



Swansea University  
Prifysgol Abertawe



## Cronfa - Swansea University Open Access Repository

---

This is an author produced version of a paper published in:  
*Scientific Reports*

Cronfa URL for this paper:  
<http://cronfa.swan.ac.uk/Record/cronfa36405>

---

### **Paper:**

Zhang, W., Wanchoo, A., Ortiz-Urquiza, A., Xia, Y. & Keyhani, N. (2016). Tissue, developmental, and caste-specific expression of odorant binding proteins in a eusocial insect, the red imported fire ant, *Solenopsis invicta*. *Scientific Reports*, 6(1)  
<http://dx.doi.org/10.1038/srep35452>

---


This item is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence. Copies of full text items may be used or reproduced in any format or medium, without prior permission for personal research or study, educational or non-commercial purposes only. The copyright for any work remains with the original author unless otherwise specified. The full-text must not be sold in any format or medium without the formal permission of the copyright holder.

Permission for multiple reproductions should be obtained from the original author.

Authors are personally responsible for adhering to copyright and publisher restrictions when uploading content to the repository.

<http://www.swansea.ac.uk/library/researchsupport/ris-support/>

# SCIENTIFIC REPORTS



OPEN

## Tissue, developmental, and caste-specific expression of odorant binding proteins in a eusocial insect, the red imported fire ant, *Solenopsis invicta*

Received: 15 June 2016  
Accepted: 28 September 2016  
Published: 21 October 2016

Wei Zhang<sup>1,2</sup>, Arun Wanchoo<sup>2</sup>, Almudena Ortiz-Urquiza<sup>2</sup>, Yuxian Xia<sup>1</sup> & Nemat O. Keyhani<sup>1,2</sup>

Insects interact with the surrounding environment via chemoreception, and in social insects such as ants, chemoreception functions to mediate diverse behaviors including food acquisition, self/non-self recognition, and intraspecific communication. The invasive red imported fire ant, *Solenopsis invicta*, has spread worldwide, displaying a remarkable environmental adaptability. Odorant binding proteins (OBPs) are chemical compound carriers, involved in diverse physiological processes including odor detection and chemical transport. *S. invicta* contains a highly divergent 17-member OBP gene family, that includes an ant-specific expansion and the social organization implicated *Gp-9* (*OBP3*) gene. A systematic gene expression analysis of the *SiOBP* repertoire was performed across social caste (workers, male and female alates), tissues (antennae, head, thorax, and abdomen), and developmental stages (egg, larvae, and pupae), revealing that although *SiOBPs* were expressed in the antennae, the major regions of expression were in the head and thorax across all castes, and the abdomen in male and female alates. *SiOBPs* were very highly expressed in female alates and at somewhat lower levels in male alates and workers. *SiOBPs* were differentially expressed, with unique signatures in various castes and tissues, suggesting functionality of *SiOBPs* beyond olfaction. Expression patterns of *SiOBP* subgroups also showed relationships with their evolutionary relatedness.

Olfaction underlies significant aspects of individual and social behaviors in insects. Insect olfaction is mediated by a subset of specialized cuticular structures known as sensilla, most commonly found on the antennae, although sensilla can be found on other body structures as well, i.e. the maxillary palp and at the ends of various appendages<sup>1,2</sup>. Olfactory sensilla contain olfactory sensory neurons (OSNs), suspended in an aqueous lymph, that in turn harbor olfactory receptors (ORs) in their dendritic membranes. While ORs are capable of being directly affected by chemical compounds<sup>3</sup>, many odorants are hydrophobic and/or have low aqueous solubility, necessitating a means of delivering odorants to ORs or other receptors. Within this context, several classes of chemosensory ligand carrier proteins have been described, one of the major being the family of odorant binding proteins (OBPs; these proteins do not share homology to the similarly named vertebrate OBPs)<sup>4</sup>. OBPs are small molecular weight (~130–150 amino acids), soluble, ligand carrier proteins and include the pheromone binding proteins (PBPs) and the general odorant binding proteins (GOBPs). OBPs have also been classified into three groupings; classic, C-minus, and C-plus, based on the distribution and number of cysteine residues. Insect OBPs exist as gene families and a significant number have been identified either via genome sequencing and/or expression (e.g. RNA-seq) analyses, particularly those focused on the antennal transcriptome<sup>5–7</sup>. Genomic and transcriptomics analyses in several ant species strongly support the idea that both OBPs and the related chemosensory proteins (CSPs) are expressed in the antennae, however many of these proteins appear to be expressed primarily in non-chemosensory tissues<sup>8</sup>. Thus, while a number of OBPs have been linked to olfaction, the roles

<sup>1</sup>Genetic Engineering Research Center, School of Life Sciences, Chongqing University, Chongqing 400045, PR China.

<sup>2</sup>Department of Microbiology and Cell Science, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL, 32611, USA. Correspondence and requests for materials should be addressed to N.O.K. (email: keyhani@ufl.edu)

of many OBPs remain obscure and they are likely to play important roles as chemical carriers in a wide range of physiological processes<sup>9</sup>.

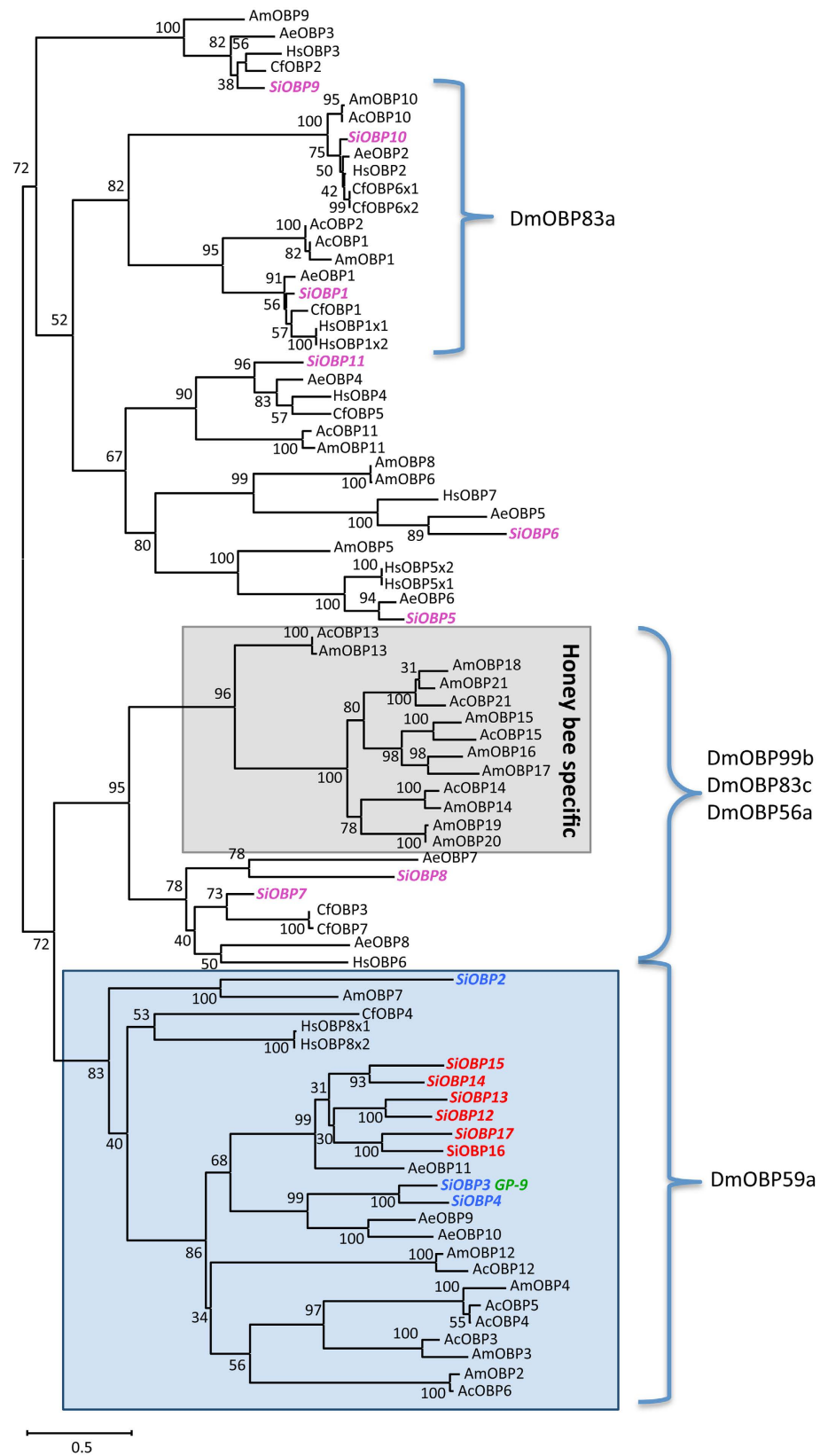
Ants are eusocial organisms having evolved a complex level of communal organization highlighted by specialization of labor and reproduction. Ant societies have been referred to as “super-organisms”, i.e. a group of cooperating animals resulting from “the combined operation of tiny and short-lived minds” leading to a colony that functions, on some levels, as a single organism<sup>10</sup>. The red imported fire ant, *Solenopsis invicta* forms complex societies that are territorial and whose members display strong nest-mate recognition, elaborate task specialization, and a multi-tiered caste system<sup>11</sup>. Workers represent a terminal differentiated form, and reproductives, i.e. winged male and female alates are produced seasonally as the colony matures, a process that can take up to five years. Depending upon their genetic background *S. invicta* colonies can have one (monogyny) or multiple queens (polygyny)<sup>12</sup>. These two differing colony organizations extend to important differences in the biology and responses of *S. invicta*, and early genetic investigations implicated variation at the *Gp-9* gene as being linked to the regulation of these divergent social behaviors<sup>13,14</sup>. *Gp-9* overexpression has also been associated with fire ant workers<sup>15</sup>, and the gene has subsequently been shown to be an odorant binding protein (SiOBP3)<sup>16</sup> and its role in mediating social organization has been challenged<sup>17,18</sup>. As the genome of *S. invicta* was described, a 13 Mb “Y-like social chromosomal” fragment (~55% of the chromosome) that includes the *Gp-9/SiOBP3* gene was found<sup>19,20</sup>. Recombination appears suppressed within this region that contains at least 616 open reading frames, and its rearrangement correlates with the divergent social phenotypes. Thus *Gp-9/SiOBP3* may form one protein product among many that contribute to certain aspects of social behavior.

Analysis of the *S. invicta* genome has indicated the presence of 17 OBPs genes and one OBP pseudogene<sup>16</sup>. Here we have performed a systematic analysis of the expression of the 16 *S. invicta* OBPs (no signal for *SiOBP17* was detected in any samples examined) across different tissues; antennae, head, thorax, and abdomen, castes; workers, and male and female alates, and developmental stages; eggs, larvae, and pupae. These data revealed that *SiOBPs* are generally more highly expressed in head and thorax tissues than in the antennae of either workers or male and female alates. *SiOBP* expression was also abundant in the abdomen of male and female alates but low in the abdomen of workers. *SiOBPs* were also differentially expressed throughout all of the developmental stages examined. *SiOBP3* (*Gp-9*) was highly expressed in female alates (head, thorax, and abdomen but lower in the antennae), in worker heads, and to a somewhat lower extent in male alates (head, thorax, abdomen > antennae). Our data provide a systematic overview of OBP expression in a social insect, suggesting their broad participation in developmental and physiological processes beyond olfaction, allowing for further functional probing of their activities.

