Molecular Elements Involved in Locust Olfaction: Gene Families in the Desert Locust *Schistocerca gregaria*

Dissertation for Obtaining the Doctoral Degree

of Natural Sciences (Dr. rer. nat.)

Faculty of Natural Sciences

University of Hohenheim

From the Institute of Physiology

Prof. Dr. H. Breer

Submitted by

Xingcong Jiang

Heze (Provinz Shandong), V.R. China

2018

Dean: Prof. Dr. Uwe Beifuß

1st reviewer: Prof. Dr. Heinz Breer

2nd reviewer: Prof. Dr. Jürgen Krieger

Submitted on:

Oral examination on: 10.08.2018

This work was accepted by the Faculty of Natural Sciences at the University of Hohenheim as “Dissertation for Obtaining the Doctoral Degree of Natural Sciences”.

UNIVERSITY OF HOHENHEIM

Annex 2 to the University of Hohenheim doctoral degree regulations for Dr. rer. nat.

Affidavit according to Sec. 7(7) of the University of Hohenheim doctoral degree regulations for Dr. rer. nat.

1. For the dissertation submitted on the topic

   Molecular Elements Involved in Locust Olfaction: Gene Families in the Desert Locust Schistocerca gregaria

   I hereby declare that I independently completed the work.

2. I only used the sources and aids documented and only made use of permissible assistance by third parties. In particular, I properly documented any contents which I used - either by directly quoting or paraphrasing - from other works.

3. I did not accept any assistance from a commercial doctoral agency or consulting firm.

4. I am aware of the meaning of this affidavit and the criminal penalties of an incorrect or incomplete affidavit.

I hereby confirm the correctness of the above declaration: I hereby affirm in lieu of oath that I have, to the best of my knowledge, declared nothing but the truth and have not omitted any information.

_____ _____

Place and Date Signature

Table of content

1 Introduction

Locusts are large herbivorous insects, which are remarkable due to their unique capability to develop strikingly different phenotypic forms depending on the local population density. While at low population densities the “solitarious” phase locusts avoid one another, at crowded conditions the “gregarious” phase locusts can form dense and highly mobile swarms, which have been feared as agricultural pests since ancient history. For this reason, but also for the purely scientific interest to unravel the mechanisms underlying the phenotypic plasticity and the reproduction behavior, locust biology has long been the subject of intense researches. Thus, it is largely accepted that the processes of phase transition and the concomitant behavioral plasticity heavily rely on chemical communication via volatile signaling compounds (Ferenz and Seidelmann, 2003). Towards a more detailed understanding of the complex processes, it is essential to elucidate the molecular and cellular basis that enables locusts to recognize the volatile signaling molecules, i.e. to identify molecular elements in the antennae of locusts that are potentially involved in the detection of odorous compounds. In this study, the desert locust *Schistocerca gregaria*, as a representative locust species, was investigated with respect to molecular elements involved in locust olfaction.

Insects utilize their hair-like cuticle appendages, so called sensilla, to receive environmental olfactory cues (Steinbrecht, 1996; Hansson and Stensmyr, 2011; Suh et al., 2014). Hydrophobic odorous molecules have to travel through the aqueous sensillum lymph before reaching the receptors residing in the chemosensory membrane of antennal olfactory neurons (Vogt et al., 1999; Leal, 2013; Suh et al., 2014). This procedure is supposed to be facilitated by odorant binding proteins (OBPs) in the sensillum lymph, an important family of soluble proteins that is capable of binding odorous molecules (Vogt and Riddiford, 1981; Pelosi et al., 2006, 2014; Vieira and Rozas, 2011). OBPs are polypeptides of 110 – 200 amino acids that fold into a globular shape forming an interior binding cavity, in which the interaction with odorous molecules takes place (Sandler et al., 2000; Tegoni et al., 2004). The

sequence of “classic OBPs” is characterized by six conserved cysteine (C) residues, a hall mark of “classic OBPs”, while “plus-C” or “minus-C OBPs” are categorized with more or less than six conserved C-residues (Xu et al., 2003; Zhou et al., 2004; Foret and Maleszka, 2006; Vieira and Rozas, 2011). OBPs are produced by auxiliary cells, which envelope the sensory neurons of a sensillum with their extended cytoplasmic processes. For many insect species, it has been reported that OBPs are expressed in the sensillum types which respond to olfactory cues such as sensilla basiconica and sensilla trichodea (Pelosi et al., 2014, 2017). In addition, expression of OBPs has also been reported for the sensillum types that were considered as “gustatory sensilla” such as sensilla chaetica (Galindo and Smith, 2001; Jeong et al., 2013).

For locusts not only the complex reproduction behavior but also the unqiue aggregation behavior is regulated by pheromonal compounds. In the past, searches for relevant pheromonal compounds were impeded by the complexity of volatile emissions from locusts (Hassanali et al., 2005), and efforts to decipher the pheromonal system controlling aggregation or reproduction behavior failed to provide a coherent picture (Pener and Simpson, 2009; Mahamat et al., 2011; Seidelmann et al., 2005). In view of the insufficient knowledge, a search for candidate pheromone receptors seems a promising approach towards to a better understanding of locust pheromone signaling. Extensive studies concerning the mechanisms underlying pheromone detection in moth and fly have uncovered an important role of “sensory neuron membrane proteins” (SNMPs) (Jin et al. 2008, Vogt et al. 2009, Li et al. 2014), which are considered as member of the CD36 family due to the transmembrane topology (Rogers et al. 1997, Nichols and Vogt 2008). Recent studies have shown that the subtype SNMP1 is specifically co-expressed with pheromone receptors in pheromone responsive neurons (Benton et al. 2007, Forstner et al. 2008). Although the functional role of SNMP1 in pheromone processing is not entirely clear (Jin et al. 2008, Vogt et al. 2009, Li et al. 2014), based on recent studies a “tunneling mechanism” has been proposed, in which the ectodomain of SNMP1 channels the pheromone molecules towards the pheromone receptors (Gomez-Diaz et al., 2016). Therefore, SNMP1 is now considered as an

indicator for pheromone responsive neurons. Thus, visualization of SNMP1 expressing cells in the locust antennae may be a first step to identify potentially pheromone responsive neurons of locusts. Furthermore, a search for odorant receptor types, which are co-expressed with SNMP1 may be a promising approach to identify candidate receptors for pheromones.

2 Results and discussion

2.1 Locust odorant binding proteins: molecular evolution, structural variation and antennal topographic expression

A detailed understanding of the molecular mechanisms underlying the chemosensory communication of locusts requires an in-depth knowledge about the molecular identity of functional elements involved in locust olfaction. This information is also of great interest in view of the phylogenetic distance between the hemimetabolous locust species and the well studied homometabolous insect species as well as for an identification of potential targets for more efficient approaches to control this agricultural pest. One of the key elements in odorant reception is the odorant binding proteins (OBPs). Towards a detailed insight into the evolutionary relationship and structural variations of odorant binding proteins (OBPs) for locusts, the sequences of 14 newly identified OBPs from *Schistocerca gregaria* (the desert locust) together with OBP sequences from three additional locust species, *Locusta migratoria* (the migratory locust), *Oedaleus asiaticus* and *Ceracris kiangsu* were subjected to thorough analyses. In a first approach, the sequences of putative OBPs were categorized based on the number of conserved C-residues. More than 30 sequences comprised six conserved C-residues, which is considered as hallmark of “classic OBPs”. 15 sequences were identified harboring more than six conserved C-residues, which are categorized as “plus-C OBPs” and one sequence with less than six conserved C-residue, categorized as “minus-C OBP”. In addition, three sequences compromised extraordinary long stretches between conserved C1 and C2, and they were categorized as “atypical OBPs”.

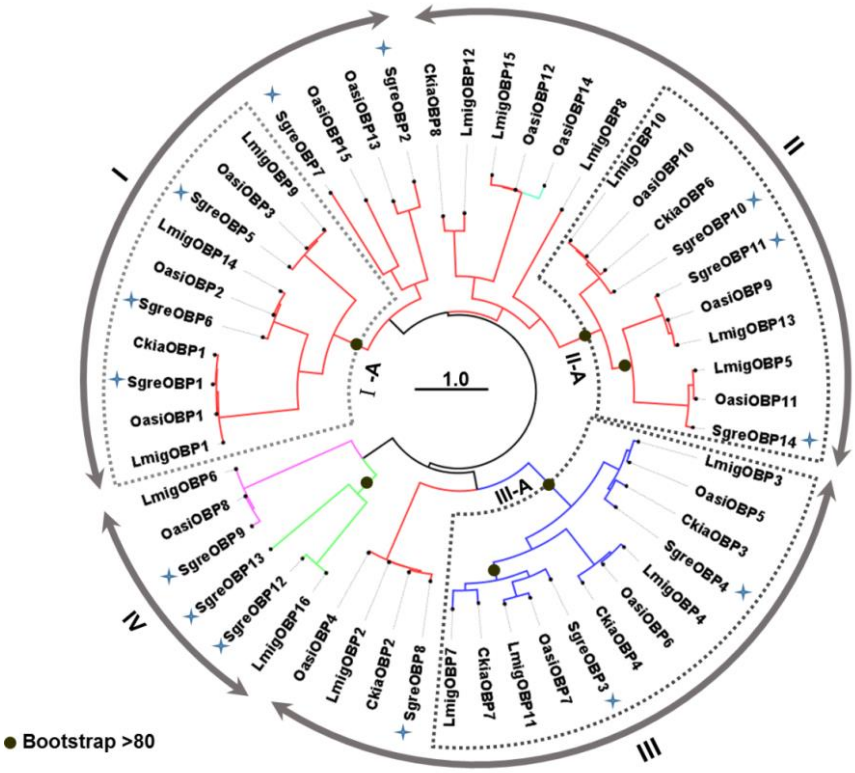


Figure 1. Phylogeny of OBPs from four locust species.
The phylogenetic tree was constructed using the maximum likelihood algorithm supported by 1000 bootstrap replicates. OBP sequences utilized to generate the tree were derived from four locust species: 14 from *Schistocerca gregaria* (SgreOBPs), 16 from *Locusta migratoria* (LmigOBPs), 15 from *Oedaleus asiaticus* (OasiOBPs) and 7 from *Ceracris kiangsu* (CkiaOBPs). Four primary families (I-IV) are denoted by arrow lines. Further classification of three subfamilies (I-A, II-A and III-A) was based on the over 80% bootstrap support at the internal node (indicated by black dots). The colored inner branches represent different OBP categories: red, classic OBPs; blue, plus-C OBPs type-A; magenta, plus-C OBPs type-B; green, atypical OBPs; cyan, minus-C OBP. Newly identified SgreOBPs are denoted by blue crosses. The tree is midpoint rooted. Scale bar represents one amino acid substitution per site.

Approaches to unravel the phylogenetic relationship of locust OBPs using the maximum likelihood algorithm revealed that the repertoire of locust OBPs could be divided into four major families (I–IV, **figure 1**). Furthermore, based on the high bootstrap support three subfamilies (I-A, II-A, and III-A) were classified (**Figure 1**). Interestingly, subfamily I-A as well as subfamily II-A represented “classic OBPs”, and each subfamily comprised three distinct groups with 3–4 orthologous OBPs from different locust species. While the sequence identities of OBPs from different groups

ranged from 28% to 35%, OBP members within each orthologous group shared a sequence identity above 80%. Together, the analyses of OBP sequences from four locust species uncovered a considerable degree of orthology and no evidence for a striking species-specific expansion within the OBP phylogeny.

2.1.1 Characterization of the OBP subfamilies I-A and II-A

Through the phylogenetic analysis of locust OBPs two conserved subfamilies I-A and II-A emerged (**Figure 1**), which appeared of particular interest since both subfamilies comprise “classic OBP” subtypes, but yet were segregated on the phylogenetic tree. This observation may imply some mutually distinct but conserved functional implications for the subtypes. Therefore, as a first step in characterizing the various OBP subtypes, systematic and comparative analyses of OBPs from subfamily I-A and subfamily II-A were performed.

2.1.1.1 Typical sequence motifs

Six consensus amino acid motifs with various widths were identified. The motif pairs 1 and 2 appeared as common motifs fitting the repertoire of subfamilies I-A and II-A, while the other four motifs selectively fit either the repertoire of subfamily I-A OBPs (motif pairs 4 and 6) or the repertoire of subfamily II-A OBPs (motif pairs 3 and 5). Conceivably, two less divergent sequence domains were represented by the motif pairs 1 and 2, spanning the domains of C2–C3 and C4–C6. In contrast, the sequence domains close to the N-terminus (42 amino acids, motif 3 and motif 4) and ahead of C4 (11–15 amino acids, motif 5 and motif 6) appeared more divergent. Besides subfamilies I-A and II-A, the common motif pairs 1 and 2 were sufficient to recapitulate the sequence information present in most of the other locust OBPs, indicating particular phylogenetic conservation of these sequence domains. As expected, the subfamily-specific motifs 3–6 failed to match any other OBP subtypes than that from subfamilies I-A and II-A, with only a fraction of OBPs in clades I and II

as an exception.

The co-existence of both common and subfamily-specific motifs in the OBP coding sequences is reminiscent of the finding that selection regimes may vary with respect to different sequence domains (Policy and Conway, 2001; Sawyer et al., 2005). It is conceivable that the subfamily-specific motifs represent sequence domains that withstand diversifying selection constraints, possibly shaped by the specialized sensillar environment, e.g. interplay with partners as well as with endogenous receptor types. In contrast, the common motifs seem to define the domains that are imposed to stabilizing selection constraints, presumably required for maintaining the common globular structures of the proteins (Pelosi et al., 2017) or for retaining conserved ligand binding sites (Yu et al., 2009).

2.1.1.2 Selection constraints and orthology relationship

The selection pressure imposing on the repertoire of locust OBPs was analyzed by recapitulating three principal concepts, e.g. the non-synonymous substitution rates (dN), the synonymous substitution rates (dS) and the ω rates (dN /dS). Calculation of the substitution rates revealed a significantly attenuated median dN level for both subfamilies I-A and subfamily II-A, in comparison with that of other OBP subtypes. However, the median dS level appeared to be not significantly different among the subfamily I-A, subfamily II-A and the other OBP subtypes. For the ω rates, the values ranging from 0 to 0.7 accounted for nearly 90% of the locust OBP repertoire, suggesting that purifying selection acts on the locust OBP repertoire in general. The ω rates larger than one were found only for a few subtypes, which possibly indicate the strength of positive selection. It was also noted that the median ω rates for subfamily I-A OBPs and subfamily II-A OBPs were significantly reduced in comparison with other OBP subtypes.

The analyzed four locust species inhabit in quite different districts: while *S. gregaria* (the desert locust) widely occurs in Africa, the Middle East and Asia and *L. migratoria* (the migratory locust) in Africa and Asia, as well as in Australia and New Zealand, the

locusts *O. asiaticus* and *C. kiangsu* (the yellow-spined bamboo locust) exist locally in North China and South China. Despite of the distinct geographic distributions, the molecular and evolutionary status remained stable for locust OBP families, especially for subfamilies I-A and II-A, which appears to be attributed to the enhanced purifying selection pressure. This scenario may be indicative for conserved chemosensory roles of subtypes belonging to the respective subfamily. In addition, a chemosensory adaptation to different habitats supposedly acts as an evolutionary driving force for the accumulation of dN substitutions in coding sequences, as a reflection of the strength of positive selection (Cicconardi et al., 2017), and several locust OBPs appear to reflect such a selection regime.

To evaluate the sequence similarity of OBPs from subfamilies I-A and II-A with OBPs from other insect species, sequences from 8 other insect species which gradually emerged in the course of insect evolution were compared. The results indicated that locust subfamily II-A OBPs were clustered in an intact clade without intermingling with any reference OBP sequences. For the subfamily I-A OBPs and their orthologs a different clustering pattern was obtained: the original profile that the three ortholog groups were clustered into a monoclade was disrupted and altered with a more complex re-clustering pattern by integrating OBPs from other species. The orthologous relationship (Theißen, 2002) of OBPs from the two locust subfamilies and OBPs from other species was also inferred. Remarkably, the number of locust subfamily I-A orthologs identified in the reference species expanded considerably, resulting in a many-to-many orthologous relationship; only *A. pisum* appeared to be an exception, likely due to a smaller size of the OBP gene family in this species (Zhou et al., 2010). By contrast, the number of locust subfamily II-A orthologs decreased significantly, consequently leading to the 1-to-many or 0-to-many orthologous relationships. Moreover, locust subfamily II-A OBPs and their orthologs appear to share a common ancestor, determined by the convergence onto a mono-phylogenetic clade. However, for locust subfamily I-A OBPs and their orthologs, the common ancestral status appeared rather ambiguous because of the absence of evident bootstrap support. Together, although the two subfamilies are subjected to

mutually similar strengthened purifying selection pressure, distinct divergent events have taken place during evolution for orthologous OBPs in other species.

2.1.1.3 Structural variations

For representative OBP subtypes from subfamilies I-A and II-A putative tertiary structures were determined by simulation procedures. In a first step, the simulated OBP backbone structures were superimposed to that of LmigOBP1, the hitherto only locust OBP subtype for which a crystal structure has been established (Zheng et al., 2015). In accordance with the phylogenetic deviation, the RMSD scores indicated a marked structural similarity among the subtypes from one subfamily. The sequences analysis of subfamily I-A OBPs revealed a striking variation concerning the C-terminal domain. Unlike LmigOBP1, which possesses a prolonged C-terminus of about 17 amino acids terming a seventh α-helix (Zheng et al., 2015), the C-terminus in SgreOBP6 and OasiOBP3 comprise 7 amino acids, which most likely form a coiled-coil strand rather than a seventh α-helix. Due to the shorter C-terminus, a groove structure emerged on the collapsed surface. Another structural variation is the enlarged interior cavity of LmigOBP1 bordered by the C-terminus, whereas the interior cavities of the other two representative OBP subtypes appear to be much smaller. For subfamily II-A OBPs the sequence analysis uncovered an aligned C-terminus but a non-aligned N-terminus, which resulted in an extension by 9 -10 amino acids for the LmigOBP10 ortholog group. Consequently, this alteration was predicted to form a coiled-coil strand at the N-terminal domain for LmigOBP10; for its two counterparts, the SgreOBP11 and OasiOBP11 at the same surface position an opening structure was visualized.

Although OBP sequences can be highly divergent, the folding of proteins to form a hydrophobic binding cavity is well conserved across insect species. So far, the structures of more than 20 OBPs have been deciphered by X-ray crystallography and/or nuclear magnetic resonance (NMR) spectroscopy (Pelosi et al., 2017). These studies revealed that the C-terminal domain, especially the length of the C-terminus

has important implications for the machinery of ligand-binding (Tegoni et al., 2004). Long C-terminal domains clearly enter the binding cavity and thus determine the shape of cavity (Sandler et al., 2000), whereas a medium-length C-terminus functions as a lid that covers the entrance to the binding cavity (Lartigue et al., 2004). In this regard, the simulated tertiary structures of locust OBPs provided some interesting characters. The three ortholog groups of subfamily I-A differed significantly in their C-terminal domain: LmigOBP1 and its orthologs have a long (17 amino acids) C-terminus forming an additional α-helix and affecting the shape of the cavity; other two ortholog groups have both a medium-length C-terminus (7 amino acids), but are divergent in the amino acid sequences. These structural variations may indicate significant differences in the mechanisms of OBP-ligand interaction among the three ortholog groups from subfamily I-A.

2.1.1.4 Topographic Expression patterns

As a first step towards an understanding of their functional implications, attempts were made to explore where on the antennae the distinct OBP subtypes may be located. Therefore, the antennal topographic expression pattern for the locust subfamilies I-A and II-A OBPs was determined, concentrating on representative subtypes, specifically OBP1, OBP5, OBP6 for subfamily I-A and OBP10, OBP11, OBP14 for subfamily II-A (**Figure 2**). For the three subtypes of subfamily I-A expression was found in cells of both sensilla basiconica and sensilla trichodea, but neither in sensilla coeloconica nor in sensilla chaetica (**Figure 2**). It is interesting to note that in *L. migratoria* the subtype OBP1, the ortholog of OBP1 in *S. gregaria,* is also expressed in sensilla basiconica and sensilla trichodea (Jin et al., 2005). These observations point to a similar sensilla-specific expression pattern of orthologous locust OBPs.

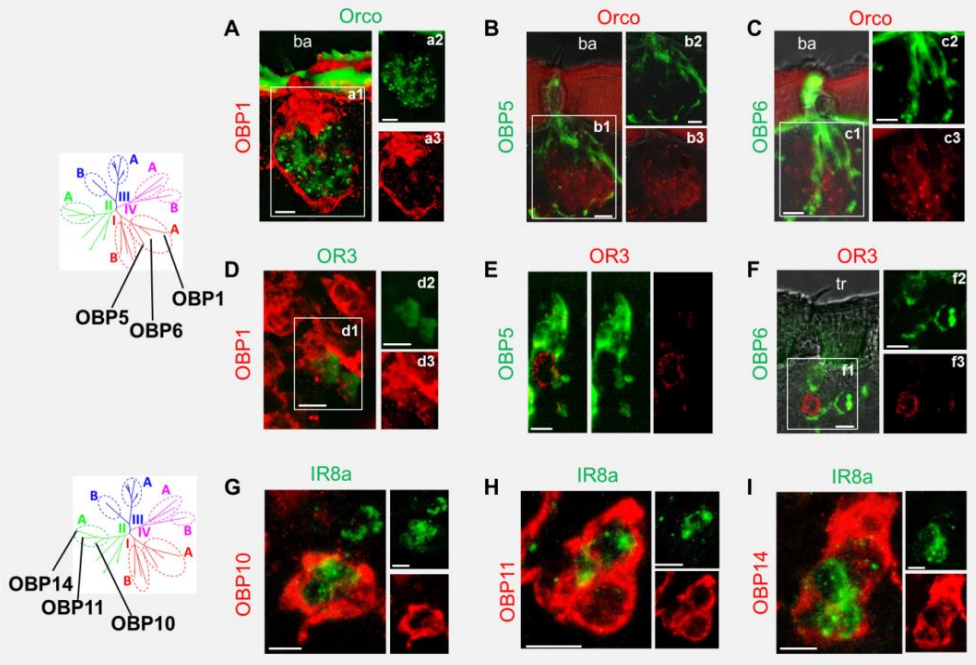


Figure 2. Cells expressing subfamily I-A and II-A SgreOBPs associate with different types of sensory neurons.
Gene-specific antisense riboprobes labeled with either biotin (green) or digoxigenin (red) were used to visualize the appropriate structures by means of two-color fluorescent *in situ* hybridization (FISH). Labeled probes for Orco and IR8a, the ubiquitous co-receptor of ORs and IRs, were used as molecular markers for sensilla basiconica and sensilla coeloconica, respectively. A labeled probe for SgreOR3 was used as a specific marker for distinct sensilla trichodea. The investigated *S. gregaria* OBP subtypes were indicated in the schematic phylogenetic tree. (**A**-**F**) Locust subfamily I-A cells are associated with OR cells. The green (**a**2-**d**2 and **f**2) and red (**a**3-**d**3 and **f**3) fluorescence channels are shown separately for the boxed area indicated in pictures of the overlaid fluorescence channels (**a**1-**d**1 and **f**1). (**G**-**I**) Subfamily II-A OBP cells are associated with IR expressing cells. OBP10, OBP11 and OBP14 expressing cells were observed to tightly envelope IR8a-positive sensory neurons, which are housed in sensilla coeloconica. Confocal images of the separated fluorescence channels are shown at a reduced-size (right) next to the overlaid fluorescence channels (left). Scale bars, 10µm.

Given that OBPs from different orthologous groups are expressed in the same sensillum type, it remained unclear whether they are co-expressed in the same cells or in distinct cells within the sensillum complex. Determining the relative topographic expression of OBP subtypes from subfamily I-A has revealed that OBP1-positive cells were present in almost all basiconic and trichoid sensilla, whereas OBP5- and OBP6-positive cells accounted only for a subpopulation of OBP1 expressing cells in the

same sensillum. Incidentally, it seemed that OBP5 and OBP6 were expressed in the same cells within the sensillum complex.

The selective expression only in sensilla basiconica and sensilla trichodea appears to be a characteristic hallmark for locust subfamily I-A OBPs and interestingly, this hold true also for their orthologs in other insects. For example, in *Drosophila melanogaste*r most of the orthologs are associated with sensilla basiconica and sensilla trichodea, analogous to their locust counterparts (Larter et al., 2016). Moreover, in the moth *Manduca sexta*, the two orthologs MsexABP2 and MsexABPx are specifically expressed in sensilla basiconica (Nardi et al., 2003). Given that locusts emerged at a much earlier stage in insect evolution than moth and fly species (Vieira and Rozas, 2011; Vogt et al., 2015), it appears conceivable that the common feature of a topographic expression only in sensilla basiconica and sensilla trichodea may represent an ancestral status. For the later emerged insect species, like moth and fly, some OBP subtypes may have evolved towards a more specific function coinciding with a constrained expression in either sensilla basiconica or sensilla trichodea (Maida et al., 2005; Larter et al., 2016).

The OBPs of subfamily II-A OBPs seemed to be expressed exclusively in the cells of sensilla coeloconica (**Figure 2**), with no evidence for an expression in any other sensillum types. Exploring the differential expression of the various subtypes revealed that OBP10 and OBP14 were found in either the same or in different set of cells. The subtypes OBP10 and OBP11 were frequently found to be co-expressed in the same cells. Since subfamily II-A OBPs and their orthologs have presumably diverged from a common ancestor, it was proposed that the orthologs may share a conserved sensillar-specific expression. In fact, the only ortholog in *Drosophila melanogaster* DmelOBP84a was found to be specifically expressed in sensilla coeloconica (Larter et al., 2016).

Interestingly, the gene encoding OBP84a retains in the genomes of most, if not all, *Drosophila* species (Cicconardi et al., 2017). Moreover, the OBP84a gene family in *Drosophila* species are subjected to purifying selection pressure (Vieira et al., 2007), and converge onto a segregated phylogenetic clade (Cicconardi et al, 2017),

resembling the locust subfamily II-A OBPs. These molecular and phylogenetic commonalities of subfamily II-A relevant OBPs may point to some similarities concerning the functional implications. In this view, it is interesting to note that single sensillum recordings from sensilla coeloconica in locusts, flies and moths revealed the response spectra that comprise ecologically relevant odorants, including organic acid, amines and plant derived odorants (Pophof, 1997; Ochieng and Hansson, 1999; Yao, 2005). Thus, it will be of particular interest to explore the potential roles of subfamily II-A OBPs and their orthologs with respect to the cognate odorants of sensilla coeloconica.

Unlike DmelOBP84a, which is ubiquitously expressed in almost all sensilla coeloconica (Larter et al., 2016), in locust the expression of subfamily II-A OBPs in sensilla coeloconica occur in a combinatorial mode. This finding is in line with the observation that in *S. gregaria* different subsets of sensilla coeloconica showed distinct response spectra to a repertoire of diagnostic odorants (Ochieng and Hansson, 1999). Together, the results indicate sensilla-specific response spectra as well as sensilla-specific repertoire of OBPs.

Further, the co-existence of more than one OBP subtypes in a sensillum may have more functional implications, due to a possible hetero- and homo-dimerization of OBPs (Andronopoulou et al., 2006), which is always accompanied by conformational changes (Wogulis et al. 2006, Mao et al., 2010). In this regard, it is interesting to note that in the presence of two OBPs, the binding affinity to cognate ligands was altered dramatically in comparison with the binding characteristics in the presence of a single OBP subtype (Qiao et al., 2011; Sun et al., 2016). Thus, the co-existence and possible interaction of multiple OBP subtypes in one sensillum may be relevant for its sensory capacity. This aspect may be particular relevant for sensilla basiconica in locusts housing up to 50 sensory neurons, which are probably tuned to a panel of different odorants (Ochieng et al., 1998; Ochieng and Hansson, 1999). Moreover, in locust the repertoire of OBP genes is much smaller than that of OR genes, with more than 140 ORs identified in *L. migratoria* (Wang et al., 2014) and at least 120 ORs identified in *S. gregaria* (Pregitzer et al., 2017). The selective expression pattern in

particular sensillum types implies that only a relatively small number of OBP subtypes are present in the sensillum lymph. One could imagine that each OBP subtype has distinct ligand specificity and the mixture contributes to the broader binding capacity.

2.1.2. Topographic expression of the OBP subtypes from other subfamilies

In view of the sensilla-specific expression of *S. gregaria* OBPs from subfamilies I-A and II-A, questions arise concerning the topographic expression of OBPs from other subfamilies. A comprehensive analysis of the expression patterns for OBPs from other subfamilies has unraveled a diverse sensilla-specific expression patterns for distinct OBP subtypes.

2.1.2.1 Subfamilies III-A, III-B and I-B

Apart from the subfamilies I-A and II-A, there is another conserved group of locust OBPs, subfamily III-A, which comprises the “plus-C type-A OBPs”. Analysis of OBP4, the representative of subfamily III-A, concerning a possible expression in four antennal sensillum types has specified the association with sensilla chaetica, but not with any of the other three sensillum types (**Figure 3**). Apart from the subfamily III-A, clade III also comprises subfamily III-B, which includes the “classic OBP” subtype OBP8 and its orthologs. Resembling the “plus-C type-A” subtype OBP4, topographic expression of OBP8 was also selectively associated with sensilla chaetica (**Figure 3**). Therefore, it seems that the OBP subtypes of clade III are specifically expressed in sensilla chaetica, and thus deviate from the spatial distribution of OBPs from subfamilies I-A and II-A.

In view of a clade-specific expression pattern, subfamily I-A and II-A OBPs in sensilla trichodea and sensilla basiconica versus subfamily III-A and III-B OBPs in sensilla chaetica, the question arises where OBPs of subfamily I-B may be expressed. Surprisingly, the two representatives of subfamily I-B, OBP2 and OBP7 were neither found in sensilla basiconica nor in sensilla trichodea, but also in sensilla chaetica

(**Figure 3**). To extend and specify the topographic expression of these OBPs in sensilla chaetica, the analyses were concentrated on the tip of the antennae where sensilla chaetica are highly enriched and only very few of the other three sensillum types exist (Ochieng et al., 1998). The notion that these subtypes were selectively expressed in sensilla chaetica was substantiated by the finding that numerous cells were visible that were positive for OBP4 (subfamily III-A), OBP8 (subfamily III-B) as well as OBP2 and OBP7 (subfamily I-B). In contrast, in the same region no evidence was found for an expression of OBPs that are expressed in other sensillum types, such as OBP5 (subfamily I-A) and OBP11 (subfamily II-A).

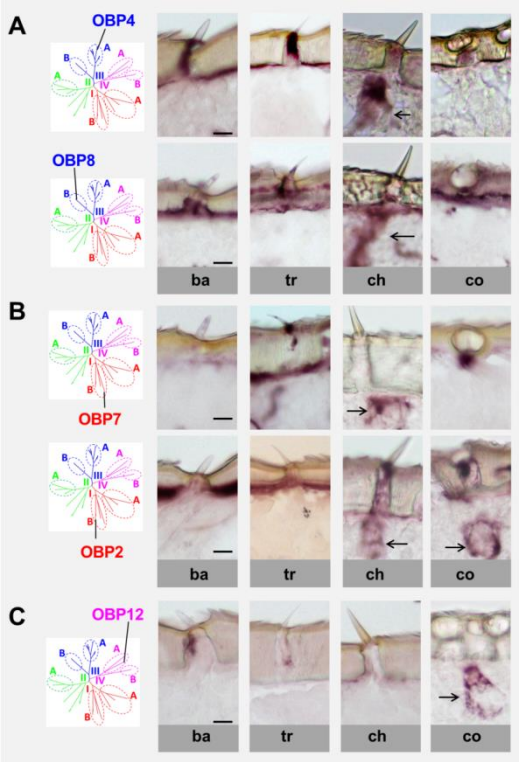


Figure 3. Sensilla-specific expression of the OBP subtypes from different subfamilies. The specific *S.gregaria* OBP subtypes studied in this analysis were indicated in the schematic phylogenetic tree. Topographic expression of OBPs was visualized by using antisense riboprobes specifically targeting distinct OBP subtypes in conjunction with chromogenic *in situ* hybridization (ISH). (**A-C**), Visualization of the labeled cells expressing distinct OBP subtypes from subfamily III-A, III-B, I-B and IV-A in four morphological types of antennal sensilla. Ba, sensilla basiconica; Tr, sensilla trichodea; Ch, sensilla chaetica; Co, sensilla coeloconica. The visible labeled structures are denoted by black arrows. Scale bars, 10µm.

Given that several OBP subtypes were expressed in sensilla chaetica, it remained unclear whether they are expressed in the same or different sets of cells. Detailed analyses revealed that OBP4 and OBP8 (clade III) are co-expressed in the same set of cells and the similar scenario occurred for OBP2 and OBP7 (subfamily I-B). Exploring the relative cellular localization of OBP subtypes from different clades uncovered that the member of subfamily III-A (OPB4) and members of subfamily I-B (OBP2 and OBP7) were widely co-expressed. In contrast, the member of subfamily

III-B (OBP8) and the members of subfamily I-B (OBP2 and OBP7) were expressed in different sets of cells. Interestingly, while OBP2 and OBP8 were expressed in different cells of the same sensillum chaeticum, OBP7 was found in different non-OBP8-positive sensilla.

Sensilla chaetica are characterized by distinct structural features, such as a thick and poreless cuticle wall, an apical pore and relatively few dendrites (Ochieng et al., 1998; Zhou et al., 2009). Consequently, this type of sensilla is regarded as specialized for the reception of tastants rather than odorants. This notion is verified for the fruit fly (Montell, 2009; Chen and Amrein, 2017; Scott, 2018) and may also hold true for locusts. Since locust sensilla chaetica are enriched on the tip of antennae and palps (Blaney and Chapman, 1969; Ochieng et al., 1998), they are supposed to play a receptive role of contact stimuli (Blaney, 1974; Blaney, 1975; Saini et al., 1995). The presence of four OBP subtypes in sensilla chaetica on the tip of antennae, therefore, implies that they may participate in the reception of gustatory stimuli. This view is supported by a recent finding that in *Drosophila* distinct OBP subtypes are expressed in “gustatory sensilla” and are indeed involved in the reception of tastant (Jeong et al., 2013). Moreover, a recent study has demonstrated that in *L. migratoria* knock-down of a sensilla chaetica-specific OBP subtype resulted in a reduced neuronal response to chemical stimuli (Zhang et al., 2017). These findings support the notion that locust OBPs are intimately involved in the reception of chemical compounds through sensilla chaetica.

Structurally, sensilla chaetica of locusts are characterized by a sensillum lymph cavity which is composed of an inner and an outer compartment (Ochieng et al., 1998; Zhou et al., 2009). In a recent study of analyzing sensilla chaetica of *L. migratoria,* the immunoreactivity for an OBP subtype was found in the non-innervated outer lumen but not in the inner compartment where the sensillum lymph is bathing the chemosensory dendrites (Yu et al., 2009). This finding has raised speculations of how the cognitive ligands may reach the chemosensory dendrites. Our finding that four distinct OBP subtypes are present in this sensillum type may give way to revisit this concept in more detail.

The observation that OBP2 (subfamily I-B) was not only expressed in sensilla chaetica but also in sensilla coeloconica was confirmed by experiments using the receptor IR8a as a specific marker for sensory neurons housed in sensilla coeloconica (Abuin et al., 2011; Guo et al., 2013). The results raised the question whether OBP subtypes from both subfamily I-B and subfamily II-A may be co-expressed in sensilla coeloconica. Using OBP10 and OBP14 as representatives for subfamily II-A, it was found that in a distinct set of cells OBP2 was indeed co-expressed with the subfamily II-A representatives. This observation further substantiates the notion that OBP2 is expressed in both sensillum types.

Sensilla coeloconica and sensilla chaetica differ markedly in both their external morphology as well as their functional implications (Montell, 2009; Rytz et al., 2013; Joseph and Carlson, 2015; Scott, 2018), nevertheless, both sensillum types share some common molecular features most remarkably the ionotropic receptor type IR25a, one of the co-receptors of divergent IRs (Abuin et al., 2011; Guo et al., 2013). Given that a multidimensional role has been assigned to the receptor IR25a (Rimal and Lee, 2018), such a multiple-function-mode may also be true for the binding proteins. Actually, it has been demonstrated that OBPs are involved in quite a variety of functional aspects (Pelosi et al., 2006; 2014; 2017). In this regard, the observation that OBP2 co-exists with other OBP subtypes in different sensillum types may suggest that OBP2 co-operates with other OBPs to fulfill the sensilla-specific functions.

2.1.2.2 Subfamily IV-A

Subfamily IV-A comprises the “atypical OBP” subtypes, which are characterized by an extraordinary long sequence of motif spanning C1 and C2. This unique structural feature suggests that the atypical subtype may fulfill a distinct functional role and thus may be expressed in a distinct cell or sensillum type. As a representative of subfamily IV-A the expression pattern for OBP12 was analyzed; OBP12-positive cells were only found in sensilla coeloconica (**Figure 3**). The sensilla specificity was verified by

demonstrating the co-localization of OBP12 cells and IR8a-positive neurons in one sensillum. Exploring the distribution of OBP12 cells relative to the cells expressing OBPs of subfamily II-A, such as OBP10 or OBP14 which are also expressed in sensilla coeloconica, revealed that they are expressed in different cells. Furthermore, a comprehensive assessment revealed that OBPs from subfamily II-A together with OBP2 from subfamily I-B are present in the majority of sensilla coeloconica, whereas the atypical subtype OBP12 from subfamily IV-A was selectively expressed in a subpopulation of sensilla coeloconica.

Due to the distinct sequence features and the topographic distribution, it is conceivable that the presence of OBP12 in some sensilla coeloconica may pronounce some unique functional capacity. In this regard, it is of particular interest to note that in locusts most of the sensilla coeloconica are tuned to leaf odors and organic acids (Ochieng and Hansson, 1999) but a subpopulation of sensilla coeloconica were found to be tuned to different stimuli, such as hygro- or thermo-stimuli (Altner et al., 1981). The functional versatility of this sensillum type may depend on specialized sensory cells with specific receptors in combination with appropriate OBPs in the sensillum lymph. Furthermore, the gene encoding the “atypical OBP” subtype OBP12 belongs to the so-called OBP59a gene family, which is well conserved across many insect species, except in Hymenoptera (Vieira and Rozas, 2011). In *Drosophila melanogaster* OBP59a was found to be exclusively expressed in sensilla coeloconica (Larter et al., 2016), analogous to its counterpart in the desert locust.

2.1.2.3 Subfamily IV-B

Based on the sequence analysis, OBPs with more than six conserved C-residues, called “plus-C OBPs”, were classified into two different categories, each category residing in a different subfamily: type-A OBPs in subfamily III-A and type-B OBPs in subfamily IV-B. Whereas type-A OBPs are expressed in sensilla chaetica (see chapter **2.1.2.1**), the topographic expression of type-B OBPs is unclear. As

representative of the “plus-C type-B” subtypes OBP9 was analyzed. On histological planes close to the antennal nerve bundle, the labeling pattern for OBP9 emerged (**Figure 4**); the labeling was apparently less associated with a distinct sensillum type as found for the other OBP subtypes. Upon closer inspection it was found that OBP9 expressing cells extended cytoplasmic processes towards the sensilla cell complexes and enclosed the associated sensory neurons. Intriguingly, when histological planes closer to the cuticle were assessed, the labeling intensity appeared to be considerably enhanced and a distinct “nest-like” labeling pattern for OBP9 emerged (**Figure 4**). This observation gives rise to the possibility that OBP9-positive cells may be associated with more than one sensillum types. This concept was scrutinized by analyzing a possible co-localization of OBP9 with markers indicative for distinct sensillum types. In a first approach OBP9-positive cells were found to tightly surrounding Orco-positive cells, a marker for sensory neurons of sensilla basiconica. A similar co-localization pattern emerged using OR3 and IR8a as specific markers for sensilla trichodea and sensilla coeloconica (**Figure 4**). Furthermore, a co-localization of cells OBP9 and OBP8 was also observed with a marker for auxiliary cells of sensilla chaetica (**Figure 4**). These results indicate an association of OBP9, a representative of “plus-C type-B” subtype, with all four antennal sensillum types.

This distribution pattern of OBP9 is unique and any functional implications remain unknown. However, the ubiquitous distribution may point to a general function of OBP9, either on its own or in an interplay with co-existing OBP partners. Such an interactive regime has recently been documented which demonstrates that an OBP complex has a broader ligand binding spectrum (Qiao et al., 2011). On the other hand, it is also conceivable that OBP9 may be involved in quite different functional aspects, e.g. differentiation and survival of cells in the sensillum complex. In this regard, recent studies have provided evidence that small soluble proteins such as CSPs may be involved in developmental processes and tissue regeneration in cockroach and honeybee (Nomura et al., 1992; Maleszka et al., 2007; Cheng et al., 2015).

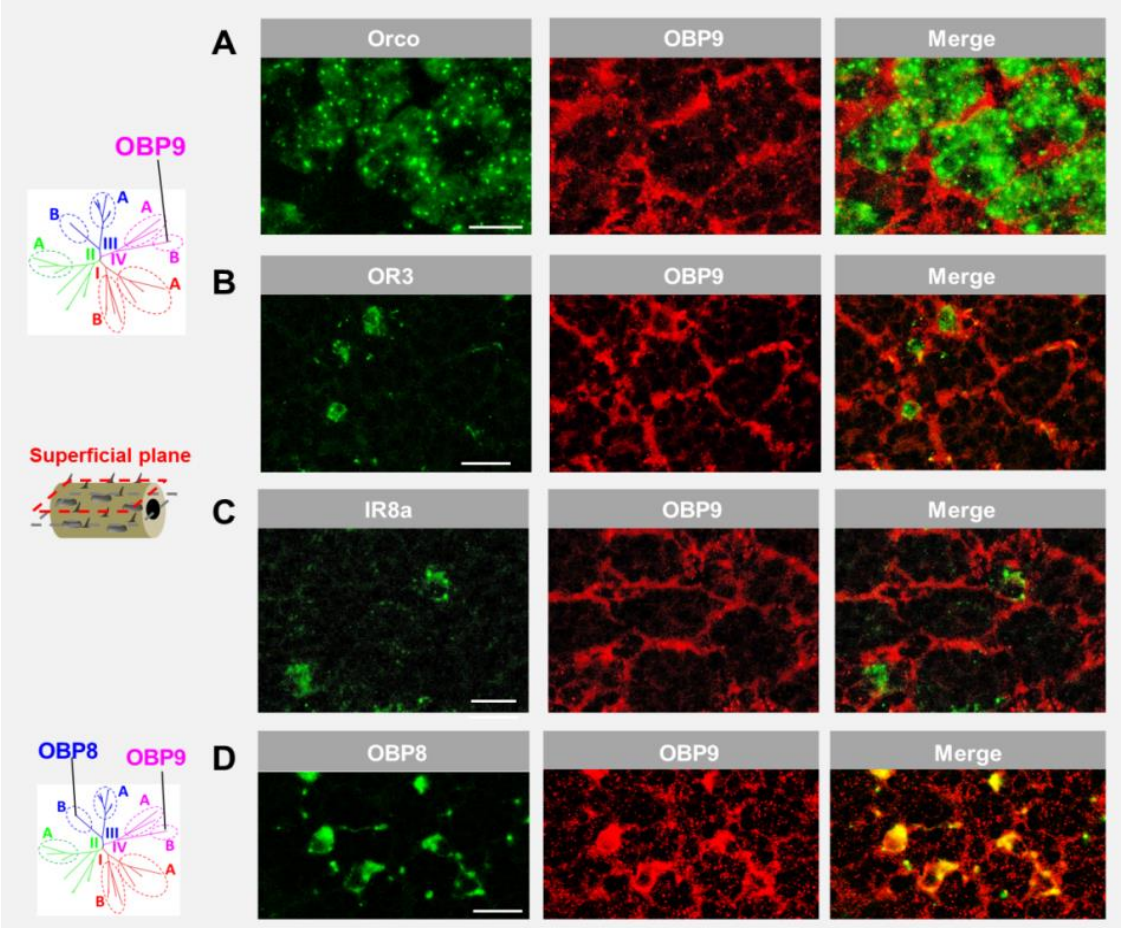


Figure 4. OBP9 expressing cells associate with four types of antennal sensilla. The relative localization of OBP9 and different marker genes indicative of specific sensillum types was analyzed by utilizing specific antisense riboprobes and the means of two-color FISH. Presented images were obtained from superficial cellular planes approaching the cuticle by performing series of horizontal sections of the antennae (diagram, left lane). (**A-C**) Orco, OR3, and IR8a were utilized as the specific molecular markers of neurons housed in sensilla basiconica, sensilla trichodea, and sensilla coeloconica, respectively. (**D**) OBP8 was used as a marker for auxiliary cells of sensilla chaetica. Scale bars, 10μm.

2.2. Search for the candidate pheromone receptors in the desert locust

In marked contrast to Lepidopteran and Dipteran species, the mechanisms of pheromone signaling in hemimetabolic insects are largely elusive, although in locusts, not only the reproduction but also the unique aggregation behavior is supposed to be controlled by specific pheromonal compounds. For some locust species, most notably the desert locust *Schistocerca gregaria*, a massive

reproduction together with an aggregation in the gregarious phase often leads to the formation huge swarms, which cause tremendous damage on agricultural crops in Africa and Asia. It is conceivable, that a targeted and efficient disruption of the reproduction and aggregation behavior could be the bases for an environmentally safe strategy to control these devastating pest insects. An essential prerequisite for such biological approaches to control locust swarm formation is a detailed knowledge about the molecular bases of pheromone signaling in locust. In the past, searches for relevant pheromonal compounds were hampered by complexity of volatile emissions of locusts (Hassanali et al., 2005) and although many pheromonal effects were reported, the findings did not provide a coherent picture for the relevant pheromones controlling reproductive and aggregation behaviors (Pener and Yerushalmi, 1998). Also, attempts to elucidate the mechanisms underlying the detection of pheromonal compounds in locust were not successful. Key elements in the process of pheromone detection are of course the receptors for pheromones, but all approaches to search for such receptors in locust, including analyses based on sequence information of pheromone receptors in moths and flies failed. During the course of deciphering the moth pheromone system, it was found that pheromone responsive neurons not only express a specific pheromone receptor protein but in addition a cell-specific protein, the so-called “sensory neuron membrane protein 1” (SNMP1) (Forstner et al., 2008). Given the fact that all direct approaches for identifying locust pheromone receptors failed, an alternative, indirect strategy was developed based on the paradigm that candidate pheromone receptors should be expressed in SNMP1-positive neurons of locust antennae. As a first step towards this approach, it was necessary to identify the locust-specific SNMP1 and to visualize which antennal neurons do express this protein. As a second step, candidate odorant receptors have to be identified by screening an antennal transcriptome database, and subsequently for representative OR candidates, the topographic expression patterns in the antennae have to be evaluated concerning a possible co-expression with SNMP1.

