Peer-reviewed version available at *Insects* **2019**, *10*, 76; doi:10.3390/insects10030076

Article

Molecular cloning and expression characterization of three suppressors of cytokine signaling genes (SOCS5, SOCS6, SOCS7) from the mealworm beetle, *Tenebrio molitor*

Bharat Bhusan Patnaik[1,2], Bo Bae Kim[1], Yong Hun Jo[1], 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
[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

Abstract: Suppressors of cytokine signaling (SOCS) influence cytokine and growth factor signaling by negatively regulating the JAK-STAT pathway. This maintains homeostasis during host immune response. However, functional characterization of SOCS family members in invertebrates is limited. In this study, we discovered the Type-I subfamily of the SOCS genes in the mealworm beetle, *T. molitor*. The full-length ORFs of TmSOCS5, TmSOCS6, and TmSOCS7 consisted of 1,389, 897 and 1,458 nucleotides, encoding polypeptides of 462, 297 and 485 amino acids, respectively, The C-terminal region of TmSOCS was highly conserved in the SH2 and SOCS box domains. Phylogenetic analysis revealed that the three SOCS genes clustered within the same sub-family and the highest amino acid identity was with the *Tribolium castaneum* SOCS genes (TcSOCS). While the expression of TmSOCS5 and TmSOCS6 was low in larval, pupal, and adult stages of the insect, TmSOCS7 showed higher expression. The expression of TmSOCS5 and TmSOCS6 was higher in larval hemocytes and adult ovary. The microbes expressed the three TmSOCS genes to varying degrees. *C. albicans* elicited the strongest response in the host with highest 15-fold expression in TmSOCS7 3 h post-inoculation. Collectively, these data suggest that the Type I TmSOCS could play a role in eliciting host immunity.

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

1. Introduction

Cytokines are secretory proteins that regulate various inflammatory responses in the cell. Most cytokines promote gene expression using the Janus kinase (JAK)-signal transducers and activators of transcription (STAT) pathway. Cytokine signaling and the JAK-STAT pathway are known to play an essential role in metazoan development and homeostasis of the host immune machinery in the case of infections [1, 2]. Considering the extensive repertoire of JAKs and STATs (four JAKs and seven STATs in humans), there seems to be a differential engagement of specific JAK-STAT pathway components in response to signaling mediated by a plethora of cytokine molecules. Hence, multiple inflammatory responses could be channeled to produce various pathological responses via the response from each cytokine receptor complex. Due to its role in the activation of pro-inflammatory responses and its adverse effects if unchecked, the JAK-STAT pathway is tightly controlled by multiple regulators in the cellular context. Proteins such as the members of suppressor of cytokine signaling (SOCS) family (specific member of the SH2-domain containing tyrosine phosphatase (SHP) family), and protein inhibitors against STATs (PIAS) act as negative regulators of the JAK-STAT pathway maintaining homeostasis during defense reactions [3-5].

Members of the SOCS family of proteins are the prime regulators of the JAK-STAT pathway. They control the inflammatory signaling pathway by acting as pseudo JAK substrates thus blocking STAT signaling and directing multiple pathway components for ubiquitin-mediated proteasomal degradation [6]. Originally, four members of this family were identified, SOCS1, SOCS2, SOCS3, and a cytokine-inducible SH2-containing protein (CISH). All of which negatively regulated cytokine induction either by competing with STAT for binding to the cytoplasmic domains of the phosphorylated receptors or by inactivating the enzyme activity of JAKs [7]. This was followed by the discovery of four new members, SOCS4, SOCS5, SOCS6, and SOCS7 based on the conserved central SH2 domain and the C-terminal SOCS-box domain. Evolutionary divergence analysis of SOCS family proteins has revealed a clear division of CISH, SOCS 1, SOCS-2, and SOCS-3 into the type II subfamily and the rest of the members into the type I subfamily [8]. SOCS4 and SOCS5 are close homologs involved in the negative feedback loop of epidermal growth factor receptor (EGFR) signaling by directing the degradation of EGFR in a ligand and E3 ubiquitin ligase c-Cbl-independent manner [4]. SOCS6 associates with and inhibits the insulin receptor and is related to cytokine-mediated insulin resistance while SOCS7 interacts with STAT5 or STAT3 thereby preventing their nuclear translocation and attenuating prolactin, growth hormone, and leptin signaling [9, 10]. All eight members of the SOCS gene family have been identified in vertebrates. Further, vertebrate –species- specific lineages have been suggested for the classification of SOCS family members due to the expansion of the type II SOCS subfamily through whole genome duplication and the apparent evolution of the rainbow trout SOCS gene family to 26 expressed genes [11].

Discovery of SOCS gene families in invertebrate phyla is limited with isolated reports from a few molluscan and arthropod species. In the *Drosophila* model, SOCS36E (64% identity to human SOCS5) was first identified and studied for its expression during embryogenesis [12] and later as a negative regulator of the JAK/STAT and EGFR pathways. The other two identified genes, SOCS44A (34% and 33% identity to human SOCS6 and SOCS7, respectively) and SOCS16D (shares 48% and 45% identity with human SOCS6 and SOCS7, respectively) showed limited involvement in the JAK/STAT cascade, although SOCS44A was considered to be a transcriptional target of STAT92E [13]. In the pacific oyster, *Crassostrea gigas,* three SOCS genes (SOCS2, SOCS5, and SOCS7) have been identified as putatively inducing NF-kB gene transcriptional activity. SOCS2 homologues with immune expression have also been identified in *Ruditapes philippinarum* [14], *Haliotis discus* [15], *Eriocheir sinensis* [16], *Procambarus clarkii* [17], and *Litopenaeus vannamei* [18]. Further, direct functional evidence suggests that SOCS6 has a role in the activation of the NF-kB signaling pathway in *E. sinensis* [19]. Among insects, the *Bombyx mori* SOCS2 homolog has been confirmed to function as a negative regulator of the JAK/STAT and ecdysteroid signaling pathways [20], while the SOCS6 homolog regulates the EGFR pathway [21]. The type I subfamily of SOCS genes have not yet been reported in insects. In the present study, we identified type I subfamily of SOCS genes (SOCS5, SOCS6, and SOCS7) in the coleopteran model insect, *Tenebrio molitor* using bioinformatics analysis. Subsequently, we cloned the SOCS homologs and characterized their evolutionary relationship using phylogenetic analysis. Finally, we have studied their expression profiles in response to with microorganisms providing insights into the putative immune function of SOCS in the insect.

2. Results and Discussion

2.1 Identification of TmSOCS homologs and molecular characterization

A local tblastn search was conducted using TcSOCS5, TcSOCS6, and TcSOCS7 to retrieve members of the TmSOCS gene family sequences from the *Tenebrio* RNAseq and EST libraries. These gene sequences were designated TmSOCS5, TmSOCS6, and TmSOCS7. The putative ORF sequences for the TmSOCS genes were identified using the gene-finding program FGENESH (http://www.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind). Next, primers were designed to clone the full-length TmSOCS sequences. The constructs (T-blunt vector + full-length insert) were sequenced, thereby validating the TmSOCS genes. The TmSOCS5 ORF was 1389 bp encoding a 462 long polypeptide (Figure 1).

Figure 1. The nucleotide and deduced amino acid sequence of the *T. molitor* SOCS5 (TmSOCS5) ORF. The SH2 (Src Homology 2) and SOCS box domain are black and blue boxed, respectively. The polypeptide binding sites, namely the phosphotyrosine binding pocket, hydrophobic binding pocket, and putative elongin B/C interaction residues are indicated by red, blue, and green text, respectively.

```
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 TTC CGC   1020
 G   T   F   L   L   R   D   S   A   Q   E   E   F   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
```

The computed molecular mass and isoelectric point of the TmSOCS5 protein are 52.8 kDa and 6.23, respectively. These results are similar to an SOCS5 homolog in *C. gigas* which had a molecular mass of 51.5 kDa and an isoelectric point of 6.40 [22]. TmSOCS5 also contains an SH2 domain (residues 301- 379) with the phosphotyrosine hydrophobic binding pocket. A C-terminal SOCS box domain (residues 381- 437) was also found

with the putative elongin B/C interaction residues. The TmSOCS6 ORF was 897 bp encoding a 297 amino acid polypeptide (Figure 2). Its calculated molecular mass was 34.5 kDa and its isoelectric point was 9.40. The TmSOCS6 full-length sequence was shorter than the SOCS6 homologue in *E. chinensis* [19], although the typical SH2 and the SOCS box domains were detected. The TmSOCS7 ORF was 1458 bp long encoding 485 amino acid polypeptide (Figure 3). Its predicted molecular mass was 56 kDa with a theoretical isoelectric point of 9.41. Similar to other SOCS family members, the phosphotyrosine, hydrophobic, and the putative elongin B/C interaction residues were present in the conserved SH2 and SOCS box domains. Both this and closely related studies revealed that the amino-terminal region of SOCS family proteins is extended with divergent sequences while the central SH2 and carboxyl-terminal SOCS box domain sequences are largely conserved. The SOCS box amino acid consensus sequence has also been found in protein families having WD-40 repeats (IPR017986), ankyrin repeats (IPR020683), and SPRY domains (IPR003877) [7]. The elongin B/C region in the SOCS box domain could assist in the negative regulation of signaling pathways by directing the target proteins for ubiquitination and proteasomal degradation [23, 24].