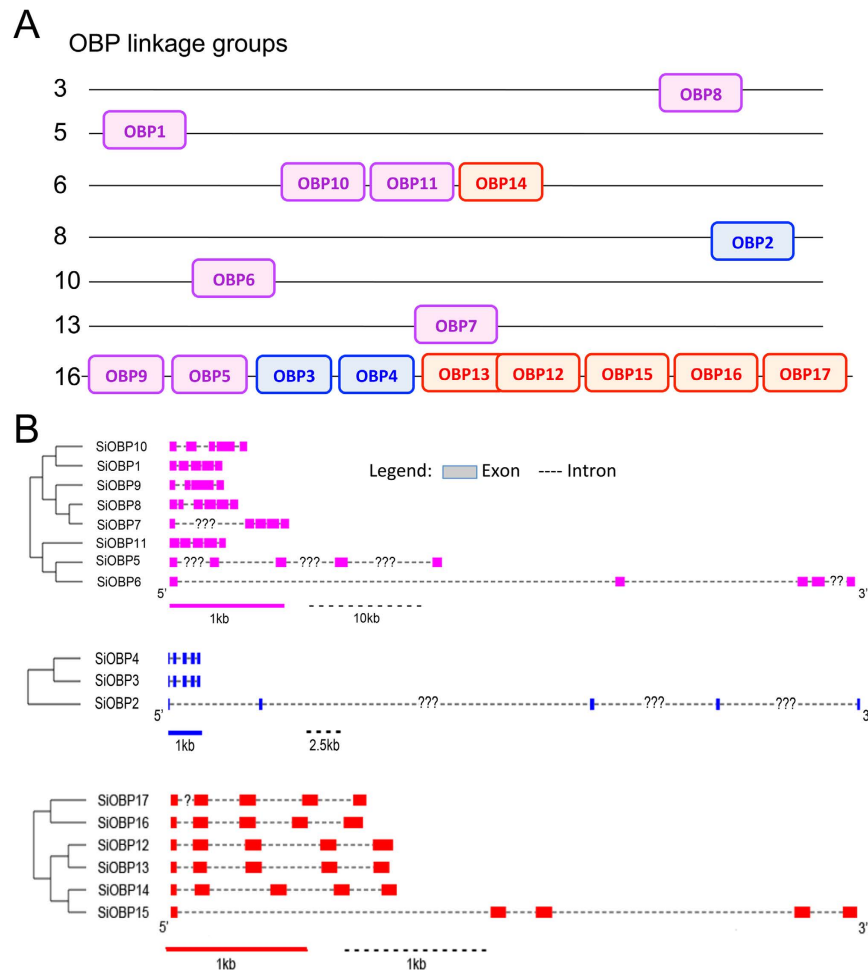
## Results

**The odorant binding protein repertoire of *Solenopsis invicta*.** A set of 17 OBP genes has been previously identified in the *S. invicta* genome (Supplemental Table S1). A limited phylogenetic tree of deduced protein sequences that includes OBPs from the available ant genomes; i.e. *Acromyrmex echinator* (New World, fungus growing, leaf cutter ant), *Camponatus floridanus* (carpenter ant), and *Harpegnathos saltator* (Indian/Jerdon jumping ant), two honey bee species; *Apis mellifera* (European honeybee) and *Apis cerana* (Asiatic honeybee), illustrates the division of the SiOBPs into two discrete major branches (Fig. 1, related *Drosophila melanogaster* OBPs are also shown). SiOBPs 1, 5, 6, 9, 10, and 11 were distributed with various orthologous OBPs from different insects (top clade in figure). On the other branch (bottom in figure), two discrete clusters could be discerned; one containing SiOBPs 7 and 8, along with other ant orthologs (e.g. AeOBP7 and CfOBPs 3 and 7) that were distinct from an apparent honey bee OBP gene expansion. In the second cluster (lower part of figure) SiOBP2 appeared distinct and distributed with orthologous sequences from the honeybee (*A. mellifera* OBP7). The remaining SiOBPs; 3 & 4, and 12 through 17, formed closely grouped members, respectively. SiOBPs 12–17 form a *S. invicta* OBP gene expansion as has been previously reported<sup>16</sup>, along with an ortholog identified in *A. echinator* (AeOBP11).

A schematic of the genomic context of the *SiOBPs* revealed their distribution within 7 linkage groups (Fig. 2A, OBPs are color coded as in Fig. 1). *SiOBPs* were found on linkage groups as follows: LG-3, *SiOBP8*; LG-5, *SiOBP1*; LG-6, *SiOBPs* 10, 11, and 14; LG-8, *SiOBP2*; LG-10, *SiOBP6*; LG-13, *SiOBP7*; LG-16, *SiOBPs* 3–5, 9, 12, 13, and 15–17. This latter linkage group (LG-16) corresponds to the Y-like social chromosome implicated as controlling two divergent forms of social organization (monogyny versus polygyny) in *S. invicta*<sup>19</sup>. *SiOBPs* 12 and 13 were overlapping in their overall genomic contexts, with the *SiOBP13* start site 5172 bp before that of *SiOBP12*. These OBPs also shared the last exon, with all others different, thus sequence corresponding the exonic region is found partially or totally in the intronic sequences of the other. Within the context of the *SiOBPs*, a correlation between phylogeny and intron/exon structure, as determined by the online Gene Structure Display Server (GSDS: <http://gsds.cbi.pku.edu.cn>) was noted (Fig. 2B). Members of the general *SiOBPs* (1, 5–11) could be further separated into two branches, one consisting of *SiOBPs* 1 and 7–10 and the other of *SiOBPs* 5, 6, and 11. All of former, with the exception of *SiOBP7* whose length of the first intron could not be determined due to incomplete sequence information, were similar in exon/intron structure, having numerous small introns with an overall coding/non-coding length <1 kb. *SiOBP11* was similar to the others as above, and three of the introns in *SiOBP5* were indeterminate, thus precluding full analyses, although it did contain at least one intron ~5 kb in length. However, *SiOBP6* was notably different than the others in this group, containing two large introns (15–38 kb), one indeterminate intron, and an overall sequence length of >60 kb. Of *SiOBPs* (i.e. 2, 3 and 4) found in the clade containing the ant-specific OBP gene expansion (but distinct from these, see Fig. 1), *SiOBPs* 3 and 4, were essentially identical in intron/exon structure (small intron, overall genomic sequence length ~1 kb), highly similar in sequence (69% identity), and tandemly arranged in linkage group 16 (Fig. 2A), strongly supporting a recent gene duplication event. In contrast, *SiOBP2* contained at least three large (9–23.5 kb) introns, one indeterminate, and an overall genomic context of >50 kb. Members of the ant-specific *SiOBP* gene expansion (*SiOBPs* 12–17) were similar



**Figure 1. Phylogenetic analyses of *S. invicta* chemosensory proteins (SiOBPs).** Limited phylogenetic analyses of *S. invicta* OBPs compared to the OBP repertoires found in the ant species, *A. echinator* (AeOBP), *C. floridanus* (CfOBP), and *H. saltator* (HsOBP), and the honey bees, *A. mellifera* (AmOBP) and *A. cerana* (AcOBP) (Genbank accession numbers given in Supplemental Table S4). The closest *D. melanogaster* homologs genetically characterized are shown after brackets. Numbers at nodes indicate bootstrap values. The tree is midpoint-rooted in the absence of a suitable out-group.

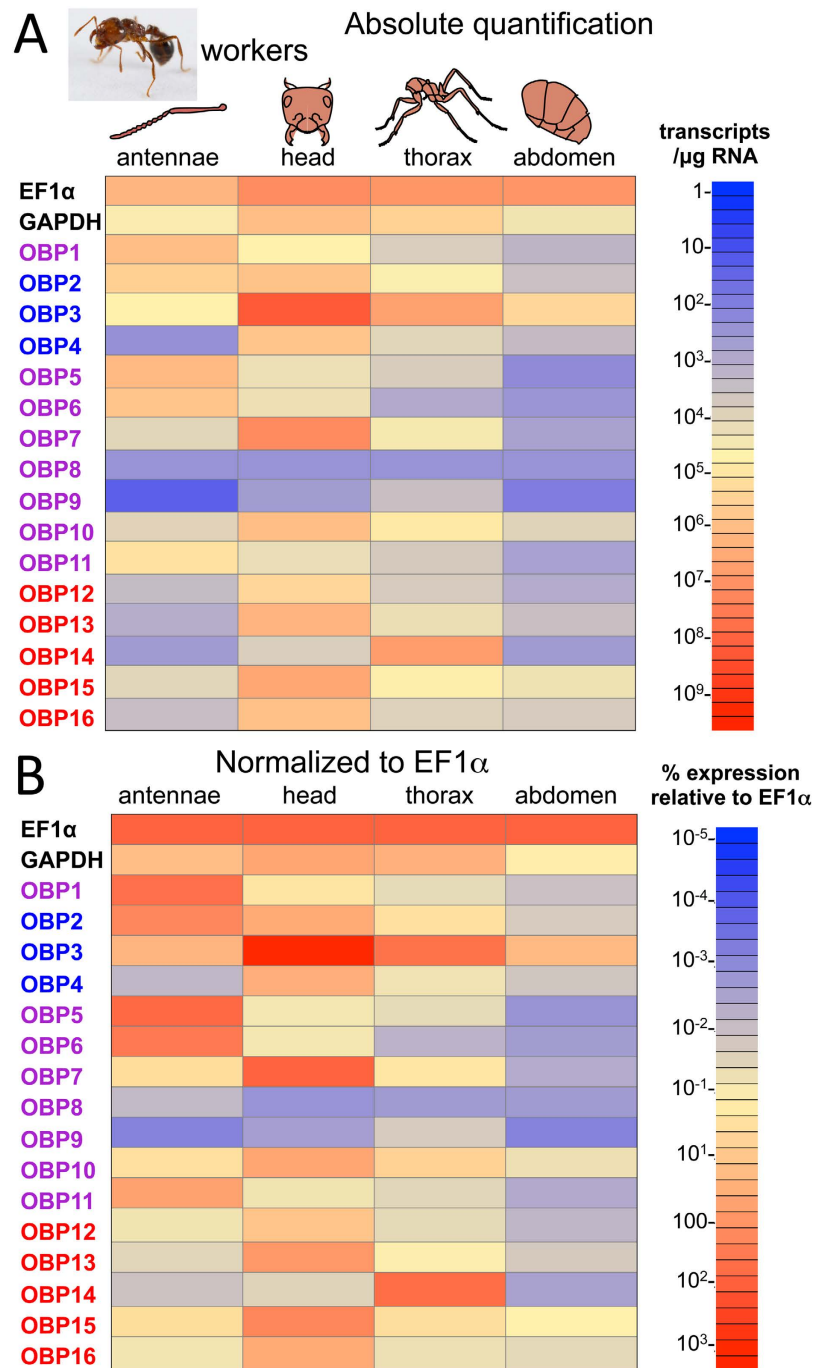


**Figure 2.** (A) Mapping of *SiOBPs* to genomic contigs. *SiOBPs* positions are not drawn to scale but indicate approximate locations on the linkage groups. *SiOBPs* 12 and 13 were overlapping in genomic context (see text for explanation). *SiOBP* exon boxes are color coded to match their phylogenetic placement as in Fig. 1 (B) Intron-exon structures of the *SiOBPs*.

to each other (with the exception of *SiOBP15*) and distinct from the other *SiOBPs*. *SiOBPs* 12, 13, 14, and 16 contained introns 100–500 bp in length (somewhat larger than seen for the *SiOBPs* 3 and 4 described above) with overall genomic sequences of ~1.5 kb. *SiOBP15* contained one intron of ~1.7 kb and another of indeterminate size, with an overall genomic context of >5 kb.