2.2.1 Identification of “Sensory Neuron Membrane Proteins” (SNMPs)

A homology-based identification strategy was employed based on sequences of SNMPs from several species to search an antennal transcriptome database from the *S. gregaria* for the sequences encoding SNMPs. This approach resulted in two transcripts that were categorized into the CD36 superfamily, a subfamily of scavenger receptor B (Acton et al., 1994; Febbraio et al., 2001; Silverstein et al., 2010), indicating the identity of SNMP homologs. To extend and specify this classification, an unrooted Neighbor-Joining phylogenetic tree was constructed choosing the SNMP sequences from the desert locust together with sequences of the species from five other insect taxa. In line with previous finding (Vogt et al., 2009), the analyzed SNMP sequences were clearly arrayed into two major clades; one of the candidate SNMP sequence of the desert locust was clustered into the SNMP1 subclade, the other sequence grouped into the SNMP2 subclade. A detailed sequence comparison was performed with SNMP1 and SNMP2 sequences from insects representing Lepidoptera, Diptera, Coleoptera, Hymenoptera and Orthoptera, revealed that SgreSNMP1 and its orthologs shared more than 40% sequences identities while only about 30% identity was found for the SNMP2 sequences. For both SgreSNMP1 and SgreSNMP2 the transmembrane topology was characterized by two transmembrane helices close to the N- and C-terminal regions, respectively resulting in a large extracellular loop comprised of 426 amino acids for SNMP1 and 431 amino acids for SNMP2. Two conserved sequence domains locating in proximity of the N-terminus and C-terminus have been reported for SNMP sequences (Vogt et al., 2009). In fact, for the SgreSNMPs four conserved sequence domains were identified; sequence domains one, two and three largely covered the previously reported conserved domains and the fourth domain is localized in proximity to the C-terminus downstream of the third domain. The seven conserved C-residues in the ectodomain of SNMPs are considered as hallmarks for SNMP1 and SNMP2 sequences.

2.2.2 Topographic localizations of SNMP1 and SNMP2

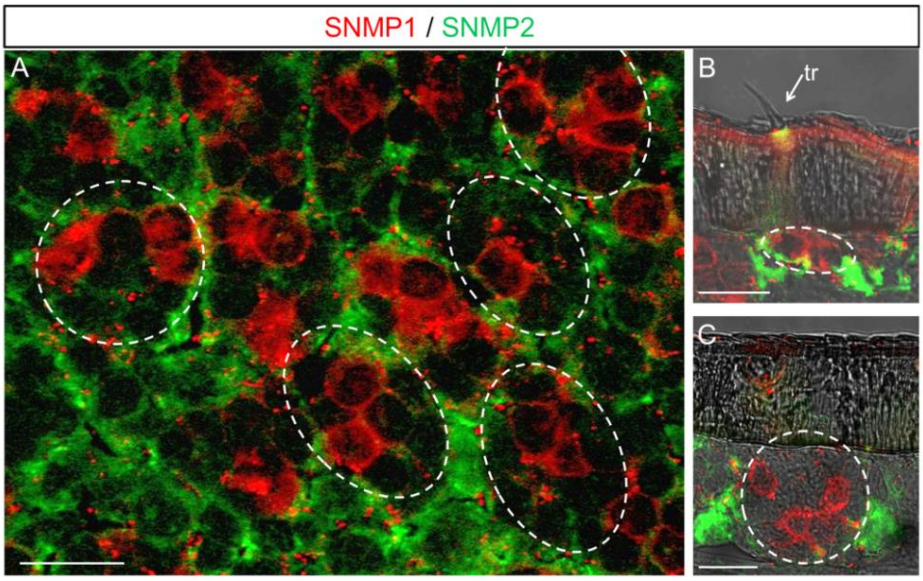


Figure 5. SgreSNMP1 and SgreSNMP2 are expressed in spatially distinct cells.
A DIG-labeled probe for SgreSNMP1 and a BIO-labeled probe for SgreSNMP2 were used in two-color FISH assays. (A) SgreSNMP1 expressing cells are encircled by SgreSNMP2 expressing cells. Dashed areas outline clusters of neurons in which some of the cells express SgreSNMP1. The sensillum trichodeum (tr, in B) and the sensillum basiconicum (in C) house SgreSNMP1 cells as well as SgreSNMP2 expressing cells; both cell types are not overlapping. SgreSNMP1 expressing cells were outlined by dash lines. Scale bar: 20 µm.

The topographic expression pattern for SNMP1 and SNMP2 was determined by means of FISH-experiments. Using probes for SNMP1 resulted in a large set of labeled SNMP1 cells in an antennal segment. Labeled cells seemed to be restricted to sensilla basiconica and sensilla trichodea; SNMP1-positive cells were found neither under sensilla coeloconica nor under sensilla chaetica. In the desert locust a typical sensillum basiconicum is innervated by up to 50 neurons, whereas a sensillum trichodium comprises no more than three sensory neurons (Ochieng et al., 1998). In line with the anatomical features, the number of SNMP1-positive cells located beneath sensilla basiconica was much higher than that of sensilla trichodea. To extend and specify SNMP1 expression in sensilla basiconica, the relative topographic expression of SNMP1 and Orco, the obligatory co-receptor of the OR types, was analyzed. SNMP1 expressing cells were found to represent a subpopulation of Orco-

positive cells, indicating that SNMP1 is indeed expressed in distinct sensory neurons of sensilla basiconia. Furthermore, the number of sensory neurons which express SNMP1 varied among distinct sensilla basiconica.

The expression of SNMP2 was previously explored in moth species and was found to be absent from sensory neurons, but rather restricted to cells flanking the sensory neurons (Forstner et al., 2008). Attempts to examine the expression of SNMP2 in the desert locust revealed that SNMP2-labeled structures appeared to be less associated with specific sensillum types. Labeling was generally distributed at the periphery of neuron clusters, the region where auxiliary cells are localized. This notion was further substantiated using Orco as marker for sensory neurons; the double labeling approaches revealed that SNMP2-positive cells were tightly encircling the Orco-positive cells, with no overlapped labeling. Therefore, as in moth species, in the desert locust SNMP2 appears to be selectively expressed in the auxiliary cells, but not in sensory neurons. This view was further supported by the observation that SNMP1-positive cells were encircled by the cells with SNMP2 (**Figure 5**).

2.2.3 Identification and classification of putative olfactory receptor types

Through a search of *S. gregaria* antennal transcriptome database, a total of 119 transcripts were identified, which encode candidate odorant receptor (OR) types; this number is comparable to that reported for the OR repertoire of the migratory locust (~142 ORs) (Wang et al., 2015). A survey of the open reading frames for the 119 identified SgreOR sequences resulted in presumptive full length sequences for 28 SgreOR types and partial sequences for 98 SgreOR types. The sequence similarity of ORs from desert locust and the migratory locust was elucidated by constructing a phylogenetic tree using 117 SgreOR sequences and 138 LmigOR sequences (**Figure 6**); rooting of the tree was based on a set of Orco sequences, a non-canonical member of OR family highly conserved among insect species. The results indicate a pronounced orthologous relationship for a large proportion of ORs, which was reflected in the convergence of orthologous ORs onto a monophyletic clade,

supported by high bootstrap evidence on the divergent node (**Figure 6**).

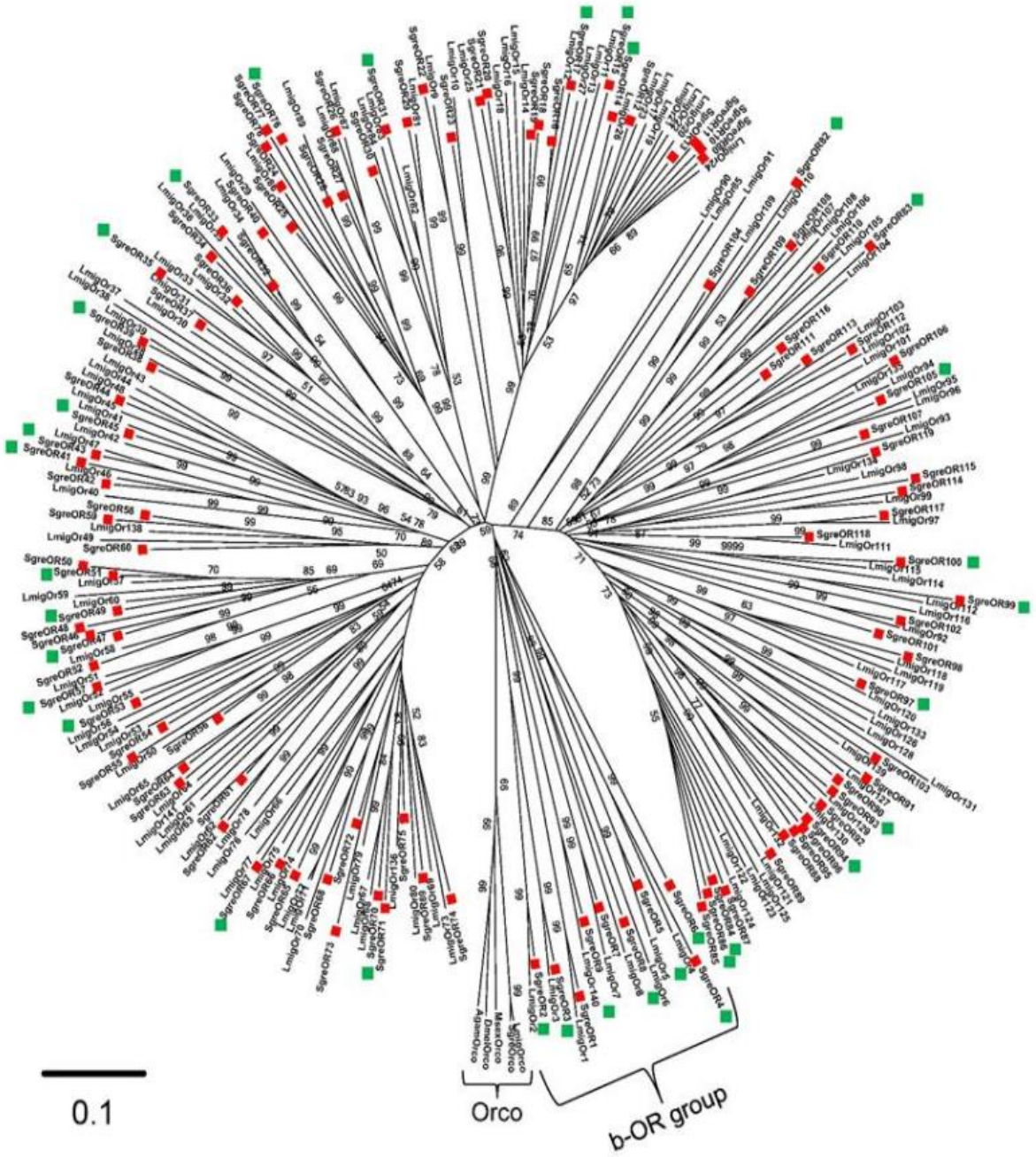


Figure 6. Evolutionary relationships among SgreORs and LmigOrs.
A neighbor-joining tree was constructed to infer the relationships among coding sequences of 117 SgreORs and 138 LmigORs. Branch lengths are proportional to the percentage of sequence differences. The scale bar indicates 10% difference. The numbers on the node indicate bootstrap support values (in %) based on 1000 replicates (only values above 50% are shown). OR-coding sequences from *S. gregaria* are denoted by red squares. SgreOR sequences further addressed in double labeling assays are highlighted with green squares.

However, for some SgreOrs and some LmigORs, a classification of direct orthologs in the other locust species was ambiguous. Albeit of the enigmatic phylogenetic relationship of OR families from two locust species, a rather clear orthologous relationship emerged for members of a “basal” OR group (b-OR group), the OR sequences which were positioned adjacent to the Orco-rooting group (**Figure 6**). The b-OR group represent a segregated phylogenetic branch and harbors 8 orthologous SgreOR / LmigOR pairs, with sequence identities ranging from 47% to 87%. Notably, a similar scenario has been reported for the pheromone receptors of moth species (Grosse-Wilde et al, 2011; Krieger et al., 2004).

2.2.4 Assessment of b-OR types

Based on the clear orthology, members from the b-OR group were considered as promising candidates for pheromone receptors of locusts. Therefore, SgreORs in this group were assessed for a possible co-expression with SNMP1. Analyses of the topographic expression profile in the antennae revealed that the numbers of cells expressing distinct SgreOR types from the b-OR group varied significantly; for example, SgreOR8 was expressed in a large number of cells whereas SgreOR6 was only expressed in a few cells. Despite of the varying number of cells, six out of seven SgreOR members from the b-OR group were found to be co-expressed with SNMP1; only SgreOR4 was not expressed in SNMP1-positive cells (**Figure 7**). Since pheromone receptors in moths and flies are specifically expressed in sensilla trichodea, it was of particular interest to determine in which sensilla type the receptor may be expressed. Analyses of the sensilla specificity for distinct SgreOR types revealed that only the receptor type SgreOR3 was specifically expressed in sensilla trichodea, whereas the other six SgreOR types were all expressed in sensilla basiconica.

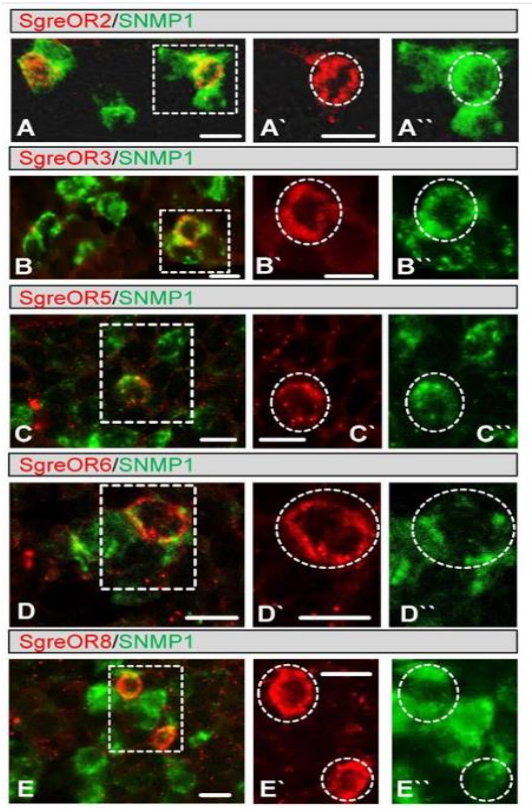


Figure 7. Co-expression of different b-OR types with SNMP1. (**A**-**E**) Two-color FISH approaches were performed on sections through antennae with antisense RNA probes for SNMP1 (A-E, green) and SgreOR2 (A, red), SgreOR3 (B, red), SgreOR5 (C, red), SgreOR6 (D, red) and SgreOR8 (E, red). The micrographs shown represent single optical planes taken from stacks of confocal images. Cells positive for SgreOR2, SgreOR3, SgreOR5, SgreOR6 or SgreOR8 co-express SNMP1. (**A`/A``-E`/E``**) Higher magnifications of the boxed areas in A, B, C, D and E are depicted in either the red or the green color channel. Cells co-expressing the relevant SgreOR types and SNMP1 are marked with dashed circles. Scale bars: 10 µm.

2.2.5 Assessment of non-b-OR types

Beyond the b-OR group, OR types belonging to other groups on the phylogenetic tree were also inspected in view of a possible co-expression with SNMP1. To this end, a total of 28 SgreOR types were selected and evaluated for their topographic expression on the antennae. Similar with the b-OR types, each of the 28 SgreOR types from the non-b-OR groups were found to be expressed in a varying number of cells. In contrast to the b-OR group in which most of the receptors were expressed in SNMP1-positive cells, only a small proportion of receptors from the non-b-OR groups (5 out of 28) was co-expressed with SNMP1; in fact the majority (23 out of 28) was found in SNMP1-negative cells. Thus, these results indicate that in addition to b-OR-types, at least five receptor types from the non-b-OR groups may also be considered promising candidates in the search for pheromone receptors.

Thus, based on the search paradigm that candidate receptors for pheromones should be expressed in SNMP1-positive cells, the results suggest that genes

encoding pheromone receptors may reside on different clades of the phylogenetic tree and some may be orthologous to other locusts whereas others are not. This observation may be relevant in view of the fact that in the desert locust specific signaling compounds may initiate and mediate the remarkable transition between the solitary and the gregarious phase (Hassanali et al., 2005). Therefore, it is conceivable that in the desert locust receptors for pheromones are not only be essential for the reproduction and aggregation processes (Seidelmann and Ferenz, 2002) but may also play a crucial role in controlling the characteristic phase transition.

3 Summary

Locusts are remarkable insects due to their unique and potentially devastating phenotypic plasticity based on the local population density. While “solitarious” phase locusts avoid one another, “gregarious” locusts can form dense and highly mobile swarms, which have been feared as agricultural pests since ancient history. For this reason alone, locust biology has long been the object of intense scientific studies; moreover, from a purely scientific perspective it is of great interest to unravel the mystery underlying the phenotypic plasticity. The unique phase transition including the behavioral plasticity heavily relies on chemical communication by means of critical volatiles. It is therefore important to elucidate the mechanisms underlying locust chemosensory communication, including the identification of molecular elements involved in recognizing odorous compounds. Towards this goal, the desert locust *Schistocerca gregaria,* as a representative locust species, was investigated in this study.

One of the key elements for recognizing odorous compounds are odorant binding proteins (OBPs). To gain insight into the repertoire of locust OBPs, genomic sequences encoding candidate OBPs from *Schistocerca gregaria* together with those from three other locust species were subjected to thorough comparative analyses. The results indicated that locust OBPs could be classified into several categories, namely, “classic OBPs”, “plus-C OBPs”, “minus-C OBPs” and “atypical OBPs” which reside in four major phylogenetic families (I to IV). With the aim to uncover distinct features of the various OBP types, the initial studies were concentrating on the conserved subfamilies I-A and II-A which comprise “classic OBPs”. The sequence analyses provided evidence for both common and subfamily-specific motifs as well as evolutionary clues based on the calculation of coden substitution rates, which suggested the effect of purifying selection pressure. The subfamily I-A comprised a much higher number of orthologous OBPs than subfamily II-A, which resulted in a distinct re-clustering patterns for subfamily I-A and subfamily II-A. Exploring the topographic expression pattern on the antennae revealed that OBPs of subfamily I-A

were selectively expressed in sensilla basiconica and sensilla trichodea, whereas OBPs of subfamily II-A were restricted to sensilla coeloconica. Furthermore, cells expressing the subtype OBP1 were present in almost all sensilla basiconica and trichodea, whereas other subtypes were only present in subpopulations. The OBPs of subfamily II-A, were expressed in distinct subpopulations of sensilla coeloconica. Analyses of representative OBPs from the remaining phylogenetic subfamilies revealed that representative subtypes from subfamily III-A and III-B were expressed in sensilla chaetica, similarly the two representatives of subfamily I-B were also expressed in this sensillum type. The selective expression of these OBPs in sensilla chaetica was substantiated by analyzing the antennal tip, which comprises numerous sensilla chaetica. The “atypical OBP” OBP12, a representative of subfamily IV-A was found to be selectively expressed in a distinct subpopulation of sensilla coeloconica, while “plus-C OBP” OBP9, from subfamily IV-B, showed a unique expression pattern and seemed to be associate with all four sensillum types. The diversity and complex sensilla- and cellular-specific distribution implies distinct functional implications of OBP subtypes in the process of chemoreception.

In locusts not only the complex reproduction behavior but also unique aggregation behavior is regulated by pheromonal compounds, however, the knowledge about the pheromone system of locusts is very limited. In view of the sparse information, a search for candidate receptors seemed a promising approach towards a better understanding of the locust pheromone system. Analyses of mechanisms underlying pheromone detection in moth and fly have uncovered an important role of the “sensory neuron membrane proteins” (SNMPs). Since SNMP1 is always co-expressed with the pheromone receptors, it is considered as a marker for pheromone responsive antennal neurons. The two types of SNMPs, the SNMP1 and SNMP2, were identified in *S. gregaria* and SNMP1 was specifically expressed in the sensory neurons of sensilla basiconica and sensilla trichodea, while SNMP2 was expressed in non-neuronal cells.

Analyzing the sequences encoding candidate odorant receptors from *Schistocerca gregaria* and *Locusta migratoria* revealed an orthologous relationship for a

considerable number of OR sequences. A remarkable relationship emerged for the so-called "basal" OR group, which comprises orthologous SgreOR / LmigOR pairs with high sequence identities. The notion that, based on the high degree of conservation these sequences may encode candidate pheromone receptors, was scrutinized by validating their possible co-expression with SNMP1. The double-labeling experiments revealed that six of the SgreOR members in the b-OR group were in fact co-expressed with SNMP1. Concerning the sensilla types, six SgreOR members were specifically located in sensilla basiconica and only one SgreOR in sensilla trichodea, the sensillum type which in moths and flies are predestined for the detection of sex pheromones. Analyzing OR types from "non-basal" groups revealed the majority of the assessed ORs were expressed in SNMP1-negative cells; however, a small number of OR types was co-expressed with SNMP1. Thus, the findings indicate that OR types from different groups may be considered as candidates for pheromone receptors.

4 Zusammenfassung

Wanderheuschrecken sind bemerkenswerte Insekten, insbesondere wegen ihrer einzigartigen, von der lokalen Populationsdichte abhängigen phänotypischen Plastizität. Dieser Phasen-Polymorphismus besteht aus der solitären Phase („Einzelphase“), die durch weitgehend ortstreue, einzeln lebende Tiere gekennzeichnet ist und der gregären Phase („Schwarmphase“), in der sich die Individuen zu großen, mobilen Schwärmen zusammenschließen. Die bei Massenauftreten verursachten Verwüstungen ganzer Landstriche sind als biblische Plagen seit der Frühgeschichte gefürchtet. Vor diesen vorwiegend sozioökonomischen Hintergründen wird die Biologie der Heuschrecken seit langer Zeit intensiv erforscht. Insbesondere gibt es ein wachsendes Interesse an der Aufklärung molekularer Mechanismen, die dem Phänomen der phänotypischen Plastizität zu Grunde liegen. Der einzigartige Phasen-Wechsel und die damit einhergehenden Verhaltensänderungen werden maßgeblich durch eine chemische Kommunikation mittels flüchtiger Signalmoleküle gesteuert. Daher ist ein besseres Verständnis der Mechanismen, die der chemosensorischen Kommunikation von Wanderheuschrecken zugrunde liegen von großem Interesse; das gilt insbesondere für die Identifizierung von molekularen Elementen, die eine Erkennung und Unterscheidung von kritischen Duftstoffen ermöglichen. Mit dieser Zielsetzung wurde im Rahmen dieser Studie die Wüstenheuschrecke *Schistocerca gregaria*, die als repräsentative Wanderheuschrecken-Spezies gilt, untersucht.

Eines der Schlüsselelemente für die Erkennung der flüchtigen Duftstoffe sind die sogen. Odorant-Bindeproteine (OBPs). Für eine Erfassung des Repertoires an OBPs bei Heuschrecken, wurden OBP-kodierende Sequenzen im Genom von *Schistocerca gregaria* und von drei verwandten Heuschrecken-Arten vergleichend untersucht. Dabei hat sich gezeigt, dass die Odorant-Bindeproteine von Heuschrecken in mehrere Kategorien eingeteilt werden können; speziell, „klassische OBPs“, „plus-C OBPs“, „minus-C OBPs“ und „atypische OBPs“, die vier phylogenetischen Familien (I-IV) zugeordnet werden können. Mit der Zielsetzung spezielle Aspekte der

verschiedenen OBP-Subtypen aufzuklären, wurden zunächst die konservierten Subfamilien I-A und II-A mit den „klassischen OBPs“ untersucht. Im Rahmen der Sequenzanalysen wurden sowohl gemeinsame als auch subfamilien-spezifische Motive identifiziert; darüber hinaus ergaben die Berechnungen von Codon-Substitutionsraten Hinweise für einen negativen Selektionsdruck. Die Subfamilie I-A umfasst eine viel größere Zahl an orthologen OBPs als Subfamilie II-A; dies resultiert in einem distinkten Re-clustering Muster für die beiden Subfamilien. Untersuchungen zur topographischen Expression auf der Antenne ergaben, dass die OBPs der Subfamilie I-A ausschließlich in Sensilla basiconica und Sensilla trichodea exprimiert wurden, während die OBPs der Subfamilie II-A nur in Sensilla coeloconica nachzuweisen waren. Außerdem hat sich gezeigt, dass Zellen welche OBP1 exprimieren in fast allen Sensilla basiconica und Sensilla trichodea vorkommen. Andere Vertreter der Subfamilie I-A kommen dagegen nur in einzelnen S. basiconica und S. trichodea vor. Auch die OBPs der Subfamilie II-A waren nur in Subpopulationen der Sensilla coeloconica exprimiert. Die Analyse von repräsentativen OBPs der übrigen phylogenetischen Subfamilien ergab, dass OBPs der Subfamilien III-A und III-B nur in Sensilla chaetica vorkamen; ähnliches gilt auch für die beiden Vertreter der Subfamilie I-B. Die Befunde einer selektiven Expression dieser OBPs in Sensilla chaetica wurden bestätigt und ergänzt durch die Analysen der Antennen-Spitzen, die eine Vielzahl an Sensilla chaetica aufweisen. Das „atypische OBP“ OBP12, ein Repräsentant der Subfamilie IV-A wurde ausschließlich in einer Subpopulation von Sensilla coeloconica nachgewiesen, während für das „plus-C OBP“ OBP9, von der Subfamilie IV-B, ein einzigartiges Expressionsmuster gefunden wurden; die Befunde sprachen dafür, dass OBP9 offenbar mit allen vier Sensillentypen assoziiert war. Insgesamt spricht die Diversität sowie die komplexe Sensillen- und Zell-spezifische Expression für distinkte Funktionen der OBP-Subtypen in dem Prozess der Chemorezeption von Heuschrecken.

Bei den Heuschrecken wird nicht nur das komplexe Reproduktionsverhalten, sondern auch das einzigartige Aggregationsverhalten durch spezielle Pheromon-Komponenten reguliert; allerdings sind die Kenntnisse hinsichtlich des Pheromon-

Systems von Heuschrecken noch sehr limitiert. Angesichts der mangelhaften Informationslage, schien die Suche nach potentiellen Rezeptoren ein erfolgversprechender Ansatz im Hinblick auf ein besseres Verständnis des Pheromon-Systems von Heuschrecken. Die umfangreichen Studien zum Mechanismus der Pheromon-Registrierung bei Motten und Fliegen haben gezeigt, dass den sogen. „sensory neuron membrane proteins" (SNMPs) eine zentrale Rolle zukommt. Da SNMP1 stets mit Pheromon-Rezeptoren co-exprimiert wird, gilt es als Marker für pheromon-reaktive Zellen in den Antennen. Die beiden SNMP-Typen, SNMP1 und SNMP2, konnten bei *Schistocerca gregaria* identifiziert werden und es hat sich gezeigt, dass SNMP1 spezifisch in sensorischen Neuronen von Sensilla basiconica und Sensilla trichodea exprimiert wird; SNMP2 dagegen in nicht-neuronalen Zellen.

Vergleichende Analysen von rezeptor-kodierenden Sequenzen im Genom von *Schistocerca gregaria* und *Locusta migratoria* ergaben eine beträchtliche Anzahl von orthologen Sequenzen. Das gilt insbesondere für die sogenannte „Basal" – Gruppe; sie umfasst orthologe Rezeptorpaare von *Schistocerca* und *Locusta*, die eine hohe Sequenz-Identität aufweisen. Angesichts der Frage, ob diese phylogenetisch konservierten Sequenzen eventuell Rezeptoren für Pheromone kodieren, wurde eine mögliche Co-Expression mit SNMP1 überprüft. In Doppel-Fluoreszenz in situ Hybridisierungsexperimenten hat sich gezeigt, dass einige Mitglieder aus der „Basal"-Gruppe tatsächlich mit SNMP1 co-exprimiert werden. Dabei werden sechs Rezeptor-Typen in Sensilla basiconica exprimiert; nur ein Rezeptor-Typ wird in Sensilla trichodea exprimiert, dem Sensillen-Typ, mit dem bei Motten und Fliegen die Registrierung von Sexualpheromonen erfolgt. Die Analysen von Rezeptortypen, die nicht zur „Basal"-Gruppe zählen ergaben, dass zwar die große Mehrzahl nicht in SNMP1-positiven Zellen exprimiert wird, allerdings konnte für einige Rezeptortypen eine Co-Expression mit SNMP1 nachgewiesen werden. Diese Befunde deuten darauf hin, dass Rezeptortypen aus verschiedenen Gruppen als Rezeptor-Kandidaten für Pheromone in Betracht kommen.

5 Reference

Abuin, L., Bargeton, B., Ulbrich, M. H., Isacoff, E. Y., Kellenberger, S., and Benton, R. (2011). Functional architecture of olfactory ionotropic glutamate receptors. *Neuron* 69, 44–60.

Acton, S. L., Scherer, P. E., Lodish, H. F., and Krieger, M. (1994). Expression cloning of SR-BI, a CD36-related class B scavenger receptor. *J. Biol. Chem.* 269, 21003–21009.

Altner, H., Routil, C., and Loftus, R. (1981). The structure of bimodal chemo-, thermo-, and hygroreceptive sensilla on the antenna of Locusta migratoria. *Cell Tissue Res.* 215.

Andronopoulou, E., Labropoulou, V., Douris, V., Woods, D. F., Biessmann, H., and Iatrou, K. (2006). Specific interactions among odorant-binding proteins of the African malaria vector *Anopheles gambiae*. *Insect Mol. Biol.* 15, 797–811.

Benton, R., Vannice, K. S., and Vosshall, L. B. (2007). An essential role for a CD36-related receptor in pheromone detection in *Drosophila*. *Nature* 450, 289–293.

Blaney, W. M. (1974). Electrophysiological responses of the terminal sensilla on the maxillary palps of *Locusta migratoria* (L.) to some electrolytes and non-electrolytes. *J. Exp. Biol.* 60, 275–293.

Blaney, W. M. (1975). Behavioural and electrophysiological studies of taste discrimination by the maxillary palps of larvae of *Locusta migratoria* (L.). *J. Exp. Biol.* 62, 555–69.

Blaney, W. M., and Chapman, R. F. (1969). The fine structure of the terminal sensilla on the maxillary palps of *Schistocerca gregaria* (Forskal) (Orthoptera, Acrididae). *Zeitschrift für Zellforsch. und mikroskopische Anat.* 99, 74–97.

Chen, Y., and Amrein, H. (2017). Ionotropic receptors mediate *Drosophila o*viposition preference through sour gustatory receptor neurons. *Curr. Biol.* 27, 2741–2750.e4.

Cheng, D., Lu, Y., Zeng, L., Liang, G., and He, X. (2015). Si-CSP9 regulates the integument and moulting process of larvae in the red imported fire ant, *Solenopsis invicta*. *Sci. Rep.* 5.

Cicconardi, F., Di Marino, D., Olimpieri, P. P., Arthofer, W., Schlick-Steiner, B. C., and Steiner, F. M. (2017). Chemosensory adaptations of the mountain fly *Drosophila nigrosparsa* (Insecta: Diptera) through genomics' and structural biology's lenses. *Sci. Rep.* 7, 43770.

Febbraio, M., Hajjar, D. P., and Silverstein, R. L. (2001). CD36: A class B scavenger receptor involved in angiogenesis, atherosclerosis, inflammation, and lipid metabolism. *J. Clin. Invest.* 108, 785–791.

Ferenz, H. J., and Seidelmann, K. (2003). Pheromones in relation to aggregation and reproduction in desert locusts. in *Physiological Entomology*, 11–18.

Foret, S., and Maleszka, R. (2006). 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.

Forstner, M., Gohl, T., Gondesen, I., Raming, K., Breer, H., and Krieger, J. (2008). Differential expression of SNMP-1 and SNMP-2 proteins in pheromone-sensitive hairs of moths. *Chem. Senses* 33, 291–299.

Galindo, K., and Smith, D. P. (2001). A large family of divergent *Drosophila* odorant-binding proteins expressed in gustatory and olfactory sensilla. *Genetics* 159, 1059–1072.

Gomez-Diaz, C., Bargeton, B., Abuin, L., Bukar, N., Reina, J. H., Bartoi, T., et al. (2016). A CD36 ectodomain mediates insect pheromone detection via a putative tunnelling mechanism. *Nat. Commun.* 7.

Grosse-Wilde, E., Kuebler, L. S., Bucks, S., Vogel, H., Wicher, D., and Hansson, B. S. (2011). Antennal transcriptome of *Manduca sexta*. *Proc. Natl. Acad. Sci.* 108, 7449–7454.

Guo, M., Krieger, J., Große-Wilde, E., Mißbach, C., Zhang, L., and Breer, H. (2013). Variant Ionotropic receptors are expressed in olfactory sensory neurons of coeloconic sensilla on the antenna of the desert locust (*Schistocerca gregaria*). *Int. J. Biol. Sci.* 10, 1–14.

Hansson, B. S., and Stensmyr, M. C. (2011). Evolution of insect olfaction. *Neuron* 72, 698–711.

Hassanali, A., Njagi, P. G. N., and Bashir, M. O. (2005). Chemical ecology of locust and related acridids. *Annu. Rev. Entomol.* 50, 223–245.

Jeong, Y. T., Shim, J., Oh, S. R., Yoon, H. I., Kim, C. H., Moon, S. J., et al. (2013). An odorant-binding protein required for suppression of sweet taste by bitter chemicals. *Neuron* 79, 725–737.

Jin, X., Brandazza, A., Navarrini, A., Ban, L., Zhang, S., Steinbrecht, R. A., et al. (2005). Expression and immunolocalisation of odorant-binding and chemosensory proteins in locusts. *Cell. Mol. Life Sci.* 62, 1156–1166.

Jin, X., Ha, T. S., and Smith, D. P. (2008). SNMP is a signaling component required for pheromone sensitivity in *Drosophila*. *Proc. Natl. Acad. Sci.* 105, 10996–11001.

Joseph, R. M., and Carlson, J. R. (2015). Drosophila chemoreceptors: a molecular interface between the chemical world and the brain. *Trends Genet.* 31, 683–695.

Krieger, J., Grosse-Wilde, E., Gohl, T., Dewer, Y. M. E., Raming, K., and Breer, H. (2004). Genes encoding candidate pheromone receptors in a moth (*Heliothis virescens*). *Proc. Natl. Acad. Sci.* 101, 11845–11850.

Larter, N. K., Sun, J. S., and Carlson, J. R. (2016). Organization and function of *Drosophila* odorant binding proteins. *Elife* 5.

Lartigue, A., Gruez, A., Briand, L., Blon, F., Bézirard, V., Walsh, M., et al. (2004). Sulfur single-wavelength anomalous diffraction crystal structure of a pheromone-binding protein from the honeybee *Apis mellifera* L. *J. Biol. Chem.* 279, 4459–4464.

Leal, W. S. (2013). Odorant reception in insects : roles of receptors , binding proteins , and degrading enzymes. *Annu. Rev. Entomol.* 58, 373–391.

Li, Z., Ni, J. D., Huang, J., and Montell, C. (2014). Requirement for *Drosophila* SNMP1 for rapid activation and termination of pheromone-induced activity. *PLoS Genet* 10, e1004600.

Mahamat, H., Hassanali, A., and Odongo, H. (2011). The role of different components of the pheromone emission of mature males of the desert locust, *Schistocerca gregaria* (Forskål) (Orthoptera: Acrididae) in accelerating maturation of immature adults. *Int. J. Trop. Insect Sci.* 20, 1–5.

Maida, R., Mameli, M., Müller, B., Krieger, J., and Steinbrecht, R. A. (2005). The expression pattern of four odorant-binding proteins in male and female silk moths, *Bombyx mori*. *J. Neurocytol.* 34, 149–163.

Maleszka, J., Forêt, S., Saint, R., and Maleszka, R. (2007). 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.

Mao, Y., Xu, X., Xu, W., Ishida, Y., Leal, W. S., Ames, J. B., et al. (2010). Crystal and solution structures of an odorant-binding protein from the southern house mosquito complexed with an oviposition pheromone. *Proc. Natl. Acad. Sci.* 107, 19102–19107.

Montell, C. (2009). A taste of the *Drosophila* gustatory receptors. *Curr. Opin. Neurobiol.* 19, 345–353.

Nardi, J. B., Miller, L. A., Walden, K. K. O., Rovelstad, S., Wang, L., Frye, J. C., et al. (2003). Expression patterns of odorant-binding proteins in antennae of the moth *Manduca sexta*. *Cell Tissue Res.* 313, 321–333.

Nichols, Z., and Vogt, R. G. (2008). The SNMP/CD36 gene family in Diptera, Hymenoptera and Coleoptera: *Drosophila melanogaster*, *D. pseudoobscura*, *Anopheles gambiae*, *Aedes aegypti, Apis mellifera*, and *Tribolium castaneum*. *Insect Biochem. Mol. Biol.* 38, 398–415.

Nomura, A., Kawasaki, K., Kubo, T., and Natori, S. (1992). Purification and localization of p10, a novel protein that increases in nymphal regenerating legs of *Periplaneta americana* (American cockroach). *Int. J. Dev. Biol.* 36, 391–398.

Ochieng, S. A., and Hansson, B. S. (1999). Responses of olfactory receptor neurones to behaviourally important odours in gregarious and solitarious desert locust, *Schistocerca gregaria*. *Physiol. Entomol.* 24, 28–36.

Ochieng, S. A., Hallberg, E., and Hansson, B. S. (1998). Fine structure and distribution of antennal sensilla of the desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae). *Cell Tissue Res.* 291, 525–536.

Pelosi, P., Iovinella, I., Felicioli, A., and Dani, F. R. (2014). Soluble proteins of chemical communication: an overview across arthropods. *Front. Physiol.* 5 AUG.

Pelosi, P., Iovinella, I., Zhu, J., Wang, G., and Dani, F. R. (2017). Beyond chemoreception: diverse tasks of soluble olfactory proteins in insects. *Biol. Rev.*

Pelosi, P., Zhou, J. J., Ban, L. P., and Calvello, M. (2006). Soluble proteins in insect chemical communication. *Cell. Mol. Life Sci.* 63, 1658–1676.

Pener, M. P., and Simpson, S. J. (2009). Locust phase polyphenism: an update. *Adv. In Insect Phys.* 36, 1–272.

Policy, S. D., and Conway, D. J. (2001). Strong diversifying selection on domains of the Plasmodium falciparum apical membrane antigen 1 gene. *Genetics* 158, 1505–1512.

Pophof, B. (1997). Olfactory responses recorded from sensilla coeloconica of the

silkmoth *Bombyx mori*. *Physiol Entomol* 22, 239–248.
Pregitzer, P., Jiang, X., Grosse-Wilde, E., Breer, H., Krieger, J., and Fleischer, J. (2017). In search for pheromone receptors: Certain members of the odorant receptor family in the desert locust *Schistocerca gregaria* (orthoptera: Acrididae) are co-expressed with SNMP1. *Int. J. Biol. Sci.* 13.
Qiao, H., He, X., Schymura, D., Ban, L., Field, L., Dani, F. R., et al. (2011). Cooperative interactions between odorant-binding proteins of *Anopheles gambiae*. *Cell. Mol. Life Sci.* 68, 1799–1813.
Rimal, S., and Lee, Y. (2018). The multidimensional ionotropic receptors of *Drosophila melanogaster*. *Insect Mol. Biol.* 27, 1–7.
Rogers, M. E., Sun, M., Lerner, M. R., and Vogt, R. G. (1997). Snmp-1, a novel membrane protein of olfactory neurons of the silk moth *Antheraea polyphemus* with homology to the CD36 family of membrane proteins. *J. Biol. Chem.* 272, 14792–14799.
Rytz, R., Croset, V., and Benton, R. (2013). Ionotropic receptors (IRs): chemosensory ionotropic glutamate receptors in *Drosophila* and beyond. *Insect Biochem. Mol. Biol.* 43, 888–897.
Saini, R. K., Rai, M. M., Hassanali, A., Wawiye, J., and Odongo, H. (1995). Semiochemicals from froth of egg pods attract ovipositing female *Schistocerca gregaria*. *J. Insect Physiol.* 41, 711–716.
Sandler, B. H., Nikonova, L., Leal, W. S., and Clardy, J. (2000). Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. *Chem. Biol.* 7, 143–151.
Sawyer, S. L., Wu, L. I., Emerman, M., and Malik, H. S. (2005). Positive selection of primate TRIM5 identifies a critical species-specific retroviral restriction domain. *Proc. Natl. Acad. Sci.* 102, 2832–2837.
Scott, K. (2018). Gustatory processing in *Drosophila melanogaster*. *Annu. Rev. Entomol.* 63, 15–30.
Seidelmann, K., and Ferenz, H. J. (2002). Courtship inhibition pheromone in desert locusts, *Schistocerca gregaria*. *J. Insect Physiol.* 48, 991–996.
Seidelmann, K., Warnstorff, K., and Ferenz, H. J. (2005). Phenylacetonitrile is a male specific repellent in gregarious desert locusts, *Schistocerca gregaria*. *Chemoecology* 15, 37–43.
Silverstein, R. L., Li, W., Park, Y. M., and Rahaman, S. O. (2010). Mechanisms of cell signaling by the scavenger receptor CD36: implications in atherosclerosis and thrombosis. *Trans. Am. Clin. Climatol. Assoc.* 121, 206–20.
Steinbrecht, R. A. (1996). Structure and function of insect olfactory sensilla. *Ciba Found. Symp.* 200, 158–174.
Suh, E., Bohbot, J. D., and Zwiebel, L. J. (2014). Peripheral olfactory signaling in insects. *Curr. Opin. Insect Sci.* 6, 86–92.
Sun, X., Zeng, F. F., Yan, M. J., Zhang, A., Lu, Z. X., and Wang, M. Q. (2016). Interactions of two odorant-binding proteins influence insect chemoreception. *Insect Mol. Biol.* 25, 712–723.
Tegoni, M., Campanacci, V., and Cambillau, C. (2004). Structural aspects of sexual attraction and chemical communication in insects. *Trends Biochem. Sci.* 29,

257–264.
Theißen, G. (2002). Orthology: secret life of genes. *Nature* 415, 741.
Vieira, F. G., and Rozas, J. (2011). Comparative genomics of the odorant-binding and chemosensory protein gene families across the arthropoda: Origin and evolutionary history of the chemosensory system. *Genome Biol. Evol.* 3, 476–490.
Vieira, F. G., Sánchez-Gracia, A., and Rozas, J. (2007). Comparative genomic analysis of the odorant-binding protein family in 12 *Drosophila* genomes: purifying selection and birth-and-death evolution. *Genome Biol.* 8, R235.
Vogt, R. G., Callahan, F. E., Rogers, M. E., and Dickens, J. C. (1999). Odorant binding protein diversity and distribution among the insect orders, as indicated by LAP, an OBP-related protein of the true bug *Lygus lineolaris* (Hemiptera, heteroptera). *Chem. Senses* 24, 481–495.
Vogt, R. G., Große-Wilde, E., and Zhou, J. J. (2015). 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.
Vogt, R. G., Miller, N. E., Litvack, R., Fandino, R. A., Sparks, J., Staples, J., et al. (2009). The insect SNMP gene family. *Insect Biochem. Mol. Biol.* 39, 448–456.
Vogt, R. G., and Riddiford, L. M. (1981). Pheromone binding and inactivation by moth antennae. *Nature* 293, 161–163.
Wang, X., Fang, X., Yang, P., Jiang, X., Jiang, F., Zhao, D., et al. (2014). The locust genome provides insight into swarm formation and long-distance flight. *Nat. Commun.* 5.
Wang, Z., Yang, P., Chen, D., Jiang, F., Li, Y., Wang, X., et al. (2015). Identification and functional analysis of olfactory receptor family reveal unusual characteristics of the olfactory system in the migratory locust. *Cell. Mol. Life Sci.* 72, 4429–4443.
Wogulis, M., Morgan, T., Ishida, Y., Leal, W. S., and Wilson, D. K. (2006). The crystal structure of an odorant binding protein from *Anopheles gambiae*: Evidence for a common ligand release mechanism. *Biochem. Biophys. Res. Commun.* 339, 157–164.
Xu, P. X., Zwiebel, L. J., and Smith, D. P. (2003). Identification of a distinct family of genes encoding atypical odorant-binding proteins in the malaria vector mosquito, *Anopheles gambiae*. *Insect Mol. Biol.* 12, 549–60.
Yao, C. A. (2005). Chemosensory coding by neurons in the coeloconic sensilla of the *Drosophila a*ntenna. *J. Neurosci.* 25, 8359–8367.
Yu, F., Zhang, S., Zhang, L., and Pelosi, P. (2009). Intriguing similarities between two novel odorant-binding proteins of locusts. *Biochem. Biophys. Res. Commun.* 385, 369–374.
Zhang, L., Li, H., and Zhang, L. (2017). Two olfactory pathways to detect aldehydes on locust mouthpart. *Int. J. Biol. Sci.* 13, 759–771.
Zheng, J., Li, J., Han, L., Wang, Y., Wu, W., Qi, X., et al. (2015). Crystal structure of the *Locusta migratoria* odorant binding protein. *Biochem. Biophys. Res. Commun.* 456, 737–742.
Zhou, J. J., Huang, W., Zhang, G. A., Pickett, J. A., and Field, L. M. (2004). 'Plus-C'

odorant-binding protein genes in two Drosophila species and the malaria mosquito *Anopheles gambiae*. *Gene* 327, 117–129.

Zhou, J. J., Vieira, F. G., He, X. L., Smadja, C., Liu, R., Rozas, J., et al. (2010). Genome annotation and comparative analyses of the odorant-binding proteins and chemosensory proteins in the pea aphid *Acyrthosiphon pisum*. *Insect Mol. Biol.* 19, 113–122.

Zhou, S. H., Zhang, S. G., and Zhang, L. (2009). The chemosensilla on tarsi of *Locusta migratoria* (Orthoptera: Acrididae): distribution, ultrastructure, expression of chemosensory proteins. *J. Morphol.* 270, 1356–1363.

6 Publication and author contribution

Publication 1

Jiang, X., Pregitzer, P., Grosse-Wilde, E., Breer, H., and Krieger, J. Identification and characterization of two “sensory neuron membrane proteins” (SNMPs) of the desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae). Journal of Insect Science, 2016, 16(1): 33; 1-10.

