates. *Homo sapiens* SOCS1 (Type II member) was used as an outgroup. ♦ indicates *T. molitor* SOCS5, SOCS6, and SOCS7 proteins.

3.3. Developmental Expression of TmSOCS Genes

We used qRT-PCR analysis to quantify TmSOCS5, TmSOCS6, and TmSOCS7 mRNA expression across the developmental (egg, early larva, late larva, pre-pupa, pupa [days 1–7] and adult [days 1–5]) stages and within various tissues (Figure 8). We hypothesized that there may be shifts in SOCS gene expression during the transitions between developmental stages and within a single developmental stage. Genome-wide screens have earlier been used to analyze differential expression of genes during different stages of insect development to understand the mechanisms of genic interplay regulating insect development [38,39]. TmSOCS5, TmSOCS6, and TmSOCS7 gene sequences were also retrieved from the *T. molitor* larval transcriptome database. Here, we were interested to note the expression of TmSOCS during *T. molitor* developmental stages. Our results indicate that TmSOCS5 and TmSOCS6 expression levels were higher at the egg stage. Relative to the expression levels of TmSOCS5 and

TmSOCS6 in the egg stage, the mRNA expression in the larval, pupal, and adult stages of the insect were found to be lower. (Figure 8A,B). Moreover, we found a more consistent expression of TmSOCS5 and TmSOCS6 at the metamorphosing stages. In contrast, TmSOCS7 expression levels were relatively higher at the larval, pupal, and adult stages in comparison to that of the egg stage (Figure 8C). Further, in the present study, the adult insects used for the expression study were not sexed, and therefore a sex-specific relationship of TmSOCS genes could not be established. Such a relationship has been previously observed in the mosquito (*Anopheles culicifacies*), where SOCS mRNA was found to be highly expressed in the male relative to expression in female adults. It was also observed that although biases in the expression of genes begin at the larval stage, the expression of SOCS genes does not change between male and female larvae [40]. The involvement of TmSOCS gene in sexual dimorphism is more emphasized in the tissue-specific expression analysis of *T. molitor* adults. *Drosophila* SOCS36E, which has high levels of sequence identity with vertebrate SOCS5 (75% and 44% in the SH2 and SOCS-box domains, respectively), is known to be expressed during embryogenesis (especially during embryonic and imaginal disc development) [13]. Additionally, SOCS36E regulates STAT activity levels through either Cullin2 (Cul2) scaffolding protein-dependent or independent mechanisms in the egg chamber of *Drosophila* [41]. Relative expression levels of the *A. culicifacies* SOCS gene (AcSOCS) were also high at the egg stage signifying JAK-STAT control of embryogenesis [40]. Further, we assumed that SOCS5 and SOCS6 could be an important player in the embryogenesis process, but these proteins would exert limited function in the innate immunity of *T. molitor*. Preferentially, SOCS7 could be more involved in cellular homeostasis mechanisms during infection and immunity processes in the host insect. Supporting evidence reflects the involvement of SOCS7, and not SOCS5, in triggering the activation of the NF- κ B signalling pathway in the mollusc, *C. gigas* [30].