**Expression pattern of the *SiOBPs* in *S. invicta* workers.** Oligonucleotide primers were designed and validated for quantitative RT-PCR as described in the Methods section. Primer efficiency, amplification of single bands corresponding to the predicted size of each amplicon was verified (Supplementary Tables S2 and S3). All amplicons were cloned and used to construct standard curves for absolute quantification of each respective transcript number in the samples. In addition to primers designed to the 17 *SiOBPs*, primers were also designed to 2 different housekeeping genes, glyceraldehyde-phosphate dehydrogenase (GAPDH) and elongation factor 1 $\alpha$  (EF1 $\alpha$ ) for use as references. As significant variation of both genes was seen across tissues and life stages, absolute quantification of the expression of each gene was performed and the expression levels for GAPDH and EF1 $\alpha$  are presented along with that for the *SiOBPs*. Normalization to the housekeeping genes was also performed in order to allow for better comparison between tissues in certain instances. The expression distribution of *SiOBPs* within each caste or developmental stage is included as supplemental materials (Supplemental Excel File S1). These data show representations of the tissue distribution for each *SiOBP* calculated separately for workers, male and female alates, and across developmental stages by dividing the absolute expression for each sample (cell in the heat map) by the total expression for that *SiOBP* in the caste or developmental stage (Supplemental Excel File S1, “Tissue distribution %” subtab). The original calculated data (# transcript copies/1  $\mu$ g RNA) are also included.

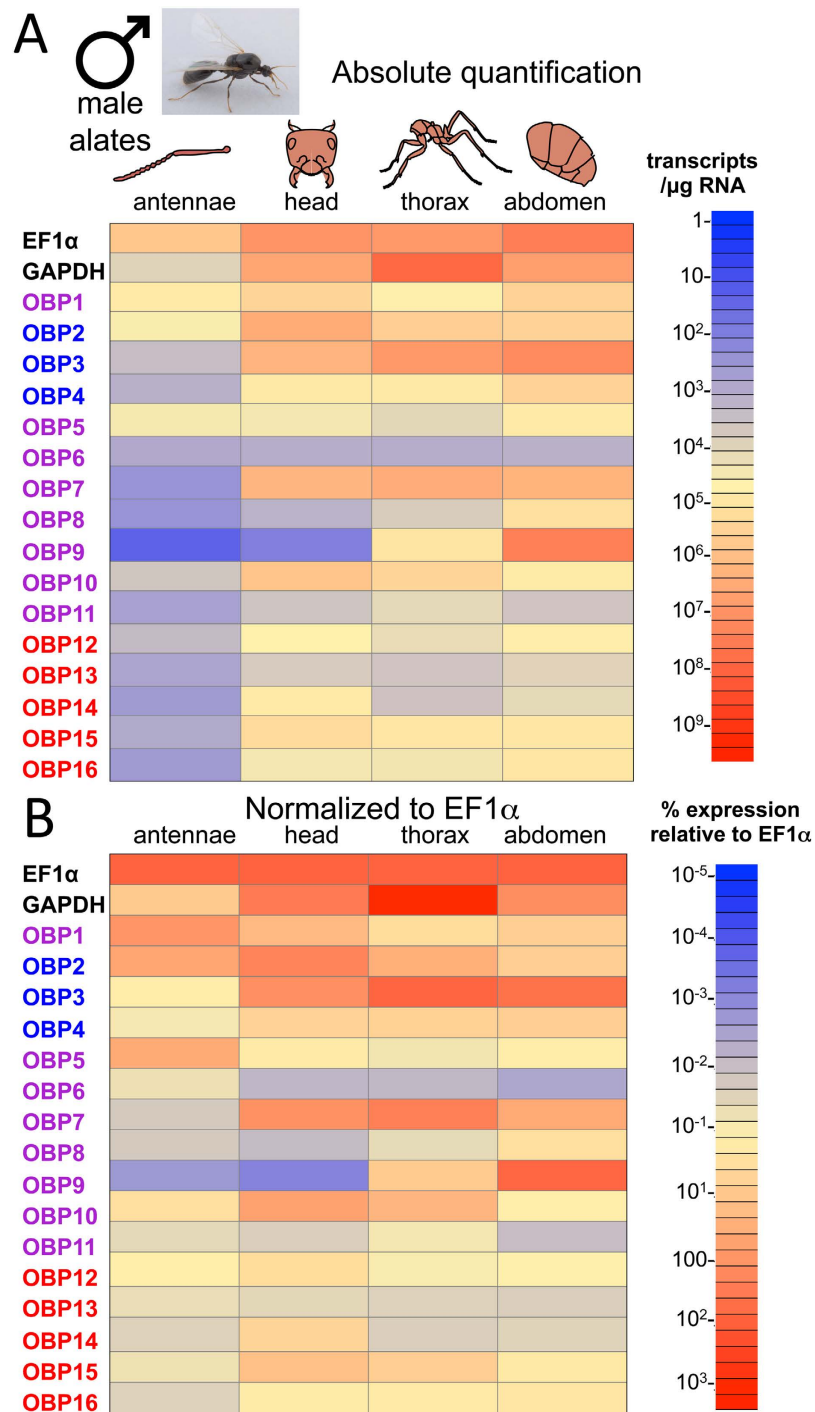
*SiOBPs* gene expression was examined in dissected tissues derived from workers that included antennae, head, thorax, and abdomen (Fig. 3). *SiOBP17* failed to yield any product including when tested against *S. invicta* genomic DNA and was eliminated from further analyses. In terms of absolute expression, *SiOBP* expression was highest in the head; with *SiOBP3* highly expressed, followed by *SiOBP7*, and then *SiOBPs* 2, 4, 10, 12, 13, 15, and 16 (Fig. 3A). Expression levels of *SiOBPs* 8 and 9 in the head were low, whereas moderate expression levels of the remaining



**Figure 3. Expression of *SiOBPs* in *S. invicta* workers.** (A) qRT-PCR was performed on RNA isolated from fire ant worker antennae, head, thorax, and abdomen. Absolute quantification values are shown. Color scale shows transcripts/ $\mu$ g total RNA used for cDNA synthesis. (B) Data in (A) normalized to EF1 $\alpha$  expression levels.

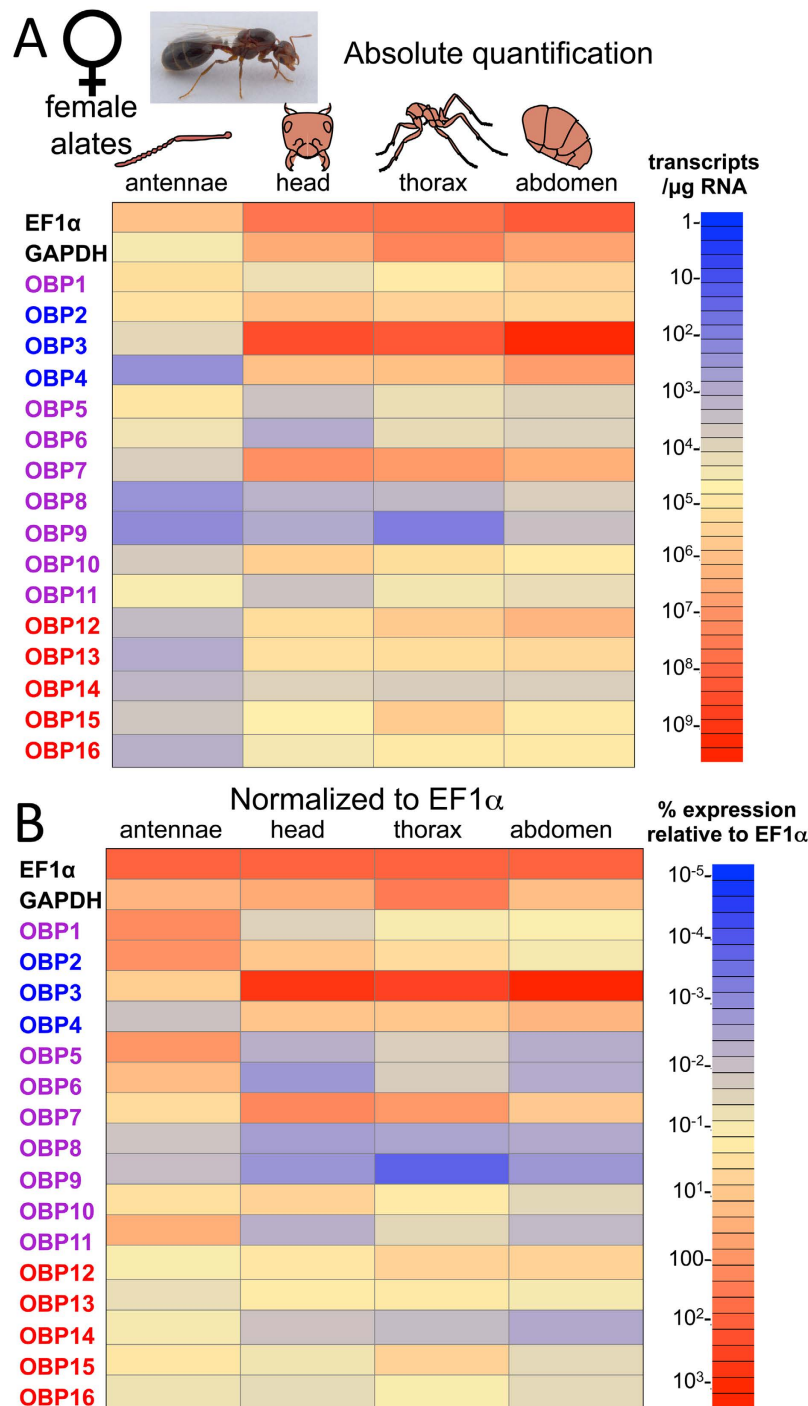
*SiOBPs* (1, 5, 6, 11, and 14) were seen in these tissues. In general, expression of the *SiOBPs* in the antennae ranged from moderate (*SiOBPs* 1, 2, 5, 6, and 11 > *SiOBPs* 3, 7, 10, and 15) to low (*SiOBPs* 4, 8, 9, 12–14, and 16). *SiOBPs* 3 and 14 were highly expressed in the thorax, with moderate expression levels of the remaining OBPs observed, with the exception of *SiOBPs* 6, 8 (in particular), 9, 11, and 12, whose expression levels were lower than the others. *SiOBPs*, in particular OBPs 1, 2, 4–9, 11, 12, and 14, were apparently poorly expressed in the abdominal section, with only moderate levels of OBP3 seen in these tissues. When data are normalized to EF1 $\alpha$  levels the general trends in each tissue remained consistent with the notable change of higher relative expression of the *SiOBPs* in the antennae (Fig. 3B). For the other tissues, normalization resulted in only minor changes in the *SiOBP* expression pattern.