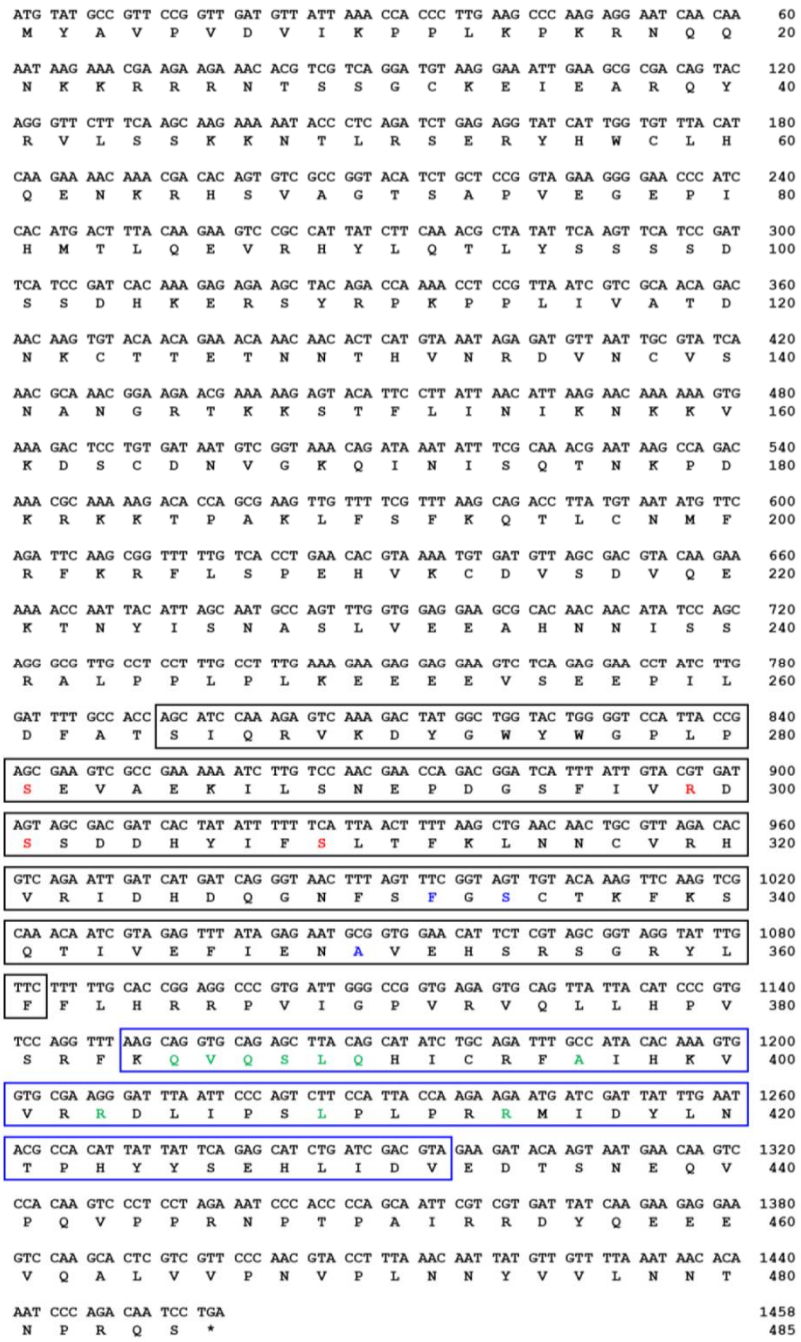
Figure 2. The nucleotide and deduced amino acid sequence of the *T. molitor* SOCS6 (TmSOCS6) ORF. Black and blue box indicates SH2 and SOCS box domains, respectively. The phosphotyrosine binding pocket, hydrophobic binding pocket, and putative elongin B/C interaction residues are indicated by red, blue, 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
```

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2019 doi:10.20944/preprints201901.0263.v1

Peer-reviewed version available at Insects 2019, 10, 76; doi:10.3390/insects10030076

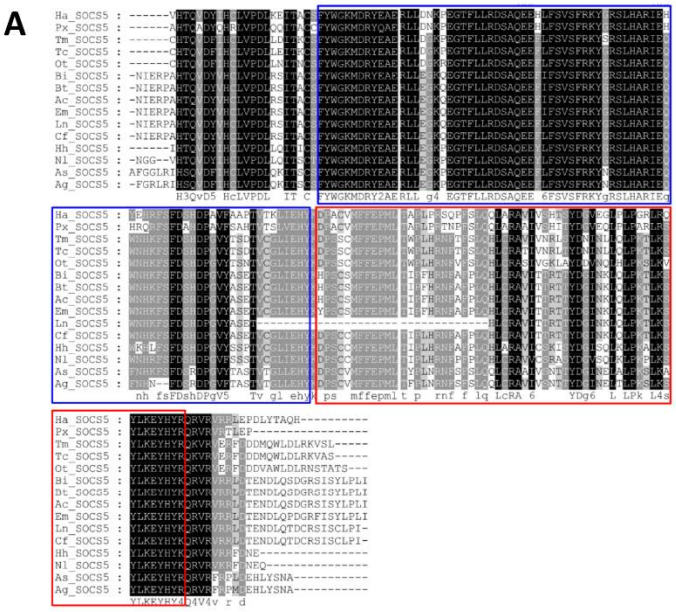
Figure 3. The nucleotide and deduced amino acid sequence of the *T. molitor* SOCS7 (TmSOCS7) ORF. The conserved SH2 and SOCS box domains are black and blue boxed, respectively. The phosphotyrosine binding pocket, hydrophobic binding pocket, and putative elongin B/C interaction residues are indicated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



The multiple sequence alignment (MSA) and percentage identity of the SOCS5 C-terminal region with insect orthologs is shown in Figure 4. The SH2 domain consists of identical residues across the species, while the SOX box domain also shows high levels of similarity (Figure 4A). TmSOCS5 is most similar to TcSOCS5 with 97% identity followed 88% identity with the dung beetle, *Onthophagus taurus* SOCS5 (OtSOCS5) (Figure 4B). In the case of TmSOCS6, MSA has less similarity in the SH2 and SOCS box domains with a maximum identity of 91% with TcSOCS6 (Figure 5). There was significantly less similarity to other insect orthologs with percent

identity ranging between 35-55%. BmSOCS6 showed similarity to lepidopteran SOCS6 gene family members and *E. sinensis* SOCS6 showed high levels of sequence similarity with other molluscan SOX gene family members [19, 20]. High levels of sequence identity and similarity were observed in the SH2 and SOCS box domains of SOCS7 homologs (Figure 6A). Interestingly, TmSOCS7 SH2 and SOCS box domains had 100% identity to TcSOCS7 followed by 74-86% identity with other represented insect species (Figure 6B). Furthermore, TmSOCS5, TmSOCS6, and TmSOCS7 were closely related to TcSOCS5, TcSOCS6, and TcSOCS7, respectively.

Figure 4. Alignment and identity of TmSOCS5. (A) Multiple sequence alignment of the SOCS5 C-terminal region with insect orthologs. The blue and red boxed sequences represent the SH2 and SOCS box domains, respectively. Identical residues in all sequences are shaded black and most common sequences are shaded grey. Deletions are indicated by dashes. (B) Percentage identity matrix of represented SOCS5 members. 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), *Eufrisea 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).



B

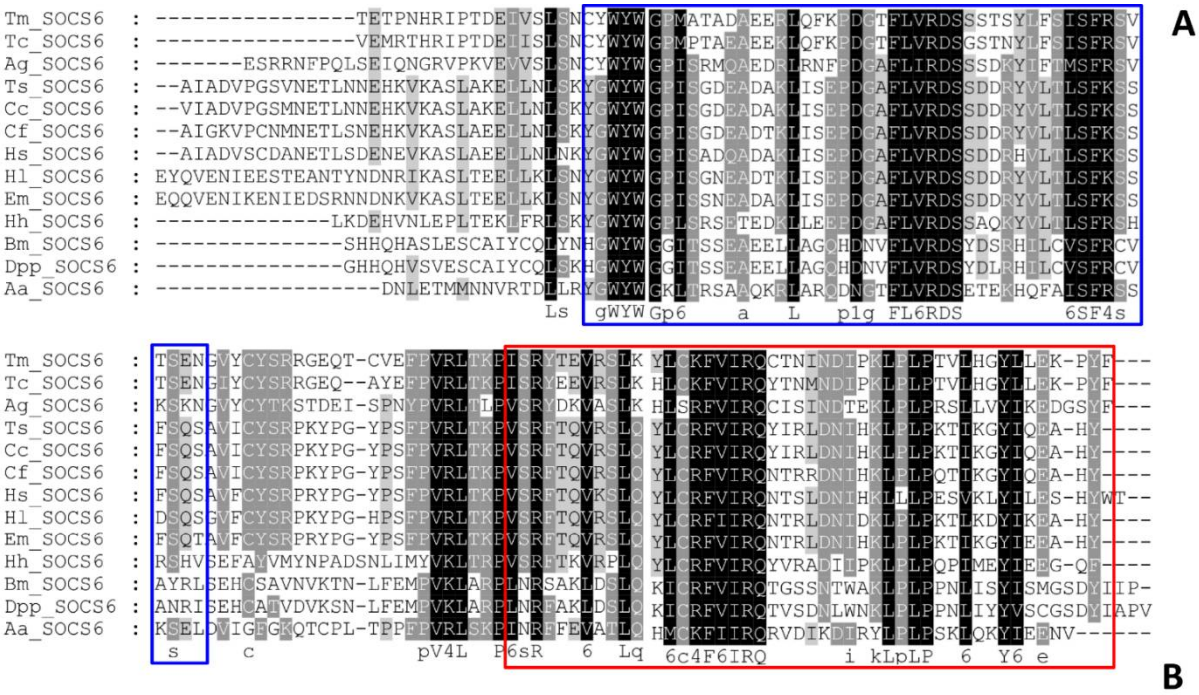
	Tm_SOCS5	Tc_SOCS5	Ot_SOCS5	Bi_SOCS5	Ac_SOCS5	Bt_SOCS5	Em_SOCS5	Ln_SOCS5	Cf_SOCS5	Hh_SOCS5	Nl_SOCS5	As_SOCS5	Ag_SOCS5	Ha_SOCS5	Px_SOCS5
Tm_SOCS5	100	97	88	78	78	78	78	78	80	82	82	77	77	70	69
Tc_SOCS5		100	88	79	79	79	79	80	81	81	81	78	77	69	70
Ot_SOCS5			100	74	74	74	74	74	77	81	80	76	75	69	69
Bi_SOCS5				100	100	100	98	96	95	85	83	77	76	69	68
Ac_SOCS5					100	100	98	96	95	85	83	77	76	69	68
Bt_SOCS5						100	98	96	95	85	83	77	76	69	68
Em_SOCS5							100	95	95	85	83	77	76	69	68
Ln_SOCS5								100	100	84	82	79	78	71	71
Cf_SOCS5									100	87	86	79	78	71	70
Hh_SOCS5										100	91	77	77	74	72
Nl_SOCS5											100	76	77	74	72
As_SOCS5												100	95	74	74
Ag_SOCS5													100	74	74
Ha_SOCS5														100	88
Px_SOCS5															100

Figure 5. Alignment and identity of TmSOCS6. (A) Multiple alignment of SOCS6 with insect orthologs. Only the C-terminal region was conserved. The blue and red boxed sequences represent the SH2 and SOCS box

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2019 doi:10.20944/preprints201901.0263.v1

Peer-reviewed version available at Insects 2019, 10, 76; doi:10.3390/insects10030076