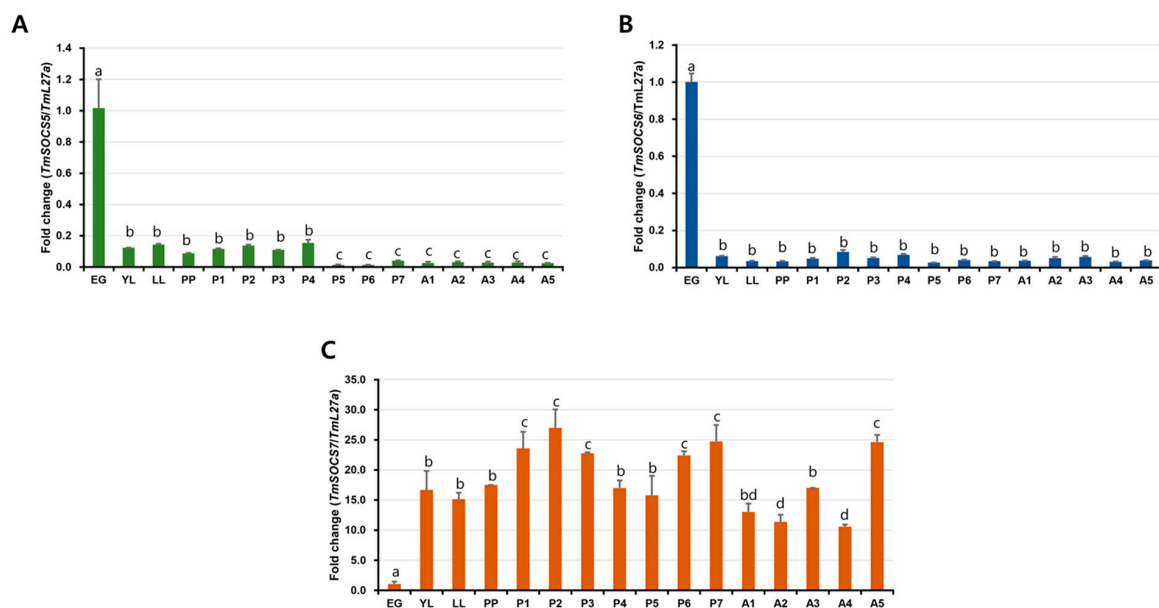


Figure 8. Developmental expression patterns of (A) SOCS5, (B) SOCS6, and (C) SOCS7 transcripts in *T. molitor*. Relative expression of SOCS mRNA in the egg (EG), early larval (YL), late larval (LL), day 1–7 pupal (P1–P7), and day 1–5 adult (A1–A5) stages. Transcript levels were analyzed by quantitative real-time PCR. RpL27a (*T. molitor*) was included as an internal control to normalize RNA levels between samples. Results represent mean \pm S.E. The experiment was performed three times with similar results. Different letters (a, b, c, d, and bd) above the bars denotes significant differences ($p < 0.05$).

3.4. Tissue-Specific Expression of *T. molitor* SOCS Genes

The tissue-specific expression analysis of *T. molitor* larvae indicated that the mRNA levels of TmSOCS5 and TmSOCS6 were higher in the hemocytes than those observed in the other tissues

(Figure 9AI,BI), while TmSOCS7 levels were elevated in Malpighian tubules (Figure 9CI). High expression levels in hemocytes were also observed for SOCS5 and SOCS7 mRNA in the pacific oyster *C. gigas* [30] and for SOCS6 mRNA in the Chinese mitten crab *E. sinensis* [20]. Hemocytes and fat bodies were the principal target tissues for the expression of SOCS6 mRNA in *B. mori* [22]. In adults, the expression levels of TmSOCS5, TmSOCS6, and TmSOCS7 transcripts were higher in the ovaries (Figure 9AII, BII, and CII). This indicates that *T. molitor* SOCS transcripts show sexually dimorphic gene expression patterns. It is interesting to observe the active involvement of TmSOCS in the ovarian processes. We understand that the female-biased genes are more conserved than the male-biased genes, and this observation is unique. Reports from the developmental expression profile of SOCS genes in the mosquito species *A. culicifacies* and *Anopheles gambiae* suggest a male bias [40,42]. Further, the sexual dimorphism of the JAK-STAT pathway substrate “STAT5” (paralogs STAT5a and STAT5b) has been analyzed in mouse models. STAT5b-deficient mice exhibit sexually dimorphic growth, with downregulation of female-specific proteins and lower gene expression in males [43]. Sex-biased genes and their transcriptional regulation have both been documented in the malaria vector *A. gambiae* using comparative genomics and transcriptomics [42]. TmSOCS transcript expression in immune-related tissues, such as the fat body and gut, was constitutive. It is possible that SOCS proteins exert different biological functions in different tissues. This was also observed for the expression of SOCS6 in the hemocytes, fat body, and Malpighian tubules of *B. mori* larvae [22]. Additionally, higher expression levels of SOCS6 in hemocytes have also been reported in other invertebrate and vertebrate species [20,43,44].