As the first author, ***Xingcong Jiang*** *contributed to develop the project concept, performed most of the experiments and acquired the data; he participated in the data interpretation and in writing the manuscript.*

Publication 2

Pregitzer, P., **Jiang, X.**, Grosse-Wilde, E., Breer, H., Krieger, J., and Fleischer, J. In search for pheromone receptors: certain members of the odorant receptor family in the desert locust *Schistocerca gregaria* (Orthoptera: Acrididae) are co-expressed with SNMP1. International journal of biological sciences, 2017, 13(7); 911-922.

As one of the co-authors, ***Xingcong Jiang*** *contributed in designing the project concept, as well as in performing the experiments and data acquisition.*

Publication 3

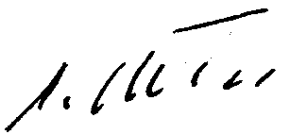
Jiang, X., Krieger, J., Breer, H., and Pregitzer, P. Distinct subfamilies of odorant binding proteins in locust (Orthoptera, Acrididae): molecular evolution, structural variation, and sensilla-specific expression. Frontiers in Physiology, 2017, 8: 734.

As the first author, ***Xingcong Jiang*** *contributed to develop the project concept, performed most of the experiments and acquired the data; he participated in the data interpretation and in writing the manuscript.*

Publication 4

Jiang, X., Ryl, M., Krieger, J., Breer, H., and Pregitzer, P. Odorant binding proteins of the desert locust *Schistocerca gregaria* (Orthoptera, Acrididae): topographic expression patterns in the antennae. Frontiers in Physiology, 2018, 9: 417.

As the first author, ***Xingcong Jiang*** *contributed to develop the project concept, performed most of the experiments and acquired the data; he participated in the data interpretation and in writing the manuscript.*



Signature of the supervisor

7 Curriculum Vita

Personal Information

Name: Xingcong Jiang

Date of Bith: 10.02.1989

Place of Birth: Heze, China

Nationality: China P. R.

Education

PhD	**Since 09.2014** University of Hohenheim Research topic: Gene families involved in odorant perception of the desert locust *Schistocerca gregaria*
Master	**09.2011 — 06.2014** China Agricultural University Major: agricultural pests and biological control
Bachelor	**09.2007 — 06.2011** Shandong Agricultural University Major: Plant and animal protection and plant quarantine

Xingcong Jiang

8 Acknowledgment

Approaching the closing mark of my PhD career, I would like to send my sincere gratitude to Prof. Heinz Breer and all the members in the lab for their receiving me as an integrate part of the group.

Prof. Heinz Breer is more than a supervisor as he appears; actually every progress I made in both science and life is promoted by his suggestions, criticisms, appreciations and encouragements, without which there would have been much more unimaginable difficulties during my PhD career. Those moments when the “face-to-face” communication occurred in the mensa, in the seminar room and in his office as well have been engraved by my mind, and bearing all the gains from these valuable experiences I do have evolved a better man than I could have become.

All the members in the “insect group” receive my sincere thankfulness. Prof. Jürgen Krieger is acknowledged for his constructive suggestions and excellent scientific experience which he willingly donate to prompt the proceedings of my research. Dr. Pablo Pregitzer behaves like my older brother who protects me from many troubles and offers enormous useful suggestions for my conundrums about science and life. Heidrun Froß is acknowledged for her excellent assistant of my research without whom my research would be heavily delayed. Sincere gratitude is given to the staff of “mice group”, as well as to the secretaries, the technicians and the helping students in the lab for offering consistent high-quality of research assistance.

I would like to show my thankfulness to all my Chinese friends in the University of Hohenheim who have offered me tremendous help, and have created many unforgettable happy moments over the past years.

All my meagre steps towards my goal are tightly bound with my families in China e.g. my mom, my brother and all my relatives who remotely exerted their spiritual powers and willingly resolved all my financial difficulties. I attribute the ultimate triumph of this “PhD war” to my families’ reliable backup. Meanwhile, I must thank any anonymous god of love for the fateful knots tying up my girlfriend Ruo Sun and me, and saving me from stumbling alone.

Last but not least, I do wish my father perceived all the fabulous achievements made by his son and the reputation won for the whole family, I do hope the preserved home for him in my heart immortalized him forever, and I do believe in peace he settled witnessing every flour flourishing as it should be.

Upon being encoded in China, I did resolve that I should receive post-transcriptional modifications in German. With these done, the migration back to China in a mature and functional status is programmed as my fate, where I believe my presence will make a difference.

9 Attachments

Journal of Insect Science (*2016*) 16(1): 33; 1–10
doi: 10.1093/jisesa/iew015
Research article

Identification and Characterization of Two "Sensory Neuron Membrane Proteins" (SNMPs) of the Desert Locust, *Schistocerca gregaria* (Orthoptera: Acrididae)

Xingcong Jiang,[1] Pablo Pregitzer,[1] Ewald Grosse-Wilde,[2] Heinz Breer,[1] and Jürgen Krieger[1,3,4]

[1]University of Hohenheim, Institute of Physiology, 70593 Stuttgart, Germany (jiangxingcong@126.com; p_pregitzer@uni-hohenheim.de; breer@uni-hohenheim.de; juergen.krieger@zoologie.uni-halle.de), [2]Department of Evolutionary Neuroethology, Max Planck Institute for Chemical Ecology, 07745 Jena, Germany (grosse-wilde@ice.mpg.de), [3]Department of Animal Physiology, Martin Luther University Halle-Wittenberg, Institute of Biology/Zoology, 06099 Halle, Germany, and [4]Corresponding author, e-mail: juergen.krieger@zoologie.uni-halle.de

Subject Editor: Dr. Russell A Jurenka

Received 22 December 2015; Accepted 4 February 2016

Abstract

Pheromone-responsive neurons of insects not only require specific receptors but in addition several auxiliary components, including the "sensory neuron membrane protein," SNMP. Accordingly, SNMP is considered as a marker for neurons responding to pheromones. For the desert locust *Schistocerca gregaria*, it is known that the behavior, including aggregation behavior and courtship inhibition, is largely controlled by pheromones. However, little is known about pheromones, their receptors, and the pheromone-responsive cells in locusts. In this study, we have identified two SNMP subtypes, SNMP1 and SNMP2, and compared their phylogenetic relationship and primary structure motifs with SNMPs from other species. Both SNMPs were found in chemosensory tissues, especially the antennae. Employing double in situ hybridization, we identified and localized the SNMP-expressing cells in the antennae. Cells expressing SNMP1 were localized to sensilla trichodea but also to sensilla basiconica, which in locust respond to pheromones. One or a few cells express SNMP1 within the multineuron clusters from sensilla basiconica, whereas the SNMP2 subtype was expressed in cells surrounding the neuron clusters, possibly supporting cells. Based on the finding that SNMP1 is expressed in distinct neurons under chemosensory sensilla, it is conceivable that these cells may represent pheromone-responsive neurons of the desert locust.

Key words: desert locust, olfaction, antenna, sensory neuron membrane protein, in situ hybridization

In order to perceive olfactory cues, insects have evolved hair-like appendages, called sensilla, over the surface of the olfactory organs. The task to sense pheromone compounds is assigned to pheromone-responsive sensilla (Almaas and Mustaparta 1991, Baker et al. 2004, Hansson and Stensmyr 2011). Pheromone-responsive sensilla from moth and fly species are innervated by one to three olfactory sensory neurons (OSNs), which project their dendrites to the lymph cavity of the sensillum (Keil and Steinbrecht 1984, Steinbrecht and Gnatzy, 1984, Clyne et al. 1997). The response of the sensory neurons to volatile compounds is determined by odorant receptors (ORs; Hallem and Carlson 2006, Touhara and Vosshall 2009, Carey et al. 2010, Wang et al. 2010) residing on the dendritic membrane of the OSN (Elmore and Smith 2001, Gohl and Krieger 2006). Receptors for pheromones are members of the OR family, but they are specifically tuned to pheromonal compounds (Große-Wilde et al. 2007, Touhara and Vosshall 2009, Sakurai et al. 2015). Thus, the specific response of sensory neurons to distinct pheromones is mediated by the receptor type, such as BmOR1 for bombykol in *Bombyx mori* and DmelOR67d for 11-cis-vaccenyl acetate in *Drosophila melanogaster* (Sakurai et al. 2004, Krieger et al. 2005, Kurtovic et al. 2007). In addition to the specific OR-type, each OSN expresses a coreceptor, called Orco that is highly conserved among insect species (Krieger et al. 2003, Benton et al. 2006, Vosshall and Hansson 2011). For an adequate response to pheromones, sensory neurons not only require the specific receptor type, but as an additional molecular element the "sensory neuron membrane protein," SNMP (Benton et al. 2007, Jin et al. 2008, Li et al. 2014). SNMP is a transmembrane protein with some homology to the mammalian CD36 receptor family (Rogers et al. 1997, Nichols and Vogt 2008). Recent studies have shown that the SNMP subtype SNMP1 is coexpressed with pheromone receptors in pheromone-responsive neurons (Benton et al. 2007, Forstner et al. 2008). Thus, although the functional implications of SNMP1 in pheromone sensing are still controversial (Jin et al. 2008, Vogt et al. 2009, Li et al. 2014), SNMP1 expression seems to be an indicator for pheromone-responsive

neurons. In contrast, the expression of the SNMP2 subtype in moth is confined to supporting cells (Forstner et al. 2008, Gu et al. 2013, Liu et al. 2013, Zhang et al. 2015).

The endeavors to decipher the mechanism underlying insect pheromone reception and transduction have mainly been concentrated on two taxa, the Lepidoptera and the Diptera. The desert locust (*Schistocerca gregaria*) is a member of the taxa Orthoptera that on the phylogenetic scale represents a remote lineage relative to Lepidoptera and Diptera (Grimaldi and Engel 2005, Song et al. 2015). Moreover, the Orthoptera undergo a hemimetabolic developmental process that substantially differs from that of holometabolous insects, which were hitherto mainly used in research studying SNMP-proteins (Nichols and Vogt 2008). Locusts are unique by their behavioral plasticity to switch between a solitary and a gregarious phase (Uvarov 1966, Hassanali et al. 2004, Simpson and Sword, 2008, Wang et al. 2014), and there is evidence indicating that locust body volatiles may act as aggregation pheromones, which are involved in shaping and maintaining the gregarious phase (Chapman 1990, Byers 1991, Heifetz et al. 1996). In addition, pheromones mediating courtship–inhibition have been observed in *Schistocerca gregaria* (Seidelmann and Ferenz 2002, Seidelmann et al. 2003) and the related species *Schistocerca americana* (Stahr et al. 2013). Thus, locust antennae are supposed to have pheromone-responsive neurons. In this context, it appeared to be of particular interest to explore whether SNMPs are in fact expressed in desert locust and how SNMPs of Orthoptera may relate to SNMPs of other species with respect to structural features and expression patterns.

Materials and Methods

Animals and Tissue Collection

Schistocerca gregaria were purchased from Bugs-International GmbH (Irsingen/Unterfeld, Germany). Antennae of locust nymphs (stages 1–5), adult male antennae, adult female antennae, adult tarsi, and adult maxillary palps were dissected using autoclaved surgical scissors and were immediately frozen in liquid nitrogen. Tissues were stored at −70°C until RNA extraction.

Reverse transcription PCR (RT-PCR)

Total RNA was isolated from frozen tissues using TRIzol reagent (Invitrogen) following the manufactures protocol. For cDNA synthesis, 1 µg of total RNA was used for reverse transcription in a total volume of 20 µl employing SuperScript™ III Reverse Transcriptase (Invitrogen). Alternatively, poly $(A)^+$ RNA was purified from 100 µg of total RNA using oligo $(dT)_{25}$ magnetic dynabeads (Invitrogen) following the recommendation of the supplier and reverse-transcribed using SuperScript™ III Reverse Transcriptase. Nonquantitative RT-PCRs were employed to determine expression patterns using gene specific primers. A primer pair matching the actin gene (AEV89776) of *S. gregaria* was used to test the integrity of the cDNA preparations.

Primer pairs were the following:

SgreSNMP1 sense: 5′ CATCCAGAACATCGACGACCT 3′
SgreSNMP1 antisense: 5′ GGAGTGTCCAGGGCTAGTATCTG 3′
SgreSNMP2 sense: 5′ CTACGCTCTTCCCACCTTTCA 3′
SgreSNMP2 antisense: 5′ ACAGCGGTCCCCCGATGATTA 3′
Actin sense: 5′ AACTGGCTTGCTGCATCCTC 3′
Actin antisense: 5′ CACATCTGCTGGAAGGTGGA 3′

PCR conditions used in RT-PCR experiments were: 94°C for 1 min 40 s, then 20 cycles with 94°C for 30 s, 65°C for 30 s and 72°C for 2 min, with a reduction in the annealing temperature by 0.5°C per cycle. This was followed by 20 further cycles at the condition of the last cycling step (annealing temperature = 55°C) and a final extension step for 7 min at 72°C. PCR products were analyzed on 1.5% agarose gels and visualized by ethidium bromide staining. PCR products were purified using the Geneclean II Kit (MP Biomedicals, LLC, IIIkrich, France), cloned into the pGEM-T plasmid (Promega, Madison, USA) and sequenced using vector or gene specific primers.

To verify that the bands amplified by the SNMP1- and SNMP2-specific primers were not resulting from amplification of genomic DNA, control experiments were conducted with templates representing cDNA syntheses reactions where the reverse transcriptase have been omitted (-RT) (Supp. Fig. 1 [online only]).

Identification and Analysis of *S. gregaria* SNMP1 and SNMP2 Sequences

Previously reported insect SNMP genes (Forstner et al. 2008, Vogt et al. 2009, Gu et al. 2013, Liu et al. 2013) were used in a tBLASTx search (cut-off value 10^{-10}) as queries to identify putative SNMP genes in a nonpublished *S. gregaria* antennal transcriptome database. The identified putative *S. gregaria* SNMP genes were then utilized as queries to identify additional SNMP genes in the antennal transcriptome using tBLASTx and BLASTp; iterative BLAST searches were completed till no more candidates were identified. Verification of the BLAST outcomes as candidate SNMP sequences was further confirmed via BLASTx searches in the NCBI nonredundant protein database.

Phylogenetic Analysis

The sequence alignment was created and processed using the programs Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/; accessed February 2016) and BioEdit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html; accessed February 2016). Transmembrane domain prediction was performed using the TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0/; accessed February 2016). An unrooted neighbour-joining tree was constructed based on a Clustal alignment using the MEGA6 software (Tamura et al. 2013) in accordance with default settings.

In Situ Hybridization

RNA in situ hybridization probes were generated as described previously (Guo et al. 2014; Yang et al. 2012). Linearized pGEM-T vectors carrying SNMP coding sequences representing the c-terminus (902bp, SNMP1; 848bp, SNMP2) were utilized to synthesize digoxigenin- (DIG) and biotin- (BIO) labeled anti-sense and sense RNA probes using the T7/SP6 RNA transcription system (Roche, Germany). The labeled riboprobe for *Schistocerca gregaria* Orco (SgreOrco) was prepared as described previously (Yang et al. 2012).

Antennae were dissected and embedded in Tissue-Tek O.C.T. Compound (Sakura Finetek Europe, The Netherlands), then 12 µm sections were thaw mounted on SuperFrost Plus slides (Menzel-Gläser, Braunschweig, Germany) at −21°C (Jung CM300 cryostat). Sections were stored at −70°C temporarily. RNA in situ hybridization was conducted as previously reported (Yang et al. 2012, Guo et al. 2014) with the following modifications. Section were fixed (4% paraformaldehyde in 0.1 M $NaHCO_3$, pH 9.5) at 4°C for 22 min. The following steps (at room temperature) were a wash for 1 min in PBS (phosphate buffered saline = 0.85% NaCl, 1.4 mM KH_2PO_4, 8 mM Na_2HPO_4, pH 7.1), incubation for 10 min in 0.2 M HCl and another wash for 1 min in PBS. After incubation for 10 min in acetylation solution (25% acetic anhydride freshly added in 0.1 M triethanolamine), sections were washed three times in PBS (3 min

each). 100 µl hybridization solution containing the labeled RNA in hybridization buffer (Yang et al. 2012) was placed onto the tissue section. A coverslip was placed on top and slides were incubated in a moister box at 60°C overnight (at least 18h). After hybridization, slides were washed twice for 30 min in 0.1x SSC at 60°C, then each slide was treated with 1 ml 1% blocking reagent (Roche) for 40 min at room temperature.

In single and double fluorescent in situ hybridization (FISH), visualization of hybridized probes was performed by using an anti-Dig AP-conjugated antibody in combination with HNPP/Fast Red (Roche) for DIG-labeled probes and an anti-biotin streptavidin horse radish peroxidase-conjugate together with fluorescein-tyramides as substrate (TSA kit, Perkin Elmer, MA) for biotin-labeled probes.

Sections from FISH experiments were analyzed with a Zeiss LSM510 Meta laser scanning microscope (Zeiss, Oberkochen, Germany). Confocal images stacks were processed by ZEN 2009 software. For an optimal view, otherwise non-altered images were adjusted in brightness and contrast and arranged using PowerPoint (Microsoft).

Results

Identification of Two SNMPs in *S. gregaria*

Based on the sequences of SNMPs from several insect species, we have performed iterative tBLASTx searches in a nonpublished *Schistocerca gregaria* antennal transcriptome database as well as tBLASTn exploitation against the NCBI database. These approaches have led to an initial identification of several sequences related to scavenger receptors; two of the sequences were classified into the CD36 superfamily, a subfamily of scavenger receptor B (Acton et al. 1994, Febbraio et al. 2001, Silverstein et al. 2010) and thereby were considered as candidate SNMP homologous. To verify the assumed classification, we initiated a phylogenetic relationship analysis of the two sequences with other SNMPs (Fig. 1). Given that SNMPs are divided into two subgroups (Vogt et al. 2009), we constructed an unrooted Neighbor-Joining tree (Fig. 1) based on 30 SNMP sequences (20 from SNMP1 and 10 from SNNP2) from six insect taxa. All the sequences were classified clearly into two orthologous groups and the two candidate SNMP sequences from *Schistocerca gregaria* were clustered into the SNMP1 and the SNMP2 subgroup, respectively. Within orthologous groups, members were arrayed into monophyletic lineages.

The transcriptome data suggested intact open reading frames for the newly identified SgreSNMP1 (1.5 kb) and SgreSNMP2 (1.6 kb), encoding proteins of 513 and 551 amino acids (aa), respectively. In order to gain more valid sequence information, we have conducted a sequence alignment analysis (Fig. 2). We selected 10 SNMP sequences from Lepidoptera, Diptera, Coleoptera and Hymenoptera insect species for comparison with SgreSNMP1 and SgreSNMP2 (Orthoptera). Homologous comparison revealed that members of SNMP1 subgroup are more conserved than those of the SNMP2 subgroup. SgreSNMP1 exhibited an overall amino acid identity of 40% with other SNMP1 members whereas around 30% identity was observed between SgreSNMP2 and other SNMP2 sequences. The sequence identity of 33% between SgreSNMP1 and SgreSNMP2 is in accordance with the overall identity between SNMP1 and SNMP2 proteins in general.

Utilizing the transmembrane topology prediction software TMHMM (Krogh et al. 2001), we found two transmembrane helixes in proximity to the N- and C-terminal regions. The predicted protein structure indicates a large extracellular loop in the SgreSNMPs (SNMP1, 426aa; SNNP2, 431aa). Earlier studies on SNMPs from other species have identified two domains with relatively high sequence conservation in the ectodomain of SNMP: the first motif extends 44 amino acids approximating the N-terminal

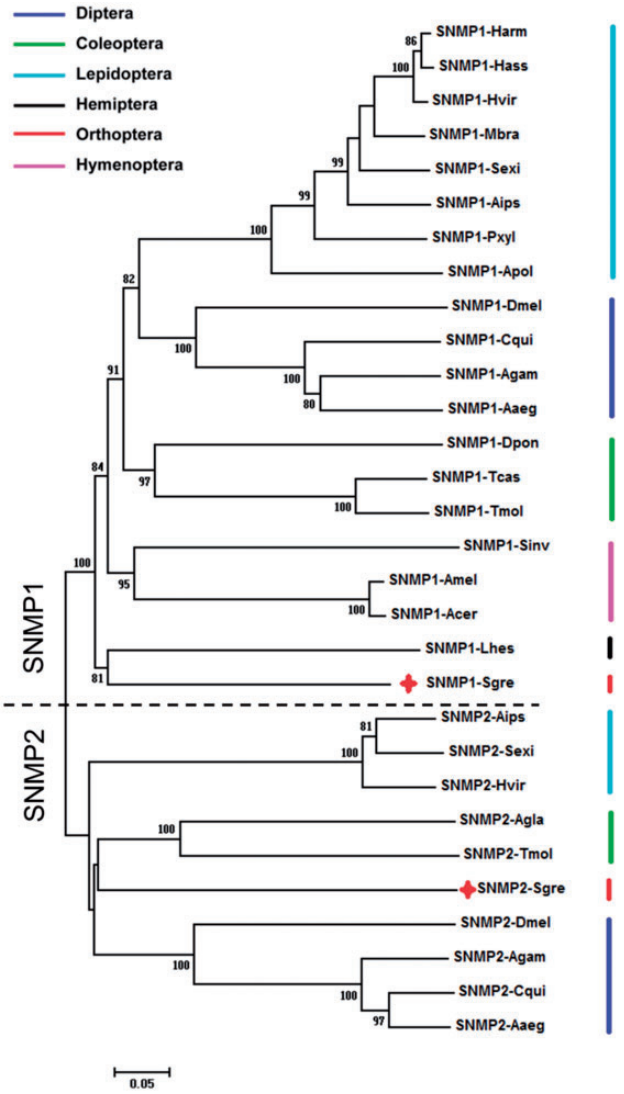


Fig. 1. Phylogenetic relationship of two putative *Schistocerca gregaria* SNMPs and 28 SNMPs from five other insect taxa. A neighbor joining tree was built using MEGA6 based on a Clustal Omega alignment of amino acid sequences deposited in Genbank. Bootstrap support values were based on 1,000 replicates and only values above 80% are shown. Branch lengths are proportional. Abbreviation and accession numbers: Aips, *Agrotis ipsilon* (SNMP1: AGF87119.1; SNMP2: AGF87120.1); Apol, *Antheraea polyphemus* (AAC47540.1); Agam, *Anopheles gambiae* (SNMP1: Q7QC49.3; SNMP2: Q7Q6R1.5); Aaeg, *Aedes aegypti* (SNMP1: Q17A88.2; SNMP2: C3U0S3.3); Amel, *Apis mellifera* (P86905.1); Acer, *Apis cerana cerana* (AGC91908.1); Agla, *Anoplophora glabripennis* (JAB63926.1); Cqui, *Culex quinquefasciatus* (SNMP1: EDS40329.1; SNMP2: AEK32389.1); Dmel, *Drosophila melanogaster* (SNMP1: AAF55863.2; SNMP2: E1JI63.1); Dpon, *Dendroctonus ponderosae* (AFI45067.1); Harm, *Helicoverpa armigera* (AAO15604.1); Hass, *Helicoverpa assulta* (ACC61201.1); Hvir, *Heliothis virescens* (SNMP1: Q9U1G3.1; SNMP2: B2RFN2.1); Lhes, *Lygus hesperus* (JAG22531.1); Mbra, *Mamestra brassicae* (AAO15603.1); Pxyl, *Plutella xylostella* (ADK66278.1); Sexi, *Spodoptera exigua* (SNMP1: AGN52676.1; SNMP2: AGN52677.1); Sinv, *Solenopsis invicta* (XP_011159295.1); Tcas, *Tribolium castaneum* (EFA02899.1); Tmol, *Tenebrio molitor* (SNMP1: AJO62245.1; SNMP2: AJO62246.1).

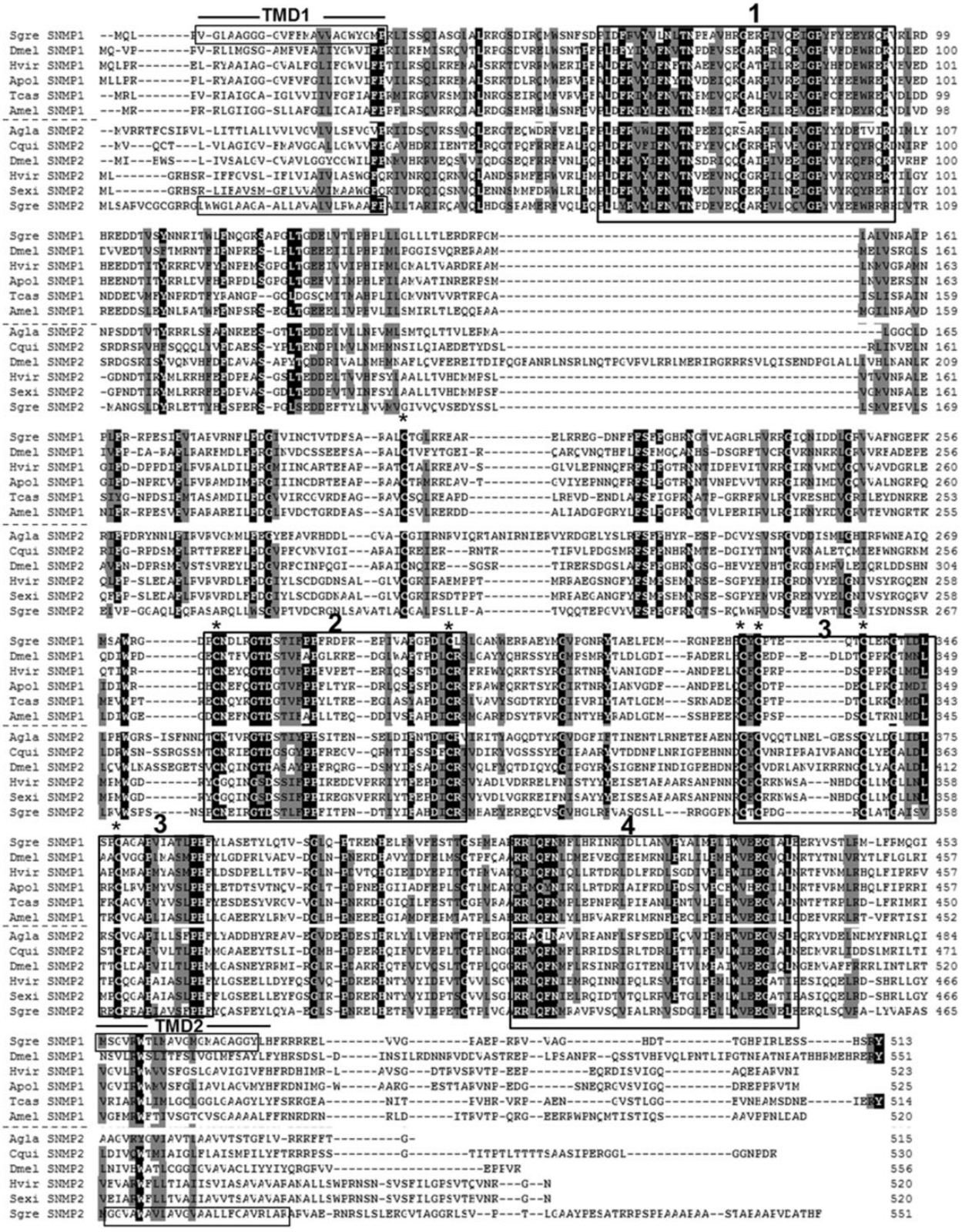


Fig. 2. Alignment of two SgreSNMP amino acid sequences with selected SNMPs from other insect species. Numbers on the right side indicate the number of last residue in a line of SNMP sequence. Residues that share more than 80% amino acid similarity are shaded in grey and identical residues are shaded in black. Predicted SgreSNMP transmembrane domains (TMD) are shown in black box. Asterisks above amino acid refer to conserved cysteine residues. Box areas below the number 1-4 propose four well conserved motifs exhibiting more than 40% sequence similarity.

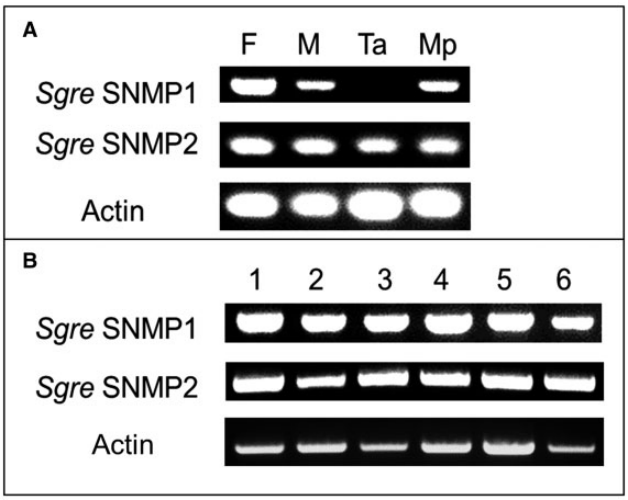


Fig. 3. Spatial and developmental expression pattern of SgreSNMP1 and SgreSNMP2. RT-PCRs were conducted with cDNA prepared from indicated tissues (A) as well as antennae from different growth stages (B). Expression of actin (AEV89776) is used as a control to reflect cDNA integrity. F, female antenna; M, male antenna; Ta, tarsus; Mp, maxillary palp; 1–5, locust nymphal stages; 6, adult locust.

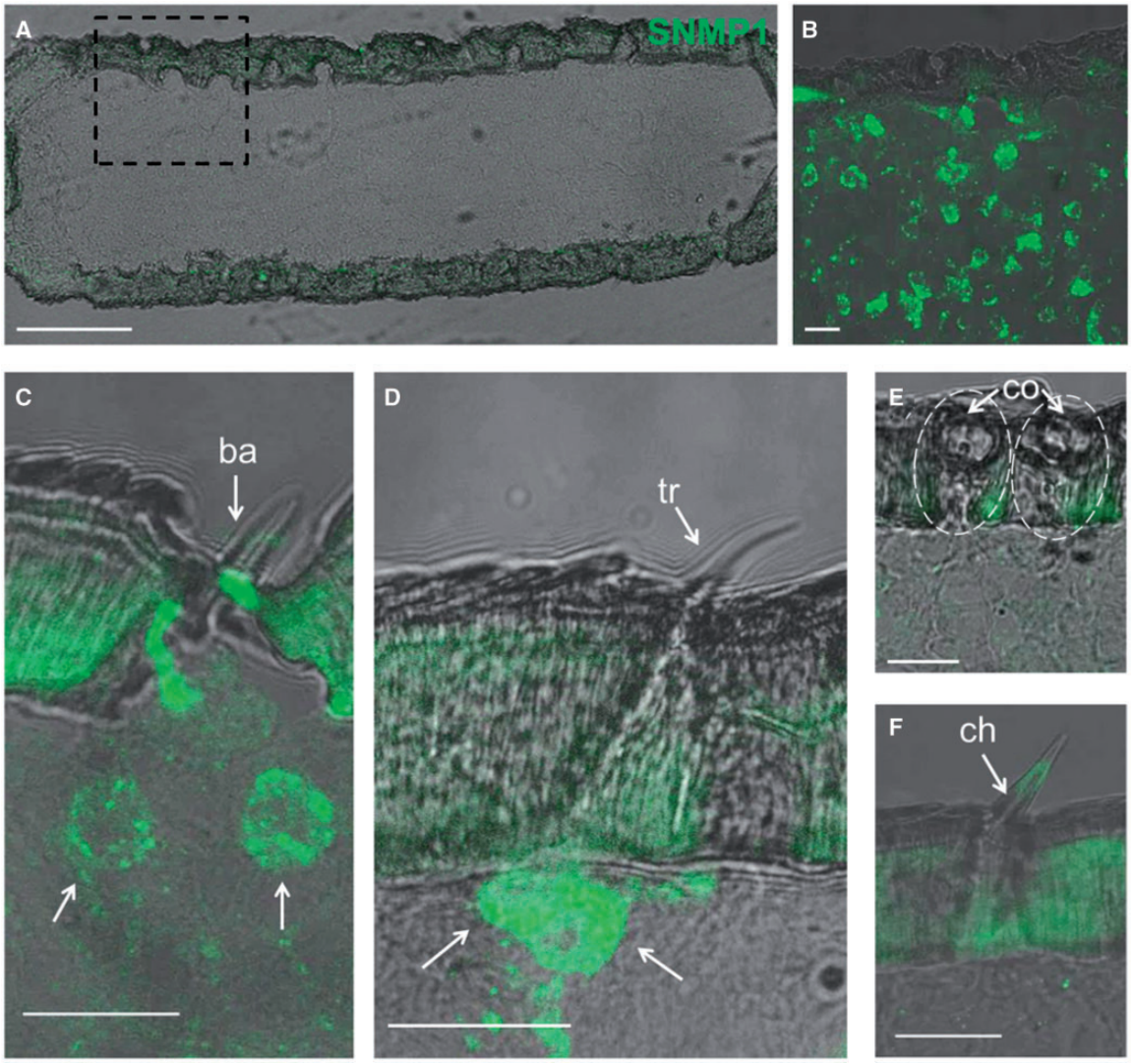


Fig. 4. Expression of SgreSNMP1 is restricted to two sensillum types on the male antenna. FISH was performed by employing a BIO-labeled probe for SgreSNMP1. (A) A section through an antennal segment at lower magnification (10×). (B) Amplified area of the dashed box in (A): numerous SgreSNMP1 expressing cells are visible. (C–F) Localization of SgreSNMP1 expressing cells in relation to four types of sensory hairs. SgreSNMP1 labeled cells (white arrows) are visible beneath a sensillum basiconicum (ba, in C) and a sensillum trichodeum (tr, in D), however, are absent from sensilla coeloconica (co, dashed area in E) and the sensillum chaeticum (ch, in F). Scale bar: A, 100 μm; B–F, 20 μm.

and the second motif contains 92 amino acids which are close to the C-terminal. For the SgreSNMPs, we identified four domains that are well conserved (over 40% sequence identity with other SNMPs); the motifs 1-3 covered the reported two conserved domains (Rogers et al. 1997); a fourth conserved motif was identified near the C-terminal following the third motif (Fig. 2). Moreover, the sequence alignment study unraveled seven conserved cysteine residues (asterisk in Fig. 2), which are located in the ectodomain (residues 190-350 in *S. gregaria*).

Expression of *S. gregaria* SNMPs

The expression profile of *SgreSNMP* genes was studied by nonquantitative RT-PCR analysis comparing several selected tissues including chemosensory organs, such as antennae, maxillary palps, and tarsi from adult locust. We found gene products of *SNMP1* and *SNMP2* in both female and male antenna, the major olfactory organ. Expression of SgreSNMP1 and SgreSNMP2 was also detected in maxillary palp, another important chemosensory organ. Using cDNA from tarsi only SgreSNMP2 was successfully amplified (Fig. 3A). Comparing antennae from different developmental stages, we found that SgreSNMP1 as well as SgreSNMP2 were expressed throughout development from the nymphal (1st to 5th) to the adult stages, with no observable variation in the expression levels (Fig. 3B).

SgreSNMP1 in Different Sensilla Types on the Antenna

To visualize where the *SNMP* genes were expressed in adult *S. gregaria* antenna, we generated riboprobes labeled with either digoxigenin (DIG) or biotin (BIO) and conducted FISH on tissue sections throughout the antenna. The specificity of the FISH results with the antisense probes was ascertained by the observation that there was no labeling in control experiments using corresponding sense probes (data not shown). Adult *S. gregaria* antennae comprise approximately 24 flagellar segments. Figure 4A depicts a longitudinal section of one segment treated with the bio-labeled SgreSNMP1 probe. Higher magnification of the area marked by the dashed box in Figure 4A revealed many intensely labeled SgreSNMP1 cells that were distributed across the antennal segment (Fig. 4B). The *S. gregaria* antennae comprise four types of sensilla, which are morphologically distinguishable (Ochieng et al. 1998) and are supposed to play different roles in chemosensation. The results of FISH-experiments indicate that labeled SgreSNMP1 cells can be assigned to sensilla basiconica (Fig. 4C) and sensilla trichodea (Fig. 4D). In *S. gregaria*, a sensillum basiconicum is innervated by up to 50 neurons, while no more than three neurons are housed in sensilla trichodea (Ochieng et al. 1998). Typically, two SgreSNMP1-positive cells were located beneath sensilla basiconica (Fig. 4C), but there were also cases with more than two SgreSNMP1 cells. Beneath sensilla trichodea, not more than two cells expressing SgreSNMP1 were found (Fig. 4D). No SgreSNMP1-positive cells were found under sensilla coeloconica (Fig. 4E) and sensilla chaetica (Fig. 4F).

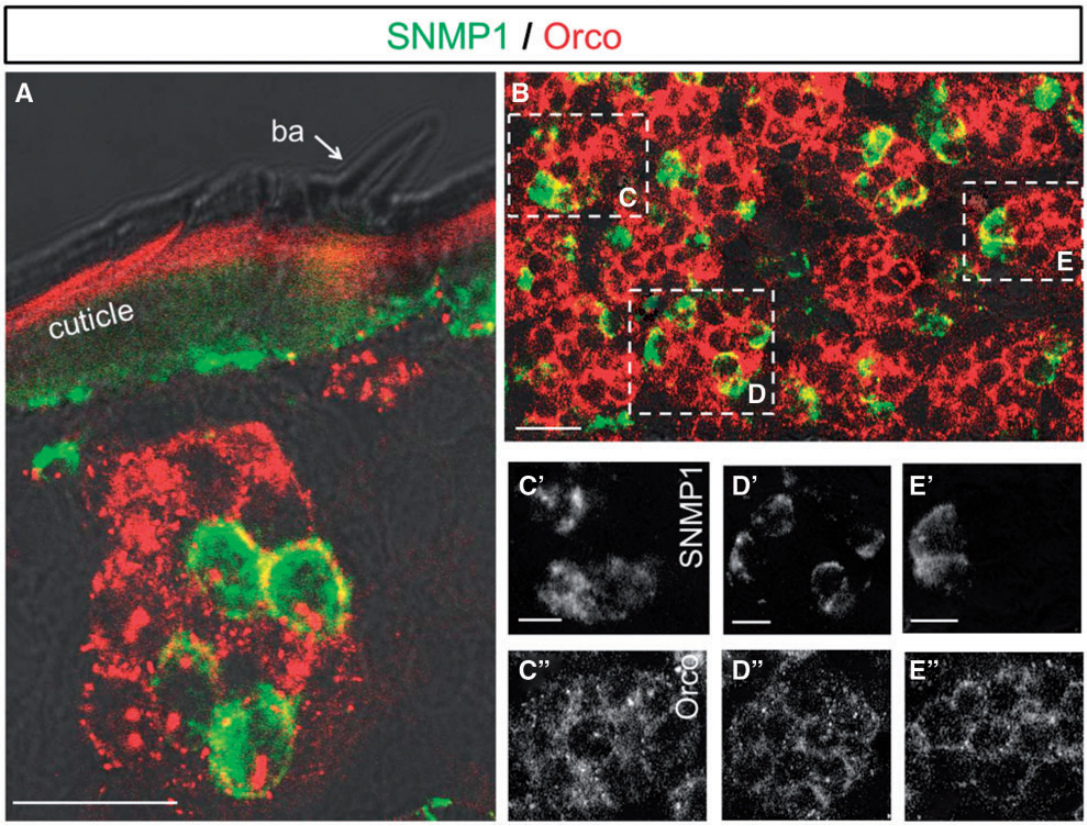


Fig. 5. Expression of SgreSNMP1 within neuron clusters on the male antenna. Two-color FISH assays were performed to visualize the expression of SgreSNMP1 and Orco in neuron clusters utilizing a BIO-labeled probe for SNMP1 and a DIG-labeled probe for Orco. (A) Some of the Orco expressing neurons in the clusters (red) also express SgreSNMP1 (green). The neuron cluster is located below a short blunt sensory hair (sensillum basiconicum, ba). (B) In almost all Orco neuron clusters one or several cells also express SgreSNMP1. (C–E) In three selected examples, the varying number of SgreSNMP1-positive cells in Orco clusters is demonstrated. The profile of SgreSNMP1 cells (C′–E′) and Orco clusters (C″–E″) were visualized by selectively blocking the red-color channel or green-color channel, respectively. Scale bar: A and B, 20 μm; C′–E″, 10 μm.

In order to explore whether the expression of SgreSNMP1 in sensilla basiconica was confined to OSNs, we performed double FISH experiments with probes for Orco and SgreSNMP1. Orco is ubiquitously coexpressed in sensory neurons expressing odorant receptors (Benton et al. 2006, Sato et al. 2008). Previous studies have shown that in the antennae of *S. gregaria* Orco expressing neurons can be assigned to sensilla basiconica and sensilla trichodea (Guo et al. 2014, Yang et al. 2012). Analyses of antennal sections treated with a specific probe for Orco revealed clusters of labeled cells; a typical example is depicted in Figure 5A showing a cluster of approximately 15 cells stained by the Orco probe (red). Out of this group, four cells were stained green by the SgreSNMP1 probe. The results indicate that only a small subgroup of the Orco-positive cells in the cluster also express SgreSNMP1. To determine what proportion of Orco-positive cells in a cluster coexpresses SgreSNMP1, tissue sections from antennal segments with numerous Orco-positive cell clusters were analyzed. The results are depicted in Figure 5B. Within each of the numerous Orco neuron clusters, the number of SgreSNMP1-positive cells varied; ranging from one to several. This is confirmed in three cases representing color-shaded profiles of SgreSNMP1 labeling (Fig. 5C'–E') and Orco labeling (Fig. 5C''–E''), separately. On the single imaging layer, the number of visible SgreSNMP1 cells ranges from 2–4, which is in line with the notion that within the multicellular clusters of the sensilla basiconica at least one cell expresses SgreSNMP1.

Localization of SgreSNMP1 and SgreSNMP2

Previous studies have shown that in moth species, SNMP2 is not expressed in sensory neurons, but in the supporting cells which surround the sensory neurons (Forstner et al. 2008). Therefore, attempts were made to explore which cell types may express SgreSNMP2. Incubating cryosections of antennal segments with a SgreSNMP2 probe led to a labeling pattern (Fig. 6A) quite different from that of SgreSNMP1 (Figs. 4 and 5). There were no labeled cells beneath the sensory hairs; instead, the periphery of neuron clusters was labeled, the region where supporting cells are located. To support this notion, we performed two-color FISH experiments with probes for Orco and SgreSNMP2; a typical result is depicted in Fig. 6B. SgreSNMP2-positive cells were

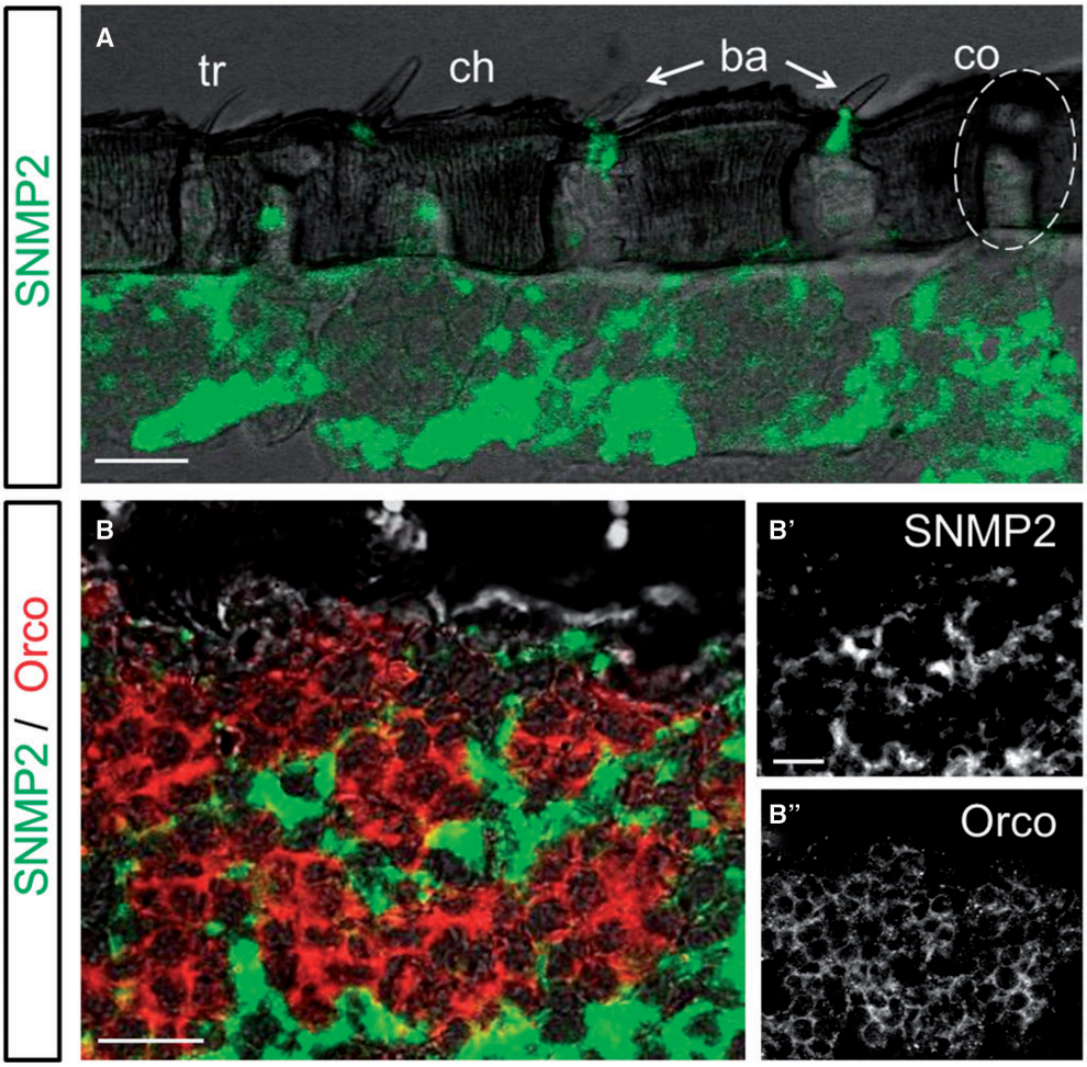


Fig. 6. SgreSNMP2 is not expressed in Orco-positive neurons. (A) FISH assays were performed on sections from male antenna using a BIO-labeled SgreSNMP2 probe. The staining for SgreSNMP2 was located relatively distant from the sensory hairs; the sensilla types are indicated; abbreviations see figure 3. (B) Employing a two-color FISH assay with a BIO-labeled probe for SNMP2 and a DIG-labeled probe for Orco revealed that the SgreSNMP2 expressing cells are located at the periphery of Orco neuron clusters. (B') and (B''): separate visualization of SgreSNMP2 cells and Orco neuron clusters by blocking red-color channel or green-color channel, respectively. Scale bar: 20 μm.