**Figure 4.** Expression of *SiOBPs* in *S. invicta* male alates. (A) qRT-PCR was performed on RNA isolated from fire ant male alate antennae, head, thorax, and abdomen. Absolute quantification values are shown. Color scale shows transcripts/ $\mu$ g total RNA used for cDNA synthesis. (B) Data in (A) normalized to EF1 $\alpha$  expression levels.

**Expression pattern of the *SiOBPs* in *S. invicta* male and female alates.** In terms of absolute expression, *SiOBPs* were poorly expressed in the male antennae, with low transcript levels seen for *SiOBPs* 4, and 6–16, and only moderate expression levels were seen for *SiOBPs* 1–3, and 5 (Fig. 4A). In contrast, male head, thorax, and abdomen regions displayed higher levels of OBP expression. High levels of *SiOBPs* 1–3, 7 and 10 were noted for the male head, with moderate levels of expression seen for *SiOBPs* 4, 5, 12, 14, 15, and 16, low levels seen for *SiOBPs* 6, 8, 11, and 13, with only barely detected levels of *SiOBP9* noted. *SiOBPs* 3 and 7 were highly expressed in the thorax, with moderate expression of *SiOBPs* 2, 9, and 10, with lower levels of *SiOBPs* 1, 4, 15, and 16 were seen in the male thorax, with *SiOBPs* 5, 6, 8, and 11 through 14 expressed at lower levels. *SiOBPs* were generally robustly expressed in the male abdomen, with the exception of *SiOBPs* 6, 8 and to a lower extent *SiOBPs* 13 and



**Figure 5. Expression of *SiOBPs* in *S. invicta* female alates.** (A) qRT-PCR was performed on RNA isolated from fire ant female alate antennae, head, thorax, and abdomen. Absolute quantification values are shown. Color scale shows transcripts/ $\mu\text{g}$  total RNA used for cDNA synthesis. (B) Data in (A) normalized to EF1 $\alpha$  expression levels.

14 whose expression levels were low. *SiOBPs* 3 and 9 were the most highly expressed, followed by *SiOBPs* 1, 2, 4, 5, 7, 8, 10, 12, 15 and 16. Normalization to EF1 $\alpha$  levels had a moderate effect in increasing the apparent expression levels of the *SiOBPs* in the male antennae (Fig. 4B), although overall expression was lower than the other tissues. Only minor changes in *SiOBP* gene pattern expressions were seen for the other male tissues examined when normalized.

Similar to male alates, absolute expression of the *SiOBPs* was lowest in the female antennae as compared to the other tissues (Fig. 5A), however, as seen in the male antennae EF1 $\alpha$  levels were also lower in the female antennal tissues as compared to the others. Low expression levels of *SiOBPs* 3, 4, 7–10, 12–15 were seen in female antennae, whereas moderate expression of *SiOBPs* 6 and 11, and somewhat higher expression levels of *SiOBPs* 1,



2, and 5 were seen in these structures. In contrast very high levels of *SiOBP3* and high levels of *SiOBPs* 2, 4, 7, 10, 12, and 13 were seen in the female alate head. The female alate head also showed moderate expression levels of *SiOBPs* 15 and 16, somewhat lower levels of expression of *SiOBPs* 1 and 14, and low levels of *SiOBPs* 5, 6, 8, 9 and 11 expression. *SiOBP* expression in the female alate thorax was similar to the head. Very high levels of *SiOBP3* were observed, with high levels of expression of *SiOBPs* 2, 4, 7, 10, 12, 13, and 15, seen in the female alate thorax. Moderate to low levels of expression of *SiOBPs* 1, 5, 6, 11, 14 and 16 were seen in these tissues. *SiOBPs* 8, and in particular 9, were expressed at low to very low levels in the female alate thorax. As seen in the head and thorax regions, *SiOBP3* was highly expressed in the female alate abdomen. In addition, *SiOBPs* 1, 2, 4, 7, 12, and 13 were highly expressed in the female abdomen, with moderate expression of *SiOBPs* 10, 15, and 16, and low expression of *SiOBPs* 5, 6, 8, 9, and 14 seen in these tissues. Normalization to *EF1 $\alpha$*  levels, increased the apparent expression levels of the *SiOBPs* in the female antennae, but had only minor effects with respect to the other tissues examined (Fig. 5B).

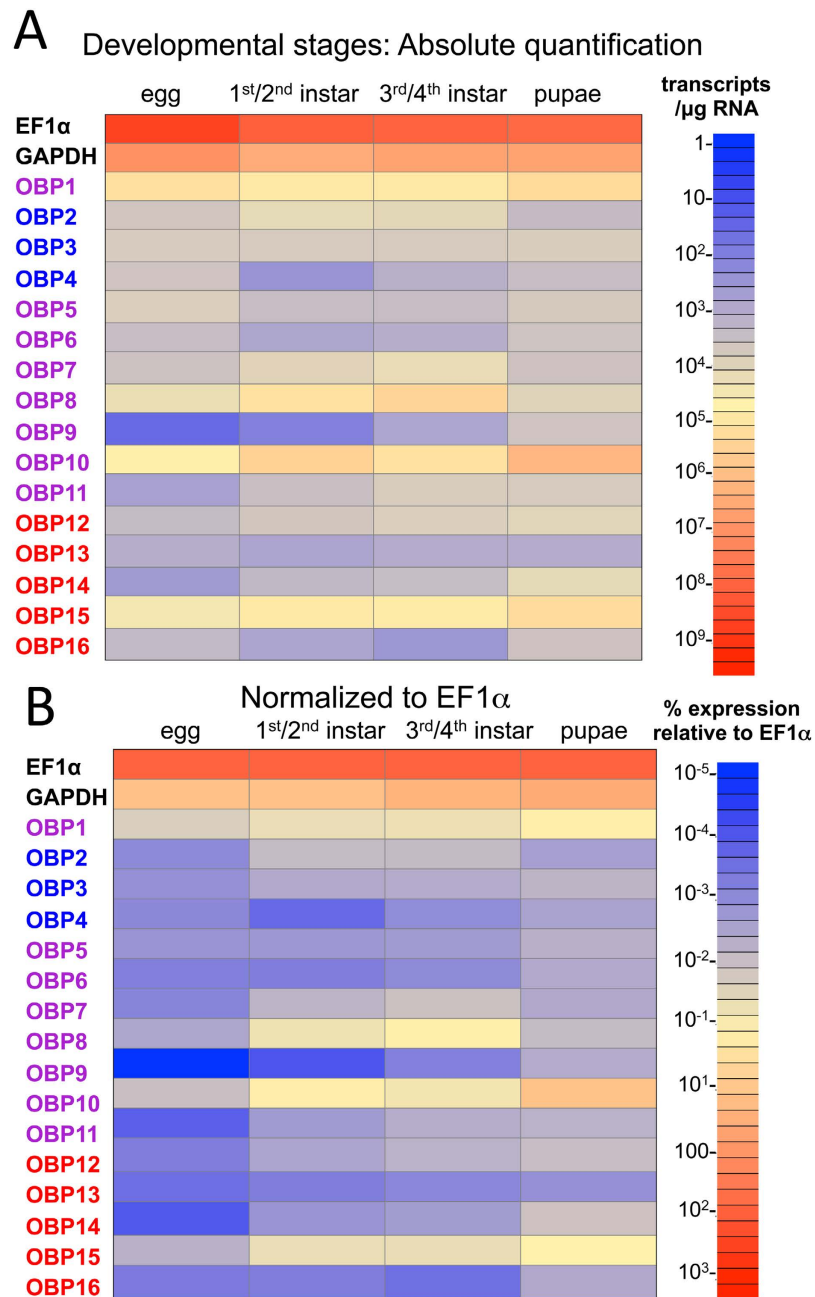
**Expression pattern of the *SiOBPs* in different developmental stages of *S. invicta*.** The expression pattern of the *SiOBPs* was examined in eggs, larvae, and pupae (Fig. 6). Larvae were further separated into early (1<sup>st</sup> and 2<sup>nd</sup>) and late (3<sup>rd</sup> and 4<sup>th</sup>) instars based on size. Absolute expression analyses indicated that *SiOBPs* 1, 8, 10, and 15 were the most highly expressed OBPs in eggs, with *SiOBPs* 2 through 7, 12 and 16, less, but still moderately expressed (Fig. 6A). Compared to the preceding mentioned genes, lower expression of *SiOBPs* 11, 13, and 14 was seen in eggs, and very low expression of *SiOBP* 9 was observed in these samples. Early instar larvae displayed high to moderate levels of *SiOBPs* 1, 8, 10, and 15. In addition, moderate expression of *SiOBPs* 2, 3, 7, and 12, low expression of *SiOBPs* 5, 11, and 14, and very low expression of *SiOBPs* 4, 6, 13, and 9 were seen in the early instar larvae. In late larvae, expression of *SiOBPs* 1, 8, 10, and 15 was moderate to high, expression of *SiOBPs* 2, 3, 7, 11, and 12 was moderate, with moderate to low levels of *SiOBPs* 4, 6, 9, 13 and 16 seen. *SiOBPs* were also found in pupal samples. *SiOBPs* 1, 10, and 15 were highly expressed, with *SiOBPs* 3, 5, 8, 9, 11, 12, 14 and 16 moderately expressed, and *SiOBPs* 2, 4, 6, 7, and 13 moderate/low in expression. In these tissues, as compared to the adults, *EF1 $\alpha$*  levels were significantly higher ( $P < 0.001$ ) and normalization to *EF1 $\alpha$*  resulted in very low relative expression levels for most of the OBPs across the developmental stages (Fig. 6B). *SiOBPs* still moderately expressed during development (after normalization) were few and included *SiOBPs* 7, 9, and 15 seen in early and late instar larvae, and *SiOBPs* 1, 10 (relatively highly expressed), and 15 in pupae.