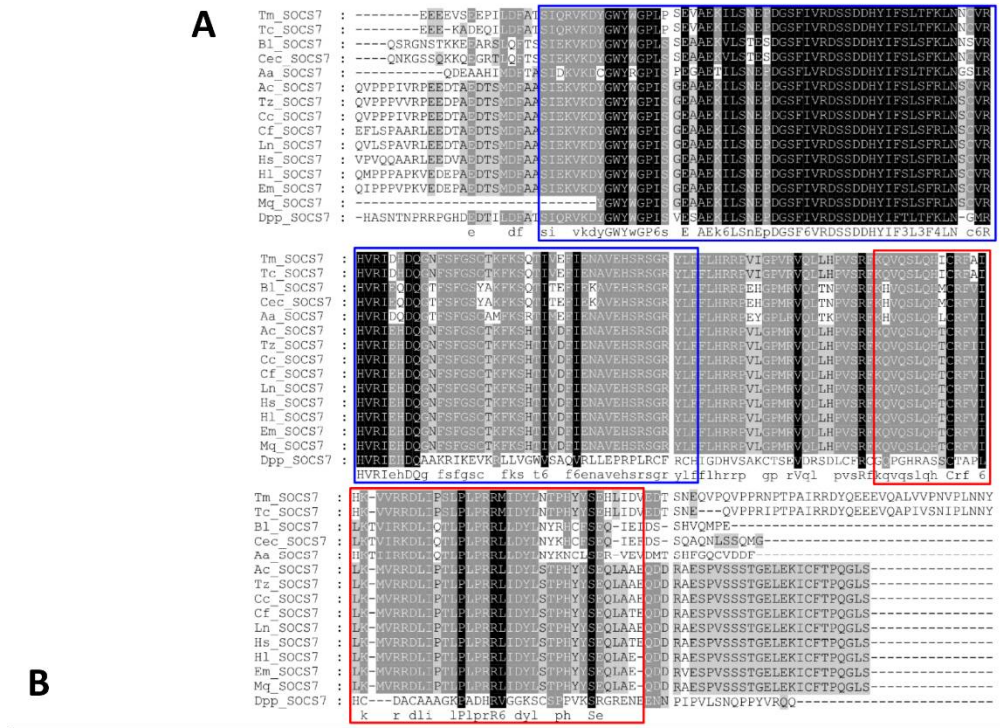
domains, respectively. Identical residues in all sequences are shaded black and common sequences are shaded grey. Deletions are indicated by dashes. (B) Percentage identity of SOCS6 members from representative insect species. Analysis was performed by clustalX2 using representative amino acid sequences 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), *Halyomorpha halys* (XP_014279265.1), *Bombyx mori* (NP_001185652.1), *Danaus plexippus plexippus* (OWR49085.1), *Aedes aegypti* (XP_001660156.3).



	Tm_SOCS6	Ln_SOCS6	Ag_SOCS5	Hl_SOCS5	Em_SOCS5	Ts_SOCS5	Cc_SOCS5	Cf_SOCS5	Hs_SOCS5	Hh_SOCS5	Aa_SOCS5	Bm_SOCS5	Dpp_SOCS5
Tm_SOCS6	100	91	54	51	52	50	50	52	49	41	43	36	35
Tc_SOCS6		100	52	49	50	51	51	51	47	41	42	37	36
Ag_SOCS6			100	48	47	48	48	48	48	42	41	39	39
Hl_SOCS6				100	90	87	87	88	82	54	39	41	41
Em_SOCS6					100	90	90	90	84	56	38	42	42
Ts_SOCS6						100	100	95	86	54	38	41	41
Cc_SOCS6							100	95	86	54	38	41	41
Cf_SOCS6								100	86	54	38	41	41
Hs_SOCS6									100	51	37	39	39
Hh_SOCS6										100	37	40	40
Aa_SOCS6											100	37	37
Bm_SOCS6												100	87
Dpp_SOCS6													100

Figure 6. Alignment and identity of TmSOCS7. (A) Multiple alignment of the SOCS7 C-terminal sequence with insect orthologs. The blue and red boxed sequences represent the SH2 and SOCS box domains, respectively. Identical residues in all sequences are shaded black and most common sequences are shaded grey. Deletions are indicated by dashes. (B) Percent identity of SOCS7 members using the amino acid sequences from representative insect species. ClustalX2 program was used to analyze the identity matrix. The sequences used for the analysis were from *Tribolium castaneum* (XP_008190646.1), *Bactocera latifrons* (JAI49983.1), *Ceratitis capitata* (JAC02138.1), *Aedes aegypti* (XP_021707758.1), *Acromyrmex echinator* (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), *Danaus plexippus plexippus* (OWR46815.1).

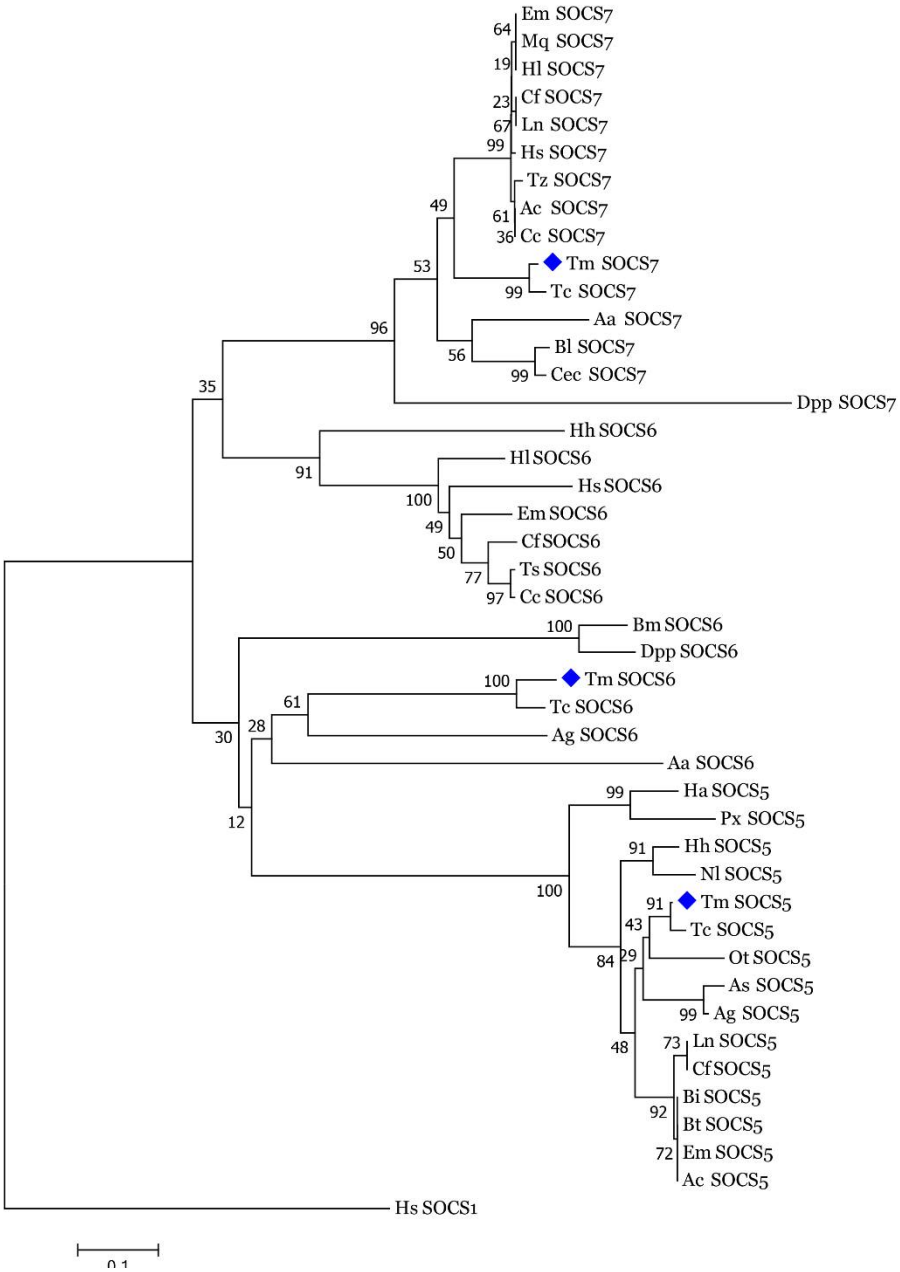


	Bl_SOCS7	Cec_SOCS7	Aa_SOCS7	Tm_SOCS7	Tc_SOCS7	Ac_SOCS7	Cc_SOCS7	Ln_SOCS7	Tz_SOCS7	Cf_SOCS7	Hs_SOCS7	Mq_SOCS7	Hl_SOCS7	Em_SOCS7	Dpp_SOCS7
Bl_SOCS7	100	98	81	77	77	79	79	79	78	79	79	78	79	79	34
Cec_SOCS7		100	81	77	77	80	80	80	79	80	80	79	80	80	34
Aa_SOCS7			100	74	74	74	74	74	74	74	74	74	74	74	35
Tm_SOCS7				100	100	85	85	85	85	85	85	86	86	86	38
Tc_SOCS7					100	85	85	85	85	85	85	86	86	86	38
Ac_SOCS7						100	100	100	99	99	99	99	99	99	38
Cc_SOCS7							100	100	99	99	99	99	99	99	38
Ln_SOCS7								100	99	99	99	99	99	99	38
Tz_SOCS7									100	99	99	99	99	99	38
Cf_SOCS7										100	100	99	99	99	38
Hs_SOCS7											100	99	99	99	38
Mq_SOCS7												100	100	100	36
Hl_SOCS7													100	100	37
Em_SOCS7														100	37
Dpp_SOCS7															100

2.2 Molecular evolutionary relationships of the TmSOCS proteins

To explore the relatedness of the *T. molitor* type I subfamily of SOCS protein, we constructed a phylogenetic tree based on the amino acid sequences of insect SOCS proteins (Figure 7). Human SOCS1 protein belonging to the type II subfamily was used as an outgroup. The phylogram showed three conspicuous clusters of type II SOCS proteins. The SOCS5 and SOCS7 proteins formed a more compact cluster than the SOCS6 cluster. As expected, TmSOCS5, TmSOCS6, and TmSOCS7 were closely related to TcSOCS5, TcSOCS6, and TcSOCS7, their respective SOCS family proteins. In addition, *Orthophagus taurus,* another coleopteran species, SOCS5 (OtSOCS5) was found to be closely related to TmSOCS5 and TcSOCS5. The TmSOCS6 cluster consisted of two sub-clusters. We had shown earlier through MSA and percent identity that the SH2 and SOCS box domains of TmSOCS6 gene members are comparatively less conserved than those of TmSOCS5 and TmSOCS7 homologs. Moreover, the longer N-terminal region of insect SOCS6 proteins, similar to that of vertebrate SOCS6 is relevant for its localization to the nucleus [25]. These results are consistent with similar reports on *E. sinensis* SOCS6 [19]. The relatedness of *Bombyx mori* SOCS6 (BmSOCS6) and *Danaus plexippus* SOCS6 (DppSOCS6) proteins is also consistent with the earlier report [21].