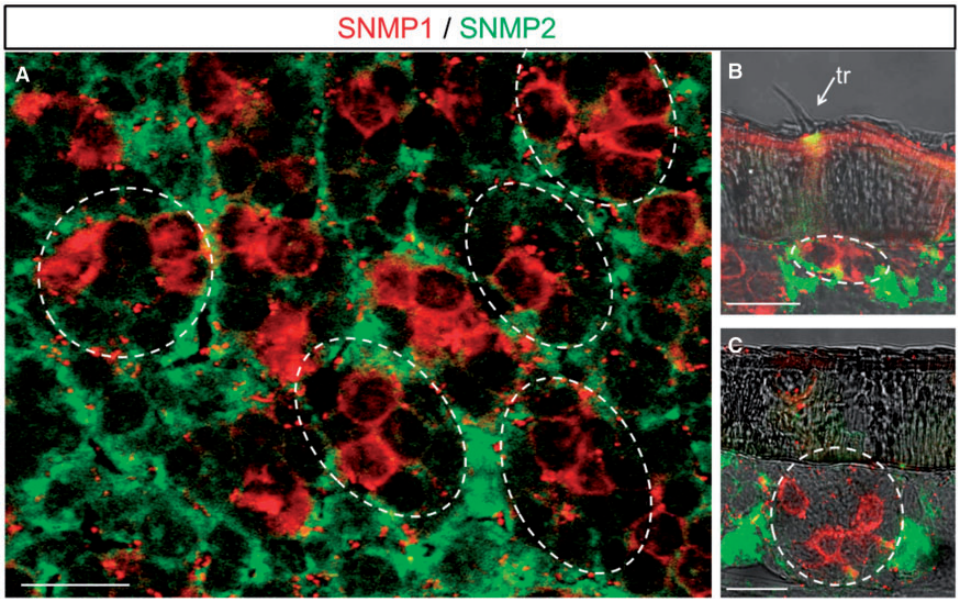


Fig. 7. SgreSNMP1 and SgreSNMP2 are expressed in spatially distinct cells. A DIG-labeled probe for SgreSNMP1 and a BIO-labeled probe for SgreSNMP2 were used in two-color FISH assays. (A) SgreSNMP1 expressing cells are encircled by SgreSNMP2 expressing cells. Dashed areas outline clusters of neurons in which some of the cells express SgreSNMP1. The sensillum trichodeum (tr, in B) and the sensillum basiconicum (in C) house SgreSNMP1 cells as well as SgreSNMP2 expressing cells; both cell types are not overlapping. SgreSNMP1 expressing cells were outlined by dash lines. Scale bar: 20 μm.

tightly encircling the Orco-positive cell clusters and none of the red cluster cells was stained green. This nonoverlapping pattern indicates that SgreSNMP2 is not expressed in neurons but appears to be expressed in supporting cells flanking neuron clusters.

Based on the results of in situ hybridization experiments, we suggested a spatially segregated expression pattern for the two SgreSNMPs. This was confirmed by two-color FISH assays using the SgreSNMP1 probe (DIG-labeled) and SgreSNMP2 probe (BIO-labeled). On tissue section of antennae, one or more SgreSNMP1-positive cells were located within a neuron cluster of neurons, identifiable as red-cells in the dashed area. The SgreSNMP2 labeling was always at the periphery of the area, never overlapping with SgreSNMP1 labeling (Fig.7A). This nonoverlapping expression profile of SgreSNMP1 and SgreSNMP2 was found beneath both sensilla trichodea (Fig. 7B) and sensilla basiconica (Fig. 7C).

Discussion

Previous studies characterizing the "sensory neuron membrane proteins" (SNMPs) focused on holometabolous insects, which undergo a complete metamorphosis from the larval to the adult stage. In this study, we have identified two SNMP subtypes of a hemimetabolous insect, the desert locust *Schistocerca gregaria*, which represents the order of Orthoptera, originating in the Carboniferous period (350–300 million years ago) and comprising more than 25,000 species, many of which are pests of economic importance (Song et al. 2015). Bioinformatics analysis of antennal transcriptome sequences from the desert locust has led to the identification of two sequences with relatively high homology to the SNMP/CD36 family (Nichols and Vogt 2008). All the identified SNMP sequences from numerous species have been categorized into two subclades (Vogt et al. 2009), and detailed analysis of the two *S. gregaria* sequences allowed to assign them to the SNMP1 and SNMP2 group, respectively. The high degree of sequence homology across SNMPs from various insect species which are far apart on the phylogenetic scale indicates a strong negative selection pressure on the primary structure of SNMPs and suggests a conserved and important function for these proteins (Nichols and Vogt 2008). The view that SNMPs fulfill a specific and important functional role is further supported by the finding that the two locust SNMP sequences were readily categorizes into one of the two subclades, indicating that each of the locust SNMP subtypes was more related to the relevant SNMP class of distantly related species than to the other SNMP subtype of *S. gregaria* protein. The conserved sequences between the members of each SNMP subfamily from different insect species, including the hemimetabolous locust suggest important physiological implications. An alignment of the previously known and the newly derived SNMP sequences revealed some characteristic motifs, such as the two transmembrane domains and a large putative extracellular loop with several cysteines that are conserved in the SNMPs as well as in the CD36 proteins and are supposed to form stabilizing disulfide bridges (Rasmussen et al. 1998, Rogers et al. 2001).

In the holometabolous moth *Agrotis ipsilon*, a characteristic age-dependent expression profile was reported for SNMP1 and SNMP2; the expression level increased quite dramatically from the day before eclosion reaching a peak at the third day (Gu et al. 2013). In contrast, in the hemimetabolous *Schistocerca gregaria* we did not find a stage where the expression was particular strong, but rather a similar level of expression for both SNMPs throughout development from nymphal to the adult stages. Visualization of cells which express SgreSNMP1 revealed that they are localized under sensilla trichoidea but also under sensilla basiconica; SNMP1 was not found to be expressed in sensilla coeloconic and sensilla chaetica. A similar spatial expression pattern has recently been reported for *Spodoptera exigua* (Liu et al. 2014), whereas in other moth species and in the fruitfly SNMP1 cells were found only at the base of sensilla trichodea but not under sensilla basiconica and sensilla chaetica (Forstner et al. 2008, Gu et al. 2013, Liu et al. 2013, Zhang et al. 2015). For the antennal sensilla of *Schistocerca gregaria*, several unique features have been described in great detail previously (Ochieng et al. 1998); most notably, the slender sensilla trichodea contain one to three sensory neurons whereas the

broader sensilla basiconica are innervated by 20 to 50 sensory neurons located below the hair base. Most of the neurons clustered in sensilla basiconica appear to be OSNs based on expression of Orco (Fig. 5). Interestingly, out of such a cluster of Orco-positive neurons only a small number (one to very few cells) expressed SNMP1. This observation indicates a molecular and probably functional heterogeneity among the olfactory neurons innervating the same sensillum, which would render the sensilla basiconica responsive to multiple chemical stimuli. The notion that SNMP1 is preferentially expressed in chemosensory neurons which are tuned to specific pheromones (Benton et al. 2007, Forstner et al. 2008) and the finding that only a small fraction of the clustered neurons express SNMP1 may indicate that these cells are responsive to pheromones. In this context, it is interesting to note that previous studies have shown that sensilla basiconica of *Schistocerca gregaria* do in fact respond to stimulation with aggregation pheromone compounds as well as to the courtship–inhibition pheromone phenylacetonitril (Hansson et al. 1996, Seidelmann and Ferenz 2002).

With respect to pheromone detection by OSN in sensilla trichodea of the desert locust, only one study reported a response of trichoid OSNs to (E,Z,)-2,6,-nonadienal. This compound was supposed as a possible female sex pheromone, but was also mentioned as a major constituent of a preferred host plant of *S. gregaria* (Ochieng and Hansson 1999).

The results of double labeling experiments indicate that in *S. gregaria* the two SNMP subtypes are expressed in different cells. The locust SNMP2 was found to be expressed in cells surrounding the Orco-positive cell clusters, suggesting that SNMP2 is not expressed in neurons but rather in the supporting cells. Thus, the spatial expression pattern of SNMP1 and SNMP2 in the hemimetabolous locust is reminiscent to that found in moth species (Forstner et al. 2008, Gu et al. 2013, Zhang et al. 2015) but different from that described for the fly (Benton et al. 2007). Although the implications of the differential expression pattern for SNMP1 and SNMP2 are unknown, the expression in different cell types, together with the low sequence similarities between the two SNMP subtypes, support the notion that SNMP1 and SNMP2 are involved in different physiological processes.

Acknowledgments

Heidrun Froß is acknowledged for excellent technical assistance. We would express gratitude to all members in the lab for helpful suggestions toward this manuscript.

Funding

This work was partly supported by the Deutsche Forschungsgemeinschaft (SPP1392) by a grant to J.K. (KR1786/4-2) and a grant to X.J. by the China Scholarship Council (CSC).

Data Deposition

The SNMP sequences reported in this paper have been deposited in Genbank under accession numbers: KU659599 (SgreSNMP1) and KU659600 (SgreSNMP2).

References Cited

Acton, S. L., P. E. Scherer, H. F. Lodish, and M. Krieger. 1994. Expression cloning of SR-BI, a CD36-related class B scavenger receptor. J. Biol. Chem. 269: 21003–21009.

Almaas, T. J., and H. Mustaparta. 1991. *Heliothis virescens*: Response characteristics of receptor neurons in sensilla trichodea type 1 and type 2. J. Chem. Ecol. 17: 953–972.

Baker, T. C., S. A. Ochieng, A. A. Cossé, S. G. Lee, J. L. Todd, C. Quero, and N. J. Vickers. 2004. A comparison of responses from olfactory receptor neurons of *Heliothis subflexa* and *Heliothis virescens* to components of their sex pheromone. J. Comp. Physiol. A 190: 155–65.

Benton, R., S. Sachse, S. W. Michnick, and L. B. Vosshall. 2006. Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors *in vivo*. PLoS Biol. 4: e20.

Benton, R., K. S. Vannice, and L. B. Vosshall. 2007. An essential role for a CD36-related receptor in pheromone detection in *Drosophila*. Nature 450: 289–293.

Byers, J. A. 1991. Pheromones and chemical ecology of locusts. Biol. Rev. 66: 347–378.

Carey, A., G. Wang, C.-Y. Su, L. J. Zwiebel, and J. R. Carlson. 2010. Odourant reception in the malaria mosquito *Anopheles gambiae*. Nature 464: 66–71.

Chapman, R. F., and A. Joern. 1990. Biology of grasshoppers. John Wiley and Sons, New York, NY.

Clyne, P., A. Grant, R. O'Connell, and J. R. Carlson. 1997. Odorant response of individual sensilla on the *Drosophila* antenna. Invert. Neurosci. 3: 127–35.

Elmore, T., and D. P. Smith. 2001. Putative *Drosophila* odor receptor OR43b localizes to dendrites of olfactory neurons. Insect Biochem. Mol. Biol. 31: 791–798.

Febbraio, M., D. P. Hajjar, and R. L. Silverstein. 2001. CD36: A class B scavenger receptor involved in angiogenesis, atherosclerosis, inflammation, and lipid metabolism. J. Clin. Invest. 108: 785–791.

Forstner, M., T. Gohl, I. Gondesen, K. Raming, H. Breer, and J. Krieger. 2008. Differential expression of SNMP-1 and SNMP-2 proteins in pheromone-sensitive hairs of moths. Chem. Senses 33: 291–299.

Gohl, T., and J. Krieger. 2006. Immunolocalization of a candidate pheromone receptor in the antenna of the male moth, *Heliothis virescens*. Invertebr. Neurosci. 6: 13–21.

Grimaldi, D., and M. S. Engel. 2005. Evolution of the insects. Cambridge University Press, Cambridge, United Kingdom.

Große-Wilde, E., T. Gohl, E. Bouché, H. Breer, and J. Krieger. 2007. Candidate pheromone receptors provide the basis for the response of distinct antennal neurons to pheromonal compounds. Eur. J. Neurosci. 25: 2364–2373.

Gu, S.-H., R.-N. Yang, M.-B. Guo, G.-R. Wang, K.-M. Wu, Y.-Y. Guo, J.-J. Zhou, and Y.-J. Zhang. 2013. Molecular identification and differential expression of sensory neuron membrane proteins in the antennae of the black cutworm moth *Agrotis ipsilon*. J. Insect Physiol. 59: 430–43.

Guo, M., J. Krieger, E. Große-Wilde, C. Mißbach, L. Zhang, and H. Breer. 2014. Variant ionotropic receptors are expressed in olfactory sensory neurons of coeloconic sensilla on the antenna of the desert locust (*Schistocerca gregaria*). Int. J. Biol. Sci. 10: 1–14.

Hallem, E. A., and J. R. Carlson. 2006. Coding of odors by a receptor repertoire. Cell 125: 143–160.

Hansson, B. S., and M. C. Stensmyr. 2011. Evolution of insect olfaction. Neuron 72: 698–711.

Hansson, B. S., S. A. Ochieng, X. Grosmaitre, S. Anton, and P. G. N. Njagi. 1996. Physiological responses and central nervous projections of antennal olfactory receptor neurons in the adult desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae). J. Comp. Physiol. A 179: 157–167.

Hassanali, A., P.G.N. Njagi, and M. O. Bashir. 2004. Chemical ecology of locust and related acridids. Annu. Rev. Entomol. 50: 223–245.

Heifetz, Y., H. Voet, and S. Applebaum. 1996. Factors affecting behavioral phase transition in the desert locust, *Schistocerca gregaria* (Forskål) (Orthoptera: Acrididae). J. Chem. Ecol. 22: 1717–1734.

Jin, X., T. S. Ha, and D. P. Smith. 2008. SNMP is a signaling component required for pheromone sensitivity in *Drosophila*. Proc. Natl. Acad. Sci. USA 105: 10996–11001.

Keil, T., and R. A. Steinbrecht. 1984. Mechanosensitive and olfactory sensilla of insects, pp. 477–516. *In* R. King and H. Akai (eds.), Insect Ultrastructure. Springer, New York, NY.

Krieger, J., O. Klink, C. Mohl, K. Raming, and H. Breer. 2003. A candidate olfactory receptor subtype highly conserved across different insect orders. J. Comp. Physiol. A 189: 519–526.

Krieger, J., E. Große-Wilde, T. Gohl, and H. Breer. 2005. Candidate pheromone receptors of the silkmoth *Bombyx mori*. Eur. J. Neurosci. 21: 2167–2176.

Krogh, A., B. Larsson, G. von Heijne, and E. L. L. Sonnhammer. 2001. Predicting transmembrane protein topology with a hidden markov model: Application to complete genomes. J. Mol. Biol. 305: 567–580.

Kurtovic, A., A. Widmer, and B. J. Dickson. 2007. A single class of olfactory neurons mediates behavioural responses to a *Drosophila* sex pheromone. Nature 446: 542–546.

Li, Z., J. D. Ni, J. Huang, and C. Montell. 2014. Requirement for *Drosophila* SNMP1 for rapid activation and termination of pheromone-induced activity. PLoS Genet. 10: e1004600.

Liu, S., Y.-R. Zhang, W.-W. Zhou, Q.-M. Liang, X. Yuan, J. Cheng, Z. R. Zhu, and Z. J. Gong. 2013. Identification and characterization of two sensory neuron membrane proteins from *Cnaphalocrocis medinalis* (Lepidoptera: Pyralidae). Arch. Insect Biochem. Physiol. 82: 29–42.

Liu, C., J. Zhang, Y. Liu, G. Wang, and S. Dong. 2014. Expression of SNMP1 and SNMP2 genes in antennal sensilla of *Spodoptera exigua* (HÜBNER). Arch. Insect Biochem. Physiol. 85: 114–126.

Nichols, Z., and R. G. Vogt. 2008. The SNMP/CD36 gene family in Diptera, Hymenoptera and Coleoptera: *Drosophila melanogaster, D. pseudoobscura, Anopheles gambiae, Aedes aegypti, Apis mellifera,* and Tribolium castaneum. Insect Biochem. Mol. Biol. 38: 398–415.

Ochieng, S. A., and B. S. Hansson. 1999. Responses of olfactory receptor neurones to behaviourally important odours in gregarious and solitarious desert locust, *Schistocerca gregaria*. Physiol. Entomol. 24: 28–36.

Ochieng, S. A., E. Hallberg, and B. S. Hansson. 1998. Fine structure and distribution of antennal sensilla of the desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae). Cell Tissue Res. 291: 525–536.

Rasmussen, J. T., L. Berglund, M. S. Rasmussen, and T. E. Petersen. 1998. Assignment of disulfide bridges in bovine CD36. Eur. J. Biochem. 257: 488–494.

Rogers, M., J. Krieger, and R. Vogt. 2001. Antennal SNMPs (sensory neuron membrane proteins) of Lepidoptera define a unique family of invertebrate CD36-like proteins. J. Neurobiol. 49: 47–61.

Rogers, M. E., M. Sun, M. R. Lerner, and R. G. Vogt. 1997. Snmp-1, a novel membrane protein of olfactory neurons of the silk moth *Antheraea polyphemus* with homology to the CD36 family of membrane proteins. J. Biol. Chem. 272: 14792–14799.

Sakurai, T., T. Nakagawa, H. Mitsuno, H. Mori, Y. Endo, S. Tanoue, Y. Yasukochi, K. Touhara, and T. Nishioka. 2004. Identification and functional characterization of a sex pheromone receptor in the silkmoth *Bombyx mori*. Proc. Natl. Acad. Sci. USA. 101: 16653–16658.

Sakurai, T., H. Mitsuno, A. Mikami, K. Uchino, M. Tabuchi, F. Zhang, H. Sezutsu, and R. Kanzaki. 2015. Targeted disruption of a single sex pheromone receptor gene completely abolishes *in vivo* pheromone response in the silkmoth. Sci. Rep. 5: 11001.

Sato, K., M. Pellegrino, T. Nakagawa, T. Nakagawa, L. B. Vosshall, and K. Touhara. 2008. Insect olfactory receptors are heteromeric ligand-gated ion channels. Nature 452: 1002–1006.

Seidelmann, K., and H. J. Ferenz. 2002. Courtship inhibition pheromone in desert locusts, *Schistocerca gregaria*. J. Insect Physiol. 48: 991–996.

Seidelmann, K., H. Weinert, and H. J. Ferenz. 2003. Wings and legs are production sites for the desert locust courtship-inhibition pheromone, phenylacetonitrile. J. Insect Physiol. 49: 1125–1133.

Silverstein, R. L., W. Li, Y. M. Park, and S. O. Rahaman. 2010. Mechanisms of cell signaling by the scavenger receptor CD36: implications in atherosclerosis and thrombosis. Trans. Am. Clin. Climatol. Assoc. 121: 206–220.

Simpson, S. J., and G. A. Sword. 2008. Locusts. Curr. Biol. 18: R364–R366.

Song, H., C. Amédégnato, M. M. Cigliano, L. Desutter-Grandcolas, S. W. Heads, Y. Huang, D. Otte, and M. F. Whiting. 2015. 300 million years of diversification: elucidating the patterns of orthopteran evolution based on comprehensive taxon and gene sampling. Cladistics 31: 621–651.

Stahr, C., A. Svatoš, and K. Seidelmann. 2013. Chemical identification, emission pattern and function of male-specific pheromones released by a rarely swarming locust, *Schistocerca americana*. J. Chem. Ecol. 39: 15–27.

Steinbrecht, R. A., and W. Gnatzy. 1984. Pheromone receptors in *Bombyx mori* and *Antheraea pernyi*. I. Reconstruction of the cellular organization of the sensilla trichodea. Cell Tissue Res. 235: 25–34.

Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30: 2725–2729.

Touhara, K., and L. B. Vosshall. 2009. Sensing odorants and pheromones with chemosensory receptors. Annu. Rev. Physiol. 71: 307–332.

Uvarov, B. P. 1966. Grasshoppers and locusts: A handbook of general acridology. Anti-Locust Research Centre, Cambridge University Press, Great Britain.

Vogt, R. G., N. E. Miller, R. Litvack, R. A. Fandino, J. Sparks, and J. Staples, R. Friedman, J. C. Dickens.. 2009. The insect SNMP gene family. Insect Biochem. Mol. Biol. 39: 448–56.

Vosshall, L. B., and B. S. Hansson. 2011. A unified nomenclature system for the insect olfactory coreceptor. Chem. Senses 36: 497–498.

Wang, G., A. F. Carey, J. R. Carlson, and L. J. Zwiebel. 2010. Molecular basis of odor coding in the malaria vector mosquito *Anopheles gambiae*. Proc. Natl. Acad. Sci. USA. 107: 4418–4423.

Wang, X., X. Fang, P. Yang, X. Jiang, F. Jiang, D. Zhao, B. Li, F. Cui, J. Wei, C. Ma, et al. 2014. The locust genome provides insight into swarm formation and long-distance flight. Nat. Commun. 5: 2957.

Yang, Y., J. Krieger, L. Zhang, and H. Breer. 2012. The olfactory co-receptor Orco from the migratory locust (*Locusta migratoria*) and the desert locust (*Schistocerca gregaria*): identification and expression pattern. Int. J. Biol. Sci. 8: 159–170.

Zhang, J., Y. Liu, W. B. Walker, S.-L. Dong, and G.-R. Wang. 2015. Identification and localization of two sensory neuron membrane proteins from *Spodoptera litura* (Lepidoptera: Noctuidae). Insect Sci. 22: 399–408.

IVYSPRING INTERNATIONAL PUBLISHER

International Journal of Biological Sciences

2017; 13(7): 911-922. doi: 10.7150/ijbs.18402

Research Paper

In Search for Pheromone Receptors: Certain Members of the Odorant Receptor Family in the Desert Locust *Schistocerca gregaria* (Orthoptera: Acrididae) Are Co-expressed with SNMP1

Pablo Pregitzer[1]✉, Xingcong Jiang[1], Ewald Grosse-Wilde[2], Heinz Breer[1], Jürgen Krieger[3] and Joerg Fleischer[3]✉

1. University of Hohenheim, Institute of Physiology (230), Stuttgart, Germany
2. Max Planck Institute for Chemical Ecology, Department of Evolutionary Neuroethology, Jena, Germany
3. Martin Luther University Halle-Wittenberg, Institute of Biology/Zoology, Department of Animal Physiology, Halle (Saale), Germany

✉ Corresponding author: Pablo Pregitzer, University of Hohenheim, Institute of Physiology (230), Stuttgart, Germany, phone: +49 711 459 22270, fax: +49 711 459 23726, email: p_pregitzer@uni-hohenheim.de and Joerg Fleischer, Martin Luther University Halle-Wittenberg, Institute of Biology/Zoology, Department of Animal Physiology, Halle (Saale), Germany, phone: +49 345 5526476, fax: +49 345 5527152, email: joerg.fleischer@zoologie.uni-halle.de

Received: 2016.11.18; Accepted: 2017.03.09; Published: 2017.07.15

Abstract

Under given environmental conditions, the desert locust (*Schistocera gregaria*) forms destructive migratory swarms of billions of animals, leading to enormous crop losses in invaded regions. Swarm formation requires massive reproduction as well as aggregation of the animals. Pheromones that are detected via the olfactory system have been reported to control both reproductive and aggregation behavior. However, the molecular basis of pheromone detection in the antennae of *Schistocerca gregaria* is unknown. As an initial step to disclose pheromone receptors, we sequenced the antennal transcriptome of the desert locust. By subsequent bioinformatical approaches, 119 distinct nucleotide sequences encoding candidate odorant receptors (ORs) were identified. Phylogenetic analyses employing the identified ORs from *Schistocerca gregaria* (SgreORs) and OR sequences from the related species *Locusta migratoria* revealed a group of locust ORs positioned close to the root, i.e. at a basal site in a phylogenetic tree. Within this particular OR group (termed basal or b-OR group), the locust OR sequences were strictly orthologous, a trait reminiscent of pheromone receptors from lepidopteran species. In situ hybridization experiments with antennal tissue demonstrated expression of b-OR types from *Schistocerca gregaria* in olfactory sensory neurons (OSNs) of either sensilla trichodea or sensilla basiconica, both of which have been reported to respond to pheromonal substances. More importantly, two-color fluorescent in situ hybridization experiments showed that most b-OR types were expressed in cells co-expressing the "sensory neuron membrane protein 1" (SNMP1), a marker indicative of pheromone-sensitive OSNs in insects. Analyzing the expression of a larger number of SgreOR types outside the b-OR group revealed that only a few of them were co-expressed with SNMP1.

In summary, we have identified several candidate pheromone receptors from *Schistocerca gregaria* that could mediate responses to pheromones implicated in controlling reproduction and aggregation behavior.

Key words: olfaction, insect, pheromone receptor, desert locust, antenna, sensory neuron membrane protein 1

Introduction

By means of various cuticular hair-like structures (sensilla) on their antennae, insects are capable of detecting a variety of volatile chemicals in their surroundings, including pheromone compounds that are used to exchange important information between conspecifics and shape distinct behaviors such as aggregation, mate finding and courtship (1,2). To date, our knowledge about the molecular mechanisms

underlying the reception of pheromones is mainly based on studies of holometabolic insects, especially lepidopteran and dipteran species (3–6). In Lepidoptera and Diptera, pheromone detection appears to be mainly assigned to trichoid sensilla (6–8), which typically harbor the dendritic processes of up to three olfactory sensory neurons (OSNs) (9–13). In these pheromone–responsive sensilla, at least one OSN responds to a pheromonal compound and comprises the molecular machinery for pheromone signaling, including a specific odorant receptor (OR) (6,14,15), the odorant receptor co-receptor (Orco) (16–19) as well as the "sensory neuron membrane protein 1" (SNMP1) (20–25).

In marked contrast to lepidopteran and dipteran species, the mechanisms of pheromone signaling in hemimetabolic insects are largely elusive although some pheromonal compounds have been identified in locusts (26–32). The hemimetabolic locusts are only distantly related to holometabolic insects (33). Moreover, some locust species (in particular the desert locust *Schistocerca gregaria*) are unique for a striking phase polyphenism including a change between a solitary and a gregarious phase (1,34). In *Schistocerca gregaria*, animals in the gregarious phase can form huge swarms that tremendously threaten agricultural crops in Africa and Asia. Swarm formation is based on massive reproduction as well as aggregation. In locusts, both reproductive and aggregation behaviors are supposed to involve pheromones (1,27,32,35,36) and evidence has been accumulated that such pheromones elicit responses in antennal sensilla (31,37–39). The antennal sensilla of locusts are categorized in distinct classes with the slender sensilla trichodea containing one to three OSNs and the broader and more massive sensilla basiconica comprising between 20 to 50 sensory neurons. As a third type, sensilla coeloconica (pegs in pits) contain one to four OSNs (40). In *Schistocerca gregaria*, OSNs in sensilla coeloconica do not express Orco or SNMP1 (41,42) but comprise variant ionotropic receptors (43) that mediate responses to distinct classes of odorants (carboxylic acids, amines, ammonia) in other insects (44,45). In contrast, based on the expression of Orco, OSNs in sensilla trichodea and sensilla basiconica of the desert locust most likely express ORs for odorant detection. Moreover, some OSNs in these types of sensilla were also found to express SNMP1 (41,42). Since SNMP1 is considered indicative of pheromone-responsive neurons in other insects (20–22,25,46,47), SNMP1-expressing OSNs in *Schistocerca gregaria* might also respond to pheromones and supposedly express appropriate pheromone-binding OR types. In view of the indicated roles of pheromones in locust aggregation and reproductive behaviors and with regard to the potential of using a blockade of pheromone signaling for insect control strategies, a better understanding of pheromone reception in the swarm-forming crop pest *Schistocerca gregaria* is highly desirable. However, until now, receptors for pheromones in locusts are still elusive and in view of the sparse information about candidate pheromone compounds, experimental approaches to identify pheromone receptors are limited. Therefore, as an initial step to identify OR types in *Schistocerca gregaria* (SgreORs) activated by pheromonal compounds, the present study was based on the paradigm that candidate pheromone receptors should be expressed in SNMP1-positive neurons of locust antennae. Towards this goal, we have screened a *Schistocerca gregaria* antennal transcriptome database for candidate OR-encoding sequences and subsequently determined their topographic expression pattern in the antennae with a particular emphasis on a possible co-expression with SNMP1.

Materials and Methods

Animals and tissue treatment

Adult *Schistocerca gregaria* were purchased from Bugs International (Irsingen/Unterfeld, Germany) and their antennae were dissected using autoclaved surgical scissors. For RNA extraction, antennae were immediately frozen in liquid nitrogen and stored at -70 °C. For in situ hybridization experiments, antennae were embedded in Tissue-Tek O.C.T. Compound (Sakura Finetek, Alphen aan den Rijn, The Netherlands).

Identification of OR-encoding sequences from *Schistocerca gregaria* by transcriptome sequencing and bioinformatical analyses

From the antennae of adult male and female desert locusts, total RNA was extracted using Trizol reagent (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer's protocol. The material was sent to the Max Planck-Genome-centre (Cologne, Germany) where a TruSeq RNA library was generated. The library was sequenced on a HiSeq2500 (Illumina, San Diego, CA, USA), generating a total of 51,151,235 paired end 100 base pair (bp) reads. The data were cleaned and trimmed by the Max Planck-Genome-centre. The results were assembled in CLC Genomics Workbench 8 (Qiagen, Venlo, The Netherlands) using the de novo assembler algorithm with default options (yet, all contigs below 300 bp size were omitted). This resulted in 55,060 contigs with an N50 of 2,223 bp. To identify candidate OR sequences, contigs were analyzed with blastx searches using databases of known OR-coding sequences in

Geneious 7 (Biomatters, Auckland, New Zealand). Transcripts with E-values below 10^{-3} were extracted and assembled with the Geneious assembler under highest similarity settings to reduce redundancy. The resultant contigs and unique sequences were manually annotated using standard blastx comparisons with the nr database (NCBI, Bethesda, MD, USA). Using both tBLASTx and BLASTp methods, the latter sequences were subsequently utilized as secondary queries to identify additional OR-encoding sequences present in the transcriptome database. Next, verification of the identified sequences as putative OR-encoding sequences was accomplished via BLASTx survey based on a NCBI non-redundant protein database.

Reverse transcription polymerase chain reaction (RT-PCR)

Total RNA was extracted from antennae of adult males and females using Trizol reagent (Thermo Fisher Scientific) following the manufacturer's protocol. Poly(A)$^+$ RNA was isolated from 100 µg total RNA with oligo(dT)25 magnetic beads (Thermo Fisher Scientific) according to the supplier's specifications and with a final elution in 30 µl H_20. Poly(A)$^+$ RNA was converted into cDNA utilizing 10 µl poly(A)$^+$ RNA elution, 4 µl first strand buffer (250 mM Tris pH 8.3, 375 mM KCl, 15 mM $MgCl_2$), 1 µl 10 mM dNTP mix, 1 µl RNaseOUT recombinant ribonuclease inhibitor, 2 µl 1,4-dithiothreitol (DTT) (0.1M), 1 µl oligo(dT)18 primer and 1 µl Superscript III reverse transcriptase (Thermo Fisher Scientific). Synthesis of cDNA was conducted at 55 °C for 50 min followed by incubation at 70 °C for 15 min. Non-quantitative RT-PCR was performed using SgreOR-specific sense and antisense primers (supplementary table 1). For PCR amplification of SgreOR-encoding sequences, the following conditions were used: 94 °C for 90 s followed by 20 cycles with 94 °C for 30 s, 50-60 °C for 30 s (thereby, the annealing temperature was decreased by 0.5 °C per cycle) and 72 °C for 90 s. Subsequently, 20 cycles with an annealing temperature of 40-50 °C were performed followed by incubation at 72 °C for 15 min. Alternatively, PCR reactions were run at 97 °C for 1 min followed by 34 cycles with 97 °C for 40 s and 68 °C for 3 min. After the last cycle, a final incubation at 68 °C for 3 min was performed. PCR products were run on 1% agarose gels and visualized by staining with ethidium bromide. PCR products of the expected size were subsequently purified using the geneclean kit (MP Biomedicals, Santa Ana, CA, USA) and cloned into the pGEM-T or the pGEM-T easy vector (Promega, Madison, WI, USA). Vectors carrying putative OR sequences were sequenced by Macrogen (Amsterdam, The Netherlands) and analyzed utilizing the Chromas Software (http://technelysium.com.au/wp/chromas; Technelysium, South Brisbane, Australia).

Phylogenetic analyses

An amino acid alignment comprising the identified OR sequences from *Schistocerca gregaria* (SgreOR; this study) and the recently described OR sequences from *Locusta migratoria* (LmigOr) (48) was created using BioEdit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) and MEGA 6.06 software (49). Thereby, two SgreORs and four LmigOrs that represented only very short sequences were not included. Next, a neighbor-joining tree was constructed based on a Clustal X alignment utilizing the MEGA 6.06 software and the default settings of this software package. Finally, the tree was rooted using sequences from several insect species (*Schistocerca gregaria*, *Locusta migratoria*, *Manduca sexta*, *Drosophila melanogaster* and *Anopheles gambiae*) encoding the odorant receptor co-receptor Orco, an unusual member of the OR family that is highly conserved across insect orders (16,17).

In situ hybridization

Digoxigenin-labeled or biotin-labeled antisense probes were synthesized from linearized pGEM-T vectors containing partial or full length coding sequences of SgreORs using the T7/Sp6 RNA transcription system (Roche Diagnostics, Mannheim, Germany) following the manufacturer's protocol. Antennae of male and female adult *Schistocerca gregaria* locusts were crosscut into two halves, embedded in Tissue-Tek O.C.T. Compound and used to make 12 µm thick longitudinal sections with a Jung CM300 or a Leica CM3050 S cryostat (Leica Microsystems, Bensheim, Germany) at -21 °C. Sections were thaw mounted on Super Frost Plus slides (Menzel, Braunschweig, Germany) and stored at -70 °C until use. With few modifications, in situ hybridization experiments were performed using the protocol described in detail previously (41,43). Briefly, sections were taken out from the -70 °C freezer and immediately transferred into fixation solution (4% paraformaldehyde in 0.1 M $NaHCO_3$, pH 9.5) for 22 min at 4 °C. Next, sections were washed in 1xPBS (0.85% NaCl, 1.4 mM KH_2PO_4, 8 mM Na_2HPO_4, pH 7.1) for 1 min, incubated in 0.2 M HCl for 10 min and washed twice in 1xPBS for 2 min each. Then sections were incubated for 10 min in acetylation solution (0.25% acetic anhydride freshly added in 0.1 M triethanolamine) followed by 3 wash steps in 1xPBS (each wash step lasted 3 min). Sections were incubated in pre-hybridization solution [5xSSC (0.75 M NaCl, 0.075 M sodium citrate, pH 7.0) and 50%

formamid] for 15 min at 4 °C. For non-fluorescent in situ hybridization, each slide was subsequently covered with 100 µl hybridization solution 1 [50% formamide, 25% H_2O, 25% Microarray Hybridization Solution Version 2.0 (GE Healthcare, Freiburg, Germany)] containing the labeled antisense RNA probe. After placing a coverslip, slides were incubated in a humid box (50% formamide) at 60 °C overnight. Visualization of digoxigenin-labeled probes in non-fluorescent in situ hybridization experiments was performed as described earlier (41) by means of an anti-digoxigenin alkaline phosphatase-conjugated antibody (Roche Diagnostics) diluted 1:750 and a substrate solution containing NBT (nitroblue tetrazolium) and BCIP (5-brom-4-chlor-3-indolyl phosphate). Tissue sections were analyzed with an Axioskop2 MOT (Carl Zeiss MicroImaging, Göttingen, Germany) equipped with an AxioCam MRc5 camera (Carl Zeiss) and AxioVision SE64 Rel. 4.9 software (Carl Zeiss).

For two-color fluorescent in situ hybridization (FISH), fixation, acetylation and hybridization were carried out as described above. Sections were hybridized with digoxigenin- and biotin-labeled probes simultaneously. However, for two-color FISH, 100 µl hybridization solution 2 (50% formamide, 2xSSC, 10% dextran sulphate, 0.2 mg/ml yeast t-RNA, 0.2 mg/ml herring sperm DNA) supplemented with labeled antisense RNA was placed per slide onto the tissue sections. Visualization of labeled probes was performed as described previously (50). In short, digoxigenin-labeled probes were visualized by the anti-digoxigenin alkaline phosphatase-conjugated antibody in combination with the HNPP fluorescent detection set (Roche Diagnostics). Incubation with the anti-digoxigenin alkaline phosphatase-conjugated antibody as well as incubation with the HNPP/Fast Red substrate were conducted overnight at 4 °C. For visualization of biotin-labeled probes, the TSA fluorescein system kit (Perkin Elmer, Waltham, MA, USA) was used. Incubation of sections with biotin-binding streptavidin conjugated to horse radish peroxidase and incubation with fluorescein-conjugated tyramides were conducted overnight at 4 °C. Sections were analyzed for fluorescent hybridization signals using a LSM 510 Meta laser scanning microscope (Carl Zeiss). Confocal image stacks were recorded from antennal segments. The micrographs shown represent selected optical planes.

Results

Identification of sequences encoding candidate ORs in *Schistocerca gregaria*

To identify sequences coding for SgreOR types, iterative tBLASTx searches in a *Schistocerca gregaria* antennal transcriptome database were performed using OR sequences from several other insect species. The putative OR-encoding sequences obtained from *Schistocerca gregaria* were further screened by BLASTx exploitations using the NCBI database. This approach led to the identification of 119 candidate SgreOR sequences of various lengths. These SgreORs were designated consecutively with Arabic numerals (SgreOR1 to SgreOR119). Naming of SgreOR sequences was conducted mainly in accordance with OR sequences of the related species *Locusta migratoria* for which 142 candidate ORs (termed LmigOrs) have been reported recently (48). Of the 119 identified candidate SgreOR sequences, 18 contained full length open reading frames (with a stop codon preceeding the presumptive start codon). Based on their length and their sequence, the open reading frames of further 6 sequences are putatively full length (they lack a stop codon preceeding the presumptive start codon). The remaining 95 SgreOR sequences presumably represent only partial sequences (more detailed information about the length of the SgreOR sequences and their GenBank accession numbers is given in supplementary table 2 and supplementary table 3). As a first step to analyze and characterize the SgreOR sequences, the deduced amino acid sequences of 117 SgreORs and 138 LmigOrs were used to generate a neighbor-joining tree (Fig. 1) (for these analyses, two SgreORs and four LmigOrs that represented only very short and partial OR sequences were excluded). Rooting of the tree was conducted with sequences from five different insect species that encode the odorant receptor co-receptor Orco, a non-canonical member of the OR family highly conserved among insects (16,17). In the resulting phylogenetic tree, ORs from both species were arranged in several distinct groups of different sizes. Within the tree, for numerous SgreORs and LmigOrs, an orthologous sequence in the other species was found. The amino acid sequence identities for orthologous SgreORs/LmigOrs pairs were up to ~90%. However, for several SgreORs, no clear orthologue from *Locusta migratoria* was identified (e.g. SgreOR89, SgreOR101 and SgreOR111). In some cases, such SgreORs were part of small groups of paralogous sequences (for instance SgreOR94, SgreOR95 and SgreOR96). Vice versa, also for a number of LmigOrs, no orthologue from *Schistocerca gregaria* was observed (such as LmigOr78, LmigOr91 and LmigOr101) and some of these sequences belonged to small groups of paralogous LmigOrs (for example LmigOr20, LmigOr21, LmigOr22).

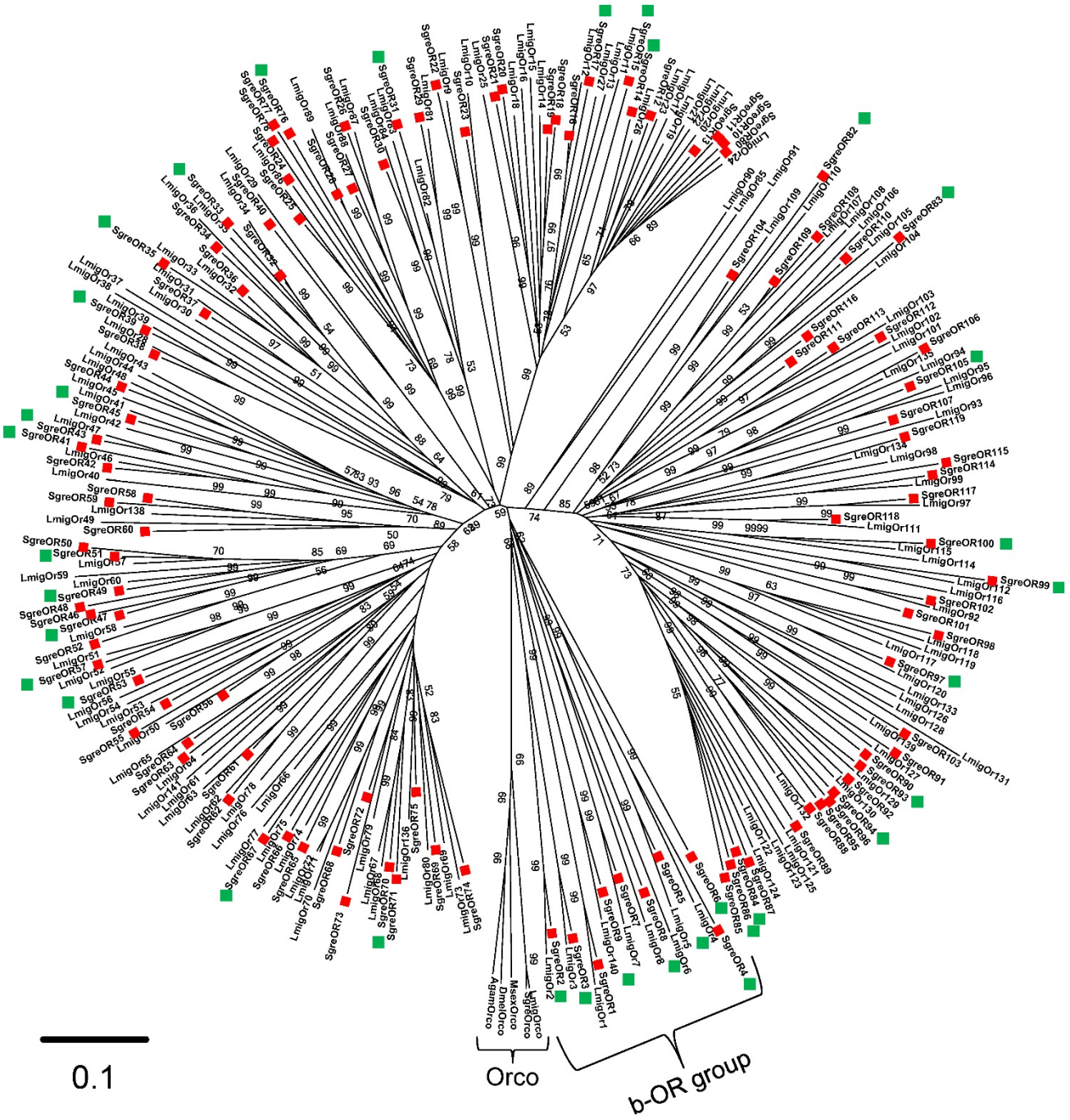


Fig. 1. Evolutionary relationships among SgreORs and LmigOrs. A neighbor-joining tree was calculated using the MEGA program. For rooting the tree, Orco-encoding sequences of different insect species were used. Calculations are based on a Clustal X alignment of the 117 SgreOR and 138 LmigOr amino acid sequences included in the tree. Branch lengths are proportional to the percentage of sequence differences. The scale bar indicates 10% difference. The numbers in the tree indicate bootstrap support values (in %) based on 1000 replicates (only values above 50% are shown). OR-coding sequences from *Schistocerca gregaria* are denoted by red squares. SgreOR sequences used in two-color FISH experiments together with a SNMP1-specific probe are highlighted with green squares.

In the tree, one group of ORs including SgreOR1 to SgreOR9, LmigOr1 to LmigOr8 and LmigOr140 appeared to be unique due to some characteristic features. In particular, this group of ORs was more closely related to the Orco-encoding sequences at the root (basis) of the tree than other SgreORs, i.e. this group was situated at a phylogenetically more basal site than the other SgreORs (Fig. 1). Therefore, this group was designated as the basal OR (b-OR) group. It comprises 8 pairs of orthologous SgreORs/LmigOrs. Within these pairs, the orthologues shared amino acid sequence identities between 47 and 87%, indicating pronounced sequence conservation across species borders. The clear separation from other OR

sequences in neighbor-joining trees and high sequence identity between species is reminescent of pheromone receptors from lepidopteran species, such as the tobacco budworm *Heliothis virescens* (3), the silkworm *Bombyx mori* (15), the tobacco hornworm *Manduca sexta* (51) or the cotton leafworm *Spodoptera littoralis* (52). However, no substantial sequence identities between b-ORs and known pheromone receptors from other insect species were found (data not shown).

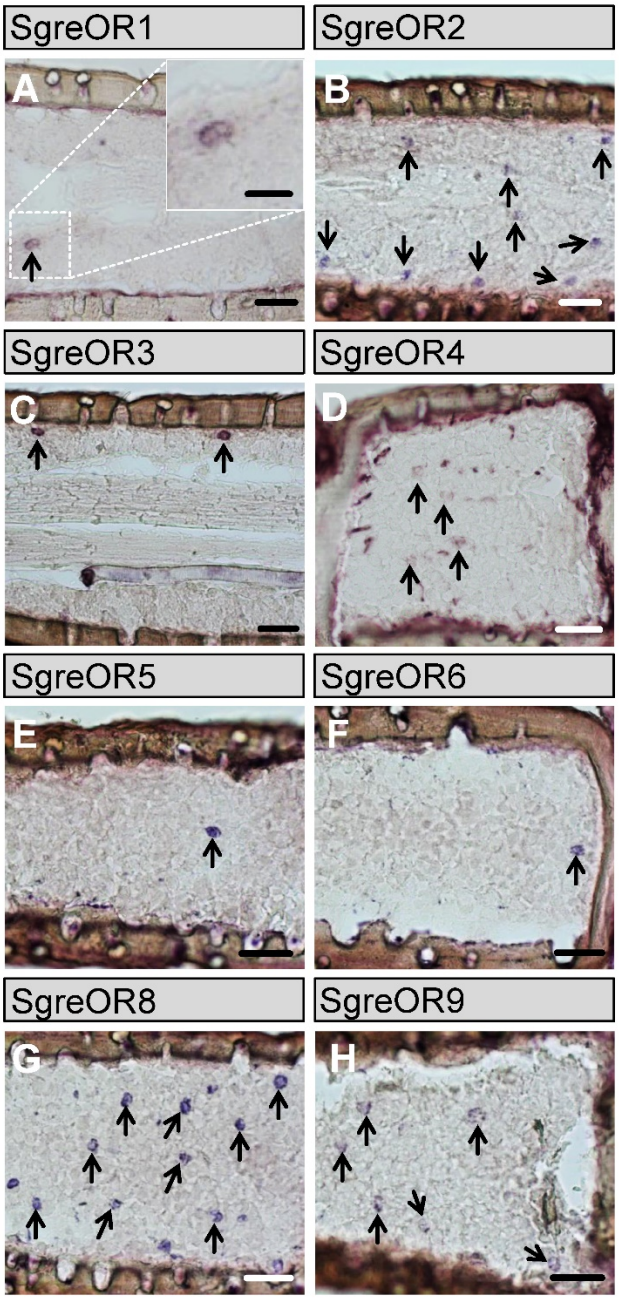


Fig. 2. Expression of the b-OR types in the antenna of *Schistocerca gregaria*. (**A**-**H**) In situ hybridizations were performed on longitudinal tissue sections of antennae with antisense RNA probes for the b-OR types indicated. Some of the stained cells are exemplarily denoted by arrows. Scale bars: 40 µm. Scale bar in the inset in A: 20 µm.