**Some evolutionarily related sub-groups of *SiOBPs* show similar expression patterns.** Although no overall relationship between the *SiOBP* phylogenetic groups and expression was seen, certain *SiOBPs* in some phylogenetic subgroups showed similar patterns of expression (Supplemental Excel File S2, these data are identical to the data shown in Figs 3–7 and Supplemental Excel File S1, but the order of the *SiOBPs* has been rearranged to match the tree topology). These analyses revealed a number of patterns or relationships between *SiOBP* expression and their phylogenetic relatedness. In terms of absolute expression and somewhat clearer when normalized to *EF1 $\alpha$*  expression, *SiOBPs* 12–16 were expressed in a similar manner across castes (male and female alates), developmental stages, and tissues, although some differences were noted in workers where expression of *SiOBPs* 14 and 15 were dissimilar from the rest (Supplemental Excel File S2). *SiOBPs* 8 and 9, but not 7, were also similarly expressed across caste and tissue, but not in the developmental stages. A pattern was seen between *SiOBPs* 5, 6, and 11 in workers, female alates, and in the developmental stages, but not in male alates, where only *SiOBPs* 6 and 11 were similar in expression. For the subgroup comprised of *SiOBPs* 2, 3, and 4, with the exception of *SiOBP3* (*Gp-9*), which was highly expressed in all castes, similar trends in expression were noted.

When the data were normalized in terms of tissue distribution, several patterns of *SiOBP* expression emerged (Supplemental Excel File S2, “Tissue distribution %” subtab). In workers, *SiOBP* subgroup 5, 6, and 11 were preferentially expressed in the antennae, whereas *SiOBPs* 2, 3, and 4 and 12–16 (with the exception of *SiOBP* 15) were preferentially expressed in the head. These patterns were not seen in the male alate, but instead various subgroup preferences, although not containing all members were seen in the abdomen expression profiles. Patterns could also be seen in the female alates, notably with *SiOBPs* 5, 6, and 11 in the antennae, *SiOBPs* 8, 9 (but not 7), 3, 4 (but not 2), and 12, 13, 16 (but not 14 and 15) preferentially found in the female abdomen.

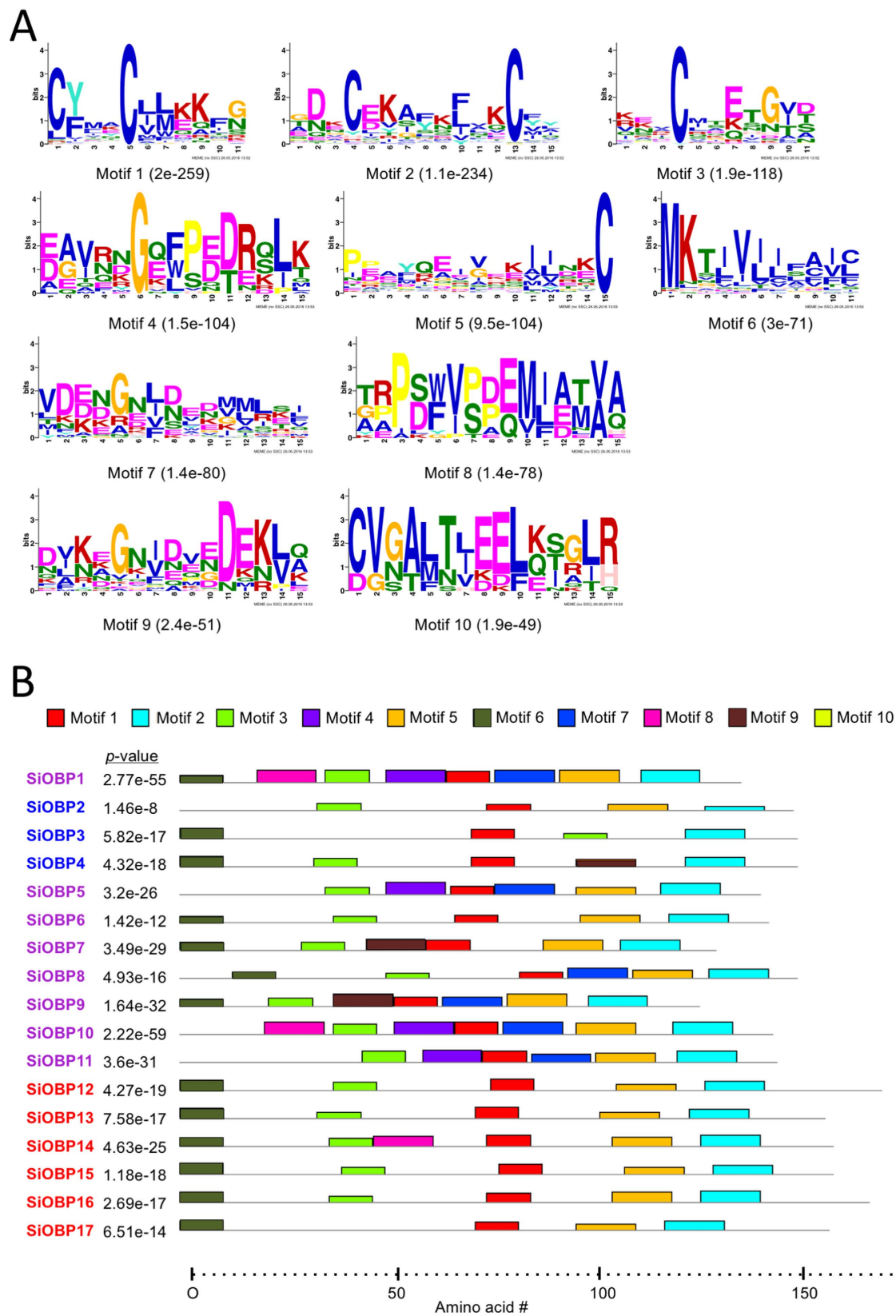
To confirm differential expression across tissues and across castes/developmental stages, four statistical analyses were performed, namely; (1) comparison of the expression of each *SiOBP* across different tissues (Supplemental Excel File S3), (2) comparison of the expression of different *SiOBPs* within each tissue (Supplemental Excel File S4), and (3) comparison of the total expression of each *SiOBPs* across tissues within each caste (Supplemental Excel File S5). These analyses showed significant (1) differential expression of OBPs within the same tissues, (2) differential expression of individual OBPs across tissues, (3) differential expression of individual as well as the entire set of OBPs across castes, and (4) differential expression of OBPs within developmental stages and between developmental stages as well as between the adult stages.

**Protein motifs identified in the *S. invicta* OBPs.** The MEME server<sup>21</sup> was used to identify conserved amino acid motifs found in the 17 *S. invicta* OBPs, along with the other homologs used to build the phylogenetic tree (82 OBPs total). In total 16 motif with  $e$ -values  $< 2e^{-15}$  were identified, with the 10 most common motif patterns, i.e. those with the lowest  $e$ -values and ascending, and their distribution within the *S. invicta* OBPs shown (Fig. 7A,B, respectively). Of the remaining 6 motifs most were not found in the *SiOBPs* and were thus excluded from the analyses. Motif 10 was not found in any of the *SiOBPs*, but was amongst the top 10 most common motifs due to its presence in the other OBPs examined. The full distribution of the 10 most common motifs across all of the OBPs examined is given in Supplemental Fig. S1. *SiOBP8* contained 8 out of the top 10 most common motifs, whereas *SiOBPs* 2, 3, and 17 contained only 4 out the top 10 motifs identified. Motifs 1 and 2 were in all *SiOBPs*, although divergence of these motifs was seen in *SiOBP2* (for both motifs 1 and 2; the height of color blocks in



**Figure 6. Developmental expression of SiOBPs.** (A) qRT-PCR was performed on RNA isolated from fire ant eggs, early and late instar and pupae. Absolute quantification values are shown. Color scale shows transcripts/ $\mu$ g total RNA used for cDNA synthesis. (B) Data in (A) normalized to EF1 $\alpha$  expression levels.

figure is proportional to the e-value confidence, i.e. the level in which the sequence highlighted is identical/similar to the consensus sequence, thus a sequence in which the motif is represented by a smaller block indicates that it has diverged the consensus), and SiOBP8 (for motif 1 only). Motifs 3 and 5 were found in all SiOBPs with the exception of motif 3 not found in SiOBP17 and motif 5 not found in SiOBP3. Motif divergence was seen in SiOBP2 (motifs 3 and 5), SiOBPs 3, 6, 8, 13, 15, and 16 (motif 3), and in addition, the motif 3 found in SiOBP3 was notably in a different position (towards the C-terminus) as compared to its position in the other SiOBPs. Motif 4 was found only in SiOBPs 1, 5, 10, and 11, whereas motif 6 was found in most SiOBPs, being absent in SiOBPs 2, 5, and 11. As this latter motif (6) was localized to the N-terminus (except for SiOBP8 which as shifted slightly towards the C-terminus), it likely reflects signal peptide similarities. Motif 7 was identified in SiOBPs 1, 5, and 8–11, whereas motif 8 was found only in SiOBPs 1, 10, and 14. Motif 9 was found in SiOBPs 4, 7, and 9, although the sequence had somewhat diverged in SiOBP4.



**Figure 7. Motif analyses of OBPs.** (A) Ten most common motifs identified in the 82 OBPs using MEME (version 4.11.2, <http://meme-suite.org/tools/meme>). The size (height) of the amino acid represents the degree of conservation of that amino acid in the consensus sequence. (B) Positions of indicated motifs in the amino acid sequences of the *SiOBPs*. Colored box height is proportional to match to consensus sequence. Motifs 1, 3, 6, and 9 are further annotated in the figure with corresponding numbers and arrows to better distinguish their identities and positions in each sequence. The complete analyses of the 82 OBPs is given in Supplemental Figure S1.

## Discussion

The generalized model of insect chemosensation involves a suite of proteins that function to relay identification of chemical ligands to neurophysiological and behavioral responses. Within the context of the initial detection of chemical molecules, the major proteins involved are thought to include sensory neuron membrane proteins (SNMPs), odorant and ionotropic receptors (ORs and IRs, respectively), gustatory receptors (GRs), the soluble odorant binding and chemosensory proteins (OBPs and CSPs), and odorant-degrading enzymes (ODEs), all housed within specialized sensilla that are found, although not exclusively, on the antennae<sup>2</sup>. The water soluble CSPs and OBPs act as ligand binding proteins to capture often hydrophobic compounds, e.g. environmental volatiles and pheromones, transporting them to the neuronal membrane receptors. In all insects examined thus far, CSPs and OBPs, exist as highly divergent gene families, and *S. invicta* appears to have one of the smaller sets of OBPs (i.e. 17–18) as compared to fruit flies, (*Drosophila melanogaster*; 61), mosquitoes (*Aedes* sp.; 64–72 and *Culex quiquefasciatus*; 53), although the honey bee (*A. mellifera*; 21), the tsetse fly (*Glossina morsitans*; 20), and the pea aphid (*Acyrtosiphon pisum*; 15) have similar numbers of OBPs in their genomes<sup>22</sup>. The best studied of these are the Lepidopteran proteins where significant information concerning distribution, expression, pheromone and ligand binding, and structural determinations have been made<sup>22–28</sup>. Genomics analyses in Lepidoptera have indicated intriguing possibilities of OBP gene loss and gain linked to behavioral shifts. For example, butterflies appear to have lost a specific OBP involved in moth pheromone detection, which the authors suggest may correlate to shifts in stimulus perception/modes of communication from olfactory in moths to visual in butterflies<sup>29</sup>. Expression (transcriptomic) studies have been the main source of OBP discovery including in the tobacco cutworm, *Spodoptera litura*, the tsetsefly, *Glossina morsitans*, winged and unwinged morphotypes of the pea aphid, and several social hymenoptera, amongst others<sup>30–33</sup>.