Figure 7. Phylogenetic analysis of TmSOCS5, TmSOCS6, and TmSOCS7 full-length ORF sequences and insect SOCS5, SOCS6, and SOCS7 family proteins. The tree was constructed using the Neighbor-Joining method in MEGA7 Numbers at the node indicate the bootstrap support of 1000 replicates. *Homo sapiens* SOCS1 (Type II member) was used as an outgroup. ◆show *T. molitor* SOCS5, SOCS6, and SOCS7.



2.3 The developmental and tissue distribution of TmSOCS gene expression

We used qRT-PCR analysis to quantify TmSOCS5, TmSOCS6, and TmSOCS7 mRNA expression in the developmental stages and in various tissues of the larval and adult stages of the insect (Figure 8). TmSOCS5 and TmSOCS6 expression was higher in the egg stage than in the larval, pupal, and adult stages. In the larval, pupal, and adult stages the mRNA expression was consistent although in the late pupal and adult stages the expression was significantly lower than during the larval and early pupal stages (Figure 9). In contrast, TmSOCS7 expression in the larval, pupal, and adult stages of *T. molitor* were much higher than during the egg stage (Figure

10). *Drosophila* SOCS36E, which has high levels of sequence identity with vertebrate SOCS5 (75% and 44% at SH2 and SOCS-box level), is known to be expressed during embryogenesis (especially during embryonic and imaginal disc development) [12]. Furthermore, SOCS36E regulates STAT activity levels through either the Cullin2 (Cul2) scaffolding protein-dependent or independent mechanisms in the egg chamber of *Drosophila* [26].

Figure 8. Developmental and Tissue-specific expression pattern of the SOCS5 transcript in *T. molitor*. (A) TmSOCS5 expression in the egg (EG), early larvae (YL), late larvae (LL), pupal day 1-7 (P1–P7), adult day 1-5 (A1-A5) stages. (B) Distribution of TmSOCS5 in the larval tissues of *T. molitor*. (C) Distribution of TmSOCS5 in the adult tissues of *T. molitor*. Tissue abbreviations are as follows: IT-integument, FB-fat body, HC- hemocytes, GT-gut, MT-Malpighian tubules, OV-ovary, TS-testis. Analysis of transcript level was carried out by quantitative real-time PCR. Data shown are the mean of the relative expression ± SE for three sets of biological replications. Different letters above the bars represent significant differences at a 95% confidence level.

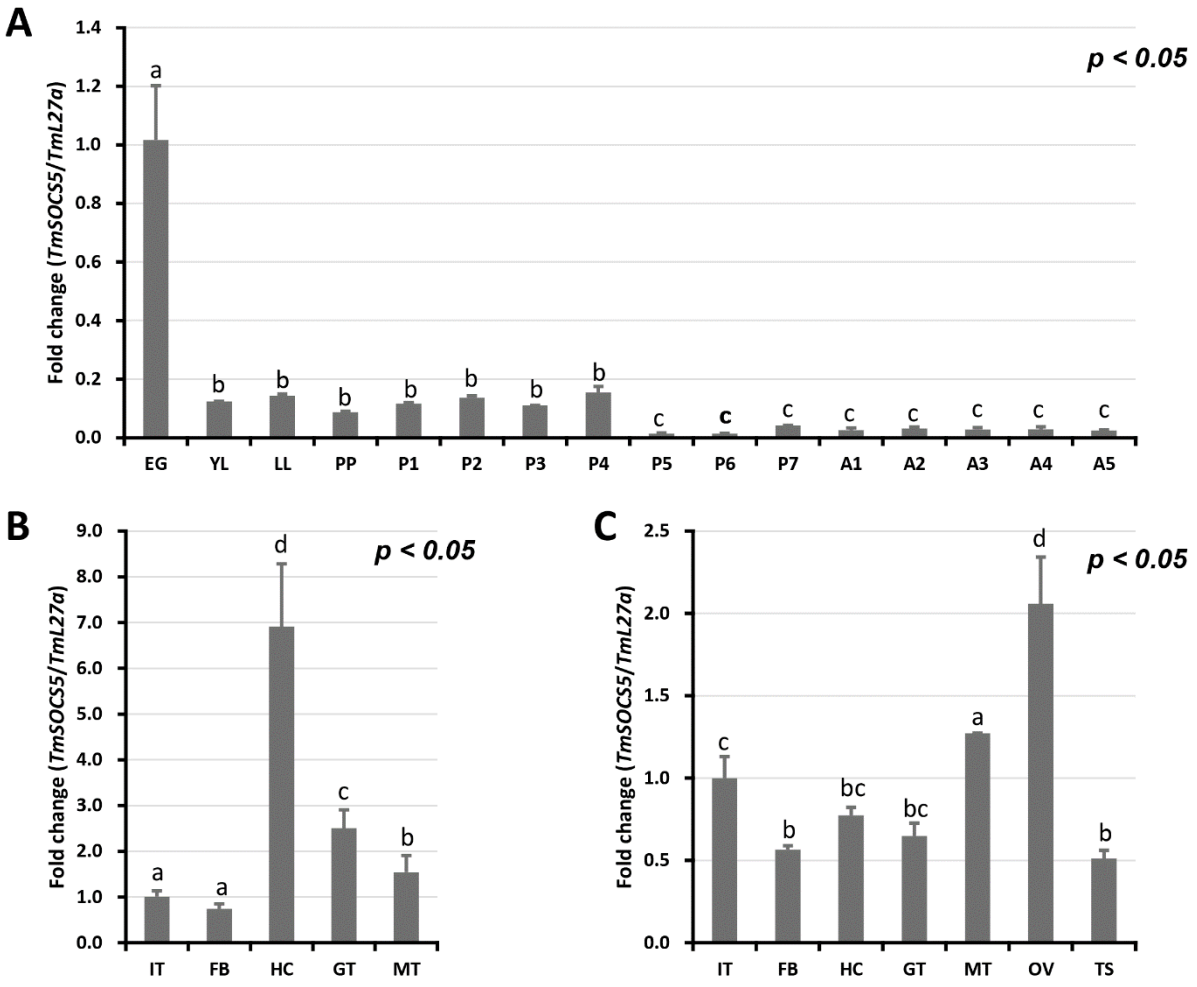


Figure 9. Developmental and Tissue-specific expression pattern of SOCS6 transcript in *T. molitor*. (A) TmSOCS6 expression in the egg (EG), early larvae (YL), late larvae (LL), pupal day 1-7 (P1–P7), adult day 1-5 (A1-A5) stages. (B) Distribution of TmSOCS6 in the larval tissues of *T. molitor*. (C) Distribution of TmSOCS5 in the adult tissues of *T. molitor*. Tissue abbreviations are as follows: IT-integument, FB-fat body, HC- hemocytes, GT-gut, MT-Malpighian tubules, OV-ovary, TS-testis. The transcript level was analyzed by quantitative real-time PCR. Data are the mean of the relative expression ± SE for three sets of biological replications. Different letters above the bars represent significant differences at a 95% confidence level.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2019 doi:10.20944/preprints201901.0263.v1

Peer-reviewed version available at Insects 2019, 10, 76; doi:10.3390/insects10030076

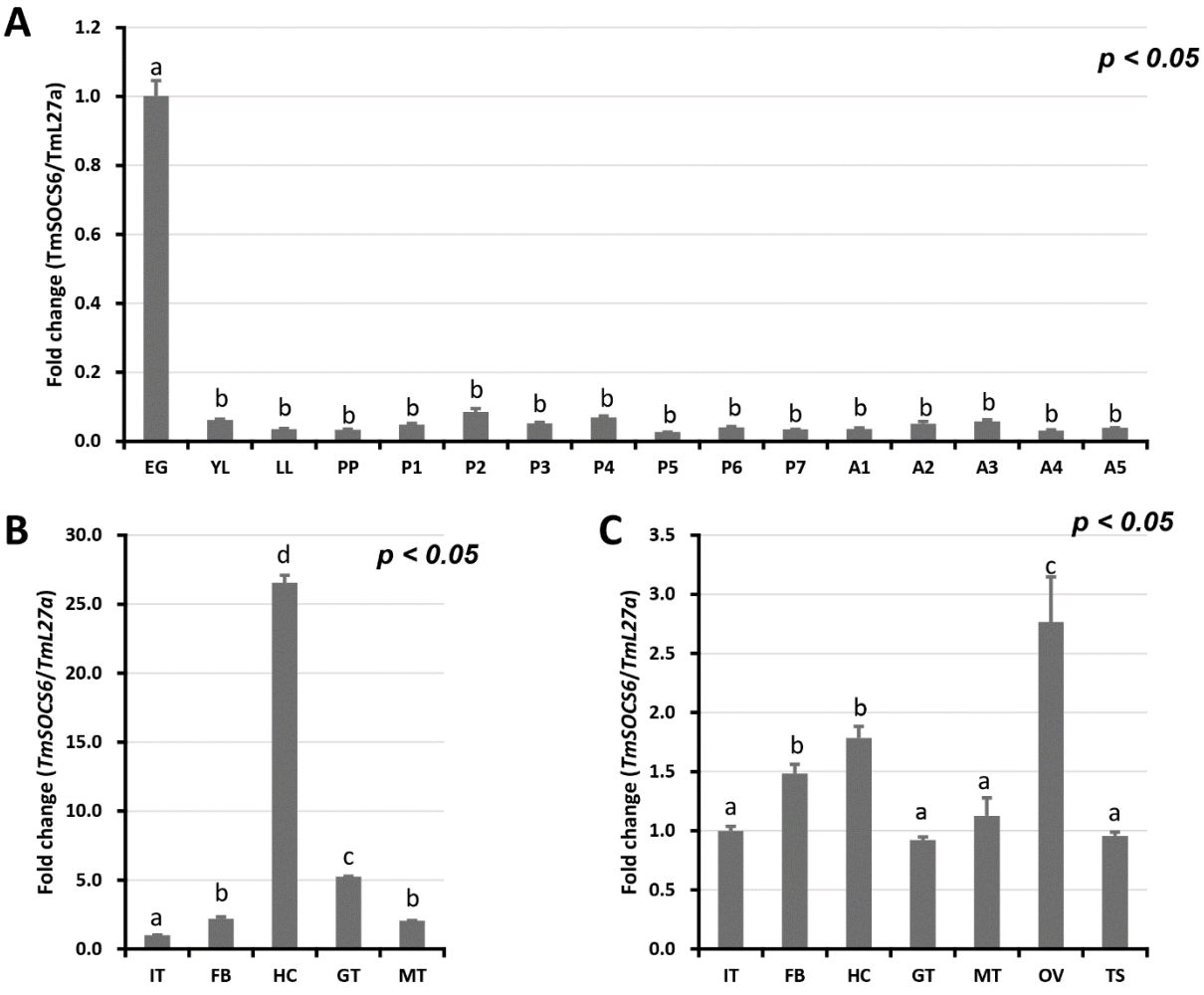
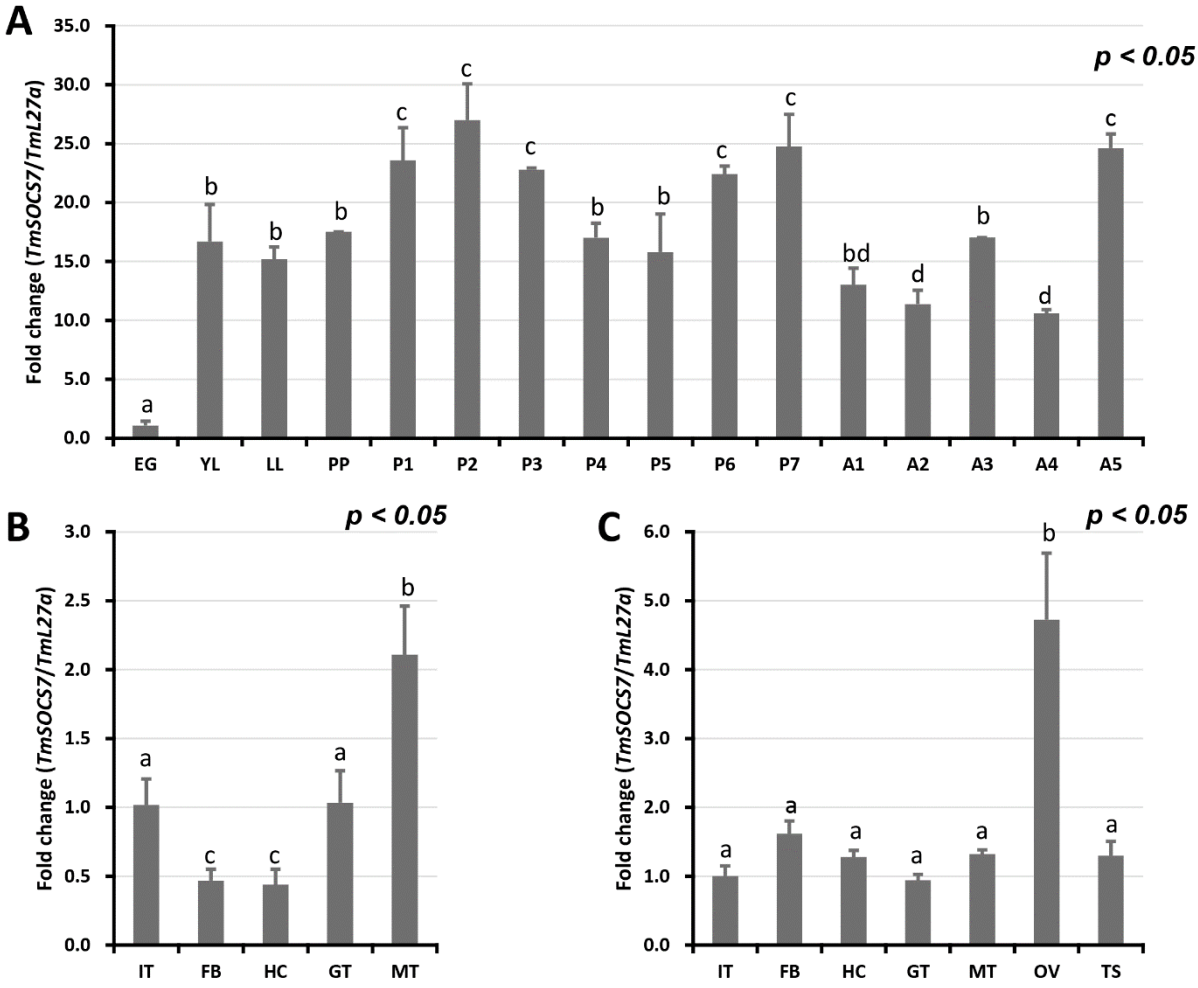