Topographic expression of b-OR genes in the antennae of *Schistocerca gregaria*

To analyze the expression of b-ORs in the antenna in more detail, PCR experiments with antennal cDNA from *Schistocerca gregaria* and specific primer pairs for all 9 members (SgreOR1 through SgreOR9) of the b-OR group of this species were conducted (data not shown). Cloning and sequencing of the resulting PCR products showed that they indeed encoded SgreOR1 to SgreOR9, further supporting the expression of these receptors in antennal tissue from *Schistocerca gregaria*.

For subsequent in situ hybridization experiments, RNA antisense probes were generated for the 9 members of the b-OR group. The results of these approaches revealed that distinct b-OR genes were expressed in varying numbers of cells in the antennae of *Schistocerca gregaria* (Fig. 2). No obvious differences were detected in the antennal expression patterns between males and females (data not shown). The receptor types SgreOR2 (Fig. 2B) and SgreOR8 (Fig. 2G) were expressed in a rather high number of cells per segment. Probes for the other b-ORs (SgreOR1, SgreOR3, SgreOR4, SgreOR5, SgreOR6 and SgreOR9) labeled only relatively small numbers of cells (Fig. 2A, Fig. 2C-F and Fig. 2H) suggesting that only few antennal cells express these receptors. We only observed very weak signals when in situ hybridization experiments were performed on antennal tissue sections with an antisense probe for SgreOR7 (data not shown).

Co-expression of b-ORs with SNMP1

As an initial step to examine whether members of the b-OR group may represent candidate pheromone receptors, we used a strategy based on the observation that pheromone-responsive OSNs in insects express SNMP1 (20–25). In this regard, we have shown recently that SNMP1 is also expressed in a subpopulation of OSNs from *Schistocerca gregaria* (42). To approach the question whether members of the b-OR group may be co-expressed with SNMP1, two-color FISH experiments were performed with a biotin-labeled probe for SNMP1 and specific digoxigenin-labeled probes for distinct b-OR types from *Schistocerca gregaria* (SgreOR2, SgreOR3, SgreOR4, SgreOR5, SgreOR6, SgreOR8 and SgreOR9). In such experiments, SNMP1 was not co-expressed with SgreOR4 (supplementary Fig. 1A). By contrast, expression of SNMP1 was observed in cells positive for SgreOR2, SgreOR3, SgreOR5, SgreOR6, SgreOR8 (Fig. 3) and SgreOR9 (supplementary Fig. 1B). For unknown reason, it was not possible to visualize expression of SgreOR1 and SgreOR7 in two-color FISH experiments. Taken together, these findings

demonstrate that most receptors of the b-OR group are co-expressed with SNMP1, suggesting that they might function as pheromone receptors.

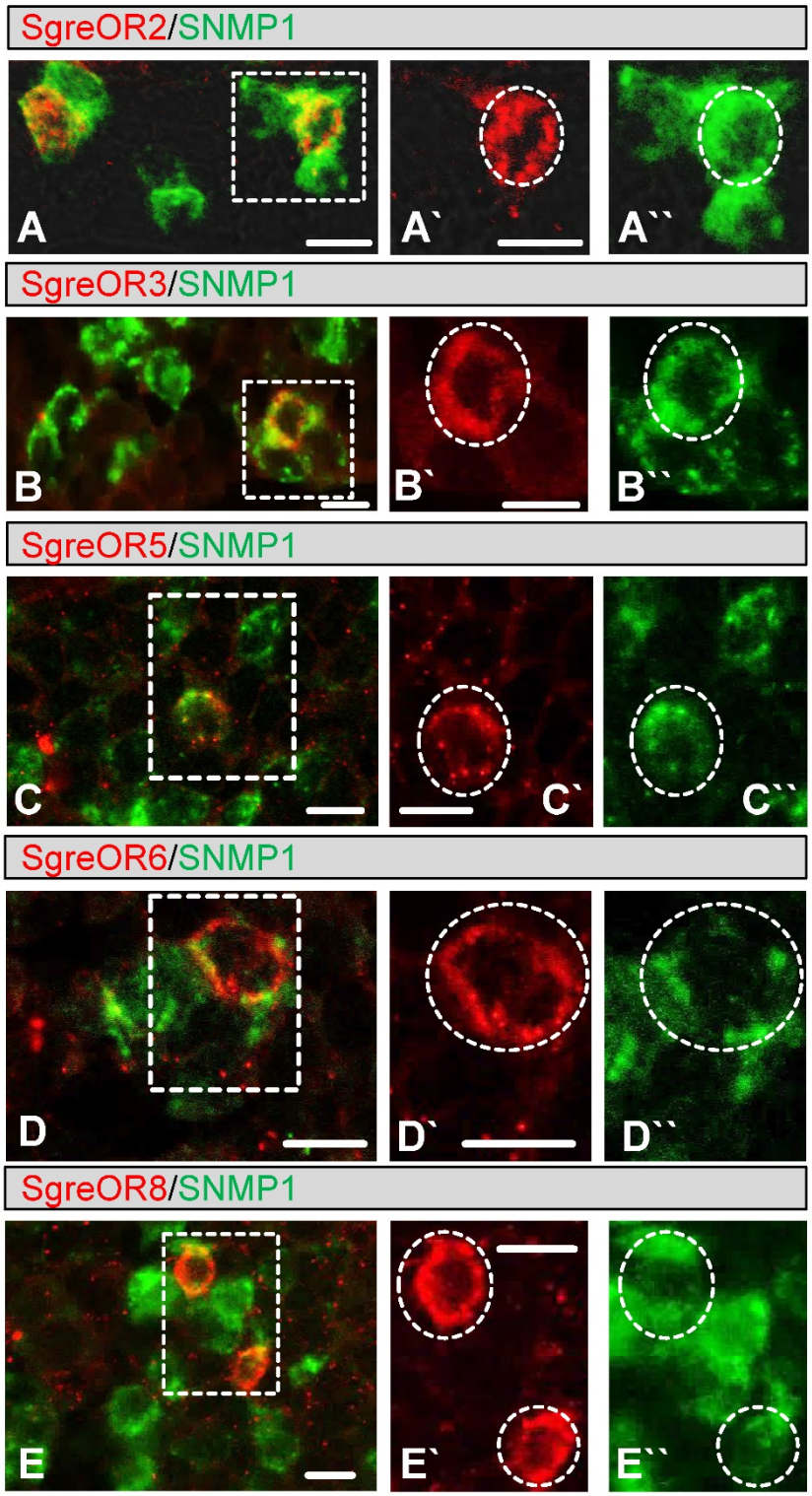


Fig. 3. Co-expression of different b-OR types with SNMP1. (**A**–**E**) Two-color FISH approaches were performed on sections through antennae with antisense RNA probes for SNMP1 (A-E, green) and SgreOR2 (A, red), SgreOR3 (B, red), SgreOR5 (C, red), SgreOR6 (D, red) and SgreOR8 (E, red). The micrographs shown represent single optical planes taken from stacks of confocal images. Cells positive for SgreOR2, SgreOR3, SgreOR5, SgreOR6 or SgreOR8 co-express SNMP1. (**A`/A``-E`/E``**) Higher magnifications of the boxed areas in A, B, C, D and E are depicted in either the red or the green color channel. Cells co-expressing the relevant SgreOR types and SNMP1 are marked with dashed circles. Scale bars: 10 µm.

In two-color FISH approaches, several SNMP1-expressing neurons were frequently found to be located in close proximity to each other. The notion that all cells of such a small cluster of SNMP1-positive neurons may express the same b-OR type was not supported by our experiments. Instead, we found that only one cell in such a cluster expresses a given receptor of the b-OR group (Fig. 3).

Sensilla-specific expression of b-OR group members

SNMP1-expressing OSNs of *Schistocerca gregaria*, are found in sensilla trichodea and in sensilla basiconica (42). Consistently, in electrophysiological recordings from *Schistocerca gregaria* antennae, responses of both sensilla trichodea and sensilla basiconica to pheromonal substances have been observed (38). In a next step, we therefore set out to explore whether distinct receptors of the b-OR group can be assigned to these types of sensilla. To address this issue, we performed two-color FISH experiments employing digoxigenin-labeled probes for distinct b-ORs (SgreOR2, SgreOR3, SgreOR5, SgreOR6, SgreOR8 and SgreOR9) and a biotin-labeled probe for Orco since OSNs beneath sensilla trichodea and sensilla basiconica express Orco (41,42). It was found that cells expressing the receptor types SgreOR2 (Fig. 4A), SgreOR8 (Fig. 4B), SgreOR5, SgreOR6 and SgreOR9 (supplementary Fig. 2) were situated in the Orco-positive neuron clusters characteristic of sensilla basiconica. By contrast, cells expressing SgreOR3 were located directly beneath trichoid sensilla (Fig. 4C) and SgreOR3-positive neurons were only found within small groups of 2-3 Orco-expressing cells, an arrangement typical for sensilla trichodea (Fig. 4D). It is worthy of note that labeling with the Orco-specific probe was generally less intense in OSNs from trichoid sensilla as compared to neurons from sensilla basiconica (Fig. 4D).

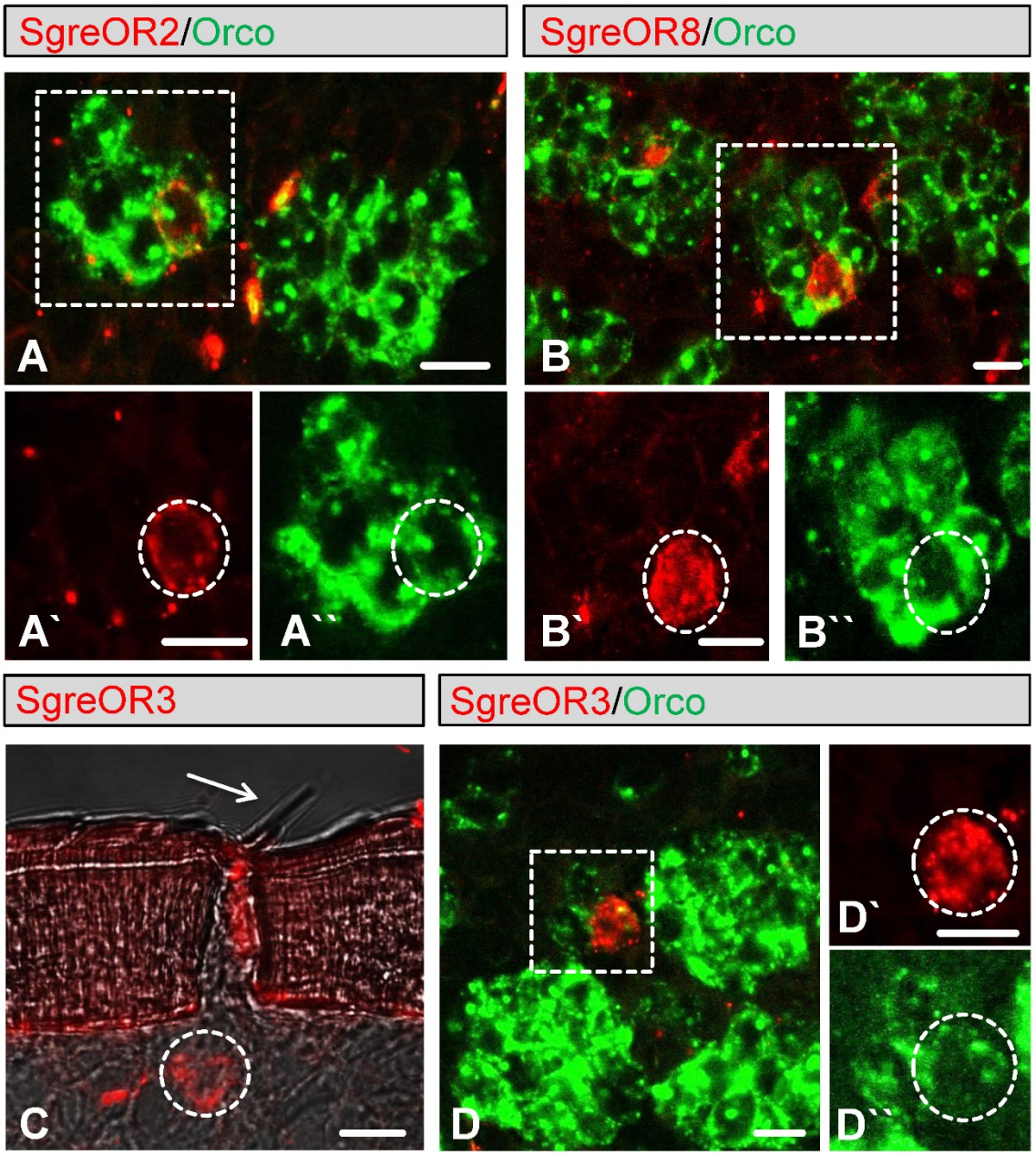


Fig. 4. Expression of distinct b-OR types in Orco-positive OSNs of sensilla basiconica or sensilla trichodea. (**A**-**B**) Two-color FISH experiments on sections through antennae were conducted with labeled antisense RNA probes for Orco (A-B, green) and SgreOR2 (A, red) or SgreOR8 (B, red). Higher magnifications of the regions indicated by dashed boxes are shown for the green and the red color channel in A`-A`` and B`-B``. Cells expressing SgreOR2 or SgreOR8 are located in the Orco-positive neuron clusters characteristic of sensilla basiconica. (**C**) FISH on antennal tissue employing an antisense RNA probe for SgreOR3 (red). A SgreOR3-positive labeled cell (dashed circle) is located directly below a sensillum trichodeum (arrow). (**D**) Two-color FISH with antisense probes for SgreOR3 (red) and Orco (green). As exemplarily depicted, SgreOR3-positive neurons were only found within small groups of Orco-expressing cells that are typical for sensilla trichodea. The area circumscribed by the dotted box in D is given in the green and the red color channel in D`-D``. Scale bars: 10 µm.

Antennal expression and co-expression with SNMP1 of SgreOR types outside the b-OR group

To compare the expression pattern of the b-OR group members with SgreOR types not belonging to this group and to get a more detailed insight into the expression of SgreORs in the antennae of *Schistocerca gregaria*, we designed specific primer pairs for 28 SgreOR types (indicated in black font color in table 1) beyond the b-OR group and used them for PCR experiments with antennal cDNA from *Schistocerca gregaria*. Cloning and sequencing of the obtained PCR products demonstrated that they encoded the relevant receptors, further supporting expression of these SgreOR types in antennae of *Schistocerca gregaria*. Next, we prepared specific RNA probes for several of these SgreOR types and conducted in situ

hybridizations. Similar to the results obtained for b-ORs (Fig. 2), these experiments revealed expression of distinct SgreOR types in different numbers of cells (Fig. 5). While a few of these SgreORs tested were expressed in a relatively high number of cells (as exemplarily shown for SgreOR17 and SgreOR35 in Fig. 5A-B), for most of them, only lower numbers of stained cells were found per antennal section (as exemplarily depicted for SgreOR31 and SgreOR76 in Fig. 5C-D).

To test a potential co-expression of receptors outside the b-OR group with SNMP1, two-color FISH experiments were conducted with probes for SNMP1 and the 28 above-mentioned SgreOR types (these are also denoted by green squares in Fig. 1) that represent different branches of the neighbor-joining tree. The findings of these approaches are summarized in table 1. For most of these SgreOR types, no co-expression with SNMP1 was detectable. This is exemplarily shown for SgreOR17, SgreOR31, SgreOR35, SgreOR76 (Fig. 5E-H) and SgreOR33, SgreOR45, SgreOR67, SgreOR82 (supplementary Fig. 3). However, for 5 of these SgreORs (SgreOR84, SgreOR86, SgreOR93, SgreOR94 and SgreOR97) co-expression with SNMP1 was observed (data not shown), indicating that there are SgreOR types outside the b-OR group that can be considered as candidate pheromone receptors. Yet, the percentage of SgreORs outside the b-OR group that are co-expressed with SNMP1 (5 out of 28 receptors tested; ~22%) was clearly lower than for the members of the b-OR group (6 out of 7 receptors tested; ~86%).

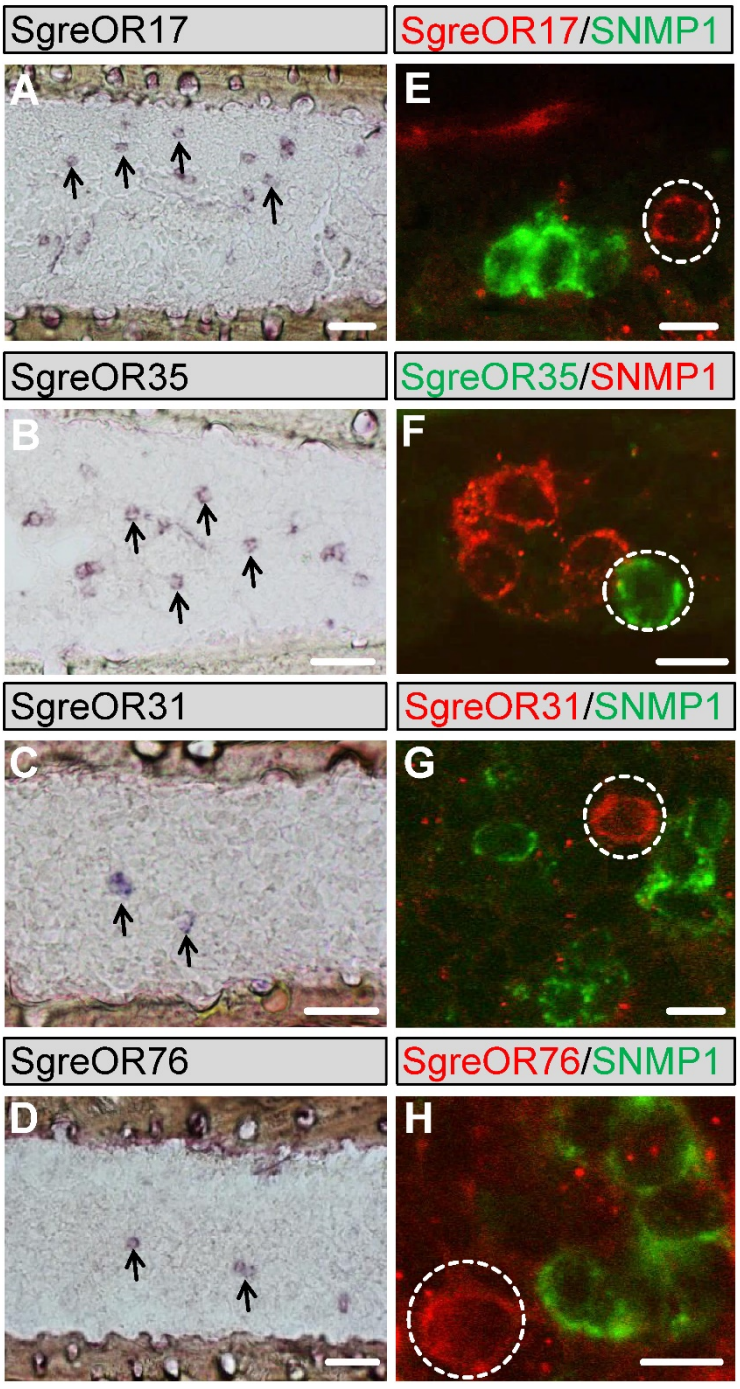


Fig. 5. Expression of SgreOR types outside the b-OR group in antennae of *Schistocerca gregaria*. (**A-D**) Non-fluorescent in situ hybridizations were conducted with probes for SgreOR17, SgreOR35, SgreOR31 and SgreOR76 on longitudinal antennal sections. Some of the stained cells are exemplarily denoted by arrows. (**E-H**) Two-color FISH experiments were performed with probes for the above-mentioned SgreOR types and SNMP1. Cells expressing these SgreORs (denoted by dashed circles) do not co-express SNMP1. Scale bars: A-C 40 µm; E-F 10 µm.

Table 1. Summary of the examined SgreOR types co-expressed or not co-expressed with SNMP1. According to the results of two-color FISH experiments, 6 of the 7 b-OR types (highlighted as *italic* in this table) tested (~86%) were found to be co-expressed with SNMP1 (only SgreOR4 was not co-expressed with SNMP1). Regarding the SgreORs outside the b-OR group, 5 of 28 receptors examined (~22%) were co-expressed with SNMP1.

co-expressed with SNMP1		not co-expressed with SNMP1		
OR2	*OR3*	*OR4*	OR14	OR15
OR5	*OR6*	OR17	OR31	OR33
OR8	*OR9*	OR35	OR39	OR41
OR84	OR86	OR43	OR45	OR47
OR93	OR94	OR49	OR51	OR53
OR97		OR57	OR67	OR70
		OR76	OR82	OR83
		OR99	OR100	OR105

Discussion

In the present study, attempts have been made to identify candidate pheromone receptors of the desert locust *Schistocerca gregaria*. So far, very little is known about pheromone signaling in locusts. Therefore,

we have used a novel approach based on the concept that pheromone-responsive OSNs in insects express SNMP1 (20,21,24). Since its discovery (46), SNMP1 has been considered as an accessory protein intimately involved in the recognition of pheromones (47,53). This view was supported by recent findings underlining the importance of SNMP1 for the sensitivity and kinetics of a pheromone detection system (25,54,55). These observations have led to the concept of a tunneling mechanism that transfers pheromones via SNMP1 to specific pheromone receptors implying an interplay between SNMP1 and receptors for pheromones (25). Consequently, ORs co-expressed with SNMP1 in antennal neurons may be considered as candidate pheromone receptors. In this study, we have identified 119 candidate OR genes from *Schistocerca gregaria*. Performing a phylogenetic comparison of the encoded OR proteins with ORs from *Locusta migratoria*, a relatively small group of receptors appeared to be unique due to its more basal position in the tree. Although the phylogenetic separation of these basal ORs (b-ORs) from other SgreORs is reminiscent of the results described for pheromone receptors in moth species (15,51), no evident sequence identities between b-ORs and pheromone receptors from other insects were found. Overall, the missing sequence identities to moth and fly pheromone receptors do not argue against a potential function of b-ORs in pheromone reception and could be due to the large phylogenetic distance between the hemimetabolic Orthoptera (including locusts) on the one hand and the holometabolic Lepidoptera and Diptera on the other hand (33).

The results of two-color FISH experiments revealed that most of the analyzed b-ORs were co-expressed with SNMP1 and may therefore be considered as candidate pheromone receptors. Although the current data strongly support a role of SNMP1 in pheromone detection (20,22,25,54,55), we are aware that verification of the SNMP1-co-expressed SgreOR types as pheromone receptors requires functional analyses. Thus, we cannot exclude the possibility that some of the SgreORs co-expressed with SNMP1 are used to detect non-pheromonal compounds. In this context, a member of the b-OR group from *Locusta migratoria* [named LmigOR3 (56); corresponding to LmigOr7 in the nomenclature used in the present study] has been recently found to be activated by food odorants (57). However, it remains unclear whether its orthologue from *Schistocerca gregaria* (SgreOR7) is co-expressed with SNMP1 since our two-color FISH experiments on antennal tissue using probes for SgreOR7 and SNMP1 were not successful.

Unlike pheromone receptors of moth and fly species (20–22,55), *Schistocerca gregaria* OSNs co-expressing SNMP1 and b-OR types were not primarily confined to trichoid sensilla. Instead, expression of most b-ORs was observed in OSNs beneath sensilla basiconica. This finding is however in accordance with previous electrophysiological recordings from *Schistocerca gregaria* antennae that demonstrated the responsiveness of both sensilla trichodea and sensilla basiconica to pheromone compounds (38).

In the desert locust, the remarkable change between the solitary and the gregarious phase seems to be controlled by signaling compounds (1). Thus, in *Schistocerca gregaria*, pheromones probably do not only play a central role in the reproduction process (31,32) but also may be essential for the phase change. In line with the existence of various pheromonal compounds for eliciting different behaviors or priming diverse physiological effects in desert locusts, our two-color FISH experiments revealed further SgreOR types that alike the b-OR group members were co-expressed with SNMP1 and thus can also be considered as candidate pheromone receptors. While the b-ORs are more closely related to Orco than to other members of the SgreOR family, almost half of the putative pheromone receptors from *Schistocerca gregaria* identified in this study (SgreOR84, SgreOR86, SgreOR93, SgreOR94 and SgreOR97; table 1) belong to an OR group that is only distantly related to Orco (Fig. 1). It remains elusive why *Schistocerca gregaria* has two distinct groups of candidate pheromone receptors. Potentially, these two groups are dedicated to the detection of distinct types of pheromones.

In conclusion, uncovering candidate pheromone receptors of locusts could open the door for extended future studies including the screening for novel pheromonal compounds using heterologously expressed candidate pheromone receptors as well as the identification of receptors for previously verified pheromones (1,31,36). Finally, in this context, the identification of SgreOR types mediating the detection of pheromones involved in reproduction and/or aggregation might provide important targets for inhibiting the formation of huge locust swarms that massively threaten agricultural crops in numerous countries.

Abbreviations

b-ORs: odorant receptors of the basal group; bp: base pairs; FISH: fluorescent in situ hybridization; LmigOrs: odorant receptors of *Locusta migratoria*; ORs: odorant receptors; OSNs: olfactory sensory neurons; SgreORs: odorant receptors of *Schistocerca gregaria*; SNMP1: sensory neuron membrane protein 1.

Supplementary Material

Supplementary figures and tables.
http://www.ijbs.com/v13p0911s1.pdf

Acknowledgements

The authors would like to thank Heidrun Froß and Kathrin Dietze for excellent technical assistance.

Funding

Xingcong Jiang was supported by a grant financed by the China Scholarship Council (CSC).

Data deposition

The SgreOR sequences reported in this study have been deposited in GenBank [accession numbers KY964918 (SgreOR1) through KY965036 (SgreOR119)].

Competing Interests

The authors have declared that no competing interest exists.

References

1. Hassanali A, Njagi PGN, Bashir MO. Chemical ecology of locusts and related acridids. Annu Rev Entomol. 2005; 50: 223-45.
2. Hansson BS, Stensmyr MC. Evolution of insect olfaction. Neuron. 2011; 72: 698–711.
3. Krieger J, Grosse-Wilde E, Gohl T, Dewer YME, Raming K, Breer H. Genes encoding candidate pheromone receptors in a moth (*Heliothis virescens*). Proc Natl Acad Sci U S A. 2004;101: 11845–50.
4. Nakagawa T, Sakurai T, Nishioka T, Touhara K. Insect sex-pheromone signals mediated by specific combinations of olfactory receptors. Science. 2005; 307: 1638–42.
5. Grosse-Wilde E, Gohl T, Bouche E, Breer H, Krieger J. Candidate pheromone receptors provide the basis for the response of distinct antennal neurons to pheromonal compounds. Eur J Neurosci. 2007; 25: 2364–73.
6. Kurtovic A, Widmer A, Dickson BJ. A single class of olfactory neurons mediates behavioural responses to a Drosophila sex pheromone. Nature. 2007; 446: 542–6.
7. Almaas TJ, Mustaparta H. *Heliothis virescens*: Response characteristics of receptor neurons in Sensilla Trichodea type 1 and type 2. J Chem Ecol. 1991; 17: 953–72.
8. Baker TC, Ochieng SA, Cosse AA, Lee SG, Todd JL, Quero C, Vickers NJ. A comparison of responses from olfactory receptor neurons of *Heliothis subflexa* and *Heliothis virescens* to components of their sex pheromone. J Comp Physiol A. 2004; 190: 155–65.
9. Keil T, Steinbrecht RA. Mechanosensitive and olfactory sensilla of insects. In: King R, Akai H, eds. Insect Ultrastructure, Volume 2, New York: Springer US; 1984: 477–516.
10. Steinbrecht RA, Gnatzy W. Pheromone receptors in *Bombyx mori* and *Antheraea pernyi*. I. Reconstruction of the cellular organization of the sensilla trichodea. Cell Tissue Res. 1984; 235: 25–34.
11. Keil A. Fine structure of the pheromone-sensitive sensilla on the antenna of the hawkmoth, *Manduca sexta*. Tissue Cell. 1989; 21: 139-51.
12. Clyne P, Grant A, O'Connell R, Carlson JR. Odorant response of individual sensilla on the Drosophila antenna. Invert Neurosci. 1997; 3: 127–35.
13. Couto A, Alenius M, Dickson BJ. Molecular, Anatomical, and Functional Organization of the Drosophila Olfactory System. Curr Biol. 2005; 15: 1535–47.
14. Sakurai T, Nakagawa T, Mitsuno H, Mori H, Endo Y, Tanoue S, Yasukochi Y, Touhara K, Nishioka T. Identification and functional characterization of a sex pheromone receptor in the silkmoth *Bombyx mori*. Proc Natl Acad Sci U S A. 2004; 101: 16653–8.
15. Krieger J, Grosse-Wilde E, Gohl T, Breer H. Candidate pheromone receptors of the silkmoth *Bombyx mori*. Eur J Neurosci. 2005; 21: 2167–76.
16. Krieger J, Klink O, Mohl C, Raming K, Breer H. A candidate olfactory receptor subtype highly conserved across different insect orders. J Comp Physiol A. 2003; 189: 519–26.
17. Larsson MC, Domingos AI, Jones WD, Chiappe ME, Amrein H, Vosshall LB. Or83b encodes a broadly expressed odorant receptor essential for Drosophila olfaction. Neuron. 2004; 43: 703–14.
18. Sato K, Pellegrino M, Nakagawa T, Nakagawa T, Vosshall LB, Touhara K. Insect olfactory receptors are heteromeric ligand-gated ion channels. Nature. 2008; 452: 1002–6.
19. Wicher D, Schäfer R, Bauernfeind R, Stensmyr MC, Heller R, Heinemann SH, Hansson BS. Drosophila odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. Nature. 2008; 452: 1007–11.
20. Benton R, Vannice KS, Vosshall LB. An essential role for a CD36-related receptor in pheromone detection in Drosophila. Nature. 2007; 450: 289–93.
21. Forstner M, Gohl T, Gondesen I, Raming K, Breer H, Krieger J. Differential expression of SNMP-1 and SNMP-2 proteins in pheromone-sensitive hairs of moths. Chem Senses. 2008; 33: 291–9.
22. Jin X, Ha TS, Smith DP. SNMP is a signaling component required for pheromone sensitivity in Drosophila. Proc Natl Acad Sci U S A. 2008;105: 10996-1001.
23. Liu S, Zhang YR, Zhou WW, Liang QM, Yuan X, Cheng J, Zhu ZR, Gong ZJ. Identification and characterization of two sensory neuron membrane proteins from *Cnaphalocrocis medinalis* (Lepidoptera: Pyralidae). Arch Insect Biochem Physiol. 2013; 82: 29–42.
24. Zhang J, Liu Y, Walker WB, Dong S-L, Wang G-R. Identification and localization of two sensory neuron membrane proteins from *Spodoptera litura* (Lepidoptera: Noctuidae). Insect Sci. 2015; 22: 399–408.
25. Gomez-Diaz C, Bargeton B, Abuin L, Bukar N, Reina JH, Bartoi T, Graf M, Ong H, Ulbrich MH, Masson JF, Benton R. A CD36 ectodomain mediates insect pheromone detection via a putative tunnelling mechanism. Nat Commun. 2016; 7: 11866.
26. Fuzeau-Braesch S, Genin E, Jullien R, Knowles E, Papin C. Composition and role of volatile substances in atmosphere surrounding two gregarious locusts, *Locusta migratoria* and *Schistocerca gregaria*. J Chem Ecol. 1988; 14: 1023–33.
27. Byers JA. Pheromones and chemical ecology of locusts. Biol Rev. 1991; 66: 347–78.
28. Obeng-Ofori D, Torto B, Njagi PGN, Hassanali A, Amiani H. Fecal volatiles as part of the aggregation pheromone complex of the desert locust, *Schistocerca gregaria* (Forskal) (Orthoptera: Acrididae). J Chem Ecol. 1994; 20: 2077–87.
29. Torto B, Obeng-Ofori D, Njagi PGN, Hassanali A, Amiani H. Aggregation pheromone system of adult gregarious desert locust *Schistocerca gregaria* (forskal). J Chem Ecol. 1994; 20: 1749–62.
30. Pener MP, Yerushalmi Y. The physiology of locust phase polymorphism: an update. J Insect Physiol. 1998; 44: 365–77.
31. Seidelmann K, Ferenz H-J. Courtship inhibition pheromone in desert locusts, *Schistocerca gregaria*. J Insect Physiol. 2002; 48: 991–6.
32. Stahr C, Svatos A, Seidelmann K. Chemical identification, emission pattern and function of male-specific pheromones released by a rarely swarming locust, *Schistocerca americana*. J Chem Ecol. 2013; 39: 15–27.
33. Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, et al. Phylogenomics resolves the timing and pattern of insect evolution. Science. 2014; 346: 763-7.
34. Simpson SJ, Sword GA. Locusts. Curr Biol. 2008; 18: R364-6.
35. Heifetz Y, Voet H, Applebaum S. Factors affecting behavioral phase transition in the desert locust *Schistocerca gregaria* (Forskal) (Orthoptera: Acrididae). J Chem Ecol. 1996; 22: 1717–34.
36. Seidelmann K, Weinert H, Ferenz H-J. Wings and legs are production sites for the desert locust courtship-inhibition pheromone, phenylacetonitrile. J Insect Physiol. 2003; 49: 1125–33.
37. Hansson BS, Ochieng SA, Grosmaitre X, Anton S, Njagi PGN. Physiological responses and central nervous projections of antennal olfactory receptor neurons in the adult desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae). J Comp Physiol A. 1996; 179: 157–67.
38. Ochieng SA, Hansson BS. Responses of olfactory receptor neurones to behaviourally important odours in gregarious and solitarious desert locust, *Schistocerca gregaria*. Physiol Entomol. 1999; 24:28–36.
39. Cui X, Wu C, Zhang L. Electrophysiological response patterns of 16 olfactory neurons from the trichoid sensilla to odorant from fecal volatiles in the locust, *Locusta migratoria manilensis*. Arch Insect Biochem Physiol. 2011; 77: 45–57.
40. Ochieng SA, Hallberg E, Hansson BS. Fine structure and distribution of antennal sensilla of the desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae). Cell Tissue Res. 1998; 291: 525–36.
41. Yang Y, Krieger J, Zhang L, Breer H. The olfactory co-receptor Orco from the migratory locust (*Locusta migratoria*) and the desert locust (*Schistocerca gregaria*): identification and expression pattern. Int J Biol Sci. 2012; 8: 159–70.
42. Jiang X, Pregitzer P, Grosse-Wilde E, Breer H, Krieger J. Identification and Characterization of Two "Sensory Neuron Membrane Proteins" (SNMPs) of the Desert Locust, *Schistocerca gregaria* (Orthoptera: Acrididae). J Insect Sci. 2016; 16: 33.
43. Guo M, Krieger J, Große-Wilde E, Mißbach C, Zhang L, Breer H. Variant ionotropic receptors are expressed in olfactory sensory neurons of coeloconic sensilla on the antenna of the desert locust (*Schistocerca gregaria*). Int J Biol Sci. 2014; 10: 1–14.
44. Benton R, Vannice KS, Gomez-Diaz C, Vosshall LB. Variant ionotropic glutamate receptors as chemosensory receptors in Drosophila. Cell. 2009; 136: 149–62.
45. Rytz R, Croset V, Benton R. Ionotropic Receptors (IRs): Chemosensory ionotropic glutamate receptors in Drosophila and beyond. Insect Biochem Mol Biol. 2013; 43: 888–97.

46. Rogers ME, Sun M, Lerner MR, Vogt RG. Snmp-1, a novel membrane protein of olfactory neurons of the silk moth *Antheraea polyphemus* with homology to the CD36 family of membrane proteins. J Biol Chem. 1997; 272: 14792–9.
47. Rogers ME, Steinbrecht RA, Vogt RG. Expression of SNMP-1 in olfactory neurons and sensilla of male and female antennae of the silkmoth *Antheraea polyphemus*. Cell Tissue Res. 2001; 303: 433–46.
48. Wang Z, Yang P, Chen D, Jiang F, Li Y, Wang X, Kang L. Identification and functional analysis of olfactory receptor family reveal unusual characteristics of the olfactory system in the migratory locust. Cell Mol Life Sci. 2015; 72: 4429–43.
49. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol. 2013; 30: 2725-9.
50. Krieger J, Raming K, Dewer YME, Bette S, Conzelmann S, Breer H. A divergent gene family encoding candidate olfactory receptors of the moth *Heliothis virescens*. Eur J Neurosci. 2002; 16: 619–28.
51. Grosse-Wilde E, Kuebler LS, Bucks S, Vogel H, Wicher D, Hansson BS. Antennal transcriptome of *Manduca sexta*. Proc Natl Acad Sci U S A. 2011; 108: 7449–54.
52. Legeai F, Malpel S, Montagne N, Monsempes C, Cousserans F, Merlin C, François MC, Maibeche-Coisne M, Gavory F, Poulain J, Jacquin-Joly E. An Expressed Sequence Tag collection from the male antennae of the Noctuid moth *Spodoptera littoralis*: a resource for olfactory and pheromone detection research. BMC Genomics. 2011; 12: 86.
53. Vogt RG, Miller NE, Litvack R, Fandino RA, Sparks J, Staples J, Friedman R, Dickens JC. The insect SNMP gene family. Insect Biochem Mol Biol. 2009; 39: 448–56.
54. Li Z, Ni JD, Huang J, Montell C. Requirement for Drosophila SNMP1 for rapid activation and termination of pheromone-induced activity. PLoS Genet. 2014; 10: e1004600.
55. Pregitzer P, Greschista M, Breer H, Krieger J. The sensory neurone membrane protein SNMP1 contributes to the sensitivity of a pheromone detection system. Insect Mol Biol. 2014; 23: 733-42.
56. Xu H, Guo M, Yang Y, You Y, Zhang L. Differential expression of two novel odorant receptors in the locust (*Locusta migratoria*). BMC Neurosci. 2013; 14: 50.
57. You Y, Smith DP, Lv M, Zhang L. A broadly tuned odorant receptor in neurons of trichoid sensilla in locust, *Locusta migratoria*. Insect Biochem Mol Biol. 2016; pii: S0965-1748(16)30155-2.

frontiers in Physiology

ORIGINAL RESEARCH
published: 26 September 2017
doi: 10.3389/fphys.2017.00734

Distinct Subfamilies of Odorant Binding Proteins in Locust (Orthoptera, Acrididae): Molecular Evolution, Structural Variation, and Sensilla-Specific Expression

Xingcong Jiang[1], *Jürgen Krieger*[2], *Heinz Breer*[1] *and Pablo Pregitzer*[1]*

[1] *Institute of Physiology, University of Hohenheim, Stuttgart, Germany,* [2] *Department of Animal Physiology, Institute of Biology/Zoology, Martin Luther University Halle-Wittenberg, Halle, Germany*

Odorant binding proteins (OBPs) play an important role in insect olfaction, facilitating transportation of odorant molecules in the sensillum lymph. While most of the researches are concentrated on Lepidopteran and Dipteran species, our knowledge about Orthopteran species is still very limited. In this study, we have investigated OBPs of the desert locust *Schistocerca gregaria*, a representative Orthopteran species. We have identified 14 transcripts from a *S. gregaria* antennal transcriptome encoding SgreOBPs, and recapitulated the phylogenetic relationship of SgreOBPs together with OBPs from three other locust species. Two conserved subfamilies of classic OBPs have been identified, named I-A and II-A, exhibiting both common and subfamily-specific amino acid motifs. Distinct evolutionary features were observed for subfamily I-A and II-A OBPs. Surface topology and interior cavity were elucidated for OBP members from the two subfamilies. Antennal topographic expression revealed distinct sensilla- and cellular- specific expression patterns for SgreOBPs from subfamily I-A and II-A. These findings give first insight into the repertoire of locust OBPs with respect to their molecular and evolutionary features as well as their expression in the antenna, which may serve as an initial step to unravel specific roles of distinct OBP subfamilies in locust olfaction.

Keywords: locust, *Schistocerca gregaria*, odorant binding protein, evolution, structure, sensilla

OPEN ACCESS

Edited by:
Shuang-Lin Dong,
Nanjing Agricultural University, China

Reviewed by:
Paolo Pelosi,
University of Pisa, Italy
Dan-Dan Zhang,
Lund University, Sweden
Long Zhang,
China Agricultural University, China

***Correspondence:**
Pablo Pregitzer
p_pregitzer@uni-hohenheim.de

Specialty section:
This article was submitted to Invertebrate Physiology, a section of the journal Frontiers in Physiology

Received: *20 July 2017*
Accepted: *11 September 2017*
Published: *26 September 2017*

Citation:
Jiang X, Krieger J, Breer H and Pregitzer P (2017) Distinct Subfamilies of Odorant Binding Proteins in Locust (Orthoptera, Acrididae): Molecular Evolution, Structural Variation, and Sensilla-Specific Expression. Front. Physiol. 8:734. doi: 10.3389/fphys.2017.00734

INTRODUCTION

In insects, the process of olfactory signal processing begins in hair-like cuticle appendages, called sensilla, located mainly on the antennae and palps (Steinbrecht, 1996; Hansson and Stensmyr, 2011; Suh et al., 2014). Olfactory sensory neurons (OSNs) project their dendrites into the lumen of the sensillar hairs, which is filled with sensillum lymph (Hansson and Stensmyr, 2011; Suh et al., 2014). The hydrophobic odorant molecules enter the sensillum via the porous cuticle and have to pass the aqueous lymph till reaching the chemosensory membrane of the sensory neurons (Vogt et al., 1999; Leal, 2013; Suh et al., 2014). This passage is thought to be mediated by small soluble proteins enriched in the sensilla lymph, the so called odorant binding proteins (OBPs), which are produced and secreted by accessory cells (Pelosi et al., 2006, 2017). OBPs are polypeptides comprised of

∼110–200 amino acids; usually they exhibit a considerable degree of sequence divergence. Based on the number of conserved cysteine (C)-residues, several subtypes are discriminated. Whereas, the pattern of six conserved C-residues represents a hallmark of classic OBPs (Pelosi et al., 2006), OBPs with more or with less C-residues are designated as plus-C and minus-C OBPs (Zhou et al., 2004; Foret and Maleszka, 2006). In addition, atypical OBPs have been classified which may originate from a fusion of two classic OBPs (Xu et al., 2003; Vieira and Rozas, 2011). Typically, the tertiary structure of insect OBPs consists of six α-helices forming an interior binding cavity. This structure is maintained and stabilized by disulfide bridges formed by conserved C-residues (Leal et al., 1999; Scaloni et al., 1999; Sandler et al., 2000). However, OBP structures with more than six helices have been reported (Horst et al., 2001; Lagarde et al., 2011). It is also proposed that the C-terminal domain that is variable in length can spatially interfere with the interior binding cavity and thus may affect the ligand binding mechanism (Damberger et al., 2000; Horst et al., 2001; Tegoni et al., 2004; Pelosi et al., 2017).

Most of our current knowledge of insect OBPs is based on studies of species from the taxa Lepidoptera and Diptera (Hekmat-Scafe et al., 2002; Vogt et al., 2002; Leal, 2013; Pelosi et al., 2017). The desert locust, *Schistocerca gregaria* is a representative of the taxa Orthoptera, which is quite distant from the orders Lepidoptera and Diptera on the phylogenetic scale (Wheeler et al., 2001; Vogt et al., 2015) and as hemimetabolous insects their developmental process differ significantly from that of holometabolous insects. Very little is known about OBPs in Orthoptera; only a limited number of sequences have recently been reported for a few locust species: *Locusta migratoria* (Ban et al., 2003; Xu et al., 2009; Yu et al., 2009), *Oedaleus asiaticus* (Zhang et al., 2015), and *Ceracris kiangsu*. Information about the expression of locust OBPs in the olfactory sensilla is limited to LmigOBP1, which was found to be expressed in sensilla trichodea and sensilla basiconica (Jin et al., 2005). Concerning another olfactory sensillum type, the sensilla coeloconica, a possible expression of OBPs has rarely been documented even in holometabolous insect species (Larter et al., 2016). Incidentally, the crystal structure of locust OBPs has only been resolved for LmigOBP1, which establishes a unique seven-α-helices structure (Zheng et al., 2015). The possibility of structural differences between locust OBPs is still an open question.

In order to extend our knowledge about OBPs in Orthopteran locust species, in the current study we have performed a systematic characterization of locust OBPs with respect to the molecular evolution, structural variation and sensilla-specific expression. Based on the OBP sequences of *S. gregaria* newly identified from an antennal transcriptome and the documented OBP sequences from other locust species, we conducted a phylogenetic analysis of the current locust OBP repertoire. The emerging two subfamilies of classic OBPs were compared for sequence divergence, selection pressure and variation of the predicted tertiary structure in detail. Analysis of the topographic expression pattern revealed that the molecular and phylogenetic distinctness between the two subfamilies are accompanied by a sensilla-specific expression pattern.