Although their names, i.e. OBPs and CSPs, suggest functions in olfaction, it has long been recognized that these proteins do not function exclusively in the antennae of insects and are likely involved in many processes aside from olfaction<sup>4,9</sup>. Indeed, current ideas suggest that only a subset of OBPs are directly involved in olfaction, notable amongst these being the pheromone binding proteins (PBPs, considered a sub-group within the OBPs). However, the high degree of divergence seen in OBPs renders extrapolation of results from one insect species to another difficult. A comparative analysis of the *S. invicta* OBPs with that of the honeybee, *A. mellifera* (and as partially recapitulated here), revealed several branch point separations of the OBPs into distinct groups<sup>16</sup>. These analyses indicate at least two distinct OBP groupings; SiOBPs 1, 5, 6, and 9–11 are found in one group and have conserved shared orthologs with various ant OBPs. SiOBPs 2–4, 7, 8, and 12–17 are found within a second grouping, with SiOBPs 2, 3, and 4 having conserved *A. mellifera* orthologs, and SiOBPs 12–17 appearing to represent an *S. invicta*-gene expansion, with one shared homolog in the ant, *A. echinator*. Interspersed within the second overall grouping there exists a divergent *Apis* spp.-specific OBP gene expansion (AmOBPs 13–20), with SiOBPs 7 and 8, the most closely related to these. With respect to the bee OBPs, AmOBPs 13–21 are found as a tandem genomic array with AmOBPs 14–21 designated as C-minus OBPs due to the absence of two of the six usually conserved cysteine residues<sup>34</sup>. With the exception of SiOBP14, the other members of the fire ant expansion were found on the same linkage group contig (lg16), along with SiOBPs 3, 4, 5, and 9.

Protein amino acid sequence analyses of 82 OBPs related to the *S. invicta* repertoire allowed for the identification of a set of common motifs. The highly conserved cysteine residues, characteristic of OBPs were found in motifs 1, 2, 3, and 5, and these corresponded to the most widely conserved motifs found in most of the SiOBPs. However, in a number of cases a specific motif appears to have been lost in a particular lineage, e.g. motif 3 in SiOBP17, divergence of motifs 2 and 5 in SiOBP2. Motif 6, while not part of the conserved cysteine containing motifs, was also widely distributed, likely because it corresponds to a conserved signal peptide processing sequence. However, this motif was absent in SiOBPs 2, 5, 10, and 11, and while these latter proteins also contain signal sequences, these appear to have diverged significantly enough that they no longer resemble the others. As expected, the distribution of many motifs mirrored the phylogenetic tree, although a few notable exceptions were seen. SiOBPs 3 and 4, although very similar in their exon-intron structure and phylogenetic placement, showed differing numbers and organization of motifs. SiOBP14 while closely sharing related motifs found in SiOBP2 12, 13, 15–17, also contained one motif (8) found only in SiOBPs 1 and 10 (amongst the SiOBPs). Similarly, SiOBPs 7 and 9, shared a common motif (9) not present in the other SiOBPs. Some of these motifs may contribute to either substrate binding and/or specificity or to interactions with downstream partners, e.g. receptors, whether odorant or otherwise.

Our data indicate, that in fire ant workers, while a distinct subset of SiOBPs (SiOBPs 1, 2, 5, 6, 11 > SiOBPs 7, 10, and 15) are expressed in the antennae, higher levels are generally seen in worker heads, somewhat less in the thorax, and notably lower in the abdomen. SiOBP expression in the antennae of female fire ant alates was similar to workers, whereas expression in male alate antennae was noticeably lower. The low absolute expression levels of the OBPs seen in the antennae are in general agreement with observations of the low incidence of OBPs in ant antennae and the suggestion that ant olfaction relies more on the functions of CSPs<sup>35–37</sup> or other proteins, an example of the latter being an apolipoprotein-like protein identified as being expressed in *S. invicta* antennae<sup>38</sup>. This conclusion also agrees with expression analysis of the OBP repertoire in *A. mellifera* that indicate that OBPs are broadly expressed, and indeed, may only play a minor role in olfaction in some insects<sup>34</sup>. However, some caution should be taken in this interpretation in that particularly in workers, normalization to EF1 $\alpha$  levels indicates robust expression of a subset of SiOBPs in the antennae. This analysis is consistent with a recent transcriptomics study of several ant species including *Cerapachys biroi*, *Camponatus floridanus*, and *Harpegnathos saltator*, which indicated that subsets of both CSPs and OBPs are expressed in the antennae, suggesting that CSPs have not replaced OBPs in ant olfaction<sup>8</sup>. After normalization, SiOBPs 1, 2, 5 and 6 showed the highest relative expression levels in antennae across all castes, with no strongly caste-specific SiOBP expression seen. Low expression of SiOBPs 4, 8, and 9 was seen in all antennae regardless of caste. Mass spectrometry analyses have indicated that SiOBPs 2 and 15, both of which our data indicate are moderately expressed in these tissues, can be found in the



antennae of *S. invicta* workers<sup>16,37</sup>, leaving open the possibility of some *SiOBPs* playing a role in antennal chemical compound ligand binding.

A significant literature exists regarding *SiOBP3* (*Gp-9*), originally described as a “social organization gene” due to the observation of distinct alleles that correlated with whether colonies were monogyne (single queen) or polygyne (multiple queens) in nature<sup>13,14,39</sup>. However, the link between *SiOBP3* and sociality has been met with some controversy<sup>17,18</sup>. Recent evidence suggests that this gene is part of a much larger chromosomal segment of ~13 Mb, comprising of more than half the chromosome and containing >600 ORFs, that has undergone distinct rearrangement events in the two alternate *S. invicta* social forms (monogyne versus polygyne) and whose recombination is suppressed<sup>19</sup>. As noted above, this region (corresponding to linkage group 16) contains not only *SiOBP3*, but also *SiOBPs* 4, 5, 9, 12, 13, and 15–17, thus significantly expanding the range of chemical compounds and the OBPs that may play a role in mediating this specific aspect of social organization. Our data show robust expression of *SiOBP3* in workers, and male and female alates, with very high expression in worker heads, and throughout female alates. Interestingly, relative expression was lowest in antennae regardless of caste and low expression of *SiOBP3* was seen throughout the developmental stages examined, i.e. eggs, larvae, and pupae.

Each caste showed a distinct overall *SiOBP* expression pattern, with the differential expression of a number of specific proteins. *SiOBPs* were generally preferentially expressed in male and female alates throughout the body, i.e. head, thorax, and abdomen, as compared to the antennae. In contrast, *SiOBP* expression was generally lower in the abdomen (and antennae) of workers, and more robustly expressed in the head and thorax. These data suggest generalized roles for most *SiOBPs* and the likelihood that a number of these proteins are hemolymph proteins that function as ligand carriers of endogenous ligands, e.g. hormones and signaling molecules. A number of *SiOBPs* were expressed in a sex-specific manner. Expression of *SiOBP4* in the thorax of female alates was very high, whereas neither workers nor male alates expressed this gene at such high levels. *SiOBP4* was also notably absent in antennal fractions. Conversely, *SiOBP9* was highly expressed in the male alate abdomen, but noticeably absent in both female alates and workers. *SiOBPs* 13, 15, and 16 were noticeably more highly expressed in worker heads than female and male alate counterparts, and *SiOBP14* was highly expressed in the thoraces of workers but expressed at moderate to low levels in female and male alates, respectively.

In terms of absolute expression levels, a number of *SiOBPs* were expressed at moderate levels during the various developmental stages examined, indicating that these proteins may have some functioning during development. It should be noted that no attempt was made to discriminate between whether the larvae corresponded to workers or male/female alates visual inspection of the colony indicated less than 2% winged alates, and it is likely that the majority of the larvae examined corresponded to workers. With respect to the developmental stages examined, and in contrast to adult tissues, normalization to EF1 $\alpha$  levels had a significant effect on the relative expression levels of the *SiOBPs*, primarily due to the fact that EF1 $\alpha$  levels were very high in these samples. These analyses revealed low relative expression of most *SiOBPs* across the developmental stages examined. Notable exceptions included *SiOBP1* found during all the developmental stages, *SiOBPs* 8, 10, and 15, found in early and late instars, and *SiOBPs* 10 (highest expressed *SiOBP* detected) and 15 in pupae. These data indicate the potential for some OBPs to function as important developmentally linked protein. Such a finding has been seen for a number of CSPs, e.g. AmCSP5 and SiCSP9, implicated in mediating embryonic development in *A. mellifera* and larval molting in *S. invicta*, respectively<sup>40,41</sup>. It is tempting to speculate that some OBPs act as carriers of developmental signaling molecules and further studies probing such functions are warranted. The distribution of OBPs beyond antennal and during developmental, i.e. larval and pupal stages, has also been reported for the parasitoid wasp, *Sclerodermus* sp., although it is not clear whether the full set of these proteins in this insect has been identified<sup>5</sup>. These data do support a model in which OBPs are likely general binding proteins, with only a subset involved in antennal binding of ligands.

A subset of *D. melanogaster* OBPs have been functionally examined using RNAi mediated knockdown<sup>42</sup>. *SiOBPs* 2–4 and 12–17 appear to be orthologs of the *D. melanogaster* DmOBP59a, an OBP whose expression while not readily detectable in fruit fly antennal extracts, appears to mediate altered behavior responses to a broad range of odorants. *SiOBPs* 1 and 10 were similar to DmOBP83a, an OBP expressed in fly antennal extracts and appears to mediate behavioral responses to a narrow range of chemical odorants including citral, acetophenone, *I*-carvone, and 2-heptanone<sup>42</sup>. Notably, whereas *SiOBP1* is expressed in *S. invicta* antennae, particularly in workers, *SiOBP10* is expressed in the head (all castes) and during development (larvae and pupae). *SiOBPs* 7 and 8, correspond to the DmOBP99b/83c/56a family, two of which (56a and 99b) are robustly expressed in *D. melanogaster* antennal extracts. DmOBPs 56a and 83c appear to mediate responses to a small number of odorants in a sex-specific manner, whereas DmOBP99b RNAi knockdowns displayed a broader altered response pattern<sup>42</sup>. Neither *SiOBP7* nor 8 were strongly expressed in ant antennae; however, *SiOBP7* was very highly expressed in worker heads and throughout the bodies of male and female alates. *SiOBP8*, however, was poorly expressed in most tissues examined. In *Drosophila*, RNAi knockdown also resulted in altered antennal electrophysiological responses for several of the OBPs examined, an effect also seen for an OBP (OBP2) in the cotton aphid, *A. gossypii*<sup>43</sup>.