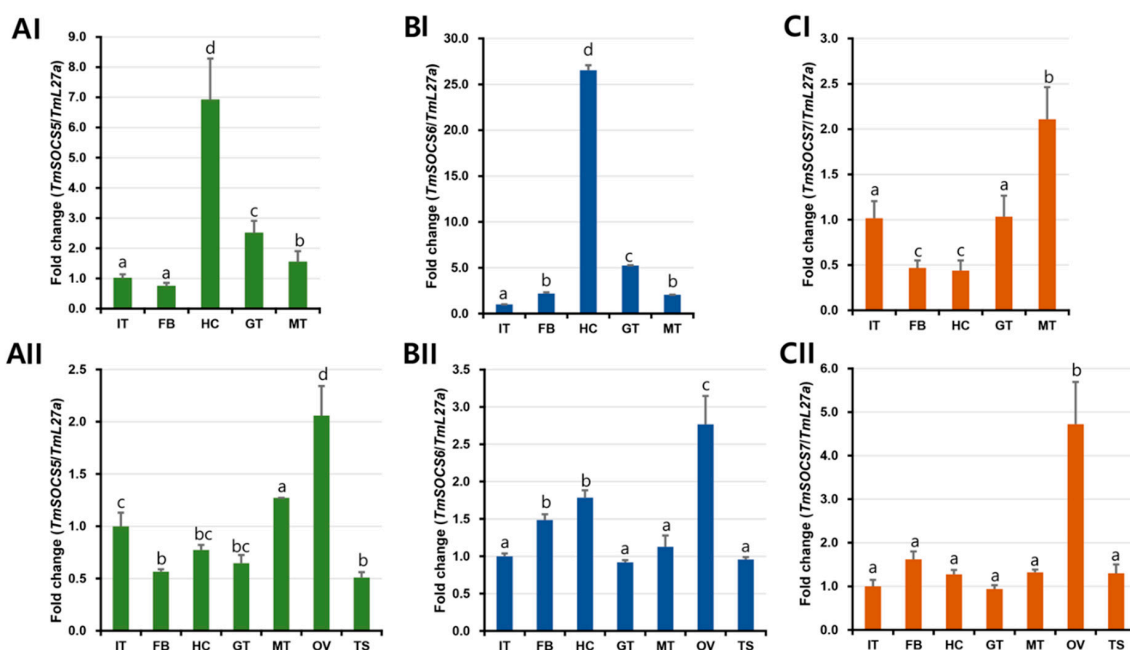


Figure 9. Tissue distribution of TmSOCS transcripts in *T. molitor*: (A) Distribution of TmSOCS5 transcripts in larval (AI) and adult tissues (AII); (B) distribution of TmSOCS6 transcripts in larval (BI) and adult tissues (BII); (C) distribution of TmSOCS7 transcripts in larval (CI) and adult tissues (CII). Tissues are abbreviated as follows: IT, integument; FB, fat body; HC, hemocytes; GT, gut; MT, Malpighian tubules; OV, ovary; TS, testis. RpL27a (*T. molitor*) was included as an internal control to normalize RNA levels among samples. Vertical bars represent mean \pm S.E. Different letters (a, b, c, d, and bc above bars) indicate significant differences among groups ($p < 0.05$).

3.5. Expression of TmSOCS Genes after Immune Stimulation

In a time-course study conducted over a 2-day period, we examined TmSOCS5, TmSOCS6, and TmSOCS7 mRNA expression in the hemocytes, fat body, and gut tissues of *T. molitor* larvae