MATERIALS AND METHODS

Identification of *S. gregaria* OBP Transcripts

A *S. gregaria* antennal transcriptome database was generated comprising a total of 55,060 contigs with an N50 of 2,223 bp. The strategy of homology-mining was adopted to identify the candidate OBP transcripts. We retrieved documented OBPs from different insect species including *Anopheles gambiae* (AgamOBPs, Diptera), *Apis mellifera* (AmelOBPs, Hemiptera), *Drosophila melanogaster* (DmelOBPs, Diptera), *Tribolium castaneum* (TcasOBPs, Coleoptera), *Acyrthosyphon pisum* (ApisOBPs, Hemiptera), *Bombyx mori* (BmorOBPs, Lepidoptera) (Vieira and Rozas, 2011), *Blattella germanica* (BgerOBPs, Blattaria) (Niu et al., 2016), and *Zootermopsis nevadensis* (ZnevOBPs, Isoptera) (Terrapon et al., 2014), as well as from three other locust species, including *L. migratoria* (LmigOBPs) (Ban et al., 2003; Yu et al., 2009), *O. asiaticus* (OasiOBPs) (Zhang et al., 2015), and *C. kiangsu* (CkiaOBPs). Using the collected sequences as queries, we conducted a local tBLASTx search on BioEdit 7.2.5 against the transcriptome database with an *E*-value $< 10^{-5}$. Annotation of the screened contigs was inspected by performing tBLASTx and BLASTp search against non-redundant (nr) protein database in NCBI (Bethesda, MD, USA). The extracted contigs which putatively encode OBPs were in turn used as new queries to identify additional candidates using tBLASTx and BLASTp methods. Open reading frames in the identified OBP transcripts were inspected by Genamics Expression (Hamilton, New Zealand). Accession numbers for the newly identified SgreOBPs and other locust OBPs are deposited in the Supplementary Material.

Characterization of Consensus Amino Acid Motifs

Signatures of sequence divergence underlying locust subfamily I-A and II-A OBPs were addressed by identifying consensus amino acid motifs. Toward that goal, the online MEME SUITE v. 4.11.2 (http://meme-suite.org/tools/meme) was used (Bailey et al., 2009), with the default setting (motif width: 6–50 amino acids; motif distribution: zero or one occurrence per sequence). The output comprised six consensus motifs which was ascertained to be sufficient to recapitulate the sequence information of subfamily I-A and II-A. The identified six motifs were also utilized to target sequences of the locust OBP repertoire to obtain the motif match degree (match *E*-value) using MAST module (Motif Alignment and Search Tool) implemented in MEME SUITE. The motif match *E*-value assesses statistical significance of the consensus motif toward a targeted sequence based on its log likelihood level and the occurrence frequencies of background amino acids. The default statistical significant threshold setting was e^{-5}.

Phylogenetic Analysis

The OBP amino acid sequences from four hitherto documented locust species were utilized to recapitulate the phylogenetic relationship: 16 from *L. migratoria*, 15 from *O. asiaticus*, 7 from *C. kiangsu* and the currently identified 14 candidates

from *S. gregaria*. Amino acid sequences of OBPs from the four locust species are deposited in the Supplementary Material. The predicted signal peptide (SP) on the N-terminal domain was deleted before the sequences being further investigated due to two reasons: (1) SP is cut off in post-translational modification when the protein is secreted into the sensillum lymph; (2) SP exhibits a certain degree of sequence divergence but may contain limited bio-information (Vieira et al., 2007). Prediction of SP was based on SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/) (Petersen et al., 2011). Multiple sequence alignments were conducted by MAFFT v. 7 (http://mafft.cbrc.jp/alignment/server/) using the algorithm E-INS-I, which is accuracy favored and is suitable for sequences with multiple conserved domains (Katoh and Standley, 2013). After the alignment, Gblocks v. 0.91b (http://molevol.cmima.csic.es/castresana/Gblocks_server.html) was used to inspect the poorly aligned sites and divergent regions (Castresana, 2000). To search an optimal amino acid substitution model, we chose the Find Best Protein Model implemented in MEGA 6.0 which performs a comprehensive parametric assessment (e.g., BIC scores, AICc value, lnL value) (Tamura et al., 2013). The Whelan and Goldman model (WAF), discrete GAMMA distribution (G) and an assumed fraction of evolutionary invariable sites (I) was considered to describe the substitution best. RAxML v. 8.2.9 implemented in the CIPRES Science Gateway (https://www.phylo.org/) was used for the locust OBP phylogeny construction (Miller et al., 2012; Stamatakis, 2014). A search of best scoring maximum likelihood tree (-f a) was launched, supported by 1,000 rapid bootstrap iterations (autoMRE based bootstopping criterion). The generated maximum likelihood tree was graphically edited by FigTree v. 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/). A similar strategy was applied to analyze the phylogenetic relationship between locust OBPs and OBPs from eight other insect species. In brief, SignalP, MAFFT, and Gblocks were used to prepare the multiple sequence alignment; RAxML was responsible for building the maximum likelihood tree (-f a, 1,000 iteration) using the proposed best fitting substitution model (WAG+G+I) by MEGA.

Selection Constraint on Locust OBP Repertoire

The nucleotide coding sequences underlying the locust OBP repertoire (see Supplementary Material) were aligned in accordance with the multiple sequence alignment from the above mentioned phylogenetic analysis using TranslatorX (http://translatorx.co.uk/). The sequence order of alignment was guided by the constructed phylogenetic tree mentioned above. The signatures of selection regime acting on sequences of the locust OBP phylogeny were estimated by resolving three principle concepts: the non-synonymous substitution rate (dN), synonymous substitution rate (dS) and the ω rate (dN/dS). Toward that, HyPhy batch program was utilized which implements maximum likelihood estimate and post-likelihood ratio test (Kosakovsky Pond et al., 2005). A local fit model (MG94xREV_3x4 substitution model) was adopted (Kosakovsky Pond et al., 2009), and each single branch in the locust OBP phylogeny was assigned with a unique set of dN and dS values, assuming the branch-to-branch variant ω rates. To support the local fit model, we additionally conducted a coarse estimate of the ω rate using the alternative global fit model, assuming invariable ω rate shared by different phylogenetic branches. A likelihood ratio test compared the results obtained from two distinct models, and strongly favored the local fit model ($P = 10^{-3}$). Normality distribution of dN, dS, and the ω rates was assessed by D'Agostino-Pearson test, and the statistical difference was evaluated by non-parametric Mann-Whitney *U*-test. GraphPad Prism 5.0 was used to analyze the data and generate the diagrams (San Diego, CA, USA).

Synthesis of Riboprobes For *in Situ* Hybridization

The coding sequences of six SgreOBPs from locust OBP subfamily I-A and II-A were amplified, sequenced and then cloned into the pGEM-T vectors (Invitrogen) for subsequent transcription. Linearized pGEM-T vectors carrying SgreOBPs coding sequences were utilized to synthesize digoxigenin (Dig) and biotin (Bio) labeled anti-sense and sense RNA probes using the T7/SP6 RNA transcription system (Roche, Germany). The sense (s) and antisense (as) primers used for amplication of the SgreOBP sequences were:

SgreOBP1 s, ctgggacgtcaacatgaaact;
SgreOBP1 as, aatgcacgaactaccaggctg;
SgreOBP5 s, ggccgcgccgtcttctcataagga;
SgreOBP5 as, cggccctggcgcagcacctgcatt;
SgreOBP6 s, acagcacaccaccgtcacac;
SgreOBP6 as, ggtgcttgcttgaagaggcac;
SgreOBP10 s, gcgtatcacccggctgtgta;
SgreOBP10 as, agtctcacctctgccagcga;
SgreOBP11 s, tggaccgcacgacaacaaca;
SgreOBP11 as, cgatagcgtatgcccttcac;
SgreOBP14 s, ctgttgggtgcagtcctgtt;
SgreOBP14 as, gtcgtgacagctcctccactg

In Situ Hybridization

Antennae of adult *S. gregaria* were dissected and embedded in Tissue-Tek O.C.T. Compound (Sakura Finetek Europe, The Netherlands). Cryosections at 12 μm were thaw mounted on SuperFrost Plus slides (Menzel-Gläser, Braunschweig, Germany) at −21°C (Jung CM300 cryostat). RNA *in situ* hybridization (ISH) was conducted as previously reported (Yang et al., 2012; Guo et al., 2013; Jiang et al., 2016). Section were fixed (4% paraformaldehyde in 0.1 M $NaHCO_3$, pH 9.5) at 4°C for 22 min. The following consecutive steps were conducted at room temperature: a wash for 1 min in PBS (phosphate buffered saline = 0.85% NaCl, 1.4 mM KH_2PO_4, 8 mM Na_2HPO_4, pH 7.1), an incubation for 10 min in 0.2 M HCl, another wash for 1 min in PBS, an incubation for 10 min in acetylation solution (0.25% acetic anhydride freshly added in 0.1 M triethanolamine) and washes for three times in PBS (3 min each). Sections were prehybridized for 1 h at 60°C in hybridization buffer (50% formamide, 5× SSC, 50 μg/ml heparin, and 0.1% Tween-20). 100 μl hybridization solution containing the labeled RNA in hybridization buffer was placed onto the tissue section. A

coverslip was placed on top and slides were incubated in a moister box at 60°C overnight (18–20 h). After hybridization, slides were washed twice for 30 min in 0.1× SSC at 60°C, then each slide was treated with 1 ml 1% blocking reagent (Roche) for 40 min at room temperature.

Visualization of Dig-labeled probe hybridizations was achieved by using an anti-Dig alkaline phosphatase (AP) conjugated antibody (1:500, Roche) and NBT/BCIP substrate. Antennal sections were analyzed on a Zeiss Axioskope2 microscope (Zeiss, Oberkochen, Germany) equipped with Axiovision software. For two-color FISH visualization of hybridized probes was performed by using an anti-Dig AP-conjugated antibody in combination with HNPP/Fast Red (Roche) for Dig-labeled probes and an anti-biotin streptavidin horse radish peroxidase-conjugate together with fluorescein-tyramides as substrate (TSA kit, Perkin Elmer, MA, USA) for Bio-labeled probes. Sections from FISH experiments were analyzed with a Zeiss LSM510 Meta laser scanning microscope (Zeiss, Oberkochen, Germany). Confocal images stacks were processed by ZEN 2009 software. The pictures shown represent projections of optical planes selected from confocal image stacks. For clear data presentation, images were only adjusted in brightness and contrast. Antennal sections of both male and female antennae were analyzed using each generated probe. No obvious difference between sexes regarding the labeling intensity and labeling pattern was observed. Thus, only the images of male antenna were adopted in this study.

Structure Modeling and Electrostatic Potential

In silico simulation of OBP tertiary structure was performed by I-TASSER server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) (Roy et al., 2010), which implements the iterative template threading refinement making full use of established homologous protein structures. PyMol was used to visualize the simulated protein tertiary structures (DeLano, 2002). The molecular surface was solvent excluded and the solvent radius was set 1.4 as default. APBS plug (Unni et al., 2011) implemented in PyMol was employed to calculate the surface electrostatic potentials in the range of −6 to 6 kT/e, and was presented as blue-red hue gradient.

RESULTS

Identification, C-Skeleton Pattern and Phylogenetic Relationship of Locust OBPs

Toward an identification of OBPs from *S. gregaria* and a comprehensive characterization of OBPs in locust species, we have performed a homology-based data mining of an antennal transcriptome which resulted in 14 transcripts putatively encoding SgreOBPs. Subsequently, a multiple sequence alignment was conducted addressing the amino acid sequences of the newly identified SgreOBPs together with hitherto documented OBPs from three other locust species: 16 from *L. migratoria*, 15 from *O. asiaticus* and 7 from *C. kiangsu*. Several OBP subtypes could be categorized based on the number of conserved C-residues (**Figure 1A**). First, 33 OBPs were classified as classic OBPs comprising six conserved C-residues, the hallmark of classic OBPs. Second, 15 OBPs were categorized in two types of plus-C OBPs harboring more than six conserved C-residues. Finally, only one minus-C OBP with less than six conserved C-residue and three atypical OBPs with extraordinary long stretches between conserved C1 and C2 were identified.

As a next step, we analyzed the phylogenetic relationship of the locust OBP repertoire by constructing a phylogenetic tree utilizing the maximum likelihood algorithm and bootstrap iterations. The emerging picture indicated that the repertoire of locust OBPs can be divided into four major families (I–IV), which apparently split at the internal nodes (**Figure 1B**). We further classified three additional subfamilies (I-A, II-A, and III-A), based on the presence of higher bootstrap support (above 80%) on the divergent nodes. It is noteworthy that subfamily I-A and II-A both represent classic OBPs and each subfamily apparently comprise three distinct groups with 3–4 orthologous OBPs from different locust species (**Figure 1B**). Within each subfamily, the sequence identity between OBPs from different groups ranged from 28 to 35%; OBP members within each ortholog group exhibit generally above 80% sequence identity. Incidentally, plus-C OBPs type-A converged onto a subfamily III-A and segregated from their counterparts plus-C OBPs type-B and classic OBPs. Together, the data indicate a considerable degree of orthology in the OBP repertoires across the four analyzed locust species and no marked species-specific expansion within the OBP phylogeny.

Elucidation of Subfamily-Specific Consensus Amino Acid Motifs

To better elucidate the clustering regime of individual subfamilies, we analyzed the consensus amino acid motifs characteristics underlying subfamily I-A and II-A OBPs. The local consensus motifs were calculated by recapitulating repeatedly occurring sequence patterns along OBP sequences. Six consensus motifs with various widths were identified and localized at distinct positions (**Figure 2**). The motif 1 and motif 2 appeared as common motifs in all OBPs of both subfamilies, whereas the other four motifs specifically fit either the repertoire of subfamily I-A OBPs (motif combination 4 and 6) or the repertoire of subfamily II-A OBPs (motif combination 3 and 5). Therefore, two less divergent sequence domains were unraveled by the presence of motif combination 1 and 2, spanning the domains of C2–C3 and C4–C6. In contrast, the sequence domains close to the N-terminus (42 amino acids, motif 3 and motif 4) and ahead of C4 (11–15 amino acids, motif 5 and motif 6) appeared to be more divergent.

Utilizing the six identified consensus motifs in **Figure 2** we have quantified the sequence divergence for the locust OBP repertoire at a local motif scale (Figure S1). Apart from subfamilies I-A and II-A, the common motif 1 and motif 2, especially the latter, recapitulate sequence information present in many of the other locust OBPs analyzed (*E*-value below 10^{-5}) indicating particular phylogenetic conservation of these regions. Not surprisingly, the subfamily-specific motifs 3–6 failed to match OBP members (*E*-value above 10^{-5}) other than subfamily

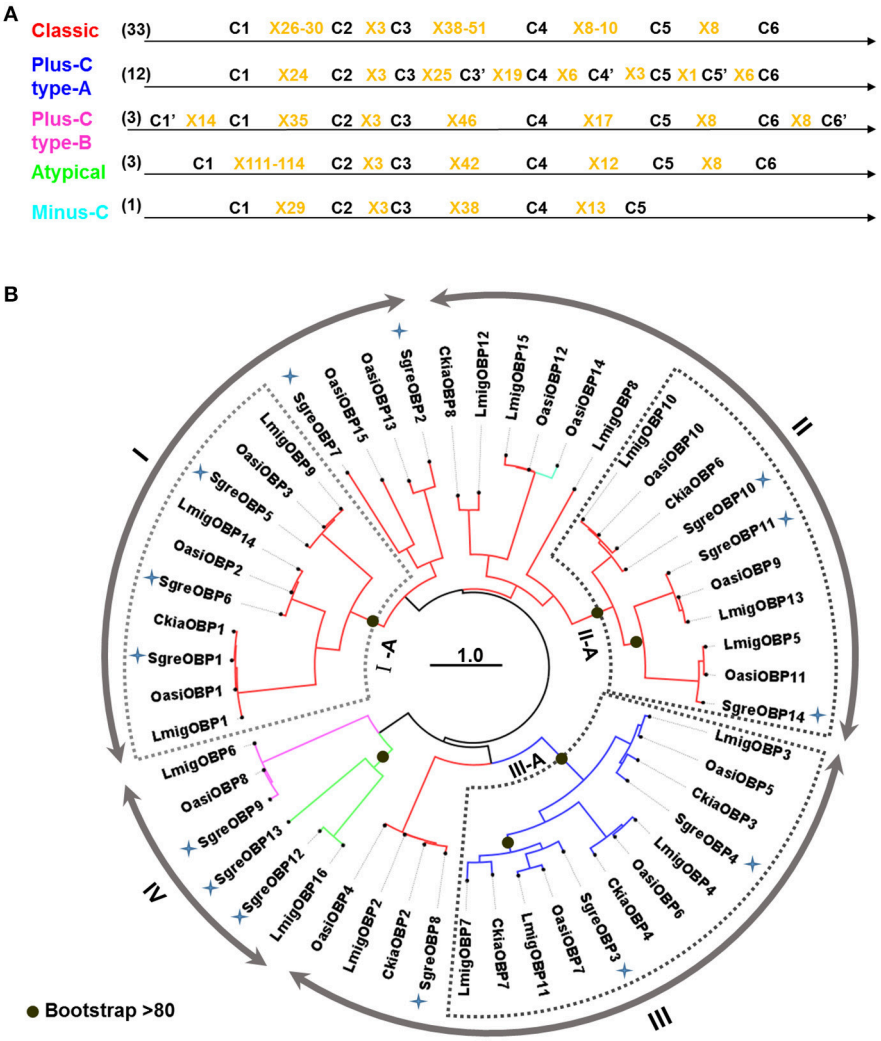


FIGURE 1 | C-residue skeletons and phylogeny of OBPs from four locust species. **(A)** OBPs subtypes were categorized based on the number of conserved C-residues. C-skeleton patterns are based on the multiple sequence alignment of 52 OBP sequences from four locust species. C-residues conserved in all OBPs are shown as C1-C6 in black characters; additional C-residues conserved in the two plus-C OBP types are shown as C'; amino acid between two C-residues are shown as X plus the number of amino acid. The number of each OBP subtype is given in the parenthesis. **(B)** The phylogenetic tree was constructed using the maximum likelihood algorithm supported by 1,000 bootstrap replicates. OBP sequences utilized to generate the tree were derived from four locust species: 14 from *Schistocerca gregaria* (SgreOBPs), 16 from *Locusta migratoria* (LmigOBPs), 15 from *Oedaleus asiaticus* (OasiOBPs) and 7 from *Ceracris kiangsu* (CkiaOBPs). Four primary families (I-IV) are denoted by arrow lines. Further classification of three subfamilies (I-A, II-A, and III-A) was based on the over 80% bootstrap support at the internal node (indicated by black dots). Inner branches in different colors represent OBP subtypes in **(A)**: red, classic OBPs; blue, plus-C OBPs type-A; magenta, plus-C OBPs type-B; green, atypical OBPs; cyan, minus-C OBP. Newly identified SgreOBPs are denoted by blue crosses. The tree is midpoint rooted. Scale bar represents one amino acid substitution per site.

I-A and II-A OBPs, despite a small number of OBPs in family I and family II (Figure S1). Taken together, the motif analysis unraveled the presence of both stabilized and diversified domains residing on the global sequences.

Selection Pressure and Orthology Evolution of Locust Subfamily I-A and II-A

The appearance of two distinct conserved subfamilies in the locust OBP phylogeny, coupled with the clustering pattern of different ortholog groups is presumably a consequence of particular selection regimes. To prove this notion, we have tried to quantify the strength of selection pressure acting on genes encoding the locust OBP repertoire. We analyzed three principal concepts which reflect the selection pressure, namely, the non-synonymous substitution rates (dN), the synonymous substitution rates (dS) and the ω rates (dN /dS) (**Figure 3**). We found a significantly reduced median dN level for both subfamily I-A (dN = 0.030, $U = 60$, $p = 0.016$, Mann-Whiteny U-test) and subfamily II-A (dN =0.028, $U = 60$, $p = 0.016$, Mann-Whiteny U-test), in comparison with that of other OBP members (dN = 0.085, **Figure 3A**). However, the median dS level appeared to be quite similar among subfamily I-A (dS = 0.12, $p = 0.154$,

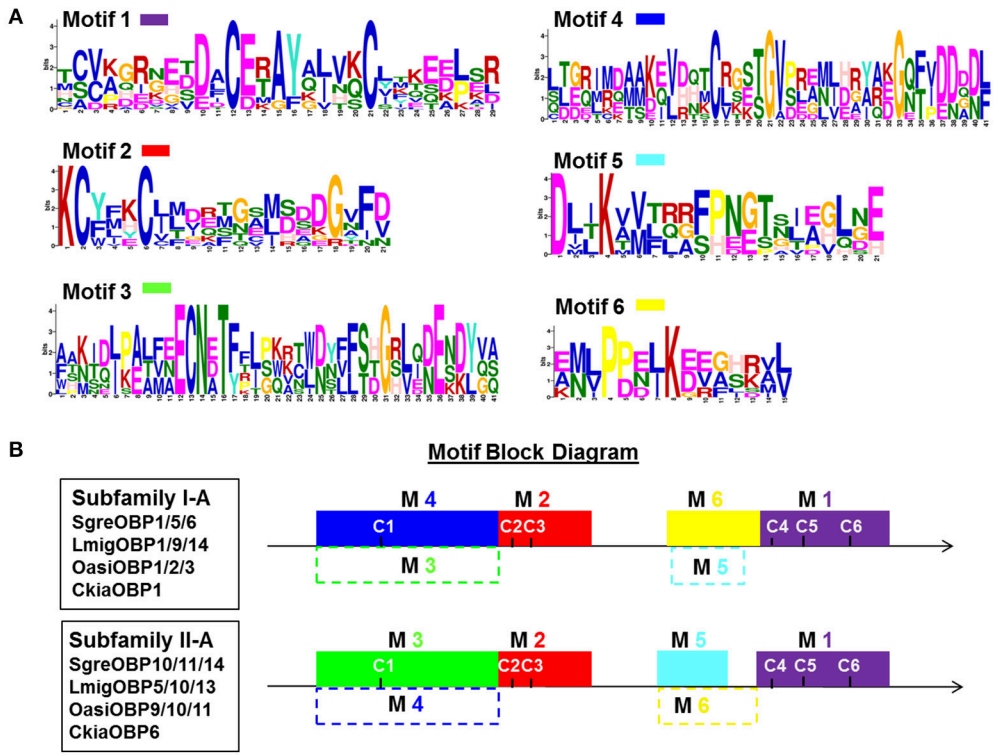
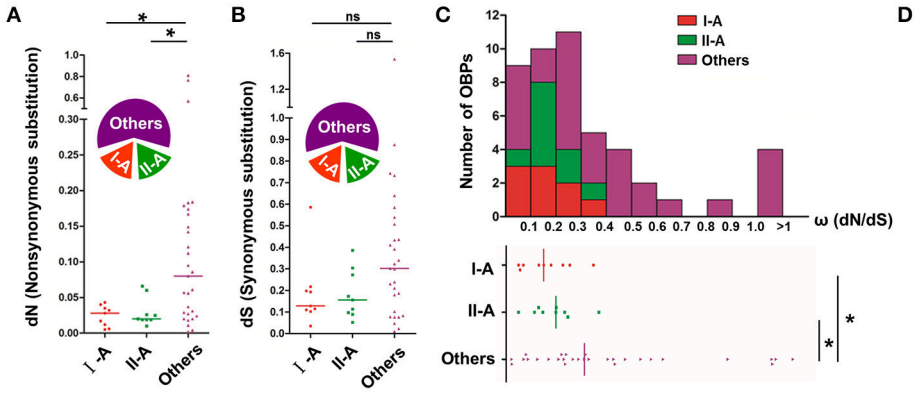


FIGURE 2 | Identification and position of consensus amino acid motifs for subfamily I-A and II-A OBPs. **(A)** Six amino acid motifs with various widths were identified *de novo* to recapitulate the subfamily I-A and II-A OBP sequence signature (classification see **Figure 1**). The height of an amino acid character is proportional to the degree of conservation in the consensus sequences. **(B)** Position of identified consensus motifs (M1–M6) in the polypeptide chain of subfamily I-A and II-A OBPs. C1–C6 indicate the position of the conserved C-residues. Motif 1 and motif 2 adequately match the repertoire of OBP sequences in both subfamily I-A and II-A. In contrast, motifs combination M4 (blue) and M6 (yellow) specifically match subfamily I-A, whereas motifs M3 (green) and M5 (cyan) are specific for subfamily II-A OBP sequences. Dash lined blocks indicates unfitness of a particular motif to the target sequences (*E*-value above e^{-5}; default statistical significant level). Obtained *E*-values for each motif are given in Figure S1.

$U = 88.5$, Mann-Whiteny *U*-test), subfamily II-A (dS = 0.16, $U = 86$, $p = 0.131$, Mann-Whiteny *U*-test) and the other OBP members (dS = 0.31, **Figure 3B**). For the ω rates, the values ranged from 0 to 0.7 for nearly 90% of locust OBPs (**Figure 3C**), which is indicative of purifying selection acting on locust OBP repertoire in general. For a few exceptions, ω rates larger than one were found which may indicate a positive selection. Notably, median ω rates for OBPs of subfamily I-A ($\omega = 0.18$, $U = 63$, $p = 0.021$, Mann-Whiteny *U*-test) and subfamily II-A ($\omega = 0.22$, $U = 69$, $p = 0.036$, Mann-Whiteny *U*-test) were significantly reduced in comparison with other OBP members in the phylogeny ($\omega = 0.35$, **Figure 3C**).

Exposed to a similar selection regime, we wondered if orthologous OBPs in other species would undergo similar divergent events in relation to the two locust OBP subfamilies. To address the issue, we made a phylogenetic analysis of the two locust OBP subfamilies and the reference OBPs derived from 8 other insect species which gradually emerged in the course of insect evolution. The analysis revealed that locust subfamily II-A OBPs remained on an intact clade without intermingling with reference OBP genes on the newly constructed phylogenetic tree (Figure S3). A different result was obtained for the subfamily I-A: the original clustering relationship of ortholog groups in locust phylogeny was disrupted and altered with a complex re-clustering pattern integrating reference OBPs. The orthologous relationship (Theißen, 2002) of OBPs between the two locust subfamilies and other species was also inferred. It is found that the number of locust subfamily I-A orthologous OBPs in the inspected insect species expanded considerably, and exhibited a many-to-many orthologous relationship with locust subfamily I-A (**Figure 3D**), with *A. pisum* as apparent exception likely due to a smaller OBP gene repertoire (Zhou et al., 2010). In contrast, the number of locust subfamily II-A orthologous OBPs in other species apparently decreased, and displayed a 1-to-many or 0-to-many orthologous relationship with locust subfamily II-A (**Figure 3D**). Moreover, it was found that locust subfamily II-A OBPs and their orthologous OBPs may share a common ancestor verified by the convergence of a mono phylogenetic clade with the bootstrap support above 70% at the basal divergent node (Figure S3). However, the common ancestral status for locust subfamily I-A OBPs and their orthologous OBPs appeared ambiguous because of the absence of evident bootstrap support (Figure S3). In sum, our results provide evidence that locust subfamily I-A and II-A OBPs are subject to mutually similar strengthened



D

Ortholog of Subfamily I-A		Ortholog of Subfamily II-A
OBP3/4/5/6/15/17 19/20/23/24/26/ 28/58/83	Agam	OBP60
OBP1/5/6/11/13	Amel	-
OBP19a/56d/56e/ 69a/76a/83a/83b	Dmel	OBP84a
OBP1/16/17/18/ 35/36	Tcas	-
OBP7/10	Apis	-
OBP15/17/18/20/ 25/26/27/31	Bmor	OBP38
OBP1/2/3/4/5/6 7/8	Znev	OBP15

FIGURE 3 | Selection constraints and orthology evolution of locust subfamily I-A and II-A. **(A,B)** Locust subfamily I-A and II-A OBPs exhibit a reduced dN rate, but a similar dS rate in comparison with the other locust OBPs. "Others" include those OBPs that do neither belong to subfamily I-A nor to subfamily II-A. The relative fraction included in each OBP group is illustrated by the wedges diagrams. Non-synonymous substitution rates (dN) and synonymous substitution rates (dS) were calculated across the locust OBP repertoire. The median level is indicated by lines. *$p < 0.05$; ns, $p > 0.05$; two-tailed Mann-Whitney *U*-test. Detailed data of this analysis are given in Figure S2. **(C)** Proportional distribution of ω rates for the locust OBP repertoire. The majority of OBPs (~90%) fall into a ω range of 0–0.7. Yielded ω ratios (dN and/or dS ≠ 0) and the median level are displayed at the bottom. *$p < 0.05$; two-tailed Mann-Whitney *U*-test. Nucleotide sequences utilized in this analysis are given in the supplementary material. **(D)** Orthologs of locust subfamily I-A and II-A OBPs in seven other insect species. It is noted that the complete genome has been sequenced for the seven inspected species, namely, *Anopheles gambiae* (Agam), *Apis mellifera* (Amel), *Drosophila melanogaster* (Dmel), *Tribolium castaneum* (Tcas), *Acyrthosyphon pisum* (Apis), *Bombyx mori* (Bmor), and *Zootermopsis nevadensis* (Znev). Orthology assignment was obtained by using EggNOG 4.5.1 which performed a hierarchical orthologous annotation (Huerta-Cepas et al., 2016). The criteria *E*-value for assessing orthologous relationship of locust subfamily I-A is set to e^{-20}, while e^{-10} for subfamily II-A. Short bar denotes that there are no appropriate hints that could be assigned as orthologous OBPs. Nomenclature of OBPs for the seven inspected insect species conforms to Vieira and Rozas (2011) and Terrapon et al. (2014).

purifying selection, whereas distinct divergent events occur during evolution of their orthologous OBPs in other species.

Prediction of Tertiary Structures for OBPs in Subfamily I-A and II-A

The intriguing sequence and evolutionary characteristics underlying locust subfamily I-A and II-A OBPs inspired us to explore the possible concurrent variation of their tertiary structures. Therefore, we have simulated the tertiary structures for OBP members from both two subfamilies covering different ortholog groups and locust species. Parametric estimates toward the accuracy and reliability of the structure prediction was scrutinized, which permitted to investigate structural variation as an exploratory trial. To unravel structural variation between the two subfamilies, we superimposed the backbone structures of those simulated OBPs to LmigOBP1, the hitherto only established crystal structure for the locust OBP repertoire (Zheng et al., 2015). The averaged RMSD score obtained by imposing subfamily II-A OBPs to LmigOBP1 (2.8) doubled that of imposing subfamily I-A OBPs to LmigOBP1 (1.39 in average, Figure S4), indicating an enhanced structural similarity within one subfamily.

Multiple sequence alignment of subfamily I-A OBPs revealed a striking variation on the C-terminal domain (Figure S4). It is known that LmigOBP1 has a prolonged C-terminus with ~17 amino acids to form a seventh α-helix (Zheng et al., 2015). In contrast, the C-terminus in OasiOBP3 and SgreOBP6 is shortened to a 7 amino acids motif and most likely constitute a coiled-coil strand instead of a seventh α-helix (**Figure 4**); a groove emerged on the collapsed surface due to the shortened C-terminus. The electrostatic potential pattern varies greatly at a global surface scale as well as on the local C-terminal surface scale (cyan dash line, **Figures 4A,C,E**). Another striking structural difference is the enlarged cavity of LmigOBP1 bordered by the prolonged C-terminus, whereas the cavity for the other two counterparts, representative of different ortholog groups shrinks to some extent (white dash line, **Figures 4B,D,F**). Unlike subfamily I-A, the multiple sequence alignment of subfamily II-A OBPs exhibited an aligned C-terminus but an unaligned N-terminus, namely, an extra extension of a 9–10 amino acids motif in the LmigOBP10 ortholog group (Figure S4). Correspondingly, this alteration was predicted to result in a coiled-coil structure on the N-terminal domain for LmigOBP10; at the same surface position, an opening structure was observed on its two counterparts, the OasiOBP11 and SgreOBP11 (**Figures 5A,C,E**). Apart from that, the surface electrostatic potential profile seems to vary slightly, both at the global surface scale and at the local N-terminal surface scale (cyan dash line, **Figures 5C,E**), regardless of the extra N-terminal coil present on LmigOBP10. However, the interior cavity could be enriched with negative potentials (LmigOBP10 and SgreOBP11,**Figures 5B,F**), or with positive potentials (OasiOBP11, **Figure 5C**).

Topographic Expression Patterns of SgreOBPs from Subfamily I-A and II-A

To approach this question, whether locust subfamily I-A and II-A OBPs may be expressed in different sensillum types and different cells, we set out to unravel the expression patterns of SgreOBPs from the two locust subfamilies in sensilla on the antenna, the major olfactory organ. By adopting RNA *in*

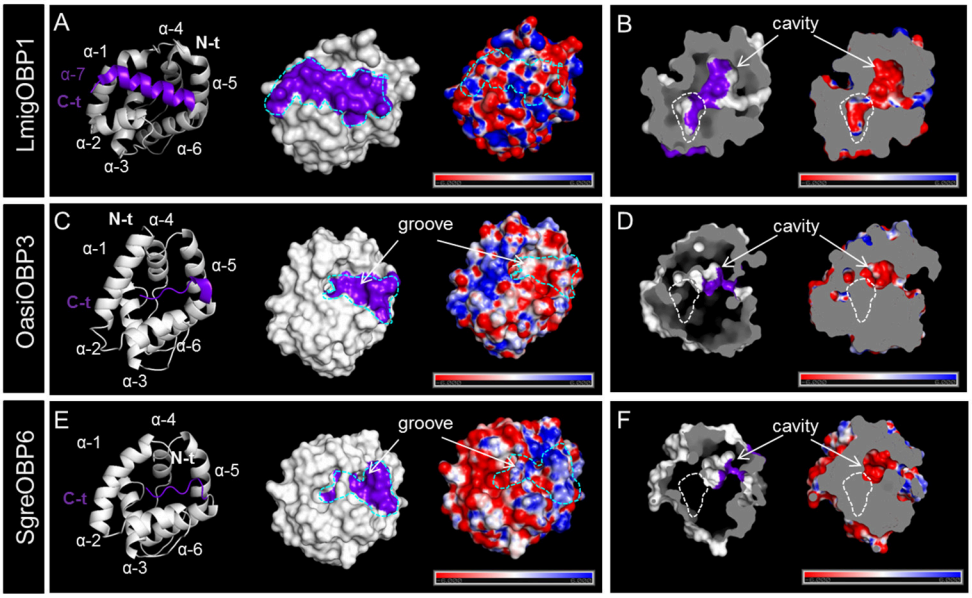


FIGURE 4 | Variations in the C-terminal domain and the interior cavity of subfamily I-A OBPs. **(A,C,E)** Comparison of the backbone structures, surface topologies and surface potentials of LmigOBP1, OasiOBP3 and SgreOBP6 which represent the three different ortholog groups in subfamily I-A. The C-terminal domains (see also Figure S4A) are highlighted in purple on both the backbone structures (left) and the molecular surfaces (middle). The dash line in cyan sketches the surface topology of the C-terminal domain (middle and right). Left and middle: an additional α-helix (α-7) is formed by the prolonged C-terminus in LmigOBP1 (Zheng et al., 2015) **(A)**. Instead of a seventh α-helix, the shortened C-terminus in OasiOBP3 **(C)** and SgreOBP6 **(E)** are likely to constitute a groove structure on the collapsed surface. Right: a map of electrostatic potential on the molecular surface. The electrostatic potential pattern of LmigOBP1 **(A)**, OasiOBP3 **(C)** and SgreOBP6 **(E)** varies greatly at a global surface scale as well as on the local C-terminal surface scale. **(B,D,F)** Depiction of the interior cavity (left) which is bordered by C-terminal domain (highlighted in purple) and the corresponding electrostatic potential map (right). The assumed enlarged interior cavity in LmigOBP1 **(B)** relative to OasiOBP3 **(D)** and SgreOBP6 **(F)** is outlined with a white dash line. Electrostatic potential was calculated in the range of −6 to 6 kT/e and was presented as blue-red hue gradient. Blue, negative potential; red, positive potential; k, Boltzmann's constant; T, temperature; e, charge of an electron.

situ hybridization (ISH) on antennal sections using specific OBP probes, we acquired a strikingly sensilla-specific expression pattern for SgreOBPs in the two subfamilies. For SgreOBP1, SgreOBP5 and SgreOBP6, the members of subfamily I-A, we found alike expression in the cells of both sensilla basiconica and sensilla trichodea (**Figure 6**). In contrast, none of the subfamily I-A SgreOBPs was expressed in sensilla coeloconica or sensilla chaetica. Conversely, for members of subfamily II-A SgreOBPs, namely, SgreOBP10, OBP11, and OBP14, the expression was found to be restricted to the cells of sensilla coeloconica; there was no evidence for an expression in cells of any other sensillum type (**Figure 6**). The notion that a similar expression pattern is conserved for orthologous OBPs from other locust species is supported by the finding that LmigOBP1 is specifically expressed in sensilla basiconica and sensilla trichodea of *L. migratoria* (Jin et al., 2005), alike its ortholog in *S. gregaria*, the SgreOBP1.

Thus, an apparent sensilla-specific expression pattern for each locust OBP subfamily emerged. To extend and specify this aspect, the expression of OBP subtypes was compared with the expression of sensilla-specific receptor types. The odorant receptor co-receptor Orco and the ionotropic receptor (IR) type IR8a are ubiquitous co-receptors expressed in insect OSNs, either together with ligand-specific ORs or with IRs, and are considered as general markers for sensilla basiconica/sensilla trichodea and sensilla coeloconica, respectively (Yang et al., 2012; Guo et al., 2013). As a marker specific for distinct sensilla trichodea, the expression of the sensilla-specific receptor type OR3 in *S. gregaria* was monitored (Pregitzer et al., 2017). We designed riboprobes labeled by either Dig or Bio, which specifically targeted the distinct sensory neuron markers and SgreOBPs of the two subfamilies. Subsequently, two-color fluorescent *in situ* hybridization (FISH) experiments were performed to visualize the expressing cells (Figure S5). The results indicated that SgreOBPs of subfamily I-A are expressed in cells located in sensilla basiconica; these cells extended cytoplasmic processes and enclosed clusters of Orco expressing neurons. Similarly, SgreOBPs of subfamily I-A were found to be expressed in cells located in sensilla trichodea, as characterized by their close association with OR3 expressing OSNs. In the sensilla coeloconica, characterized by the IR8a-positive neurons, the neurons were found to be engulfed by cells which express OBPs of the subfamily II-A.

Although, our data demonstrated that SgreOBPs from different ortholog groups in each subfamily are expressed in the same sensillum type, it remained unclear to what extent they are expressed in the same set of sensilla and whether they are co-expressed in the same cells within a distinct sensillum. To resolve this question, we performed two-color FISH on sections through the antenna of *S. gregaria* using riboprobes targeting SgreOBPs from different ortholog groups. The results for SgreOBPs in

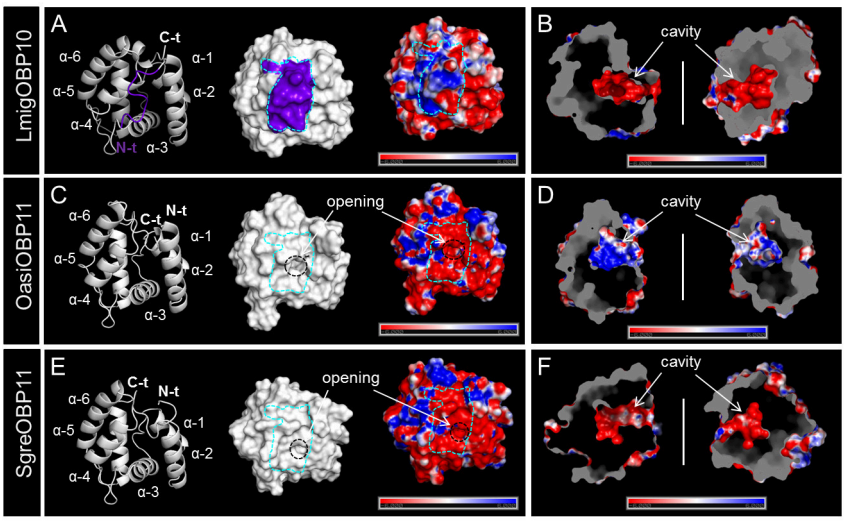


FIGURE 5 | The surface topologies and interior cavities of subfamily II-A OBPs. **(A,C,E)** Comparison of the backbone structures, surface topologies, and surface potentials of LmigOBP10, OasiOBP11, and SgreOBP11 which represent the three different ortholog groups in subfamily II-A. Left: the prolonged N-terminus in LmigOBP10 (see Figure S4B) was predicted to form a short coiled-coil shown on the backbone structure (highlighted in purple, **A**), but was absent from OasiOBP11 **(C)** and SgreOBP11 **(E)**. Middle: the N-terminal domain of LmigOBP10 was plotted on the surface and sketched by a cyan dash line **(A)**. The N-terminal domain of LmigOBP10 was labeled on the same surface position for OasiOBP11 **(C)** and SgreOBP11 **(E)**. The visible opening structure is denoted by a black circle for OasiOBP11 **(C)** and SgreOBP11 **(E)**. Right: a map of electrostatic potential on the molecular surface. Generally similar electrostatic potential pattern is observed among LmigOBP10 **(A)**, OasiOBP11 **(C)** and SgreOBP11 **(E)**. **(B,D,F)** A symmetric presentation of the interior cavity with the electrostatic potential. Electrostatic potential was calculated in the range of −6 to 6 kT/e and was presented as blue-red hue gradient. Blue, negative potential; red, positive potential; k, Boltzmann's constant; T, temperature; e, charge of an electron.

subfamily I-A indicate that SgreOBP1 was expressed in a cell population present in almost all basiconic and trichoid sensilla, whereas SgreOBP5 and SgreOBP6 were expressed only in a much smaller subset of cells than SgreOBP1 in the same sensillum (**Figure 7**). These differences became apparent in both horizontal sections giving a view onto superficial cellular layer (no cytoplasmic process expected, **Figure 7A**) as well as in longitudinal sections which allowed a view into deeper layers (cytoplasmic process expected, **Figure 7B**) of the antenna. Unlike SgreOBP1-positive cells which could be visualized both at the superficial and the deeper cellular layer, most of SgreOBP5- and SgreOBP6-positive cells appeared to be restricted to the superficial cellular layer close to the cuticle; slim cytoplasmic processes stretched to deeper cellular layers. Incidentally, there was evidence that SgreOBP5 and SgreOBP6 were expressed in the same set of cells of a sensillum (**Figures 7E,F**).

In contrast to the subfamily I-A, for subfamily II-A we did not find any evidence for an OBP subtype that was ubiquitously expressed in coeloconic sensilla (Figure S5). This result has led to the notion that particular OBP members of subfamily II-A may be specifically expressed in subsets of coeloconic sensilla. In fact, we frequently observed that expression of SgreOBP10 and SgreOBP14 were restricted to different cells in sensilla coeloconica (**Figures 8A,B**). For the subtypes SgreOBP11 and SgreOBP14 a co-expression in the same cells or expression in different cells were observed at a similar rate (**Figures 8C,D**). For the subtypes SgreOBP10 and SgreOBP11 it was frequently observed that they were co-expressed in the same cells (**Figure 8E**), indeed, more often than an expression in different cells (**Figure 8F**). Moreover, we verified the spatially separated expression of SgreOBPs from subfamily I-A and II-A (Figure S6), consistent with the results in **Figures 6**, **7**. Taken together, the results unravel a characteristic subfamily-dependent cellular expression pattern for different OBP subtypes.

DISCUSSION

The complex behavior of locust species, including the unique switch between a solitarious phase and a gregarious phase, is strongly based on a sophisticated chemical communication system (Pener and Yerushalmi, 1998; Hassanali et al., 2005; Wang and Kang, 2014). Great efforts have been made to unravel the chemical cues and underlying chemosensory mechanisms in mediating locust enigmatic behavior (Heifetz et al., 1996; Anton et al., 2007). Out of these efforts, a variety of olfactory genes, including gene families encoding odorant receptors and candidate pheromone receptors have recently been identified (Guo et al., 2011; Wang et al., 2015; Pregitzer et al., 2017). Since much less was known about their counterparts which deliver the olfactory signal molecules to the receptors, the OBPs, this study was concentrating on a systematic analysis of locust OBPs with respect to their molecular evolution as well as on an evaluation of predicted protein structures for OBP subtypes and their expression pattern in stinct sensillum types.

The in-depth analysis of locust OBP sequences uncovered the presence of both common and specific amino acid

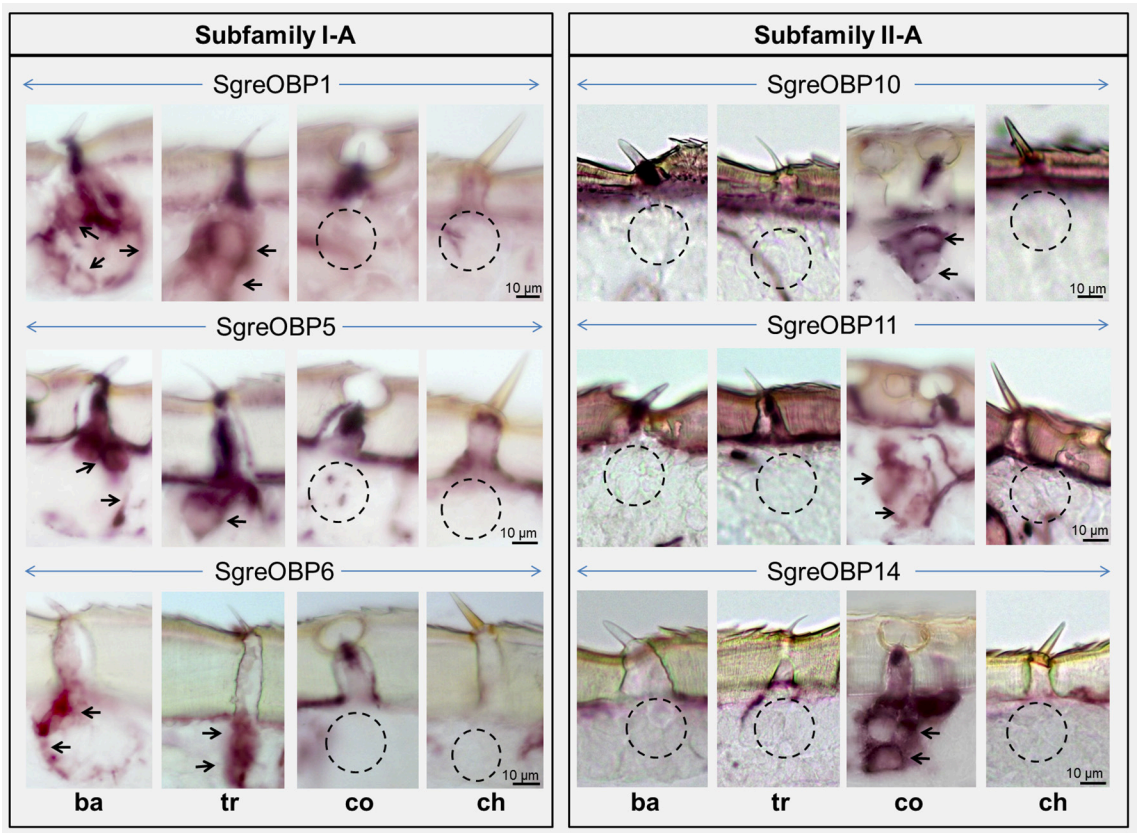


FIGURE 6 | Sensilla-specific expression of subfamily I-A and II-A OBPs in the antenna of *S. gregaria*. Antisense riboprobes which specifically target the SgreOBPs were used to visualize the appropriate structures by means of chromogenic *in situ* hybridization (ISH). SgreOBP1, SgreOBP5, and SgreOBP6 are representing three different ortholog groups of subfamily I-A, whereas SgreOBP10, SgreOBP11, and SgreOBP14 are representing three different ortholog groups of subfamily II-A. Labeling obtained with probes for subfamily I-A SgreOBPs was restricted to sensilla basiconica (ba) and sensilla trichodea (tr), but was absent in sensilla coeloconica (co) and sensilla chaetica (ch). Labeling obtained with probes for subfamily II-A SgreOBPs was detected only in sensilla coeloconica (co), but was absent in the other three sensillum types. Black arrows indicate the visible OBP labeling while black circles denote the absence of OBP labeling.

motifs (**Figure 2**). The common motifs adequately recapitulate sequence information in most of the locust OBPs, while specific motifs selectively represent locust OBP subfamilies which may contribute to the clustering of sequences on the phylogenetic tree (**Figure 1**). The mixed common and specific motif profile is reminiscent of the findings that selection regimes may vary among different sequence domains (Policy and Conway, 2001; Sawyer et al., 2005). The subfamily specific motifs define sequence domains that apparently withstand diversifying selection constraints, presumably shaped by the sensilla environment, including their likely interplay-partner, the endogenous receptor types (Figure S5). In contrast, the common motifs define sequence domains that appear to share similar stabilizing selection constraints, presumably required for the maintenance of the common globular structures of the proteins (Pelosi et al., 2017), or for retaining the conserved ligand binding sites (Yu et al., 2009).