We have conducted a systematic expression profiling of the *SiOBP* gene family in a eusocial insect. Patterns of *SiOBP* gene expression were found to show some relationships to their evolutionary relatedness, i.e. phylogeny, although no absolute overall correlation was observed. Preferential expression of phylogenetically related *SiOBPs* subgroups were seen that was caste dependent, whereas others showed tissue and/or developmental preferences, although normalized expression during development was very low for most *SiOBPs*. The small size of the fire ant OBP-gene family and their relatively low expression in antennae suggest that these proteins may only be nominally involved in direct olfaction and/or olfactory coding. Instead, our data support a model in which these proteins act as ligand carriers throughout the body, some likely functioning to carry specific molecules mediating specific processes, whereas others may be more general in nature. This does not preclude the participation of individual OBPs in aspects of olfaction, e.g. by binding and/or sequestering endogenously produced pheromones or other semiochemicals and/or ultimately acting during their release. However, our data do indicate that the term

“odorant-binding protein” is likely to be an unfortunate designation for many of these proteins and that a more generic term such as “ligand-carrier proteins” would be more apt. Our data provide a basis for future work aimed towards better understanding the diverse functions of OBPs. Determining the chemical compounds bound by the OBPs and the protein targets and/or interacting partners of these proteins, e.g. membrane receptors, is needed to further characterize the physiological role(s) of the *SiOBPs*.

## Methods

**Insects and experimental samples.** *S. invicta* colonies were collected from the field and maintained in Fluon-coated and/or talcum powder dusted trays as described<sup>44</sup>. Three independent colonies were sampled. Colonies were assessed to be polygyne due to the presence of multiple-queens in the founding colony and via sequencing of the full-length cDNA of *SiOBP3/Gp-9* in which both *Gp-9B* and *Gp-9b* alleles were detected. Colonies were fed with 300 mM sucrose and freeze dried *Galleria mellonella* larvae and kept at room temperature with ~70% humidity and 16:8 dark:light photoperiod. For each colony, around 200–500 workers, males and female alates were collected separately, immersed in RNALater (Invitrogen, Thermo Fisher Scientific, Waltham, MA) and dissected under a stereomicroscope into four sections; antenna, head (without antennae), thorax and abdomen. Approximately 200–500 larvae and pupae collected for the various development stages. Eggs were collected from mated queens collected from colonies within 24 hours and samples were immediately suspended in RNALater and stored at –80 °C until RNA extraction. Similarly, small larvae, large larvae and pupa were collected and immersed in RNALater before RNA extractions. Adult stages were not sampled as same-age cohorts and larvae were distinguished by their size only and likely reflect a mixture of minor and major workers, as well as potential male and female alates. Early instars (1<sup>st</sup> and 2<sup>nd</sup>) were selected based on size and color with early instars (1<sup>st</sup> and 2<sup>nd</sup>) displaying some melanization and late (3<sup>rd</sup> and 4<sup>th</sup>) clear to whitish. Generally the early instars were smaller than 1 mm ( $\leq 1$  mm), and late instars (3<sup>rd</sup> and 4<sup>th</sup>) were between 1.1 and 4.5 mm.

**RNA preparation and cDNA library construction.** Samples (~100 mg) were ground in liquid nitrogen and then mixed with Trizol reagent (1 ml) and total RNA was extracted following the manufacturer’s instructions (Invitrogen, Life Sciences). Genomic DNA in samples was digested using TURBODNase (Invitrogen) according to the manufacturer’s instruction. Total RNA quality and quantity were analyzed by agarose gel electrophoresis and via NanoDrop 2000 spectrophotometric analyses. Quantification of RNA concentrations in samples was performed using a Qubit H 2.0 fluorometer (Invitrogen, Carlsbad, CA). Finally, 2  $\mu$ g total RNA from different tissues were used to construct cDNA libraries using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA) in terms of standard protocol.

**Absolute qualification of cDNAs.** The sequences of the *S. invicta* odorant binding proteins (*SiOBPs*) were downloaded from NCBI (Supplemental Table S1). Absolute qualification of all OBPs was performed as previous described<sup>45</sup>. Briefly, specific qRT-PCR primers were designed using Beacon designer 8.13 (Palo Alto, CA, USA) and synthesized by Invitrogen Ltd (Carlsbad, CA). The list of primers used is given in Supplemental Table S2. Primers were used to amplify the respective regions of the *SiOBPs* using an *S. invicta* cDNA library as described above as the template. After purification, the PCR generated fragments were ligated into the pGEMT vector (Promega Corp., Madison, WI). Positive clones were identified by colony PCR and plasmids were isolated and the integrity of the inserts verified by sequencing (Eton Biosciences, San Diego, CA). Plasmid concentrations were quantified using Qubit dsDNA BR Assay Kit (Invitrogen, Carlsbad, CA). RT-PCR primer sets were validated for amplicon size, optimal ratios,  $T_m$ , and efficiency (Supplemental Table S3). The amplification efficiency (E) was calculated using the slope of a linear regression determined by the Ct values (axe Y) and the concentration of cDNA in logarithmic scale (axe X). Given the slope value, the Efficiency (E) value was calculated using the formula:

$$E = 10^{1/\text{slope}} - 1 \quad (1)$$

Plasmid constructs and optimized the  $T_m$  values were used to construct expression standard curves, using serial dilution of the plasmid templates ( $10^{-5}$  to  $10^{-9}$ ). The number of transcript copies was calculated using the molecular weight of plasmids and their empirically determined concentrations using the following formula:

$$\text{Copy number} = (6.02 \times 10^{23}) \times (\text{amount ng} \times 10^{-9}) / (\text{DNA length} \times 660) \quad (2)$$

The standard curves and empirically determined  $C_t$  values as derived from the q-RT-PCR experiments were used to calculate the absolute transcript expression values for the OBPs in the various samples. For qRT-PCR, the 2x SYBR Green qPCR Master Mix (Biotools, Houston TX) was used for the reactions. *S. invicta* cDNA libraries from different tissues were diluted 40 times and 5  $\mu$ L of diluted cDNA were used as template in a 15  $\mu$ L reaction volume. Each reaction included 1  $\times$  Master Mix, 5  $\mu$ L template and 200 nM of each of gene specific primer pairs. Real time q-PCR reactions were performed for at least two RNA preparations from each tissue sample. Controls and internal references included primers designed to the EF1 $\alpha$  and GAPDH genes of *S. invicta* that were used for initial normalization of the template amount and whose absolute copy numbers were also quantified. The PCR reactions were performed using a Eco Real-Time qPCR System (Illumina, San Diego, CA) with a thermo-profile of one cycle of 95 °C 5 min, 95 °C 2 min, then 45 cycles of 95 °C 15 s, and 60/59 °C 45 s, followed by a melting curve analysis from 55 to 95 °C.

**Data analyses.** An analysis of variance (ANOVA) was conducted to evaluate potential significant differences in expression levels among tissues and castes/developmental stages. The data, number of copies per  $\mu$ g of RNA, were transformed logarithmically in order to correct normality and unequal variances. Means were compared with LSD (Least significant differences) test and the Duncan’s new multiple range test (MRT). Statistical analyses were performed using the IBM SPSS Statistics, version 21 (Armonk, NY).



**Phylogenetic and motif analyses.** A total of 82 OBPs from *S. invicta*, *A. echinator*, *C. floridanus*, and *H. saltator*, *A. mellifera* and *A. cerana* were used for phylogenetic tree construction and motif analyses (Supplemental Table S4). Putative N-terminal signal peptides were identified using Signal P (<http://www.cbs.dtu.dk/services/SignalP/>). Exon-intron splice positions were identified using the online Gene Structure Display Server (GSDS: <http://gsds.cbi.pku.edu.cn>). Amino acid motifs were identified using MEME (version 4.11.2, <http://meme-suite.org/tools/meme>). Amino acid multiple sequence alignments (MSAs) were generated with five different alignment algorithms: MUSCLE, PRANK, MAFFT, CLUSTALW and PROBCONS<sup>46–50</sup>. For each MSA the best fitting model of amino acid substitution was estimated with MEGA 6.0<sup>51</sup>. MEGA optimizes the tree topology search starting with a Neighbor Joining tree and uses the likelihood function and three model criteria BIC (Bayesian Information Criterion), AIC (Akaike information criterion) and LnL (log likelihood) to find the best fitting model of amino acid substitution. MEGA consistently chose the model LG + G + I with 5 rates categories (G = Gamma shape parameters and I = proportion of invariant sites). The different amino acid MSAs were analysed with PhyML 3.0 available at the phylogeny.fr platform<sup>52,53</sup> to see the alignment method yielding the most likely tree (i.e. ability to best capture phylogenetic signal) given the LG amino acid substitution model and estimating G and I<sup>54</sup>. The PRANK algorithm gave the best tree (lowest LnL, parsimony score and tree size), which was re-built using RaxML at the CIPRES Science Gateway<sup>55,56</sup>, and implementing the LG amino acid substitution model with 5 rates of categories. G and I were estimated, branch lengths optimized and branch support calculated by bootstrapping. RaxML was allowed to halt bootstrapping automatically. The software MEGA 6.0 was used to draw the tree.