after infection with *E. coli*, *S. aureus*, or *C. albicans*. The fat body and hemocytes influence the production of inducible immune effectors and phagocytes in insects, and the gut is involved in specific immune reactions. Therefore, these are the preferred tissues for studies of antimicrobial responses and phagocytosis related to innate immunity in insects. TmSOCS5 expression in the fat body was highest at 12 h post-infection with the fungus *C. albicans* (Figure 10A). In the gut tissues (Figure 10B) and hemocytes (Figure 10C), TmSOCS5 expression was induced 6 and 9 h after fungus challenge, respectively. A similar expression profile was noted for TmSOCS6 in the fat body (Figure 10D), gut (Figure 10E), and hemocytes (Figure 10F) of *T. molitor* larvae after microorganism challenge. TmSOCS6 expression was significantly upregulated ($p < 0.05$) in hemocytes at 3 h post-infection. Generally, immune expression at the early stage following infection is useful for mounting an appropriate response to the pro-inflammatory cytokines induced in most tissues; however, in *Bombyx mori*, SOCS6 induction in hemocytes was delayed and was significantly higher at 24 and 48 h post-challenge with the fungus *Beauveria bassiana* and *E. coli*, respectively [22]. In the fat body, however, BmSOCS6 was significantly induced at 24 h post-infection with *E. coli*, which is in agreement with our results. The differential expression profile of SOCS6 mRNA in *T. molitor* and *B. mori* hemocytes after exposure to two different fungal pathogens may explain the role of *C. albicans* as a potent immune elicitor for regulatory genes in insects. Additionally, the relative mRNA expression of EGFR pathway-related genes including *fkhr*, *gsk3*, *ras*, and *erk*, was strongly induced 4 h after injection with recombinant BmSOCS6 protein [22]. Regulation of the EGFR pathway by *Drosophila* SOCS44A (34% identity with human SOCS6 gene) and *B. mori* SOCS6 demonstrate its importance in the control of developmental and pathophysiological processes [22,45]. TmSOCS7 was significantly upregulated in the tested tissues ($p < 0.05$) after challenge with the Gram-negative bacteria *E. coli*, the Gram-positive bacterium *S. aureus*, and the fungus *C. albicans* (Figure 10G–I). As previously observed, *C. albicans* induced a 15-fold increase in TmSOCS7 levels 3 h post-infection. Although studies examining the function of SOCS7 are limited, one study demonstrated that this protein influences STAT3 and STAT5 nuclear translocation [46]. As STAT3 is a transcriptional regulator of IFN- β and interleukin 6, it is understood that SOCS7 participates in the interferon regulatory pathway. The expression of SOCS5, SOCS6, and SOCS7 post-infection with intracellular and extracellular bacterial pathogens has been studied in tongue sole (*Cynoglossus semilaevis*) [47]. CsSOCS7 mRNA was expressed at 6 h when compared with the expression profile of CsSOCS5 and CsSOCS6 mRNA post-infection with bacteria. Further, differential expression of SOCS mRNA was also observed. These results, combined with our observations, suggest that all SOCS genes exhibit enhanced expression post-challenge with microorganisms; however, these expression mechanisms are differentially regulated.