The four locust species tackled in this study differ significantly in their geographic distribution. While *S. gregaria* (the desert locust) occurs in Africa, the Middle East and Asia and *L. migratoria* (the migratory locust) in Africa and Asia, but also in Australia and New Zealand, the locusts *O. asiaticus* and *C. kiangsu* (the yellow-spined bamboo locust) appear to live locally in North China and South China. Nevertheless, a molecular and evolutionary stabilized status can be assigned to locust OBP subfamily I-A and II-A that appear to be subject to purifying selection pressure (**Figure 3C**), indicative for conserved chemosensory roles. In addition, the chemosensory adaptation to different habitats supposedly implies positive selection constraints (Cicconardi et al., 2017), and several of the locust OBPs appear to reflect such a selection regime (**Figure 3C**).

For the locust OBP subfamily I-A, the selective expression in two distinct sensillum types, sensilla basiconica, and sensilla trichodea, appears to be a characteristic hallmark (**Figure 6** and Figure S5). This feature is also found for OBPs from other species, which are orthologous of locust OBPs subfamily I-A (**Figure 3D**). For example, in *Drosophila melanogaster*, most of the subfamily I-A orthologous OBPs are associated with sensilla basiconica and sensilla trichodea, similar to their locust counterparts. It was found that DmelOBP83a and DmelOBP83b were associated with sensilla basiconica and sensilla trichodea, while DmelOBP69a and DmelOBP76a seemed to be restricted to sensilla trichodea

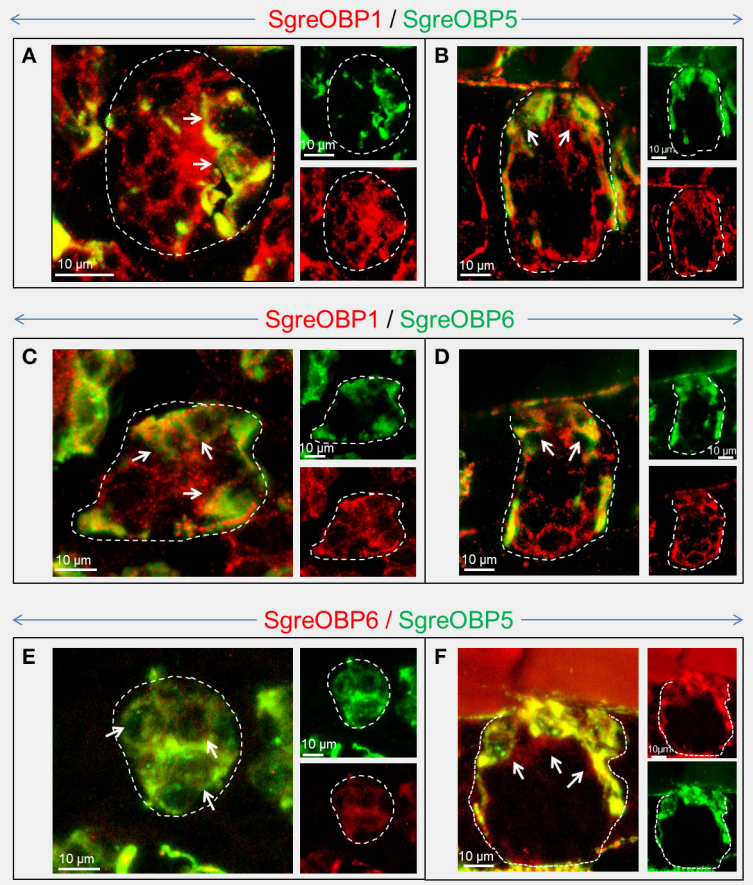


FIGURE 7 | Visualization of cells expressing distinct subtypes of subfamily I-A SgreOBPs. The confocal images show the co-localization of three SgreOBPs from subfamily I-A in the cellular compartment of a sensillum basiconicum. Cells expressing distinct subtypes of subfamily I-A SgreOBPs were visualized by two-color FISH employing subtype specific antisense riboprobes. Confocal images of the overlaid green and red fluorescence channel are shown at higher magnification on the left, the red and green fluorescent channels are shown separately at lower magnification on the right. Cells that are apparently assigned to the cell-cluster belonging to one sensillum basiconicum are outlined in a white dash line. Cells that co-express two distinct OBP subtypes are indicated by white arrows. **(A,C,E)** A horizontal perspective of the superficial cellular layer close to the cuticle is shown where the cytoplasmic processes exhibited by subfamily I-A SgreOBP-positive cells are less likely to be visualized. **(B,D,F)** A longitudinal perspective of a deep layer beneath the cuticle is shown where the cytoplasmic processes are likely to be visualized. **(A–D)** It is noted that a smaller number of cells are labeled in green compared to the number of cells labeled in red.

(Larter et al., 2016). However, for a few orthologous OBPs such as DmelOBP56d an extra sensillar expression has been reported (Larter et al., 2016). The concept of a sensilla-specific expression pattern for orthologous OBPs of locust subfamily I-A is also supported by the finding in the moth *Manduca sexta,* where two orthologous OBPs of locust subfamily I-A, named MsexABP2 and MsexABPx, are specifically expressed in sensilla basiconica (Nardi et al., 2003). Since the Orthopteran locust species emerged at a much earlier stage than the moth and fly species during the insect species divergence (Vieira and Rozas, 2011; Vogt et al., 2015), it is conceivable that a dual expression of subfamily I-A OBPs in both sensilla basiconica and sensilla trichodea may represent an ancestral status. In insect species like moths and flies, which emerged later in evolution, some OBP subtypes may have evolved towards a more specific function and expression in either sensilla basiconica or sensilla trichodea (Maida et al., 2005; Larter et al., 2016).

Our analysis suggests that the locust OBPs of subfamily II-A and their orthologous OBPs in other species have originated from a common ancestor (Figure S3), and may share a sensilla coeloconica specific expression pattern (**Figure 6**, Figure S5). In *Drosophila melanogaster*, DmelOBP84a, the only orthologous OBP of locust subfamily II-A is actually among the few OBPs that have been reported to be specifically expressed in sensilla coeloconica (Larter et al., 2016). Interestingly, the gene encoding OBP84a is retained in most, if not all, *Drosophila* species genomes (Cicconardi et al., 2017). Moreover, the OBP84a ortholog group in *Drosophila* species withstands apparent purifying selection pressure (Vieira et al., 2007) and converges onto a segregated phylogenetic clade (Cicconardi et al., 2017), which is very similar to the locust OBP subfamily II-A. These molecular and phylogenetic commonalities may point to some similarities with regard to their functional roles. In this regard, it is interesting to note that single sensillum recordings from sensilla coeloconica

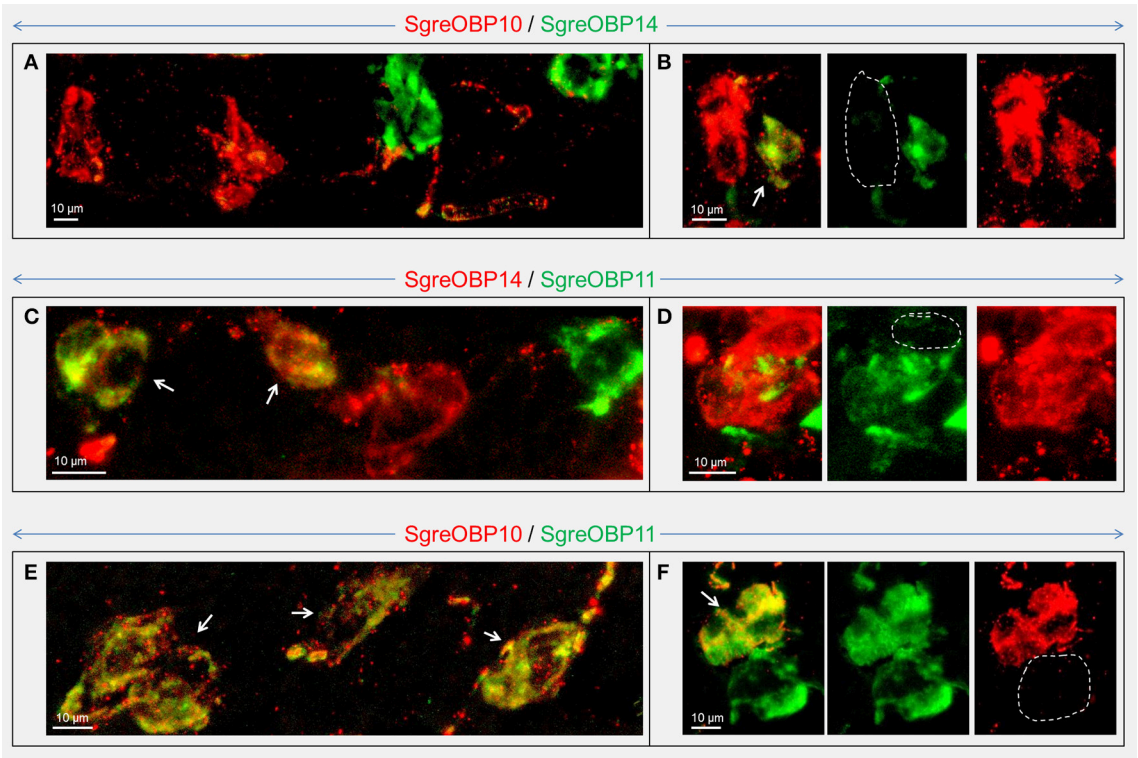


FIGURE 8 | Visualization of cells expressing distinct subtypes of subfamily II-A SgreOBPs. Cells expressing distinct subtypes of subfamily II-A SgreOBPs were visualized by two-color FISH employing combinations of subtype specific antisense riboprobes. The dash line indicates the absence of OBP labeling at the particular area. **(A)** A representative confocal image demonstrating expression of SgreOBP10 and SgreOBP14 in separate cells. **(B)** Rarely cells could be observed that co-expressed SgreOBP10 and SgreOBP14 (white arrows). **(C,D)** For the combination of SgreOBP14 and SgreOBP11 co-expression (white arrows) was observed at a similar rate as a separate expression of the two OBPs in different sensilla. **(E)** SgreOBP10 and SgreOBP11 were frequently found to be co-expressed in the same cells (white arrows). **(F)** Only few cells were detected that selectively expressed only one of the OBP subtypes. **(B,D,F)** Confocal images show the overlaid fluorescence channels (left) as well as the separated green and red fluorescence channels (middle and right) on the same magnification.

of locust, flies and moths have revealed a response spectrum confined to certain ecologically important odorants, including organic acid, amines and plant derived odorants (Pophof, 1997; Ochieng and Hansson, 1999; Yao, 2005). Thus, it will be of particular interest to unravel a potential role of locust subfamily II-A OBPs and their orthologs in other species for the detection of cognate odorants in sensilla coeloconica. While concentrating on OBPs of subfamily II-A, we are aware that sensilla coeloconica may also comprise OBPs of other phylogenetic clades.

Unlike DmelOBP84a, which is broadly expressed in almost all sensilla coeloconica (Larter et al., 2016), the OBPs of the locust subfamily II-A are expressed in sensilla coeloconica in a combinatorial mode (**Figure 8**). This is in line with the previous finding that different subsets of sensilla coeloconica in *S. gregaria* showed individual response spectra to a repertoire of odorants (Ochieng and Hansson, 1999), suggesting a sensilla-specific response spectrum and sensilla-specific repertoire of odorant sensing proteins. Thus, it is conceivable that a distinct combination of OBPs in a sensillum coeloconicum (**Figure 8**) may correlate with particular endogenous IR types.

Although amino acid sequences of OBPs can be highly divergent, the folding of proteins forming a hydrophobic pocket is well conserved across insect species; in fact to date the structures of more than 20 OBPs have been solved by X-ray crystallography and/or nuclear magnetic resonance (NMR) spectroscopy (Pelosi et al., 2017). The results of these studies revealed that the C-terminal domain, especially the length of the C-terminus has important implications on the mechanism of ligand-binding (Tegoni et al., 2004). Long C-terminus apparently enter the binding pocket and determine the shape of the cavity (Sandler et al., 2000), medium-length C-terminus act as a lid covering the entrance to the binding pocket (Lartigue et al., 2004). In view of these findings, simulation of the putative tertiary structures of locust OBPs revealed some interesting features. The three ortholog groups of subfamily I-A significantly differ in their C-terminal domain. LmigOBP1 and its orthologs have a long (17 aa) C-terminus, long enough to form an extra α-helix and thus affecting the shape of the cavity (**Figure 4**, Figure S4); other two ortholog groups have both a medium size C-terminus (7 aa), however, significantly different in the amino acid sequence. These observations may suggest significant differences in the mechanisms of OBP/ligand interaction among the three ortholog groups in subfamily I-A.

The results of this study indicate that in a considerable number of sensilla at least two OBP subtypes are co-expressed (**Figures 7**, **8**). This is of particular interest, since hetero- and homo-dimerization of OBPs have been reported *in vitro* (Andronopoulou et al., 2006), which is accompanied by a set of conformational changes (Wogulis et al., 2006; Mao et al., 2010). Although the underlying mechanisms are still elusive, there is evidence that electrostatic interaction at short range forming the salt bridges may contribute to specific protein-protein interaction (Sheinerman et al., 2000; Kumar and Nussinov, 2002). In locusts, the patch of charged residues buried on the OBP-interface (**Figures 4**, **5**) is likely to provide hot spots for protein-protein interactions. In addition, changes of the OBP tertiary structure has been demonstrated as a consequence of pH changes in the environment (Zubkov et al., 2005; Pesenti et al., 2008). This notion may also fit for locust OBPs since an intermingled distribution of both negative and positive charged residues was observed by elucidating a map of electrostatic potentials (**Figures 4**, **5**). The presence of multiple OBP subtypes and their possible interaction may have functional implications for the binding capacity of the olfactory system. In fact, recent binding assays have shown that in the presence of two OBPs the binding affinity to cognate ligands altered considerably compared to the binding characteristics of a single OBP type (Qiao et al., 2011; Sun et al., 2016). This notion may be particular relevant with respect to sensilla basiconica of locusts, which house up to 50 sensory neurons responding to a variety of different odorants (Ochieng et al., 1998; Ochieng and Hansson, 1999), and the fact that the number of OBP genes is much smaller than the size of the OR gene family in locusts, encoding more than 140 ORs in *L. migratoria* (Wang et al., 2014) and at least 120 ORs in *S. gregaria* (Pregitzer et al., 2017). The selective sensilla expression pattern implies that a small number of OBP subtypes are present in the sensillum lymph (**Figure 7**, Figure S6). Assuming that each OBP subtype has distinct ligand specificity, the mixture may provide a much broader binding spectrum. A possible combinatorial mode of OBP participation in locust olfaction is an interesting aspect for future studies.

AUTHOR CONTRIBUTIONS

HB, JK, XJ, and PP conceived the study. XJ conducted the experiments. HB, JK, XJ, and PP interpreted the results. XJ and PP drafted the preliminary manuscript. HB and JK refined and approved the final manuscript.

FUNDING

The author XJ is supported by a grant from Chinese Scholarship Council (CSC) with the award number 201406350032.

ACKNOWLEDGMENTS

We are grateful to Dr. Ewald Grosse-Wilde (Max Planck Institute for Chemical Ecology) for the help with generating the antennal transcriptome sequence data base and initial bioinformatics analysis. We thank Heidrun Froß for her excellent technical assistance, Rosolino Bumbalo for his suggestions for the data analysis and Christian Heidel for suggestions and comments on the structure modeling.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fphys.2017.00734/full#supplementary-material

REFERENCES

Andronopoulou, E., Labropoulou, V., Douris, V., Woods, D. F., Biessmann, H., and Iatrou, K. (2006). Specific interactions among odorant-binding proteins of the African malaria vector *Anopheles gambiae*. *Insect Mol. Biol.* 15, 797–811. doi: 10.1111/j.1365-2583.2006.00685.x

Anton, S., Dufour, M. C., and Gadenne, C. (2007). Plasticity of olfactory-guided behaviour and its neurobiological basis: lessons from moths and locusts. *Entomol. Exp. Appl.* 123, 1–11. doi: 10.1111/j.1570-7458.2007.00516.x

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME Suite: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–208. doi: 10.1093/nar/gkp335

Ban, L., Scaloni, A., D'Ambrosio, C., Zhang, L., Yahn, Y., and Pelosi, P. (2003). Biochemical characterization and bacterial expression of an odorant-binding protein from *Locusta migratoria*. *Cell. Mol. Life Sci.* 60, 390–400. doi: 10.1007/s000180300032

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552. doi: 10.1093/oxfordjournals.molbev.a026334

Cicconardi, F., Di Marino, D., Olimpieri, P. P., Arthofer, W., Schlick-Steiner, B. C., and Steiner, F. M. (2017). Chemosensory adaptations of the mountain fly *Drosophila nigrosparsa* (Insecta: Diptera) through genomics' and structural biology's lenses. *Sci. Rep.* 7:43770. doi: 10.1038/srep43770

Damberger, F., Nikonova, L., Horst, R., Peng, G., Leal, W. S., and Wüthrich, K. (2000). NMR characterization of a pH-dependent equilibrium between two folded solution conformations of the pheromone-binding protein from *Bombyx mori*. *Protein Sci.* 9, 1038–1041. doi: 10.1110/ps.9.5.1038

DeLano, W. (2002). Pymol: an open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr.* 700.

Foret, S., and Maleszka, R. (2006). 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

Guo, M., Krieger, J., Große-Wilde, E., Mißbach, C., Zhang, L., and Breer, H. (2013). Variant ionotropic receptors are expressed in olfactory sensory neurons of coeloconic sensilla on the antenna of the desert locust (*Schistocerca gregaria*). *Int. J. Biol. Sci.* 10, 1–14. doi: 10.7150/ijbs.7624

Guo, W., Wang, X., Ma, Z., Xue, L., Han, J., Yu, D., et al. (2011). *CSP* and *takeout* genes modulate the switch between attraction and repulsion during behavioral phase change in the migratory locust. *PLoS Genet.* 7:e1001291. doi: 10.1371/journal.pgen.1001291

Hansson, B. S., and Stensmyr, M. C. (2011). Evolution of insect olfaction. *Neuron* 72, 698–711. doi: 10.1016/j.neuron.2011.11.003

Hassanali, A., Njagi, P. G. N., and Bashir, M. O. (2005). Chemical ecology of locust and related acridids. *Annu. Rev. Entomol.* 50, 223–245. doi: 10.1146/annurev.ento.50.071803.130345

Heifetz, Y., Voet, H., and Applebaum, S. W. (1996). Factors affecting behavioral phase transition in the desert locust, *Schistocerca gregaria* (Forskål) (Orthoptera: Acrididae). *J. Chem. Ecol.* 22, 1717–1734. doi: 10.1007/BF02272410

Hekmat-Scafe, D. S., Scafe, C. R., McKinney, A. J., and Tanouye, M. A. (2002). Genome-wide analysis of the odorant-binding protein gene family in *Drosophila melanogaster*. *Genome Res.* 12, 1357–1369. doi: 10.1101/gr.239402

Horst, R., Damberger, F., Luginbühl, P., Güntert, P., Peng, G., Nikonova, L., et al. (2001). NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. *Proc. Natl. Acad. Sci. U.S.A.* 98, 14374–14379. doi: 10.1073/pnas.251532998

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., et al. (2016). EGGNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44, D286–D293. doi: 10.1093/nar/gkv1248

Jiang, X., Pregitzer, P., Grosse-Wilde, E., Breer, H., and Krieger, J. (2016). Identification and characterization of two "sensory neuron membrane proteins" (SNMPs) of the desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae). *J. Insect Sci.* 16:33. doi: 10.1093/jisesa/iew015

Jin, X., Brandazza, A., Navarrini, A., Ban, L., Zhang, S., Steinbrecht, R. A., et al. (2005). Expression and immunolocalisation of odorant-binding and chemosensory proteins in locusts. *Cell. Mol. Life Sci.* 62, 1156–1166. doi: 10.1007/s00018-005-5014-6

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kosakovsky Pond, S. L., Frost, S. D. W., and Muse, S. V. (2005). HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21, 676–679. doi: 10.1093/bioinformatics/bti079

Kosakovsky Pond, S. L., Poon, A. F. Y., and Frost, S. D. W. (2009). "Estimating selection pressures on alignments of coding sequences," in *The Phylogenetic Handbook: A Practical Approach To Phylogenetic Analysis And Hypothesis Testing,* eds P. Lemey, M. Slaemi, and A. M. Vandamme (Cambridge: Cambridge University Press), 419–490.

Kumar, S., and Nussinov, R. (2002). Close-range electrostatic interactions in proteins. *Chembiochem.* 3, 604–617. doi: 10.1002/1439-7633(20020703)3:7<604::AID-CBIC604>3.0.CO;2-X

Lagarde, A., Spinelli, S., Tegoni, M., He, X., Field, L., Zhou, J. J., et al. (2011). The crystal structure of odorant binding protein 7 from *Anopheles gambiae* exhibits an outstanding adaptability of its binding site. *J. Mol. Biol.* 414, 401–412. doi: 10.1016/j.jmb.2011.10.005

Larter, N. K., Sun, J. S., and Carlson, J. R. (2016). Organization and function of *Drosophila* odorant binding proteins. *Elife* 5:e20242. doi: 10.7554/eLife.20242

Lartigue, A., Gruez, A., Briand, L., Blon, F., Bézirard, V., Walsh, M., et al. (2004). Sulfur single-wavelength anomalous diffraction crystal structure of a pheromone-binding protein from the honeybee *Apis mellifera* L. *J. Biol. Chem.* 279, 4459–4464. doi: 10.1074/jbc.M311212200

Leal, W. S. (2013). 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

Leal, W. S., Nikonova, L., and Peng, G. (1999). Disulfide structure of the pheromone binding protein from the silkworm moth, *Bombyx mori*. *FEBS Lett.* 464, 85–90. doi: 10.1016/S0014-5793(99)01683-X

Maida, R., Mameli, M., Müller, B., Krieger, J., and Steinbrecht, R. A. (2005). The expression pattern of four odorant-binding proteins in male and female silk moths, *Bombyx mori*. *J. Neurocytol.* 34, 149–163. doi: 10.1007/s11068-005-5054-8

Mao, Y., Xu, X., Xu, W., Ishida, Y., Leal, W. S., Ames, J. B., et al. (2010). Crystal and solution structures of an odorant-binding protein from the southern house mosquito complexed with an oviposition pheromone. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19102–19107. doi: 10.1073/pnas.1012274107

Miller, M. A., Pfeiffer, W., and Schwartz, T. (2012). "The CIPRES science gateway: enabling high-impact science for phylogenetics researchers with limited resources," in *Proceedings of the 1st Conference of the Extreme Science and Engineering Discovery Environment: Bridging from the Extreme to the Campus and Beyond* (Chicago), 1–8.

Nardi, J. B., Miller, L. A., Walden, K. K. O., Rovelstad, S., Wang, L., Frye, J. C., et al. (2003). Expression patterns of odorant-binding proteins in antennae of the moth *Manduca sexta*. *Cell Tissue Res.* 313, 321–333. doi: 10.1007/s00441-003-0766-5

Niu, D. J., Liu, Y., Dong, X. T., and Dong, S. L. (2016). Transcriptome based identification and tissue expression profiles of chemosensory genes in *Blattella germanica* (Blattaria: Blattidae). *Comp. Biochem. Physiol. D Genomics Proteomics* 18, 30–43. doi: 10.1016/j.cbd.2016.03.002

Ochieng, S. A., Hallberg, E., and Hansson, B. S. (1998). Fine structure and distribution of antennal sensilla of the desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae). *Cell Tissue Res.* 291, 525–536. doi: 10.1007/s004410051022

Ochieng, S. A., and Hansson, B. S. (1999). Responses of olfactory receptor neurones to behaviourally important odours in gregarious and solitarious desert locust, *Schistocerca gregaria*. *Physiol. Entomol.* 24, 28–36. doi: 10.1046/j.1365-3032.1999.00107.x

Pelosi, P., Iovinella, I., Zhu, J., Wang, G., and Dani, F. R. (2017). Beyond chemoreception: diverse tasks of soluble olfactory proteins in insects. *Biol. Rev. Camb. Philos. Soc.* doi: 10.1111/brv.12339. [Epub ahead of print].

Pelosi, P., Zhou, J. J., Ban, L. P., and Calvello, M. (2006). Soluble proteins in insect chemical communication. *Cell. Mol. Life Sci.* 63, 1658–1676. doi: 10.1007/s00018-005-5607-0

Pener, M. P., and Yerushalmi, Y. (1998). The physiology of locust phase polymorphism: an update. *J. Insect Physiol.* 44, 365–377. doi: 10.1016/S0022-1910(97)00169-8

Pesenti, M. E., Spinelli, S., Bezirard, V., Briand, L., Pernollet, J. C., Tegoni, M., et al. (2008). Structural basis of the honey bee PBP pheromone and pH-induced conformational change. *J. Mol. Biol.* 380, 158–169. doi: 10.1016/j.jmb.2008.04.048

Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi: 10.1038/nmeth.1701

Policy, S. D., and Conway, D. J. (2001). Strong diversifying selection on domains of the *Plasmodium falciparum* apical membrane antigen 1 gene. *Genetics* 158, 1505–1512.

Pophof, B. (1997). Olfactory responses recorded from sensilla coeloconica of the silkmoth *Bombyx mori*. *Physiol. Entomol.* 22, 239–248. doi: 10.1111/j.1365-3032.1997.tb01164.x

Pregitzer, P., Jiang, X., Grosse-Wilde, E., Breer, H., Krieger, J., and Fleischer, J. (2017). In search for pheromone receptors: certain members of the odorant receptor family in the desert locust *Schistocerca gregaria* (Orthoptera: Acrididae) are co-expressed with SNMP1. *Int. J. Biol. Sci.* 13, 911–922. doi: 10.7150/ijbs.18402

Qiao, H., He, X., Schymura, D., Ban, L., Field, L., Dani, F. R., et al. (2011). Cooperative interactions between odorant-binding proteins of *Anopheles gambiae*. *Cell. Mol. Life Sci.* 68, 1799–1813. doi: 10.1007/s00018-010-0539-8

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738. doi: 10.1038/nprot.2010.5

Sandler, B. H., Nikonova, L., Leal, W. S., and Clardy, J. (2000). Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. *Chem. Biol.* 7, 143–151. doi: 10.1016/S1074-5521(00)00078-8

Sawyer, S. L., Wu, L. I., Emerman, M., and Malik, H. S. (2005). Positive selection of primate *TRIM5*α identifies a critical species-specific retroviral restriction domain. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2832–2837. doi: 10.1073/pnas.0409853102

Scaloni, A., Monti, M., Angeli, S., and Pelosi, P. (1999). Structural analysis and disulfide-bridge pairing of two odorant-binding proteins from *Bombyx mori*. *Biochem. Biophys. Res. Commun.* 266, 386–391. doi: 10.1006/bbrc.1999.1791

Sheinerman, F. B., Norel, R., and Honig, B. (2000). Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struct. Biol.* 10, 153–159. doi: 10.1016/S0959-440X(00)00065-8

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Steinbrecht, R. A. (1996). Structure and function of insect olfactory sensilla. *Ciba Found. Symp.* 200, 158–174.

Suh, E., Bohbot, J. D., and Zwiebel, L. J. (2014). Peripheral olfactory signaling in insects. *Curr. Opin. Insect Sci.* 6, 86–92. doi: 10.1016/j.cois.2014.10.006

Sun, X., Zeng, F. F., Yan, M. J., Zhang, A., Lu, Z. X., and Wang, M. Q. (2016). Interactions of two odorant-binding proteins influence insect chemoreception. *Insect Mol. Biol.* 25, 712–723. doi: 10.1111/imb.12256

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Tegoni, M., Campanacci, V., and Cambillau, C. (2004). Structural aspects of sexual attraction and chemical communication in insects. *Trends Biochem. Sci.* 29, 257–264. doi: 10.1016/j.tibs.2004.03.003

Terrapon, N., Li, C., Robertson, H. M., Ji, L., Meng, X., Booth, W., et al. (2014). Molecular traces of alternative social organization in a termite genome. *Nat. Commun.* 5:3636. doi: 10.1038/ncomms4636

Theißen, G. (2002). Orthology: secret life of genes. *Nature* 415:741. doi: 10.1038/415741a

Unni, S., Huang, Y., Hanson, R. M., Tobias, M., Krishnan, S., Li, W. W., et al. (2011). Web servers and services for electrostatics calculations with APBS and PDB2PQR. *J. Comput. Chem.* 32, 1488–1491. doi: 10.1002/jcc.21720

Vieira, F. G., and Rozas, J. (2011). Comparative genomics of the odorant-binding and chemosensory protein gene families across the arthropoda: origin and evolutionary history of the chemosensory system. *Genome Biol. Evol.* 3, 476–490. doi: 10.1093/gbe/evr033

Vieira, F. G., Sánchez-Gracia, A., and Rozas, J. (2007). Comparative genomic analysis of the odorant-binding protein family in 12 *Drosophila* genomes: purifying selection and birth-and-death evolution. *Genome Biol.* 8:R235. doi: 10.1186/gb-2007-8-11-r235

Vogt, R. G., Callahan, F. E., Rogers, M. E., and Dickens, J. C. (1999). Odorant binding protein diversity and distribution among the insect orders, as indicated by LAP, an OBP-related protein of the true bug *Lygus lineolaris* (Hemiptera, heteroptera). *Chem. Senses* 24, 481–495. doi: 10.1093/chemse/24.5.481

Vogt, R. G., Große-Wilde, E., and Zhou, J. J. (2015). 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

Vogt, R. G., Rogers, M. E., Franco, M., and Sun, M. (2002). A comparative study of odorant binding protein genes: differential expression of the PBP1-GOBP2 gene cluster in *Manduca sexta* (Lepidoptera) and the organization of OBP genes in *Drosophila melanogaster* (Diptera). *J. Exp. Biol.* 205, 719–744.

Wang, X., Fang, X., Yang, P., Jiang, X., Jiang, F., Zhao, D., et al. (2014). The locust genome provides insight into swarm formation and long-distance flight. *Nat. Commun.* 5:2957. doi: 10.1038/ncomms3957

Wang, X., and Kang, L. (2014). Molecular mechanisms of phase change in Locusts. *Annu. Rev. Entomol.* 59, 225–244. doi: 10.1146/annurev-ento-011613-162019

Wang, Z., Yang, P., Chen, D., Jiang, F., Li, Y., Wang, X., et al. (2015). Identification and functional analysis of olfactory receptor family reveal unusual characteristics of the olfactory system in the migratory locust. *Cell. Mol. Life Sci.* 72, 4429–4443. doi: 10.1007/s00018-015-2009-9

Wheeler, W. C., Whiting, M., Wheeler, Q. D., and Carpenter, J. M. (2001). The phylogeny of the extant hexapod orders. *Cladistics* 17, 113–169. doi: 10.1111/j.1096-0031.2001.tb00115.x

Wogulis, M., Morgan, T., Ishida, Y., Leal, W. S., and Wilson, D. K. (2006). The crystal structure of an odorant binding protein from *Anopheles gambiae*: evidence for a common ligand release mechanism. *Biochem. Biophys. Res. Commun.* 339, 157–164. doi: 10.1016/j.bbrc.2005.10.191

Xu, P. X., Zwiebel, L. J., and Smith, D. P. (2003). Identification of a distinct family of genes encoding atypical odorant-binding proteins in the malaria vector mosquito, *Anopheles gambiae*. *Insect Mol. Biol.* 12, 549–560. doi: 10.1046/j.1365-2583.2003.00440.x

Xu, Y.-L., He, P., Zhang, L., Fang, S.-Q., Dong, S.-L., Zhang, Y.-J., et al. (2009). Large-scale identification of odorant-binding proteins and chemosensory proteins from expressed sequence tags in insects. *BMC Genomics* 10:632. doi: 10.1186/1471-2164-10-632

Yang, Y., Krieger, J., Zhang, L., and Breer, H. (2012). The olfactory co-receptor Orco from the migratory locust (*Locusta migratoria*) and the desert locust (*Schistocerca gregaria*): identification and expression pattern. *Int. J. Biol. Sci.* 8, 159–170. doi: 10.7150/ijbs.8.159

Yao, C. A. (2005). Chemosensory coding by neurons in the coeloconic sensilla of the *Drosophila* antenna. *J. Neurosci.* 25, 8359–8367. doi: 10.1523/JNEUROSCI.2432-05.2005

Yu, F., Zhang, S., Zhang, L., and Pelosi, P. (2009). Intriguing similarities between two novel odorant-binding proteins of locusts. *Biochem. Biophys. Res. Commun.* 385, 369–374. doi: 10.1016/j.bbrc.2009.05.074

Zhang, S., Pang, B., and Zhang, L. (2015). Novel odorant-binding proteins and their expression patterns in grasshopper, *Oedaleus asiaticus*. *Biochem. Biophys. Res. Commun.* 460, 274–280. doi: 10.1016/j.bbrc.2015.03.024

Zheng, J., Li, J., Han, L., Wang, Y., Wu, W., Qi, X., et al. (2015). Crystal structure of the *Locusta migratoria* odorant binding protein. *Biochem. Biophys. Res. Commun.* 456, 737–742. doi: 10.1016/j.bbrc.2014.12.048

Zhou, J. J., Huang, W., Zhang, G. A., Pickett, J. A., and Field, L. M. (2004). "Plus-C" odorant-binding protein genes in two *Drosophila* species and the malaria mosquito *Anopheles gambiae*. *Gene* 327, 117–129. doi: 10.1016/j.gene.2003.11.007

Zhou, J. J., Vieira, F. G., He, X. L., Smadja, C., Liu, R., Rozas, J., et al. (2010). Genome annotation and comparative analyses of the odorant-binding proteins and chemosensory proteins in the pea aphid *Acyrthosiphon pisum*. *Insect Mol. Biol.* 19, 113–122. doi: 10.1111/j.1365-2583.2009.00919.x

Zubkov, S., Gronenborn, A. M., Byeon, I. J. L., and Mohanty, S. (2005). Structural consequences of the pH-induced conformational switch in A. polyphemus pheromone-binding protein: mechanisms of ligand release. *J. Mol. Biol.* 354, 1081–1090. doi: 10.1016/j.jmb.2005.10.015

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physiology

ORIGINAL RESEARCH
published: 17 April 2018
doi: 10.3389/fphys.2018.00417

Odorant Binding Proteins of the Desert Locust *Schistocerca gregaria* (Orthoptera, Acrididae): Topographic Expression Patterns in the Antennae

Xingcong Jiang[1], Miriam Ryl[1], Jürgen Krieger[2], Heinz Breer[1] and Pablo Pregitzer[1]**

[1] Institute of Physiology, University of Hohenheim, Stuttgart, Germany, [2] Department of Animal Physiology, Institute of Biology/Zoology, Martin Luther University Halle-Wittenberg, Halle, Germany

Odorant binding proteins (OBPs) enriched in the sensillum lymph are instrumental in facilitating the transfer of odorous molecules to the responsive receptors. In Orthopteran locust species, an in-depth understanding of this important soluble protein family is still elusive. In a previous study, we have demonstrated that the repertoire of locust OBPs can be divided into four major clades (I–IV) on the phylogenetic scale and for representatives of subfamily I-A and II-A a distinct sensilla-specific expression pattern was determined. In this study, by focusing on a representative locust species, the desert locust *Schistocerca gregaria*, we have explored the antennal topographic expression for representative OBPs of other subfamilies. First, subtypes of subfamily III-A and III-B were exclusively found in sensilla chaetica. Then, a similar expression pattern in this sensillum type was observed for subfamily I-B subtypes, but with a distinct OBP that was expressed in sensilla coeloconica additionally. Moreover, the atypical OBP subtype from subfamily IV-A was expressed in a subpopulation of sensilla coeloconica. Last, the plus-C type-B OBP subtype from subfamily IV-B seems to be associated with all four antennal sensillum types. These results profile diversified sensilla-specific expression patterns of the desert locust OBPs from different subfamilies and complex co-localization phenotypes of distinct OBP subtypes in defined sensilla, which provide informative clues concerning their possible functional mode as well as a potential interplay among OBP partners within a sensillum.

Keywords: locust, *Schistocerca gregaria*, odorant binding protein, sensilla, topographic expression

OPEN ACCESS

Edited by:
Shuang-Lin Dong,
Nanjing Agricultural University, China

Reviewed by:
Paolo Pelosi,
Università degli Studi di Pisa, Italy
Jin Zhang,
Max Planck Institute for Chemical Ecology (MPG), Germany

****Correspondence:***
Xingcong Jiang
jiangxingcong@126.com
Pablo Pregitzer
p_pregitzer@uni-hohenheim.de

Specialty section:
This article was submitted to Invertebrate Physiology, a section of the journal Frontiers in Physiology

Received: *28 February 2018*
Accepted: *04 April 2018*
Published: *17 April 2018*

Citation:
Jiang X, Ryl M, Krieger J, Breer H and Pregitzer P (2018) Odorant Binding Proteins of the Desert Locust Schistocerca gregaria (Orthoptera, Acrididae): Topographic Expression Patterns in the Antennae. Front. Physiol. 9:417. doi: 10.3389/fphys.2018.00417

INTRODUCTION

Insects utilize hair-like cuticle appendages, so called sensilla, to receive environmental olfactory signals (Steinbrecht, 1996; Hansson and Stensmyr, 2011; Suh et al., 2014). Hydrophobic odorous molecules have to travel through the aqueous sensillum lymph before reaching the receptors residing in the chemosensory membrane of olfactory neurons in the antennae (Vogt et al., 1999; Leal, 2013; Suh et al., 2014). This passage is supposed to be facilitated by odorant binding proteins (OBPs) in the sensillum lymph, an important soluble protein family that is capable to accommodate and transfer odorant molecules (Vogt and Riddiford, 1981; Pelosi et al., 2006, 2014;

Vieira and Rozas, 2011). OBPs are short polypeptides of approximately 110–200 amino acids that fold into a globular shape forming an interior binding cavity, where the interaction with odorous molecules takes place (Sandler et al., 2000; Tegoni et al., 2004). The sequence of classic OBPs is characterized by six conserved cysteine (C) residues, a hall mark of classic OBPs; plus-C or minus-C OBPs are categorized with more or less than six conserved C-residues (Xu et al., 2003; Zhou et al., 2004; Foret and Maleszka, 2006; Vieira and Rozas, 2011). OBPs are produced by auxiliary cells which envelope the sensory neurons by their extended processes. The enrichment of OBPs in the sensillum types that respond to olfactory cues has been reported for many insect species (Pelosi et al., 2014, 2017). Beyond the olfactory sensilla, OBP expression has also been found in the sensilla that are seemingly dedicated to gustatory cues (Galindo and Smith, 2001; Jeong et al., 2013). Incidentally, besides the sensilla-specific expression in the chemosensory organs, like the antennae, OBPs are also expressed in other tissues of which the functional connotations seem to be less associated with chemical communication (Pelosi et al., 2017).

Schistocerca gregaria, the desert locust, represents a model organism of the Orthopteran order, which emerged much earlier than the Lepidopteran and Dipteran orders on the evolutionary scale (Wheeler et al., 2001; Vogt et al., 2015). Locusts are characterized by a hemimetabolous life circle and a population density dependent behavioral plasticity, which involves the perception of behavioral relevant semiochemicals (Pener and Yerushalmi, 1998; Hassanali et al., 2005; Guo et al., 2011; Wang and Kang, 2014). For locust species an in-depth understanding of the OBP family from either molecular or cellular perspective is still elusive (Ban et al., 2003; Jin et al., 2005; Jiang et al., 2009; Xu et al., 2009; Yu et al., 2009). Previously, we have conducted a comprehensive sequence analysis of the OBP families from *Schistocerca gregaria* and three other locust species which classifies locust OBPs into several categories, e.g., classic, plus-C type-A, plus-C type-B, minus-C and atypical OBPs. Based on the phylogenetic relationship locust OBPs reside within four major phylogenetic clades. Concentrating on the two OBP subfamilies I-A and II-A, which comprise the classic OBP subtypes, we have found a characteristic sensilla-specific expression pattern for the desert locust OBP representatives in the antennae (Jiang et al., 2017). In the present study, we set out to explore the antennal topographic expression of desert locust OBPs from the remaining subfamilies on the phylogenetic tree.

MATERIALS AND METHODS

Animals and Tissue Collection

The desert locust *Schistocerca gregaria* reared on the gregarious phase were purchased from Bugs-International GmbH (Irsingen/Unterfeld, Germany). Antennae of adult male and adult female were dissected using autoclaved surgical scissors and were immediately frozen in liquid nitrogen. Tissues were stored at −70°C before subsequent RNA extraction.

RNA Extraction and Reverse Transcription PCR (RT-PCR)

Total RNA was extracted from the frozen tissues using TRIzol reagent (Invitrogen) following the protocol recommended by the manufacturer. The poly $(A)^+$ RNA was purified from 100 μg of total RNA using oligo $(dT)_{25}$ magnetic dynabeads (Invitrogen) conforming to the recommendation of the supplier. The generated mRNA was reverse transcribed to cDNA in a total volume of 20 μl employing SuperScript™ III Reverse Transcriptase (Invitrogen). PCR conditions used in RT-PCR experiments were: 94°C for 1 min 40 s, then 20 cycles with 94°C for 30 s, 60°C for 30 s and 72°C for 2 min, with a reduction in the annealing temperature by 0.5°C per cycle, which was followed by a further cycles (20 times) on the condition of the last cycling step (annealing temperature was 50°C) and a final extension step for 7 min at 72°C. The sense (s) and antisense (as) primer pairs used for amplification of the desert locust OBP coding sequences were:

OBP2 s, atggccagccattgccacgccacc
OBP2 as, ttctccggatttcctaaactccgc
OBP3 s, atgctgctggcagcccccgcaaagg
OBP3 as, ctttttcctgatcaagcatccacc
OBP4 s, cctgtggcgacacttggtggccg
OBP4 as, gcctttagccatcatccccctt
OBP7 s, cgatgtgcttcgtcggtgggtgat
OBP7 as, acgtcgttctcgtcggactctgga
OBP8 s, agactcgccaacccgccaca
OBP8 as, ttctgacggggcgtgtggga
OBP9 s, gccacagtccggtgcagcat
OBP9 as, aatctggtcgctgacgcact
OBP12 s, acaactcttgcagccatgaagtgg
OBP12 as, tccacttcttgttcccatactggt
OBP13 s, gagctgaggtaatgaagagggtca
OBP13 as, cctgcacattcagatccaagcagc

The primer pairs against other desert locust OBP subtypes were given in (Jiang et al., 2017).

Synthesis of Riboprobes for *in Situ* Hybridization

PCR products of the desert locust OBP coding sequences were sequenced and then cloned into pGEM-T vectors (Invitrogen) for the subsequent *in vitro* transcription. The linearized pGEM-T vectors consisting of desert locust OBP coding sequences were utilized to synthesize both sense and antisense riboprobes labeled with digoxigenin (Dig) or biotin (Bio) using the T7/SP6 RNA transcription system (Roche, Germany). The synthesis procedure stringently followed the protocol provided by the manufacturer.

In Situ Hybridization

Antennae of adult *Schistocerca gregaria* were dissected and embedded in Tissue-Tek O.C.T. Compound (Sakura Finetek Europe, Netherlands). Cryosections with a 12 μm-thickness were thaw mounted on SuperFrost Plus slides (Menzel-Gläser, Braunschweig, Germany) at −21°C (Jung CM300 cryostat). RNA *In situ* hybridization was performed as previously reported (Yang et al., 2012; Guo et al., 2013; Jiang et al., 2016, 2017). In brief,

the cryosections were firstly fixed (4% paraformaldehyde in 0.1 M $NaHCO_3$, pH 9.5) at 4°C for 22 min, followed by a series of treatments at room temperature: a wash for 1 min in PBS (phosphate buffered saline = 0.85% NaCl, 1.4 mM KH_2PO_4, 8 mM Na_2HPO_4, pH 7.1), an incubation for 10 min in 0.2 M HCl, another wash for 1 min in PBS, an incubation for 10 min in acetylation solution (0.25% acetic anhydride freshly added in 0.1 M triethanolamine) and washes for three times in PBS (3 min each). Afterward, the sections were pre-hybridized for 1 h at 60°C bathed in hybridization buffer (50% formamide, 5x SSC, 50 μg/ml heparin, and 0.1% Tween-20). A volume of 150 μl hybridization solution containing experiment riboprobes in hybridization buffer was evenly applied onto the tissue section. A coverslip was placed on top and slides were incubated in a moister box at 60°C overnight (18–20 h). After hybridization, slides were washed twice for 30 min in 0.1x SSC at 60°C, then each slide was treated with 1 ml 1% blocking reagent (Roche) for 35 min at room temperature.