Figure 10. Developmental and Tissue-specific expression pattern of SOCS7 transcript in *T. molitor*. (A) TmSOCS7 expression in the egg (EG), early larvae (YL), late larvae (LL), pupal day 1-7 (P1–P7), adult day 1-5 (A1-A5) tissues. (B) Distribution of TmSOCS6 in the larval tissues of *T. molitor*. (C) Distribution of TmSOCS5 in the adult tissues of *T. molitor*. Tissue abbreviations are as follows: IT-integument, FB-fat body, HC- hemocytes, GT-gut, MT-Malpighian tubules, OV-ovary, TS-testis. The relative fold-change expression of TmSOCS7 is the mean of three independent measurements ± SE. Different letters above the bars represent significant differences at a 95% confidence level.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2019 doi:10.20944/preprints201901.0263.v1

Peer-reviewed version available at Insects 2019, 10, 76; doi:10.3390/insects10030076



Spatial expression analysis, found that the mRNA levels of TmSOCS5 and TmSOCS6 were higher in the hemocytes of *T. molitor* larvae compared to the other tissues, while TmSOCS7 levels were elevated in Malphigian tubules. The ovarian expression of all three TmSOCS transcripts was higher in *T. molitor* adults. The constitutive expression of the TmSOCS transcripts in other immune tissues such as the fat body and gut was not appreciable. It is possible that SOCS proteins might have different biological functions in the tissues. This is especially true when we consider the expression of SOCS6 in the hemocytes, fat body, and Malphigian tubules of *B. mori* larvae [21]. The greater expression level of SOCS6 expression in the hemocytes has also been reported in other invertebrate and vertebrate species [19, 27, 28].

2.4. Expression of TmSOCS genes after immune stimulation

Considering the expression of TmSOCS5, TmSOCS6, and TmSOCS7 in the immune tissues, we analyzed their temporal expression profiles in the hemocytes, fat body, and gut tissues of *T. molitor* larvae after challenge with pathogens. The induction of TmSOCS5 expression in the fat body tissue was highest 12 h post immune challenge with the fungus *C. albicans* (Figure 11A). In gut (Figure 11B) and hemocyte (Figure 11C) tissues, TmSOCS5 expression was induced 9 h and 6 h after the fungus challenge, respectively. A similar TmSOCS6 expression profile was found in the fat body (Figure 12A), gut (Figure 12B), and hemocytes (Figure 12C) of *T. molitor* larvae after challenge with the pathogen. Once again the expression level was significantly upregulated ($p < 0.05$) in the hemocytes at 3 h post infection (p.i). Generally immune expression at an early stage may be useful in mounting an appropriate response to pro-inflammatory cytokines in most tissues. SOCS6 expression was upregulated following microbial challenge in *B. mori* hemocytes. Moreover, the relative mRNA expression of EGFR pathway related genes such as *fkhr*, *gsk3*, *ras* and *erk* was strongly induced 4 h after injection

with recombinant BmSOCS6 protein [21]. Regulation of the EGFR pathway by *Drosophila* SOCS44A (34% identity with human SOCS6 gene) and *B. mori* SOCS6 consistently demonstrate its control of developmental and pathophysiological processes [21, 29]. TmSOCS7 showed a significant upregulation ($p < 0.05$) after challenge with the Gram negative bacteria *E. coli*, the Gram positive bacteria *S. aureus*, and the fungus *C. albicans* in the tested tissues (Figure 13). As with other SOCS, *C. albicans* was able to induce a 15-fold increase in TmSOCS7 3 h post-infection. Although studies focusing on the function of SOCS7 are limited, one study has demonstrated that it is a check on STAT 3 and STAT5 nuclear translocation [30]. STAT3 being one of the transcriptional regulators of IFN-β and interleukin 6, it might be speculated that SOCS7 participates in the interferon regulatory pathway.

Figure 11. Expression of TmSOCS5 in the fat body (A), gut (B) and hemocytes (C) following challenge of the host with *E. coli, S. aureus and C. albicans*. Analysis of the expression was done using real-time PCR using RpL27a as the normalizing control. The relative fold-change expression of TmSOCS5 is the mean of three independent measurements ± SE. Asterisk indicates a significant difference of $p < 0.05$ between the challenged and control (PBS) groups at the same time point.