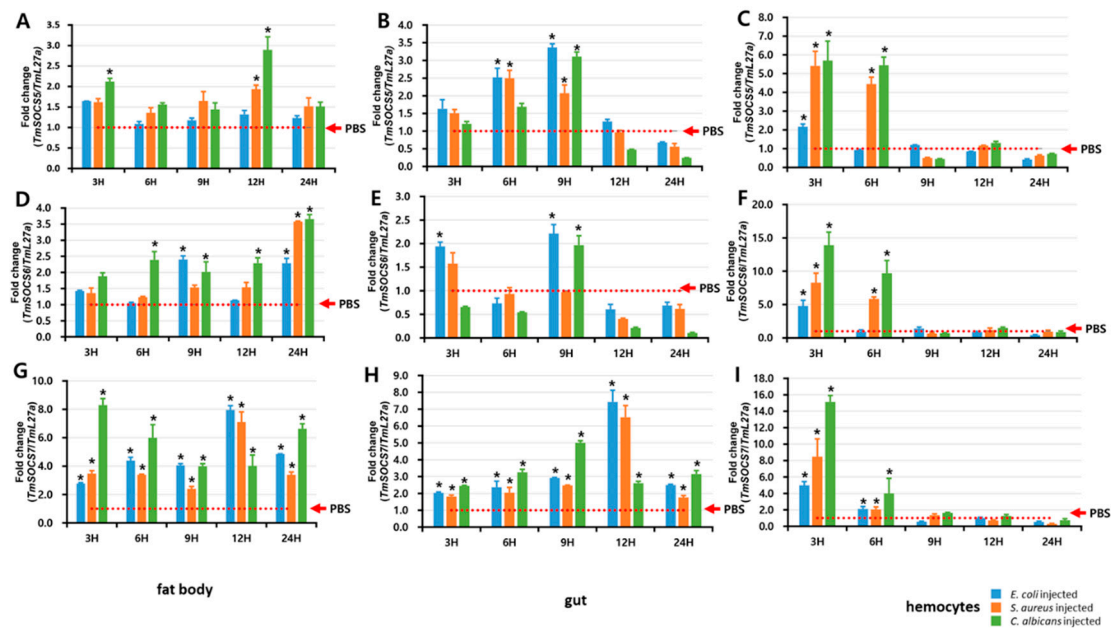


Figure 10. Expression of TmSOCS5, TmSOCS6, and TmSOCS7 mRNA in the fat body (A,D,G), gut (B,E,H), and hemocytes (C,F,I) of *T. molitor* following challenge with *E. coli*, *S. aureus*, and *C. albicans*. Expression was analyzed by real-time PCR using RplL27a as the control for normalization. For each time point, the expression level in the PBS-injected control was set to 1, and this is represented by a dotted line. Values are the mean of three independent measurements \pm S.E. ($n = 3$). * $p < 0.05$.

4. Conclusions

This study advances our knowledge of insect immunity by identifying and characterizing three type I SOCS gene family members (SOCS5, SOCS6, and SOCS7) in *T. molitor*. We screened *T. molitor* RNA-Seq and Genome-Seq datasets for type I SOCS family members and we examined the evolutionary relationships to other proteins in the type I and type II subfamilies. The observed upregulation of TmSOCS transcripts in the immune-related tissues of *T. molitor* after microbial challenge suggests that they are critical for immune reactions. In the future, we plan to conduct RNA interference experiments to study the involvement of type I SOCS family members in the regulation of key cytokine regulatory pathways. The currently available information regarding insect SOCS genes is very limited, and this study extends the repertoire of possible negative regulators involved in maintaining cellular homeostasis in insects.

Author Contributions: Conceptualization, supervision and administration of the research study, I.S.B. and Y.H.J.; conducting experiments, data analysis and data curation, B.B.P. and B.B.K.; writing of the manuscript, B.B.P. and Y.H.J.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ghoreschi, K.; Laurence, A.; O’Shea, J.J. Janus kinases in immune cell signaling. *Immunol. Rev.* **2009**, *228*, 273–287. [[CrossRef](#)] [[PubMed](#)]
2. Harrison, D.A. The Jak/STAT pathway. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, a011205. [[CrossRef](#)] [[PubMed](#)]
3. You, M.; Yu, D.H.; Feng, G.S. Shp-2 tyrosine phosphatase functions as a negative regulator of the interferon-stimulated JAK-STAT pathway. *Mol. Cell. Biol.* **1999**, *19*, 2416–2424. [[CrossRef](#)]

4. Kario, E.; Marmor, M.D.; Adamsky, K.; Citri, A.; Amit, I.; Amariglio, N.; Rechavi, G.; Yarden, Y. Suppressors of cytokine signaling 4 and 5 regulate epidermal growth factor receptor signaling. *J. Biol. Chem.* **2005**, *280*, 7038–7048. [[CrossRef](#)] [[PubMed](#)]
5. Seif, F.; Khoshmirisafa, M.; Aazami, H.; Mohsenzadegan, M.; Sedighi, G.; Bahar, M. The role of JAK-STAT signaling pathway and its regulators in the fate of T helper cells. *Cell Commun. Signal.* **2017**, *15*, 23. [[CrossRef](#)] [[PubMed](#)]
6. Hatakeyama, S. Ubiquitin-mediated regulation of JAK-STAT signaling in embryonic stem cells. *Jakstat* **2012**, *1*, 168–175. [[CrossRef](#)] [[PubMed](#)]
7. Starr, R.; Willson, T.A.; Viney, E.M.; Murray, L.J.; Rayner, J.R.; Jenkins, B.J.; Gonda, T.J.; Alexander, W.S.; Metcalf, D.; Nicola, N.A.; et al. A family of cytokine-inducible inhibitors of signaling. *Nature* **1997**, *387*, 917–921. [[CrossRef](#)] [[PubMed](#)]
8. Hilton, D.J.; Richardson, R.T.; Alexander, W.S.; Viney, E.M.; Willson, T.A.; Sprigg, N.S.; Starr, R.; Nicholson, S.E.; Metcalf, D.; Nicola, N.A. Twenty proteins containing a C-terminal SOCS box form five structural classes. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 114–119. [[CrossRef](#)] [[PubMed](#)]
9. Liongue, C.; Sertori, R.; Ward, A.C. Evolution of cytokine receptor signaling. *J. Immunol.* **2016**, *197*, 11–18. [[CrossRef](#)]
10. Li, L.; Gronning, L.M.; Anderson, P.O.; Li, S.; Edvardsen, K.; Johnston, J.; Kioussis, D.; Shepherd, P.R.; Wang, P. Insulin induces SOCS-6 expression and its binding to the p85 monomer of phosphoinositide 3-kinase, resulting in improvement in glucose metabolism. *J. Biol. Chem.* **2004**, *279*, 34107–34114. [[CrossRef](#)] [[PubMed](#)]
11. Martens, N.; Uzan, G.; Wery, M.; Hooghe, R.; Hooghe-Peters, E.L.; Gertler, A. Suppressor of cytokine signaling 7 inhibits prolactin, growth hormone, and leptin signaling by interacting with STAT5 or STAT3 and attenuating their nuclear translocation. *J. Biol. Chem.* **2005**, *280*, 13817–13823. [[CrossRef](#)] [[PubMed](#)]
12. Wang, B.; Wangkahart, E.; Secombes, C.J.; Wang, T. Insights into the evolution of the suppressors of cytokine signaling (SOCS) gene family in vertebrates. *Mol. Biol. Evol.* **2018**. [[CrossRef](#)] [[PubMed](#)]
13. Karsten, P.; Hader, S.; Zeidler, M.P. Cloning and expression of *Drosophila* SOCS36E and its potential regulation by the JAK/STAT pathway. *Mech. Dev.* **2002**, *117*, 343–346. [[CrossRef](#)]
14. Rawlings, J.S.; Rennebeck, G.; Harrison, S.M.W.; Xi, R.; Harrison, D.A. Two *Drosophila* suppressors of cytokine signaling (SOCS) differentially regulate JAK and EGFR pathway activities. *BMC Cell Biol.* **2004**, *5*, 38. [[CrossRef](#)] [[PubMed](#)]
15. Lee, Y.; Choi, J.-Y.; Oh, C.; Kang, D.-H.; Choi, S.-Y.; Heo, G.-J.; Lee, J.; De Zoysa, M. Molecular cloning and characterization of SOCS-2 from Manila clam *Ruditapes philippinarum*. *Fish Shellfish Immunol.* **2014**, *36*, 453–458. [[CrossRef](#)]
16. De Zoysa, M.; Lee, J. Suppressor of cytokine signaling 2 (SOCS-2) homologue in disk abalone: Cloning, sequence characterization and expression analysis. *Fish Shellfish Immunol.* **2009**, *26*, 500–508. [[CrossRef](#)] [[PubMed](#)]
17. Zhang, Y.; Zhao, J.; Zhang, H.; Gai, Y.; Wang, L.; Li, F.; Yang, J.; Qiu, L.; Song, L. The involvement of suppressors of cytokine signaling 2 (SOCS2) in immune defense responses of Chinese mitten crab *Eriocheir sinensis*. *Dev. Comp. Immunol.* **2010**, *34*, 42–48. [[CrossRef](#)] [[PubMed](#)]
18. Zhu, B.; Dai, L.; Yu, Y.; Wang, D.; Peng, T.; Liu, C. A role of suppressor of cytokine signaling 2 in the regulation of ecdysteroid signaling pathway in *Procambarus clarkii*. *J. Exp. Zool.* **2016**, *325A*, 441–452. [[CrossRef](#)] [[PubMed](#)]
19. Wang, S.; Song, X.; Zhang, Z.; Li, H.; Lu, K.; Yin, B.; He, J.; Li, C. Shrimp with knockdown of LvSOCS2, a negative feedback loop regulator of JAK/STAT pathway in *Litopenaeus vannamei*, exhibit enhanced resistance against WSSV. *Dev. Comp. Immunol.* **2016**, *65*, 289–298. [[CrossRef](#)] [[PubMed](#)]
20. Qu, C.; Xu, Q.; Lu, M.; Wang, F.; Liu, Z.; Liu, D.; Yang, W.; Yi, Q.; Wang, L.; Song, L. The involvement of suppressor of cytokine signaling 6 (SOCS6) in immune response of Chinese mitten crab *Eriocheir sinensis*. *Fish Shellfish Immunol.* **2018**, *72*, 502–509. [[CrossRef](#)]
21. Abbas, M.N.; Zhu, B.-J.; Kausar, S.; Dai, L.-S.; Sun, Y.-X.; Tian, J.W.; Liu, C.-L. Suppressor of cytokine signaling 2-12 regulates antimicrobial peptides and ecdysteroid signaling pathways in *Bombyx mori* (Dazao). *J. Insect Physiol.* **2017**, *103*, 47–56. [[CrossRef](#)] [[PubMed](#)]

22. Abbas, M.N.; Kausar, S.; Sun, Y.-X.; Tian, J.W.; Zhu, B.-J.; Liu, C.-L. Suppressor of cytokine signaling 6 can enhance epidermal growth factor receptor signaling pathway in *Bombyx mori* (Dazao). *Dev. Comp. Immunol.* **2018**, *81*, 187–192. [[CrossRef](#)] [[PubMed](#)]
23. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
24. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [[CrossRef](#)] [[PubMed](#)]
25. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)] [[PubMed](#)]
26. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [[PubMed](#)]
27. Yaffe, H.; Buxdorf, K.; Shapira, I.; Ein-Gedi, S.; Zvi, M.M.; Fridman, E.; Moshelion, M.; Levy, M. LogSpin: A simple, economical and fast method for RNA isolation from infected or healthy plants and other eukaryotic tissues. *BMC Res. Notes* **2012**, *5*, 45. [[CrossRef](#)]
28. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)] [[PubMed](#)]
29. Jo, Y.H.; Kim, Y.J.; Park, K.B.; Seong, J.H.; Kim, S.G.; Park, S.; Noh, M.Y.; Lee, Y.S.; Han, Y.S. TmCactin plays an important role in Gram-negative and -positive bacterial infection by regulating expression of 7AMP genes in *Tenebrio molitor*. *Sci. Rep.* **2017**, *7*, 46459. [[CrossRef](#)]
30. Li, J.; Zhang, Y.; Zhang, Y.; Liu, Y.; Xiang, Z.; Qu, F.; Lu, Z. Cloning and characterization of three suppressors of cytokine signaling (SOCS) genes from the Pacific oyster, *Crassostrea gigas*. *Fish Shellfish Immunol.* **2015**, *44*, 525–532. [[CrossRef](#)]
31. Croker, B.A.; Kiu, H.; Nicholson, S.E. SOCS regulation of the JAK/STAT signaling pathway. *Semin. Cell Dev. Biol.* **2008**, *19*, 414–422. [[CrossRef](#)] [[PubMed](#)]
32. Duncan, S.A.; Baganizi, D.R.; Sahu, R.; Singh, S.R.; Dennis, V.A. SOCS proteins as regulators of inflammatory responses induced by bacterial infections: A review. *Front. Microbiol.* **2017**, *8*, 2431. [[CrossRef](#)]
33. Alexander, W.S.; Hilton, D.J. The role of suppressors of cytokine signaling (SOCS) proteins in regulation of the immune response. *Ann. Rev. Immunol.* **2004**, *22*, 503–529. [[CrossRef](#)] [[PubMed](#)]
34. Bullock, A.N.; Rodriguez, M.C.; Debreczeni, J.E.; Songyang, Z.; Knapp, S. Structure of the SOCS4-Elongin B/C complex reveals a distinct SOCS box interface and the molecular basis for SOCS-dependent EGFR degradation. *Structure* **2007**, *15*, 1493–1504. [[CrossRef](#)] [[PubMed](#)]
35. Feng, Z.-P.; Chandrasekharan, I.R.; Low, A.; Speed, T.A.; Nicholson, S.E.; Norton, R.S. The N-terminal domains of SOCS proteins: A conserved region in the disordered N-termini of SOCS4 and 5. *Proteins* **2011**, *80*, 946–957. [[CrossRef](#)]
36. Chandrasekharan, I.R.; Mohanty, B.; Linossi, E.M.; Dagley, L.F.; Leung, E.W.; Murphy, J.M.; Babon, J.J.; Nicholson, S.E.; Norton, R.S. Structure and Functional characterization of the conserved JAK interaction region in the intrinsically disordered N-terminus of SOCS5. *Biochemistry* **2015**, *54*, 4672–4682. [[CrossRef](#)] [[PubMed](#)]
37. Hwang, M.N.; Min, C.H.; Kim, H.S.; Lee, H.; Yoon, K.A.; Park, S.Y.; Lee, E.S.; Yoon, S. The nuclear localization of SOCS6 requires the N-terminal region and negatively regulates Stat3 protein levels. *Biochem. Biophys. Res. Commun.* **2007**, *360*, 333–338. [[CrossRef](#)]
38. Chen, M.-S.; Liu, S.; Wang, H.; Cheng, X.; El Bouhssini, M.; Whitworth, R.J. Massive shift in gene expression during transitions between developmental stages of the Gall Midge, *Mayetiola destructor*. *PLoS ONE* **2016**, *11*, e0155616. [[CrossRef](#)] [[PubMed](#)]
39. Du, W.; Zeng, F. Identification of development-related genes in the ovaries of adult *Harmonia axyridis* (Pallas) lady beetles using a time-series analysis by RNA-seq. *Sci. Rep.* **2016**, *6*, 39109. [[CrossRef](#)] [[PubMed](#)]
40. Dhawan, R.; Gupta, K.; Kajla, M.; Kumar, S.; Gakhar, S.K.; Kakani, P.; Choudhury, T.P.; Gupta, L. Molecular characterization of SOCS gene and its expression analysis on *Plasmodium berghei* infection in *Anopheles culicifacies*. *Acta Trop.* **2015**, *152*, 170–175. [[CrossRef](#)]
41. Monahan, A.J.; Starz-Gaiano, M. Socs36E limits STAT signaling via Cullin2 and a SOCS-box independent mechanism in the *Drosophila* egg chamber. *Mech. Dev.* **2015**, *138*, 313–327. [[CrossRef](#)]

42. Magnusson, K.; Mendes, A.M.; Windbichler, N.; Papathanos, P.A.; Nolan, T.; Dottorini, T.; Rizzi, E.; Christophides, G.K.; Crisanti, A. Transcription regulation of sex-biased genes during ontogeny in the malaria vector *Anopheles gambiae*. *PLoS ONE* **2011**, *6*, e21572. [[CrossRef](#)] [[PubMed](#)]
43. Zhang, M.; Xiao, Z.Z.; Sun, L. Suppressor of cytokine signaling 3 inhibits head kidney macrophage activation and cytokine expression in *Scophthalmus maximus*. *Dev. Comp. Immunol.* **2011**, *35*, 174–181. [[CrossRef](#)] [[PubMed](#)]
44. Silver, D.L.; Geisbrecht, E.R.; Montell, D.J. Requirement for JAK/STAT signaling throughout border cell migration in *Drosophila*. *Development* **2005**, *132*, 3483–3492. [[CrossRef](#)] [[PubMed](#)]
45. Maehr, T.; Vecino, J.L.G.; Wadsworth, S.; Wang, T.H.; Secombes, C.J. Four CISH paralogues are present in rainbow trout *Oncorhynchus mykiss*: Differential expression and modulation during immune responses and development. *Mol. Immunol.* **2014**, *62*, 186–198. [[CrossRef](#)] [[PubMed](#)]
46. Noguchi, S.; Yamada, N.; Kumazaki, M.; Yasui, Y.; Iwasaki, J.; Naito, S.; Akao, Y. Socs7, a target gene of microRNA-145, regulates interferon- β induction through STAT3 nuclear translocation in bladder cancer cells. *Cell Death Dis.* **2013**, *4*, e482. [[CrossRef](#)] [[PubMed](#)]
47. Hao, L.X.; Sun, L. Comparative analysis of the expression patterns of eight suppressors of cytokine signaling in tongue sole, *Cynoglossus semilaevis*. *Fish Shellfish Immunol.* **2016**, *55*, 595–601. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).