## References

- Keil, T. A. Sensory cilia in arthropods. *Arthropod Structure & Development* **41**, 515–534, doi: 10.1016/j.asd.2012.07.001 (2012).
- Leal, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol.* **58**, 373–391, doi: 10.1146/annurev-ento-120811-153635 (2013).
- Carraher, C. *et al.* Towards an understanding of the structural basis for insect olfaction by odorant receptors. *Insect Biochem. Mol. Biol.* **66**, 31–41, doi: 10.1016/j.ibmb.2015.09.010 (2015).
- Pelosi, P., Zhou, J. J., Ban, L. P. & Calvellido, M. Soluble proteins in insect chemical communication. *Cell. Mol. Life Sci.* **63**, 1658–1676, doi: 10.1007/s00018-005-5607-0 (2006).
- Zhou, C. X., Min, S. F., Yan-LongTang & Wang, M. Q. Analysis of antennal transcriptome and odorant binding protein expression profiles of the recently identified parasitoid wasp, *Sclerodermus* sp. *Comp Biochem Phys D* **16**, 10–19, doi: 10.1016/j.cbd.2015.06.003 (2015).
- Dippel, S. *et al.* Tissue-specific transcriptomics, chromosomal localization, and phylogeny of chemosensory and odorant binding proteins from the red flour beetle *Tribolium castaneum* reveal subgroup specificities for olfaction or more general functions. *BMC Genomics* **15**, doi: Artn 1141 10.1186/1471-2164-15-1141 (2014).
- He, M. & He, P. Molecular characterization, expression profiling, and binding properties of odorant binding protein genes in the whitebacked planthopper, *Sogatella furcifera*. *Comp Biochem Phys B* **174**, 1–8, doi: 10.1016/j.cbpb.2014.04.008 (2014).
- McKenzie, S. K., Oxley, P. R. & Kronauer, D. J. C. Comparative genomics and transcriptomics in ants provide new insights into the evolution and function of odorant binding and chemosensory proteins. *BMC Genomics* **15**, doi: Artn 718 10.1186/1471-2164-15-718 (2014).
- Pelosi, P., Iovinella, I., Felicioli, A. & Dani, F. R. Soluble proteins of chemical communication: an overview across arthropods. *Frontiers in physiology* **5**, 320, doi: 10.3389/fphys.2014.00320 (2014).
- Holldobler, B. & Wilson, E. O. *The Superorganism: the beauty, elegance, and strangeness of insect societies.* (W.W. Norton & Company, 2009).
- Tschinkel, W. R. *The Fire Ants.* (Belknap Press of Harvard University Press, 2006).
- Ross, K. G. & Fletcher, D. J. C. Comparative study of genetic and social structure in two forms of the Fire Ant *Solenopsis invicta* (Hymenoptera, Formicidae). *Behav. Ecol. Sociobiol.* **17**, 349–356, doi: Doi 10.1007/Bf00293212 (1985).
- Gotzek, D., Shoemaker, D. D. & Ross, K. G. Molecular variation at a candidate gene implicated in the regulation of fire ant social behavior. *Plos One* **2**, doi: ARTN e1088 doi: 10.1371/journal.pone.0001088 (2007).
- Keller, L. & Ross, K. G. Selfish genes: a green beard in the red fire ant. *Nature* **394**, 573–575, doi: Doi 10.1038/29064 (1998).
- Liu, N. N. & Zhang, L. *CYP4AB1*, *CYP4AB2*, and *Gp-9* gene overexpression associated with workers of the red imported fire ant, *Solenopsis invicta* Buren. *Gene* **327**, 81–87, doi: 10.1016/j.gene.2003.11.002 (2004).
- Gotzek, D., Robertson, H. M., Wurm, Y. & Shoemaker, D. Odorant binding proteins of the red imported fire ant, *Solenopsis invicta*: an example of the problems facing the analysis of widely divergent proteins. *Plos One* **6**, e16289, doi: 10.1371/journal.pone.0016289 (2011).
- Gotzek, D. & Ross, K. G. Current status of a model system: the gene Gp-9 and its association with social organization in fire ants. *Plos One* **4**, doi: ARTN e7713 doi: 10.1371/journal.pone.0007713 (2009).
- Leal, W. S. & Ishida, Y. GP-9s are ubiquitous proteins unlikely involved in olfactory mediation of social organization in the red imported fire ant, *Solenopsis invicta*. *Plos One* **3**, doi: ARTN e3762 doi: 10.1371/journal.pone.0003762 (2008).
- Wang, J. *et al.* A Y-like social chromosome causes alternative colony organization in fire ants. *Nature* **493**, 664–668, doi: Doi 10.1038/Nature11832 (2013).
- Wurm, Y. *et al.* The genome of the fire ant *Solenopsis invicta*. *Proc Natl Acad Sci USA* **108**, 5679–5684, doi: 10.1073/pnas.1009690108 (2011).
- Bailey, T. L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* **2**, 28–36 (1994).
- Zhou, J. J. Odorant-binding proteins in insects. *Vitam Horm* **83**, 241–272, doi: 10.1016/S0083-6729(10)83010-9 (2010).
- Leite, N. R. *et al.* Structure of an odorant-binding protein from the mosquito *Aedes aegypti* suggests a binding pocket covered by a pH-sensitive “Lid”. *Plos One* **4**, e8006, doi: 10.1371/journal.pone.0008006 (2009).
- White, S. A., Briand, L., Scott, D. J. & Borysik, A. J. Structure of rat odorant-binding protein OBP1 at 1.6 angstrom resolution. *Acta Crystallogr D* **65**, 403–410, doi: 10.1107/S090744490900420x (2009).
- Biessmann, H. *et al.* The Anopheles gambiae odorant binding protein 1 (AgamOBP1) mediates indole recognition in the antennae of female mosquitoes. *Plos One* **5**, e9471, doi: 10.1371/journal.pone.0009471 (2010).
- Lagarde, A. *et al.* Crystal structure of a novel type of odorant-binding protein from *Anopheles gambiae*, belonging to the C-plus class. *Biochem. J.* **437**, 423–430, doi: 10.1042/BJ20110522 (2011).
- Zhang, T. T. *et al.* Structure, Binding Characteristics, and 3D Model Prediction of a Newly Identified Odorant-Binding Protein from the Cotton Bollworm, *Helicoverpa armigera* (Hubner). *J Integr Agr* **11**, 430–438 (2012).
- Tuccori, E. & Persaud, K. C. Differential binding of the Pheromone binding protein 1 (PBP1) and General Odorant binding protein 2 (GOBP2) of *Bombyx mori* toward the main components of the pheromone blend in air. *Chem. Senses* **40**, 255–255 (2015).

29. Vogt, R. G., Grosse-Wilde, E. & Zhou, J. J. The Lepidoptera odorant binding protein gene family: Gene gain and loss within the GOBP/PBP complex of moths and butterflies. *Insect Biochem. Mol. Biol.* **62**, 142–153, doi: 10.1016/j.ibmb.2015.03.003 (2015).
30. Calvello, M. *et al.* Expression of odorant-binding proteins and chemosensory proteins in some Hymenoptera. *Insect Biochem. Mol. Biol.* **35**, 297–307, doi: 10.1016/j.ibmb.2005.01.002 (2005).
31. Gu, S. H. *et al.* Identification and comparative expression analysis of odorant binding protein genes in the tobacco cutworm *Spodoptera litura*. *Sci Rep-Uk* **5**, doi: ARTN 13800 10.1038/srep13800 (2015).
32. Liu, R. H. *et al.* Characterisations of odorant-binding proteins in the tsetse fly *Glossina morsitans morsitans*. *Cell. Mol. Life Sci.* **67**, 919–929, doi: Doi 10.1007/S00018-009-0221-1 (2010).
33. Gu, S. H. *et al.* Identification and expression profiling of odorant binding proteins and chemosensory proteins between two wingless morphs and a winged morph of the cotton aphid *Aphis gossypii* Glover. *PLoS One* **8**, doi: ARTN e73524 10.1371/journal.pone.0073524 (2013).
34. Foret, S. & Maleszka, R. Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*). *Genome Res.* **16**, 1404–1413, doi: 10.1101/gr.5075706 (2006).
35. Ozaki, M. *et al.* Ant nestmate and non-nestmate discrimination by a chemosensory sensillum. *Science* **309**, 311–314, doi: 10.1126/science.1105244 (2005).
36. Ishida, Y., Chiang, V. & Leal, W. S. Protein that makes sense in the Argentine ant. *Naturwissenschaften* **89**, 505–507, doi: 10.1007/s00114-002-0368-1 (2002).
37. Gonzalez, D. *et al.* The major antennal chemosensory protein of red imported fire ant workers. *Insect Mol. Biol.* **18**, 395–404, doi: 10.1111/j.1365-2583.2009.00883.x (2009).
38. Guntur, K. V. P. *et al.* Apolipoprotein III-like protein expressed in the antenna of the red imported fire ant, *Solenopsis invicta* Buren (Hymenoptera: Formicidae). *Arch. Insect Biochem. Physiol.* **57**, 101–110, doi: Doi 10.1002/Arch.20019 (2004).
39. Krieger, M. J. B. & Ross, K. G. Identification of a major gene regulating complex social behavior. *Science* **295**, 328–332, doi: DOI 10.1126/science.1065247 (2002).
40. Maleszka, J., Foret, S., Saint, R. & Maleszka, R. RNAi-induced phenotypes suggest a novel role for a chemosensory protein CSP5 in the development of embryonic integument in the honeybee (*Apis mellifera*). *Dev. Genes Evol.* **217**, 189–196, doi: 10.1007/s00427-006-0127-y (2007).
41. Cheng, D. F., Lu, Y. Y., Zeng, L., Liang, G. W. & He, X. F. Si-CSP9 regulates the integument and moulting process of larvae in the red imported fire ant, *Solenopsis invicta*. *Sci Rep-Uk* **5**, doi: ARTN 9245 10.1038/srep09245 (2015).
42. Swarup, S., Williams, T. I. & Anholt, R. R. Functional dissection of odorant binding protein genes in *Drosophila melanogaster*. *Genes Brain Behav* **10**, 648–657, doi: 10.1111/j.1601-183X.2011.00704.x (2011).
43. Rebijith, K. B. *et al.* RNA Interference of odorant binding protein 2 (OBP2) of the cotton aphid, *Aphis gossypii* (Glover), resulted in altered electrophysiological responses. *Appl. Biochem. Biotechnol.* **178**, 251–266, doi: 10.1007/s12010-015-1869-7 (2016).
44. Fan, Y., Pereira, R. M., Kilic, E., Casella, G. & Keyhani, N. O. Pyrokinin beta-neuropeptide affects necrophoretic behavior in fire ants (*S. invicta*), and expression of beta-NP in a mycoinsecticide increases its virulence. *PLoS One* **7**, e26924, doi: 10.1371/journal.pone.0026924 (2012).
45. Whelan, J. A., Russell, N. B. & Whelan, M. A. A method for the absolute quantification of cDNA using real-time PCR. *J. Immunol. Methods* **278**, 261–269, doi: 10.1016/S0022-1759(03)00223-0 (2003).
46. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792–1797 (2004).
47. Katoh, K. & Toh, H. Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics* **26**, 1899–1900, doi: 10.1093/bioinformatics/btq224 (2010).
48. Do, C. B., Mahabhashyam, M. S. P., Brudno, M. & Batzoglou, S. PROBCONS: Probabilistic consistency-based multiple sequence alignment. *Genome Research* **15**, 330–340 (2005).
49. Löytynoja, A. & Goldman, N. An algorithm for progressive multiple alignment of sequences with insertions. *Proceedings of the National Academy of Sciences* **102**, 10557–10562 (2005).
50. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
51. Tamura, K., Stecher, G., Peterson, D., Filipiński, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**, 2725–2729 (2013).
52. Guindon, S. *et al.* New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systemic Biology* **59**, 307–321 (2010).
53. Dereeper, A. *et al.* Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acid Research* **36**, W465–W469 (2008).
54. Le, S. Q. & Gascuel, O. LG: An Improved, General Amino-Acid Replacement Matrix. *Molecular Biology and Evolution* **25**, 1307–1320 (2008).
55. Stamatakis, A. RAXML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
56. Miller, M. A., Pfeiffer, W., Schwartz, T. & “Creating the CIPRES Science Gateway for inference of large phylogenetic trees” in, N., New Orleans, LA pp 1–8. in *Gateway Computing Environments Workshop (GCE)*. 1–8 (2010).

## Acknowledgements

This work was supported by funds from the University of Florida, Institute of Food and Agricultural Sciences and NSF grant IOS-1557704 to NOK. Photographs of worker ants, male, and female alates were taken by AOU and NOK with kind assistance from Tyler Jones, UF-IFAS Communications Photography Department. Ant anatomy images kindly provided with permission by Robert Svensson (msitua.net).

## Author Contributions

N.O.K., A.W., A.O.-U. and W. Z. initiated the project, conceived and designed the study. W.Z., A.W. and A.O.-U. performed the samples collection, library constructions, q-RT-PCR, data processing, bioinformatic analyses, data processing, and contributed the revisions of the manuscript. N.O.K. and Y.X. wrote the manuscript, contributed to bioinformatic analyses, data processing, and data interpretation.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Zhang, W. *et al.* Tissue, developmental, and caste-specific expression of odorant binding proteins in a eusocial insect, the red imported fire ant, *Solenopsis invicta*. *Sci. Rep.* **6**, 35452; doi: 10.1038/srep35452 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016