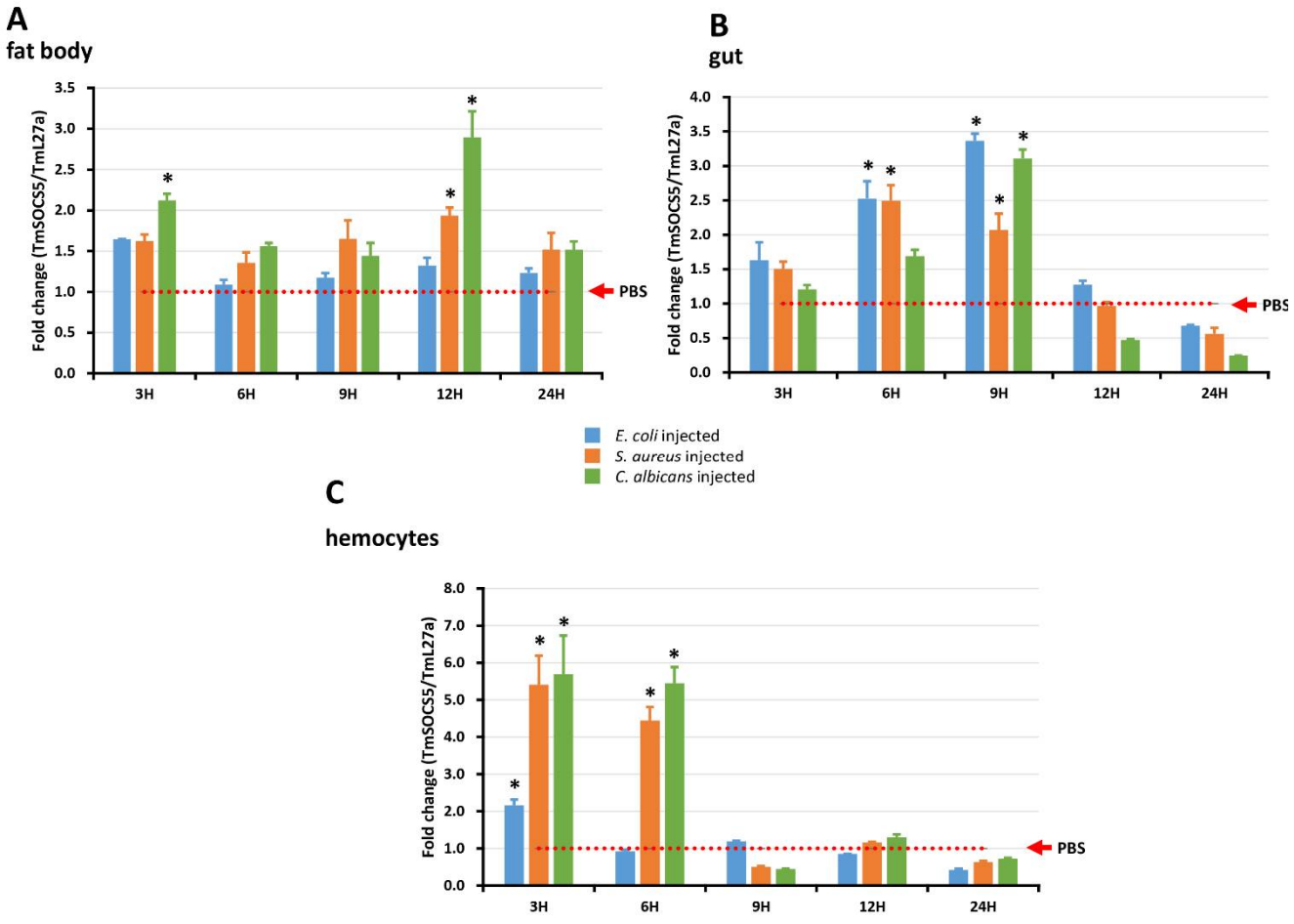


Figure12. Expression of TmSOCS6 in the fat body (A), gut (B) and hemocytes (C) following challenge of the host with *E. coli, S. aureus and C. albicans*. Analysis of the expression was done using real-time PCR using RpL27a as the normalizing control. The relative fold-change expression of TmSOCS6 is the mean of three independent measurements ± SE. Asterisk indicates a significant difference of $p < 0.05$ between the challenged and control (PBS) groups at the same time point.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2019 doi:10.20944/preprints201901.0263.v1

Peer-reviewed version available at Insects 2019, 10, 76; doi:10.3390/insects10030076

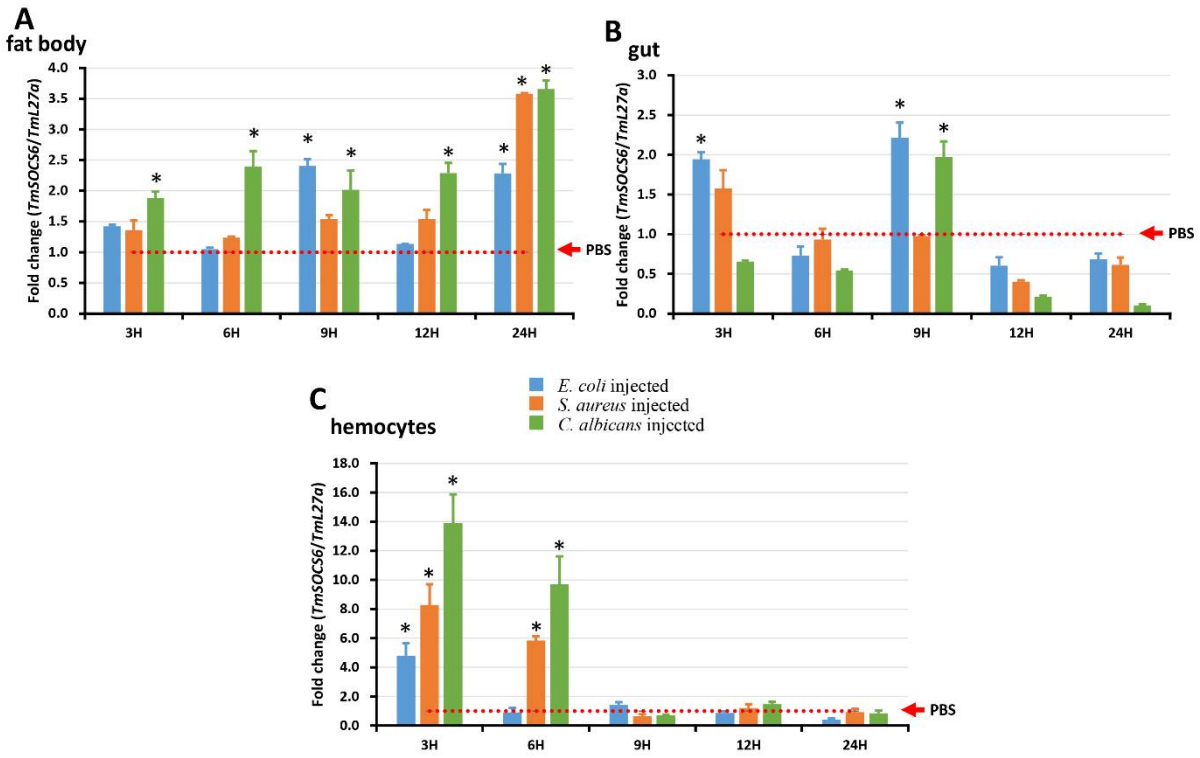
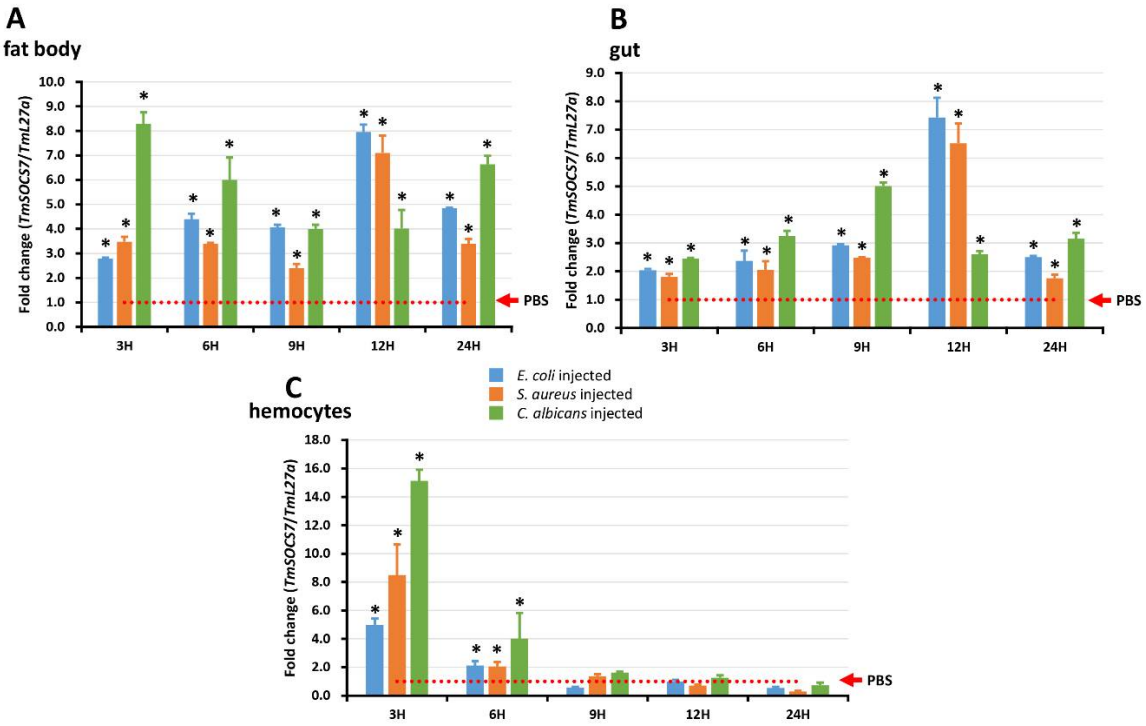


Figure 13. Expression of TmSOCS7 in the fat body (A), gut (B) and hemocytes (C) following challenge of the host with *E. coli, S. aureus and C. albicans*. Analysis of the expression was done using real-time PCR using RpL27a as the normalizing control. The relative fold-change expression of TmSOCS6 is the mean of three independent measurements ± SE. Asterisk indicates a significant difference of $p < 0.05$ between the challenged and control (PBS) groups at the same time point.



3. Materials and Methods

3.1. Insect rearing

The mealworm, *Tenebrio molitor* was reared in the laboratory with 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 4,400 ml of D.W; autoclaved at 121 °C for 15 min) at 26 ± 1 °C, 60% ± 5% relative humidity and in the dark.

3.2. Microorganisms

The Gram-negative bacteria (*Escherichia coli* strain K12), Gram-positive bacteria (*Staphylococcus aureus* strain RN4220), and the fungus (*Candida albicans*) were used for immune challenge experiments. *E. coli* and *S. aureus* were cultured overnight in Luria Bertani (LB) broth at 37 °C. *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 3,500 rpm for 10 min. The samples were then suspended in PBS and the concentrations were measured at OD_{600}. The microorganisms were diluted to 10^6 cells/µl of *E. coli* and *S. aureus*, and 5×10^4 cells/µl of *C. albicans* for immune challenge studies.

3.3. Identification and in silico characterization of TmSOCS genes

The sequences of TmSOCS5, TmSOCS6, and TmSOCS7 were retrieved from the *Tenebrio* RNAseq and EST libraries using the standalone *Tenebrio* local-blast server. Sequence retrieval was conducted by local-tblastn analysis using *Tribolium castaneum* SOCS5 (XP_015833441.1), SOCS6 (XP_008190646.1), and SOCS7 (XP_008190646.1) amino acid sequences as queries. Conserved domains were identified using the InterProScan (http://www.ebi.ac.uk/interpro/search/sequence-search) and blastx (https://blast.ncbi.nlm.nih.gov/Blast.cgi) programs. Clustal X2 [31] was used for domain-specific multiple alignment with representative SOCSs from other insects retrieved from GenBank. The percentage identity and phylogenetic analysis were performed using the Clustal X2 and MEGA7 programs [32], respectively. The evolutionary relationship of taxa was inferred using the Neighbor-Joining method [33]. The bootstrap consensus tree was inferred from 1000 replicates and the evolutionary distances computed using the Poisson correction method. Human SOCS1 belonging to the Type II subfamily of SOCS genes was used as an outgroup.

3.4. Cloning the TmSOCS ORF sequences

Based on the identified TmSOCS sequences, primers were designed to amplify the ORF regions by AccuPower® PyroHotStart Taq PCR PreMix (Bioneer, Korea) using standard polymerase chain reaction (PCR) reactions. The primers are listed in Table 1. Briefly, 1 µg of RNA was used as the starting material to synthesize cDNAs using the Oligo(dT) primers. The cDNA was diluted 20 times, 1 µl of which was used in the PCR reaction consisting of an initial denaturation 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 product thus obtained was purified using the AccuPrep® PCR Purification Kit (Bioneer, Korea) and immediately ligated into a T-Blunt vector (Solgent, Korea) The ligated PCR product was transformed into *E. coli* DH5α competent cell according to the manufacturer's instructions. After validation using colony PCR, plasmid DNA was extracted from the cells using an AccuPrep® Nano-Plus Plasmid Extraction Kit (Bioneer, Korea). The sequencing of the cloned ORF region was conducted using the M13 forward and reverse primers listed in Table 1.

Table 1. Primers used in the study.

Name	Primer sequences
TmSOCS5-cloning_Fw	5'-CCCCCTGAGATGTGATTTCC-3'
TmSOCS5-cloning_Rv	5'-AACACCGCACATGAAAACAA-3'
TmSOCS5-qPCR_Fw	5'-CGCGCCCAAAGACAAGAAAATC-3'
TmSOCS5-qPCR_Rv	5'-TTTGGTGGGCCTTCTTGTTG-3'
TmSOCS5-qPCR_Fw2	5'-TCACGTTTTCCCTGCAACAC-3'
TmSOCS5-qPCR_Rv2	5'-TGCAGGAGGTTGATGTTGTC-3'
TmSOCS6-cloning_Fw	5'-AGTGTCGGTTGTGCGTGGT-3'
TmSOCS6-cloning_Rv	5'-GCGCGATTACTAAAAGTACGG
TmSOCS6-qPCR_Fw	5'-TAAAGAGAAGCCTGCAGGACAG-3'
TmSOCS6-qPCR_Rv	5'-TCCGAGTGCTCCAAAACTTC-3'
TmSOCS7-cloning_Fw	5'-CAGTGTCTCACGATACGCTTTC-3'
TmSOCS7-cloning_Rv	5'-AGTCTCAGGATTGTCTGGGATT-3'
TmSOCS7-qPCR_Fw	5'-ATTGAAGCGCGACAGTACAG-3'
TmSOCS7-qPCR_Rv	5'-AAGTCATGTGGATGGGTTCCC-3'
TmL27a_qPCR_Fw	5'-TCATCCTGAAGGCAAAGCTCCAGT-3'
TmL27a_qPCR_Rev	5'-AGGTTGGTTAGGCAGGCACCTTTA-3'
M13_Fwd	5' –GTAAAACGACGGCCAGT -3'
M13_Rev	5'- CAGGAAACAGCTATGAC-3'

3.5. Developmental and Tissue-specific expression of TmSOCS transcripts

Total RNA was isolated from egg, early larvae (12^{th} – 15^{th} instar larvae), late larvae, pre-pupae, 1-7 days old pupae, and 1-5 days old adult insects to monitor the expression of TmSOCS5, TmSOCS6, and TmSOCS7 during development. For tissue-specific expression analysis of TmSOCS transcripts, total RNA was isolated from the hemocytes, gut, Malpighian tubules, fat body and integument of late-instar larvae. In addition, total RNA was isolated from the ovary and testis tissues of 5-day old adult *T. molitor*. The LogSpin RNA isolation method (Yaffe et al. 2012) with minor modifications was used to isolate total RNAs from tissue and whole-body samples. Briefly, the samples were homogenized using 1 ml of guanidine thiocyanate RNA Lysis buffer (20 mM EDTA, 20 mM MES buffer, 3 M guanidine thiocyanate, 200 mM sodium chloride, 40 µM, 005% tween-80, and 1 % isoamyl alcohol in D. W, pH 5.5), and centrifuged at 21,000 x g for 5 min at 4 ºC. After 1 min of incubation in absolute ethanol, the samples were transferred to silica spin columns, and centrifuged at 21,000 x g for 30 s at 4 ºC to remove debris. Following 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 the AccuPower® RT PreMix (Bioneer, Korea) and Oligo(dT)$_{12–18}$ primers on a MyGenie96 Thermal Block (Bioneer, Korea). Subsequently, quantitative real-time reverse transcriptase polymerase chain reaction (qRT-PCR) was used to analyze the developmental and tissue-specific expression of the TmSOCS genes. The qRT-PCR assay was performed in a 20 µl reaction volume containing AccuPower® 2X GreenStar qPCR Master Mix (Bioneer) and synthesized primers for TmSOCS5, TmSOCS6, and TmSOCS7 (Table 1). The cycling profile was as follows-initial denaturation at 95 ºC followed by 40 cycles with denaturation at 95 ºC and annealing at 60 ºC for 30 s. *T. molitor* ribosomal protein L27a (*TmL27a*) gene was used for normalization and the results were analyzed by the

ΔΔCt method [34].

3.6. Immune challenge studies

Healthy *T. molitor* larvae were challenged with 10^6 cells/µl of *E. coli* and *S. aureus* and 5 x 10^4 cells/µl of *C. albicans*. A phosphate buffered saline (PBS) injection group was used as a wounding control. Fat body, gut and hemocytes tissue were collected at 3, 6, 12, and 24 h to study the time-course expression of TmSOCS5, TmSOCS6, and TmSOCS7 after challenge. Total RNA isolation, cDNA synthesis, and qRT-PCR was performed as described above. To determine the expression of TmSOCS transcripts under microbial challenge, the fold-change at each time point was compared with that of the PBS-injected control. All data shown is presented as mean ± standard error. The one-way analysis of variance (ANOVA) and Tukey's multiple range tests were used to evaluate the difference between groups ($p < 0.05$).

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 expect to screen for type II SOCS family members in *T. molitor* RNA-Seq and Genome-Seq data and study the evolutionary analysis considering the type I and type II subfamilies. Furthermore, the upregulation of the TmSOCS transcripts in the immune tissues of *T. molitor* after microbial challenge suggests that they are critical for immune reaction in the host. In the future, RNA interference analysis will be conducted by us to study the involvement of type I *Tenebrio* SOCS family members in the regulation of key cytokine regulatory pathways. The information available on insect SOCS genes is currently very limited, this study provides an extended repertoire of negative regulators for maintaining cellular homeostasis in insects.

Acknowledgments

References

1. Ghoreschi, K.; Laurence, A.; O'Shea, J.J. Janus kinases in immune cell signaling. *Immunol. Rev*. **2009**, 228, 273-287.
2. Harrison, D.A. The Jak/STAT pathway. *Cold Spring Harb. Perspect. Biol.* **2012**, 4, 1-3.
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.
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**, 7038-7048.
5. Seif, F.; Khoshmirsafa, 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.
6. Hatakeyama, S. Ubiquitin-mediated regulation of JAK-STAT signaling in embryonic stem cells. *JAKSTAT* **2012**, 1, 168-175
7. 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.
8. Liongue, C.; Sertori, R.; Ward, A.C. Evolution of cytokine receptor signaling. *J. Immunol.* **2016**, 197, 11-18.

9. 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.
10. 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.
11. Wang, T.; Secombes, C.J. Rainbow trout suppressor of cytokine signaling (SOCS)-1, 2 and 3: molecular identification, expression and modulation. *Mol. Immunol*. **2008**, 45, 1449-1457.
12. 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.
13. 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.
14. 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.
15. 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.
16. 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.
17. 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.
18. Wang, S.; Song, X.; Zhang, Z.; Li, H.; Lu, Kai.; 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.**
19. 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.
20. 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.
21. 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.
22. 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.
23. 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.
24. 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.
25. Hwang, M.N.; Min, C.H.; Kim, HS.; 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.Δ
26. 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.
27. 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.

28. 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.
29. 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.
30. 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.
31. Larkin, M.A.; Blackshields, G.; Brown N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; Thompson, J.D.; Gibson, T.J.; Higgins, D.G. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, 23, 2947-2948.
32. 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.
33. Saitou, N.; Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987,** 4, 406-425.
34. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2ΔΔCT method. *Methods* **2001**, 25, 402-408.