Visualization of Dig-labeled riboprobe hybridizations was achieved by using an anti-Dig alkaline phosphatase (AP) conjugated antibody (1:500, Roche) and NBT/BCIP as substrates. Antennal sections were analyzed on a Zeiss Axioskope2 microscope (Zeiss, Oberkochen, Germany) equipped with Axiovision software. For two-color fluorescent *in situ* hybridization visualization of hybridized riboprobes was performed by using an anti-Dig AP-conjugated antibody in combination with HNPP/Fast Red (Roche) for Dig-labeled probes and an streptavidin horse radish peroxidase-conjugate together with fluorescein-tyramides as substrate (TSA kit, Perkin Elmer, Waltham, MA, United States) for biotin-labeled probes. Tissue sections in two-color FISH experiments were analyzed with a Zeiss LSM510 Meta laser scanning microscope (Zeiss, Oberkochen, Germany), and the acquired confocal images stacks were processed by ZEN 2009 software. The images presented in this paper integrate the projections of a series of optical planes selected from continuous confocal image stacks. For clear data presentation, images were only adjusted in brightness and contrast. It is noted that the images obtained via the two-color FISH approach always contained the cuticle unspecifically stained, most likely due to the intrinsic fluorescence. To clarify

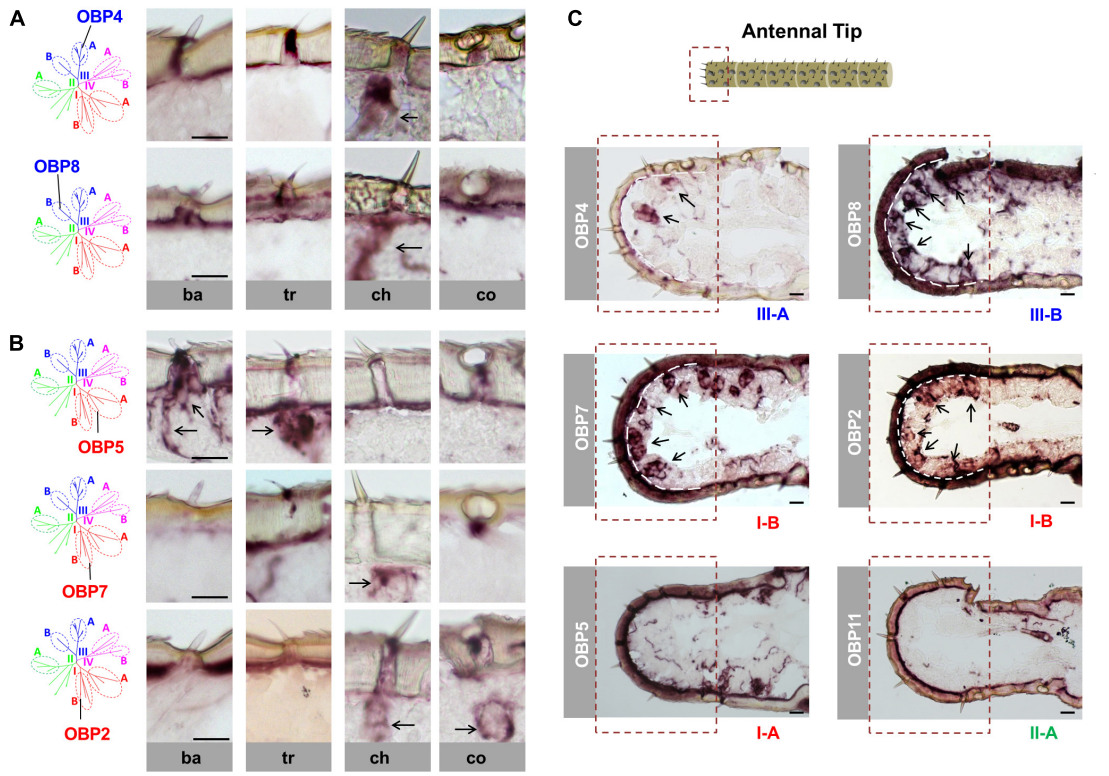


FIGURE 1 | Sensilla chaetica express OBP subtypes of two phylogenetic clades. The schematic diagram of the phylogenetic tree (left in **A,B**) was adapted from Jiang et al. (2017) where OBP families of four locust species have been analyzed. The specific *S. gregaria* OBP subtypes studied in this analysis were indicated. A detail classification of different subfamilies is illustrated in **Supplementary Figure S1**. Topographic expression of OBPs was visualized by using antisense riboprobes specifically targeting distinct OBP subtypes in conjunction with chromogenic *in situ* hybridization (ISH). **(A,B)** Visualization of the labeled cells expressing distinct OBP subtypes of subfamily III-A, III-B, I-A, and I-B in four morphological types of antennal sensilla. Ba, sensilla basiconica; Tr, sensilla trichodea; Ch, sensilla chaetica; Co, sensilla coeloconica. The visible labeled structures are denoted by black arrows. **(C)** Visualization of the cells expressing distinct OBP subtypes from different subfamilies on the tip of the antennae. Notably, sensilla chaetica are exclusively enriched on the antennal tip (Ochieng et al., 1998). The area of the antennal tip is indicated by a box with a dashed line. The visible cell clusters are denoted by black arrows, and in some images the interface between the cuticle and cellular layer is depicted as a white dashed line. The subfamily to which a distinct OBP subtype belongs is annotated below the images. Scale bars, 20 μm.

the specific fluorescent labeling, a dashed line was added to indicate the interface between the cuticle and the cellular layers. Antennal sections of both male and female were analyzed under the same experimental conditions and were tested with each generated riboprobes. There were no discernible gender dependent differences regarding to the labeling intensity as well as the labeling pattern. Therefore, only the images acquired from male antenna sections were presented in this paper.

RESULTS

Topographic Expression Patterns of OBP Subtypes From Clade I and III

A previously performed phylogenetic analysis of OBPs from four locust species revealed that the locust OBP family can be divided into four major clades consisting of three conserved subfamilies. For the two subfamilies I-A and II-A, which both comprise classic OBP subtypes, we found that the representative I-A subtypes are expressed in sensilla basiconica and sensilla trichodea, whereas the representative II-A subtypes are expressed in sensilla coeloconica (Jiang et al., 2017). In this study, we concentrated on the conserved subfamily III-A, which includes the plus-C type-A OBP subtypes that share only low sequence identities with the classic OBP subtypes. In order to explore their sensilla-specific expression pattern, we adopted the strategy of mRNA *in situ* hybridization and assessed the expression of OBP4, a representative subtype of subfamily III-A, in the four morphologically distinguishable types of antennal sensilla. The results of these approaches revealed a discernible labeling of OBP4 expressing cells in sensilla chaetica; no labeling was visible in any of the other three sensillum types (**Figure 1A**). Apart from the subfamily III-A, clade III also comprises subfamily III-B, which includes the classic OBP subtype OBP8 and its orthologs. Analysis of the expression pattern revealed that OBP8-positive cells were also exclusively enriched in sensilla chaetica, thus resembling the plus-C type-A subtype OBP4 (**Figure 1A**). Together, these results imply that OBP subtypes of the clade III are specifically expressed in sensilla chaetica and thus deviate from the distribution of OBP subtypes from subfamilies I-A and II-A (Jiang et al., 2017).

In view of a clade-specific spatial expression pattern as seen for clade III (see above) it is interesting to note that clade I comprises, besides the conserved subfamily I-A, the more divergent subfamily I-B (**Supplementary Figure S1**). Since representatives of subfamily I-A were found to be restricted to

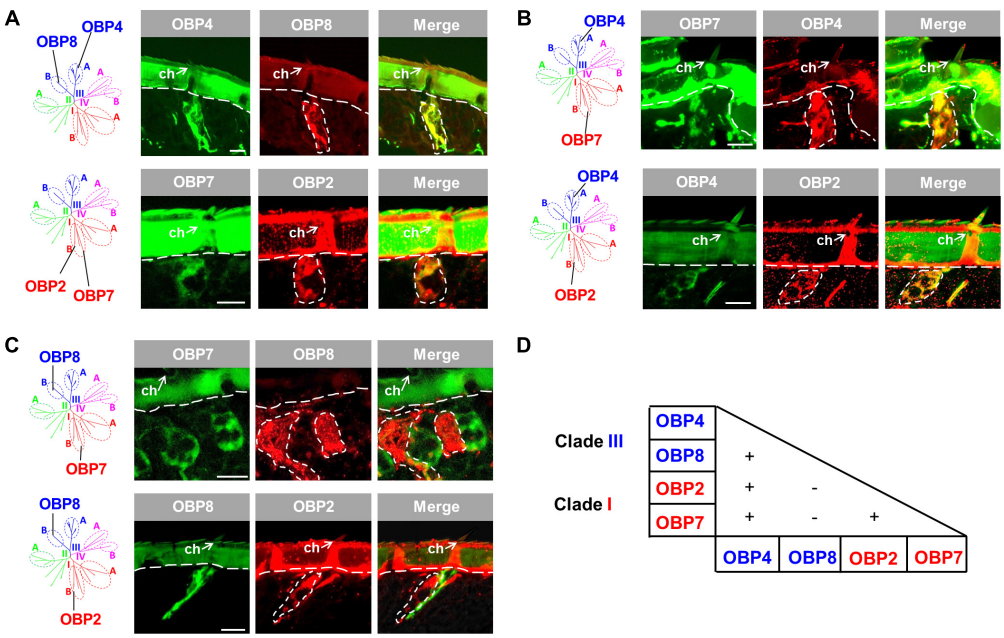


FIGURE 2 | Co-localization of four OBP subtypes from two clades in sensilla chaetica. The relative localization of OBP types was analyzed by two-color fluorescent *in situ* hybridization (FISH) using combinations of specific DIG- or biotin-labeled antisense riboprobes against distinct OBP subtypes. **(A)** OBP subtypes of the same phylogenetic clade are co-expressed in the same set of cells in sensilla chaetica (ch). OBP4 and OBP8 belong to clade III, and OBP2 and OBP7 belong to clade I. **(B)** OBP2 and OBP7 residing in subfamily I-B are co-expressed with OBP4 from subfamily III-A in the same set of cells in sensilla chaetica. **(C)** OBP2 and OBP7 residing in subfamily I-B are expressed in a different set of cells from OBP8 (subfamily III-B). It is noted that the labeling for OBP7 cells pronounces a distinct cell population in a sensillum chaeticum different from the one containing OBP8 expressing cells. In contrast, OBP2 and OBP8 labeled cells were found in the same sensillum chaeticum. The interface between the cuticle and cellular layer is depicted by a white dashed line. Distinct cell clusters visualized by the DIG-labeled probes (red) are encircled by white dashed lines. These areas are indicated also on the images showing the merged red and green fluorescence channels. **(D)** Recapitulation of the co-localization relationship among the four sensilla chaetica-positive OBP subtypes. The expression of two OBP subtypes in the same set of cells is denoted as "+", while "–" indicates expression of two OBP subtypes in different set of cells. The color code to distinguish OBP subtypes conforms to that for the phylogenetic analysis (**Supplementary Figure S1**). Scale bars, 20 μm.

sensilla basiconica and trichodea (**Figure 1B**) (Jiang et al., 2017), the question arises, whether OBPs of subfamily I-B may also be expressed in the same sensillum types. To scrutinize this notion, we have analyzed OBP2 and OBP7, the two subtypes in subfamily I-B. The results are depicted in **Figure 1B** and indicate that labeling for OBP2 and OBP7 was neither found in sensilla basiconica nor in sensilla trichodea; however, the labeling was present in sensilla chaetica and for OBP2 the labeled cells were concomitantly visible in sensilla coeloconica (**Figure 1B**). These data indicate that the topographic distribution of subfamily I-B OBPs clearly deviate from that of their counterparts of subfamily I-A and demonstrate that there is no clade-specific spatial expression pattern for members of clade I.

Previous anatomical studies have shown that sensilla chaetica are highly enriched at the tip of the antennae, a region with relatively few of the other three sensillum types (Ochieng et al., 1998). This spatial segregation of sensilla chaetica allows a more detailed analysis of the four identified OBP subtypes in this sensillum type. As shown in **Figure 1C**, numerous labeled cells were visualized using the probes for OBP4 (subfamily III-A), OBP8 (subfamily III-B) as well as OBP2 and OBP7 (subfamily I-B). In contrast, with the riboprobes for OBP subtypes that are specifically expressed in other sensillum types, such as OBP5 (subfamily I-A) and OBP11 (subfamily II-A), no discernible labeling was found at the antennal tip (**Figure 1C**).

Co-localization of OBP Subtypes From Different Subfamilies in Sensilla Chaetica

Since the four OBP subtypes reside in two different phylogenetic clades, we ask whether the different OBP subtypes are present in the same set of cells or in distinct cell populations of sensilla chaetica. To approach this question, we have generated either DIG- or BIO-labeled riboprobes for each OBP subtype and by means of two-color FISH analysis we have visualized the relative topographic localization of the labeled cells (**Figure 2**). In a first step, we have analyzed the subtypes from the same phylogenetic clade. For the two subtypes from clade III, OBP4 and OBP8, a widely overlapped labeling was found indicating that they were co-localized in the same set of cells in many, if not all, inspected sensilla chaetica (**Figure 2A**). Analysis for the two subtypes from subfamily I-B, OBP2 and OBP7, also revealed a largely overlapped labeling (**Figure 2A**). These results suggest that within clade III and subfamily I-B OBP subtypes are generally expressed in the same set of cells in sensilla chaetica. In a next step, we explored whether OBP subtypes from different clades may either be expressed in the same or a different set of cells. For the member of subfamily III-A (OBP4) and the members of subfamily I-B (OBP2 and OBP7) a largely overlapping labeling was observed (**Figure 2B**). However, for the member of subfamily III-B (OBP8) and the members of subfamily I-B (OBP2 and OBP7) no labeling overlap was found (**Figure 2C**). While labeling for OBP2 and OBP8 was found in different sets of cells of the same sensillum chaeticum, interestingly, OBP7 seemed to be present in the cells of distinct sensilla chaetica which differ from sensilla with OBP8-positive cells (**Figure 2C**). These results emphasize the complex co-localization relationship among OBP2, OBP4, and OBP8. The notion that OBP4 and OBP8 may be separately expressed in a subset of sensilla chaetica was confirmed upon a comprehensive inspection of the labeling for OBP4 and OBP8 (**Supplementary Figure S2**), indicating a broader expression scope for OBP4 in certain sensilla chaetica. In sum, the results indicate that sensilla chaetica express OBP subtypes from more than one phylogenetic clade, and co-localization of the OBP subtypes in distinct sensilla subtypes occurs in a combinatorial mode.

OBP2, Member of Subfamily I-B, Is Expressed in Sensilla Coeloconica and Chaetica

The results depicted in **Figure 1** indicate that OBP2, a subtype of subfamily I-B, may not only be expressed in sensilla chaetica (see above) but also in sensilla coeloconica. To substantiate the observation that OBP2 is in fact expressed in sensilla coeloconica, we utilized IR8a, the co-receptor of divergent IRs (Abuin et al., 2011; Guo et al., 2013), as a specific marker of sensory neurons housed in sensilla coeloconica. The results of double labeling experiments indicate that labeled OBP2 cells are tightly surrounding IR8a-positive cells in sensilla coeloconica (**Figure 3**).

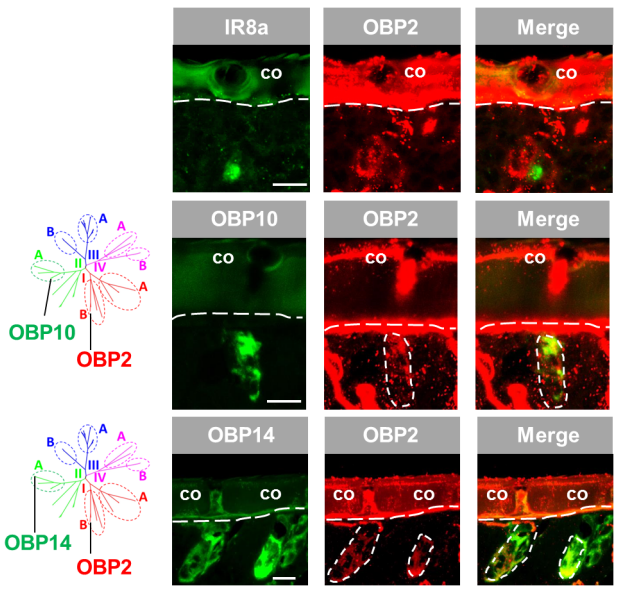


FIGURE 3 | OBP2 from subfamily I-B is expressed in sensilla coeloconica and sensilla chaetica. The relative localization of OBP2 and the marker genes indicating expression in sensilla coeloconica (co) was analyzed by utilizing antisense riboprobes targeting specific molecular elements in conjunction with two-color FISH. (Upper) OBP2 expressing cells surround a sensory neuron positive for IR8a, a specific molecular marker for sensilla coeloconica. (Middle and lower) OBP10 and OBP14 from the subfamily II-A are specifically expressed in sensilla coeloconica and are employed to mark two different sets of auxiliary cells in this sensillum type (Jiang et al., 2017). The interface between the cuticle and the cellular layer is denoted by a white dashed line. Distinct cell clusters positive for the DIG-labeled OBP2 probe (red) are encircled by white dashed lines. The position of these cell clusters is also indicated on the images showing the merged red and green fluorescence channels. Scale bars, 20 μm.

Given that in sensilla coeloconica OBP subtypes of subfamily II-A are specifically expressed, the question arises as to whether OBP2, a member of subfamily I-B, may be co-expressed with OBP subtypes of subfamily II-A. As representatives for subfamily II-A OBP10 and OBP14 were investigated. The results depicted in **Figure 3** indicate that the labeling for OBP2 indeed overlapped with that for the subfamily II-A representatives, indicating that in a set of sensilla coeloconica OBP subtypes from subfamily I-B and subfamily II-A coexist. Furthermore, the results confirm that OBP2 is in fact present in the two types of sensilla, sensilla coeloconica and sensilla chaetica.

Topographic Expression Pattern of an Atypical OBP Subtype From Subfamily IV-A

The atypical OBP subtypes converge onto the subfamily IV-A (**Supplementary Figure S1**) and are characterized by an extraordinary long span between C1 and C2 in comparison to the classic OBP subtypes (Jiang et al., 2017). This unique feature has raised the question whether atypical OBP subtypes may be expressed in specific sensillum types and/or in distinct cell populations. To approach this question, we have analyzed the expression pattern of OBP12, a subtype of subfamily IV-A. The results of labeling experiments are depicted in **Figure 4A** and indicate that OBP12 expressing cells were exclusively located in sensilla coeloconica. The sensilla specificity was subsequently confirmed by demonstrating the co-localization of OBP12 expressing cells and IR8a-positive cells in one sensillum coeloconicum (**Figure 4A**). Since OBPs of subfamily II-A are specifically expressed in sensilla coeloconica, we explored whether OBP12 may be co-localized with OBPs of subfamily II-A. Intriguingly, we found that the labeling for OBP12 cells did not overlap with the cells positive for OBP10 or OBP14 (**Figure 4B**), suggesting that OBP12 is expressed in a distinct subset of sensilla coeloconica.

It is yet unclear how many IR8a-positive neurons are surrounded by the auxiliary cells that express OBPs of subfamily

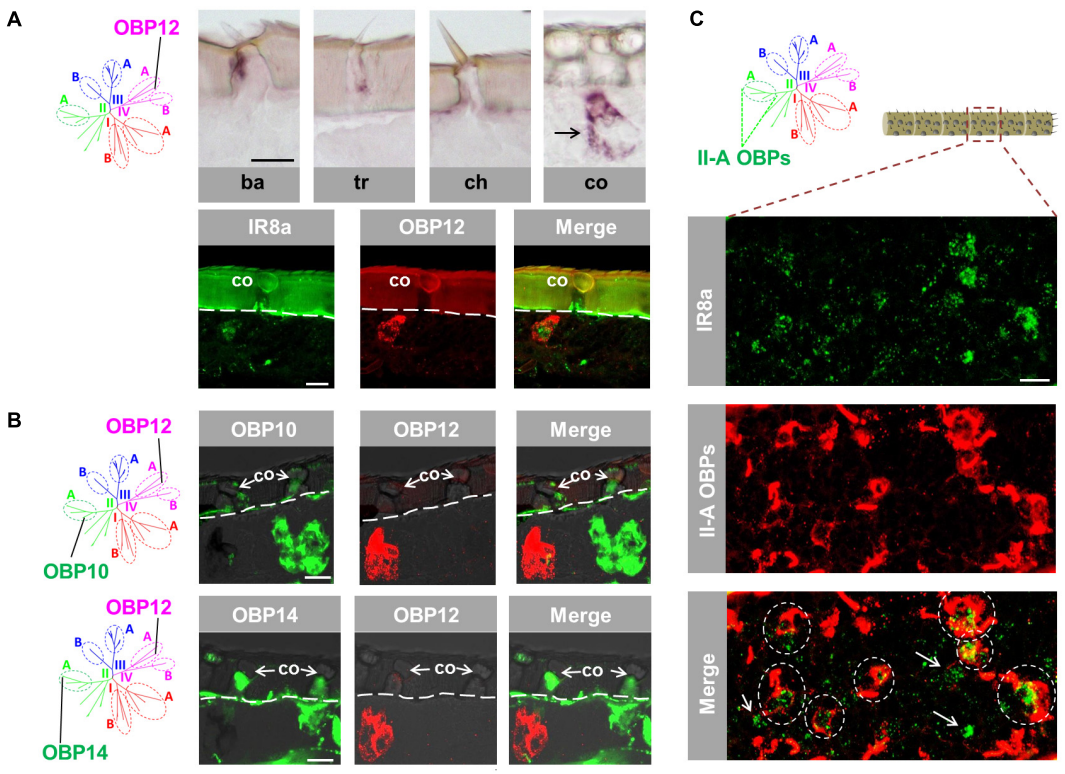


FIGURE 4 | An atypical OBP subtype pronounces a segregated subpopulation of sensilla coeloconica. **(A)** OBP12, an atypical OBP subtype residing in subfamily IV-A, is exclusively expressed in sensilla coeloconica (co). Upper panel: OBP12 expressing cells were analyzed in four morphological types of antennal sensilla using specific riboprobe by means of ISH. Labeled OBP12 cells were detected only in sensilla coeloconica and are indicated by a black arrow. Ba, sensilla basiconica; Tr, sensilla trichodea; Ch, sensilla chaetica; Co, sensilla coeloconica. Lower panel: A co-localization of OBP12 expressing cells and an IR8a-positive neuron in sensilla coeloconica was visualized by means of two-color FISH. **(B)** The labeling of OBP12-positive cells does not overlap with the labeling of cells expressing OBP10 and OBP14 from subfamily II-A. The interface between the cuticle and the cellular layer is depicted by a white dashed line. **(C)** Three OBP subtypes of subfamily II-A label the major population of auxiliary cells in sensilla coeloconica. The presented optical view was adopted from a distal antennal segment and presumably illustrates the typical association between IR8a neurons and subfamily II-A OBP cells. The utilized DIG-labeled probes representing the three ortholog groups comprised in subfamily II-A (**Supplementary Figure S1**) were generated by mixing the riboprobes against OBP10, OBP11, and OBP14, respectively, at a ratio of 1:1:1. Areas encircled by white dashed lines indicate IR8a neurons that are co-localized with auxiliary cells expressing the subfamily II-A OBPs in the same coeloconic sensillum. White arrows indicate those IR8a neurons that are presumably not associated with auxiliary cells expressing subfamily II-A OBPs. Scale bars, 20 μm.

II-A. To scrutinize this notion, double labeling experiments were performed with a probe for IR8a and a mix of riboprobes for OBP10, OBP11 and OBP14, which represent the three ortholog groups in subfamily II-A (**Supplementary Figure S1**). The results depicted in **Figure 4C** indicate that a considerable portion of IR8a-positive cells are engulfed by cells expressing OBPs of subfamily II-A (ovals in dash line). The remaining fraction of IR8a neurons seems to express non-II-A OBP subtypes, possibly OBP12. Together the results indicate that the atypical OBP subtype OBP12 is expressed in a segregated population of sensilla coeloconica.

Topographic Expression and Sensillum-Association of a Plus-C Type-B OBP Subtype

We have previously distinguished two categories of the plus-C OBPs based on the distinct conserved-C-patterns (Jiang et al., 2017). While the type-A OBP subtypes are grouped into the subfamily III-A, the type-B OBP subtypes are grouped into the subfamily IV-B (**Supplementary Figure S1**). Whereas type-A OBPs are expressed in sensilla chaetica (**Figure 1**), the expression pattern of type-B OBP subtypes is unclear. It is possible that the type-B OBPs share the sensilla specificity either with their close relatives in subfamily IV-A, e.g., OBP12, or with their type-A counterparts in subfamily III-A, e.g., OBP4. To approach this question, we have used a specific riboprobe for OBP9, a representative plus-C type-B subtype and assessed series of horizontal sections through the antennae. Upon an inspection of a deep anatomical plane close to the antennal nerve bundle, we found labeled structures for OBP9 which seemed to be less associated with a specific sensillum type, as typically found for the other OBP subtypes (**Figures 1**, **3**, **4**). Nevertheless, labeled cell bodies seemed to extend cytoplasmic processes which enclosed sensory neurons (**Figure 5A**). Interestingly, when we inspected an anatomical plane located closer to the cuticle, a more intense labeling was observed and a distinct nest-like labeling pattern for OBP9 emerged (**Figure 5B**).

The notion that OBP9 labeling seems to be associated with multiple sensillum types was scrutinized by analyzing a possible co-localization of OBP9 labeling with markers for distinct neuron types. In a first approach, Orco, the obligate co-receptor of ORs, was used to label the multiple sensory neurons in sensilla basiconica (Ochieng et al., 1998). It was found that OBP9 cells tightly surrounded the Orco-positive neuron clusters (**Figure 6**). Similarly, OR3 was used as a marker for sensilla trichodea and IR8a was used as a marker for sensilla coeloconica; it was observed that OBP9 labeling engulfed OR3- and IR8a- expressing neurons (**Figure 6**). OBP8 is considered to be specific for sensilla chaetica (**Figure 1**) and the results of double labeling experiments with OBP9 and OBP8 clearly indicated a co-localization (**Figure 6**). Together, these results indicate an association of the plus-C type-B OBP9 with all four antennal sensillum types.

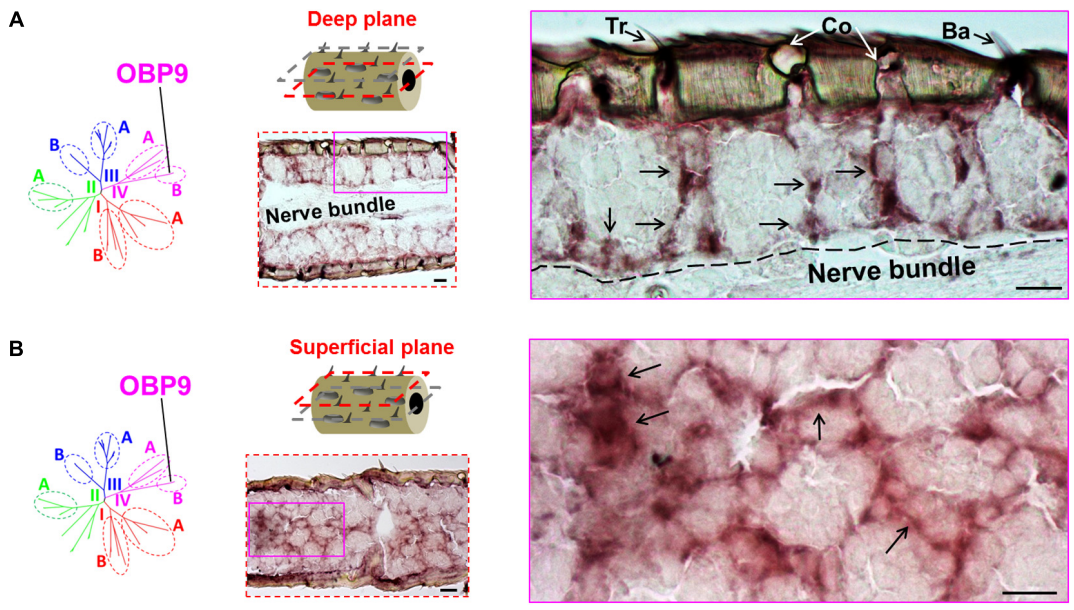


FIGURE 5 | Topographic expression of the plus-C type-B OBP9 in the antennae. The topographic expression of OBP9 was analyzed by using a specific antisense riboprobe in conjunction with ISH. **(A,B)** Labeling of OBP9 expressing cells in two different anatomical planes of the antennae. OBP9 represents the plus-C type-B OBPs that are grouped into subfamily IV-B (diagrams, left lane). Two different horizontal planes are shown to visualize the OBP9 expression pattern: the first deep plane (**A**, middle lane, red dashed frame) penetrates into the central nerve bundle; the second superficial plane (**B**, middle lane, red dashed frame) is located between the cuticle and central nerve bundle. For each plane a selected area (magenta box, middle lane) of the analyzed section is shown at a higher magnification on the right. Black arrows indicate the visible cell bodies as well as their extended processes. The border between the cellular layer and the nerve bundle is depicted by a black dashed line. Tr, sensilla trichodea; Co, sensilla coeloconica; Ba, sensilla basiconica. Scale bars, 20 μm.

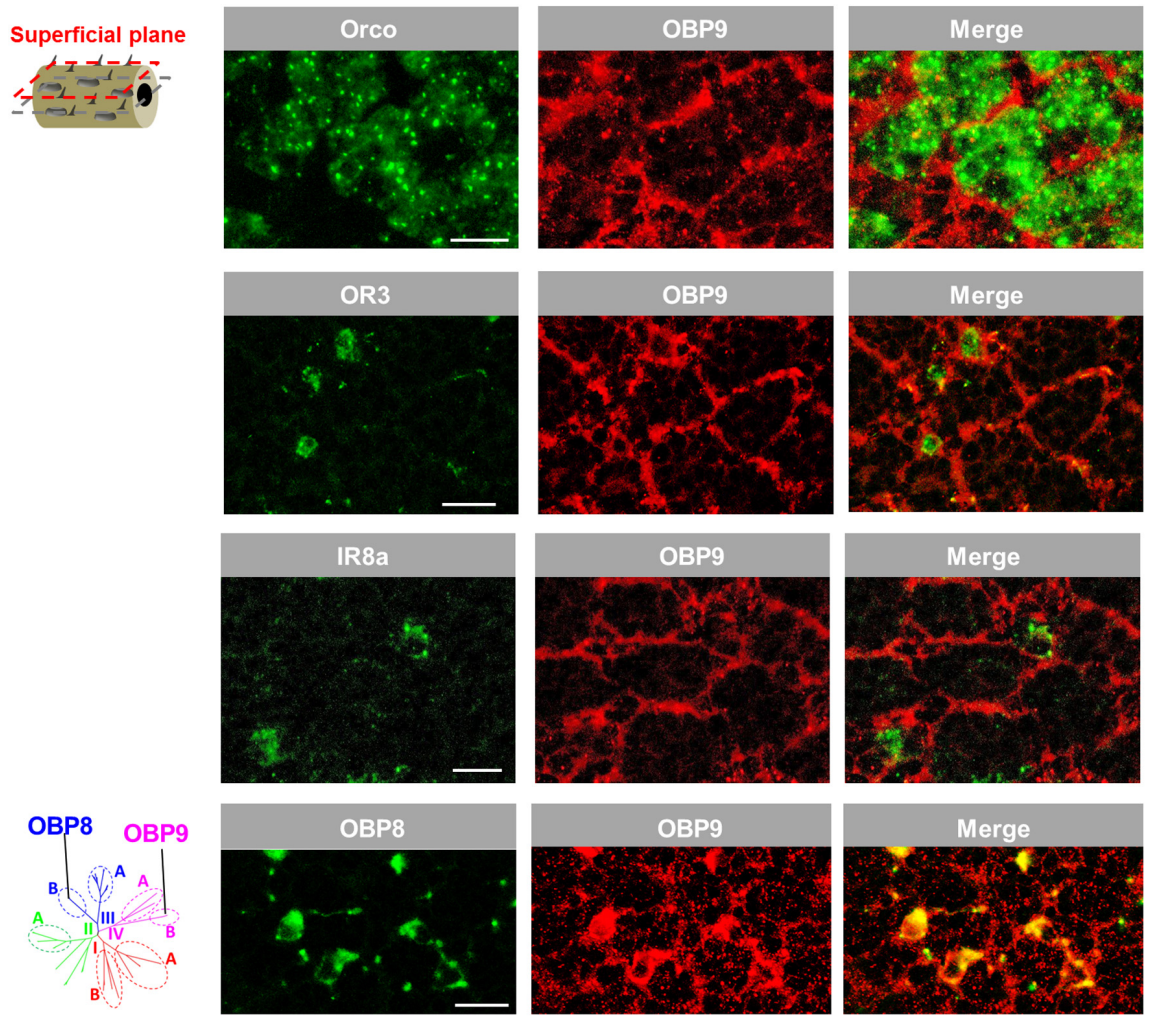


FIGURE 6 | OBP9 expressing cells associate with four types of antennal sensilla. The relative localization of OBP9 and different marker genes indicative of specific sensillum types was analyzed by utilizing specific antisense riboprobes and the means of two-color FISH. Presented images were obtained from superficial cellular planes approaching the cuticle by performing series of horizontal sections of the antennae (diagram, left lane; similar to **Figure 5B**). Orco, OR3, and IR8a were utilized as the specific molecular markers of neurons housed in sensilla basiconica, sensilla trichodea, and sensilla coeloconica, respectively. OBP8 was used as a marker for auxiliary cells of sensilla chaetica (see **Figure 1**). Scale bars, 20 μm.

DISCUSSION

Insects have evolved sensilla that are diversified in the external morphology as well as in the repertoire of molecular elements to act as versatile communication channels for environmental chemical signals (Hansson and Stensmyr, 2011; Leal, 2013; Suh et al., 2014). OBPs are considered to play an important role toward this task due to their capacity to accommodate and transfer odorous molecules. The present study, in conjunction with our previous work (Jiang et al., 2017), has concentrated on this important class of soluble proteins in the locust species *Schistocerca gregaria,* trying to decipher the principles how the multiple OBP subtypes are allocated among and within different sensillum types present on the locust antennae. The findings of this study revealed that subtypes of the desert locust OBP family display a diversified sensilla-specific expression profile and a complex co-localization phenotype in defined sensilla (**Figure 7**). Uncovering the sensillar and cellular organization pattern of distinct locust OBP subtypes may allow a first glimpse on their putative functional role as well as their potential interplay with distinct co-partners.

Our results indicate that several OBP subtypes from two phylogenetic clades are expressed in sensilla chaetica (**Figure 1**). A plus-C type-A subtype together with three classic subtypes were found to be co-expressed in a set of sensilla chaetica (**Figure 2**); this scenario is reminiscent of what was previously reported for sensilla trichodea of *Anopheles gambiae* (Schultze et al., 2013). Sensilla chaetica are characterized by distinct structural features, such as a thick and poreless cuticle wall, an apical pore and relatively few dendrites (Ochieng et al., 1998; Zhou et al., 2009); consequently, sensilla chaetica are considered as relevant for the reception of gustatory tastants rather than odorants. For the

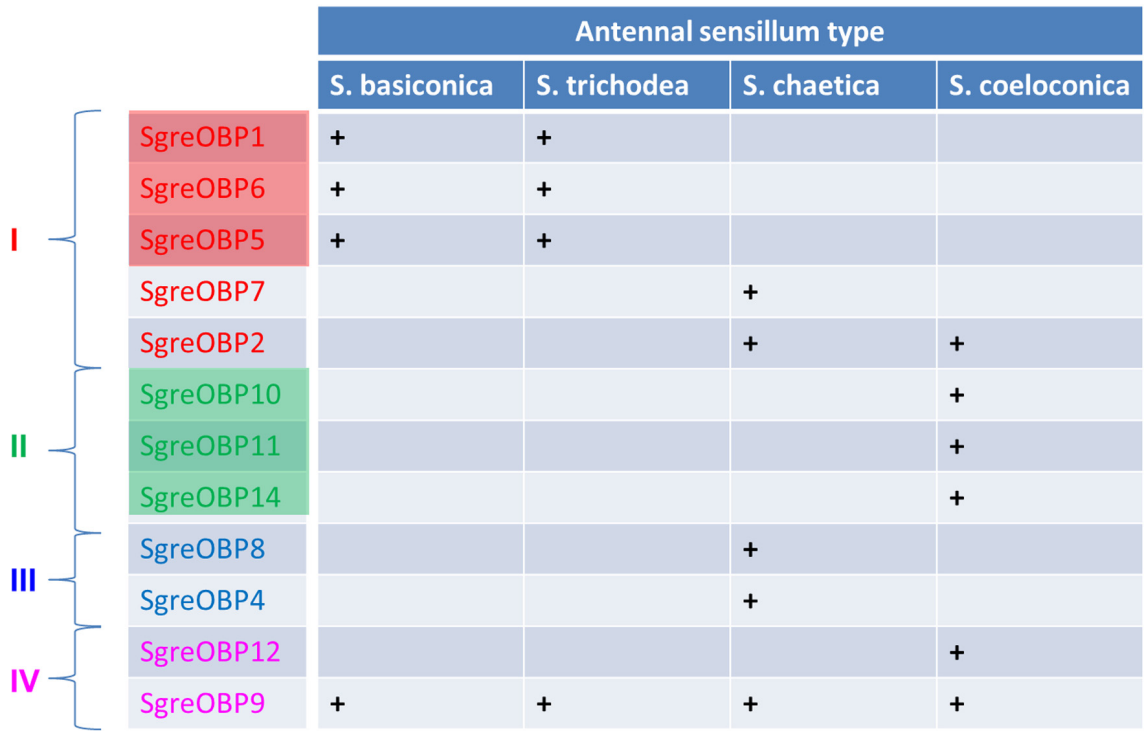


FIGURE 7 | Antennal sensilla specificity of the desert locust OBP family. A distinct OBP subtype that is ascertained to be expressed in a specific sensillum type is denoted as "+", whereas a blank field indicates the absence of particular OBP subtype in this sensillum type. The color code for individual OBPs subtypes is identical to the one used in the phylogenetic analysis (**Supplementary Figure S1**). Color shadings represent subfamily I-A and II-A, respectively.

fruit fly this view was supported by extracellular recordings, calcium imaging and behavioral assays (Montell, 2009; Chen and Amrein, 2017; Scott, 2018). This view may also hold true for sensilla chaetica in locusts which are enriched on the tip of the antennae and palps (Blaney and Chapman, 1969; Ochieng et al., 1998) and are proposed with a receptive role of contact stimuli (Blaney, 1974, 1975; Saini et al., 1995). Thus, the presence of four OBP subtypes in sensilla chaetica on the tip of the antennae (**Figure 1**) suggests that these OBPs may be tuned to mediate the reception of gustatory stimuli. This view would be analogous to the finding for *Drosophila melanogaster* where OBP subtypes expressed in gustatory sensilla are involved in the reception of tastants (Jeong et al., 2013). This is further supported by a recent study demonstrating that knock-down of a sensilla chaetica-specific OBP subtype in *Locusta migratoria* caused a reduced neuronal response to chemical stimuli (Zhang et al., 2017). This finding further supports the notion that OBPs are intimately involved in detecting chemical compounds via sensilla chaetica. Intriguingly, it has been reported that the sensilla chaetica of locust, as well as contact sensilla of other insect species, have a sensillum lymph cavity which is separated into an inner and outer compartment (Ochieng et al., 1998; Shanbhag et al., 2001; Zhou et al., 2009). In a recent study, the labeling for an OBP subtype in *Locusta migratoria* was mainly observed in the non-innervated outer lumen, but not in the inner sensillum lymph which baths the chemosensory dendrites (Yu et al., 2009); this observation has led to speculations of how the cognitive ligands may reach the chemosensory dendrites. The discovery that four distinct OBP subtypes are expressed in this sensillum type (**Figures 1**, **2**) opens the door for revisiting this aspect in more detail.

Distinct OBP subtypes from three phylogenetic clades were found to be expressed in sensilla coeloconica (**Figures 1**, **3**, **4**) (Jiang et al., 2017). Whereas OBP representatives from subfamily II-A (**Figure 4**) together with OBP2 (**Supplementary Figure S3**) were found in the majority of this sensillum type, the atypical OBP subtype OBP12 from subfamily IV-A was present in a subpopulation of sensilla coeloconica. This observation seems to coincide with a previous finding that apart from a receptive role for leaf odors and organic acids (Ochieng and Hansson, 1999), a subset of sensilla coeloconica in locusts appears to be responsive to hygro- or thermo- stimuli (Altner et al., 1981). Such a functional versatility of this sensillum type may be based on distinct sets of cells equipped with specific receptors in combination with appropriate co-partners, e.g., OBP12. Remarkably, the atypical OBP subtype OBP12 belongs to the OBP gene family OBP59a, which is conserved in many insect species, except in Hymenoptera (Vieira and Rozas, 2011). For *Drosophila melanogaster* it has recently been shown that OBP59a is specifically expressed in sensilla coeloconica (Larter et al., 2016), similar to its counterpart in the desert locust (**Figure 4**).

An unexpected finding of this study is the expression of OBP2 in two types of sensilla, sensilla coeloconica and sensilla chaetica (**Figures 1**, **3**). The two types of sensilla differ markedly in their external morphology and their functional implications

(Montell, 2009; Rytz et al., 2013; Joseph and Carlson, 2015; Scott, 2018). On the other hand, in both sensillum types some common chemosensory genes are expressed, most notably the ionotropic receptor type IR25a, one of the co-receptors of divergent IRs (Abuin et al., 2011; Guo et al., 2013). Exploring the functional mode of IR25a in *Drosophila melanogaster* has recently uncovered a multidimensional role for this receptor type (Rimal and Lee, 2018) and it is conceivable that such a versatile function may also be assigned to the OBPs. In fact, it has been proposed that OBPs may be involved in quite different functions (Pelosi et al., 2006, 2014, 2017). In this regard, the observation that OBP2 is always accompanied by a set of other OBP subtypes in a sensillum (**Figures 2**, **3**) may indicate that OBP2 operates in concert with other OBPs to fulfill the distinct functions conferred to the two types of sensilla.

One of the novel finding of this study was the discovery that the plus-C type-B subtype OBP9 is associated with the four antennal sensillum types. Although the functional implication of such a broad sensillum-association is unknown, one could imagine that OBP9, as an ubiquitous OBP, may contribute a general component for the interplay of co-localized OBP partners. Indeed, an interaction of OBP subtypes has been documented in mosquito species and the OBP complex showed a broader ligand spectrum (Qiao et al., 2011). This aspect may be of particular interest in view of the finding that in locust sensilla basiconica, with a large set of OR subtypes (Wang et al., 2015; Pregitzer et al., 2017), only a small set of OBPs is expressed (**Figure 7**). However, it can also not be excluded that OBP9 may be involved in quite different functions. In this context, it is interesting to note that in cockroach and honeybee, the chemosensory proteins, another important class of small soluble proteins, are involved in regulating tissue regeneration and embryonic development (Nomura et al., 1992; Maleszka et al., 2007; Cheng et al., 2015). Given such a broad sensillum-association, OBP9 may be involved in some general processes, such as development and/or survival of the auxiliary cells.

AUTHOR CONTRIBUTIONS

HB, JK, XJ, and PP: current study conception. XJ and MR: experiments conduction and the data acquisition. HB, JK, XJ, and PP: results interpretation. XJ and PP: preliminary manuscript composition. HB and JK: refinement and approval of final manuscript.

FUNDING

XJ was funded by a grant from China Scholarship Council (CSC) with the grant number 201406350032.

ACKNOWLEDGMENTS

We thank Heidrun Froß for her excellent technical assistance. We also thank Prof. Jörg Strotmann and Dr. Patricia Widmayer for their constructive suggestions to improve the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphys.2018.00417/full#supplementary-material

FIGURE S1 | Classification of different subfamilies of locust OBPs. The phylogenetic tree shown was adapted from a previous study analyzing phylogenetic relationship of OBP families from four locust species (Jiang et al., 2017). The branches colored in red, green, blue, and magenta represent the clade I, II, III, and IV, respectively. The classification of the subfamily I-A, II-A, and III-A was based on emergence of a higher bootstrap values on the inner divergent nodes, while other subfamilies were categorized by the emerging topologies. The subtypes belonging to desert locust OBPs were colored and denoted accordingly.

FIGURE S2 | A subset of sensilla chaetica selectively express OBP4 but not OBP8. Cells expressing the respective genes were visualized by using antisense riboprobes specifically targeting OBP4 and OBP8 and by means of two-color FISH. The position of cell clusters visualized by the DIG-labeled OBP4 probe (red) was delineated by dashed lines and is indicated in the images showing the OBP8 labeling and the merge of red and green fluorescence channels, respectively. Notably, no OBP8 labeling was detected. The interface between the cuticle and cellular layer is depicted by a white dashed line. Ch, sensilla chaetica; Ba, sensilla basiconica. Scale bar, 20 μm.

FIGURE S3 | OBP2 and OBP12 are expressed in different cells in sensilla coeloconica (co). Specific antisense riboprobes against OBP2 and OBP12 were used to visualize the expressing cells by means of two-color FISH. The interface between the cuticle and the cellular layer is depicted by a white dashed line. Scale bar, 20 μm.

REFERENCES

Abuin, L., Bargeton, B., Ulbrich, M. H., Isacoff, E. Y., Kellenberger, S., and Benton, R. (2011). Functional architecture of olfactory ionotropic glutamate receptors. *Neuron* 69, 44–60. doi: 10.1016/j.neuron.2010.11.042

Altner, H., Routil, C., and Loftus, R. (1981). The structure of bimodal chemo-, thermo-, and hygroreceptive sensilla on the antenna of *Locusta migratoria*. *Cell Tissue Res.* 215, 289–308. doi: 10.1007/BF00239116

Ban, L., Scaloni, A., D'Ambrosio, C., Zhang, L., Yahn, Y., and Pelosi, P. (2003). Biochemical characterization and bacterial expression of an odorant-binding protein from *Locusta migratoria*. *Cell. Mol. Life Sci.* 60, 390–400. doi: 10.1007/s000180300032

Blaney, W. M. (1974). Electrophysiological responses of the terminal sensilla on the maxillary palps of *Locusta migratoria* (L.) to some electrolytes and non-electrolytes. *J. Exp. Biol.* 60, 275–293.

Blaney, W. M. (1975). Behavioural and electrophysiological studies of taste discrimination by the maxillary palps of larvae of *Locusta migratoria* (L.). *J. Exp. Biol.* 62, 555–569.

Blaney, W. M., and Chapman, R. F. (1969). The fine structure of the terminal sensilla on the maxillary palps of *Schistocerca gregaria* (Forskål) (Orthoptera, Acrididae). *Z. Zellforsch Mikrosk. Anat.* 99, 74–97. doi: 10.1007/BF00338799

Chen, Y., and Amrein, H. (2017). