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

Bharat Bhusan Patnaik1,2, Bo Bae Kim1, Yong Hun Jo1, In Seok Bang3

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, SOCS1, 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.

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.
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* (AUY61566.1), *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).

**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
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).

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 seteki (KYQ56054.1), Cyphomyrmex costatus (KYM98051.1), Camponotus floridanus (EFN69786.1), Lasius niger (KM986622), Harpegnathos saltator (EFN83798.1), Habropoda laboriosa.
(KOC61557.1), Eufrisea mexicana (OAD58426.1), Melipona quadrifasciata (KOX74779.1), Danaus plexippus plexippus (OWR46815.1).

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...
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 TmSOCS6 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.
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.
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 *fkh*, *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.

**Figure 12.** 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.
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.

<table>
<thead>
<tr>
<th>Name</th>
<th>Primer sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>TmSOCS5-cloning_Fw</td>
<td>5’-CCCCCTGAGATGTGATTTC-3’</td>
</tr>
<tr>
<td>TmSOCS5-cloning_Rv</td>
<td>5’-AACACGCACATGAAAAACA-3’</td>
</tr>
<tr>
<td>TmSOCS5-qPCR_Fw</td>
<td>5’-CAGGCCAAGCAAGAAAATC-3’</td>
</tr>
<tr>
<td>TmSOCS5-qPCR_Rv</td>
<td>5’-TTTGTTGGCCCTCTTTG-3’</td>
</tr>
<tr>
<td>TmSOCS5-qPCR_Fw2</td>
<td>5’-TCACGTTCCTGCAACAC-3’</td>
</tr>
<tr>
<td>TmSOCS5-qPCR_Rv2</td>
<td>5’-TGACGAGGTGATGTTGTC-3’</td>
</tr>
<tr>
<td>TmSOCS6-cloning_Fw</td>
<td>5’-AGTGTCGTTGTGCTGGT-3’</td>
</tr>
<tr>
<td>TmSOCS6-cloning_Rv</td>
<td>5’-GCCGATTACTAAAGTACGG-3’</td>
</tr>
<tr>
<td>TmSOCS6-qPCR_Fw</td>
<td>5’-TAAAGAGAGCCCTGAGAGAC-3’</td>
</tr>
<tr>
<td>TmSOCS6-qPCR_Rv</td>
<td>5’-TCCGAGTCCTCAAAACTTC-3’</td>
</tr>
<tr>
<td>TmSOCS7-cloning_Fw</td>
<td>5’-CAGTGTCTCACGATACGCTTC-3’</td>
</tr>
<tr>
<td>TmSOCS7-cloning_Rv</td>
<td>5’-AGTCTCAGGATTGCTGGATT-3’</td>
</tr>
<tr>
<td>TmSOCS7-qPCR_Fw</td>
<td>5’-ATTGAAGCGCGACAGTACAG-3’</td>
</tr>
<tr>
<td>TmSOCS7-qPCR_Rv</td>
<td>5’-AAGTCTAGTGGATGGTCCC-3’</td>
</tr>
<tr>
<td>TmL27a_qPCR_Fw</td>
<td>5’-TCATCCTGAAGCGAAAGTCC-3’</td>
</tr>
<tr>
<td>TmL27a_qPCR_Rev</td>
<td>5’-AGGTTGGTTAGGCGACCTTTA-3’</td>
</tr>
<tr>
<td>M13_Fwd</td>
<td>5’–GTAAACGACGAGCCAGT-3’</td>
</tr>
<tr>
<td>M13_Rev</td>
<td>5’–CAGGAAACAGCTATGAC-3’</td>
</tr>
</tbody>
</table>

3.5. Developmental and Tissue-specific expression of TmSOCS transcripts

Total RNA was isolated from egg, early larvae (12th – 15th 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, Malphigian 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 °C. 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 °C 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 °C followed by 40 cycles with denaturation at 95 °C and annealing at 60 °C 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

6. Hatakeyama, S. Ubiquitin-mediated regulation of JAK-STAT signaling in embryonic stem cells. JAKSTAT 2012, 1, 168-175


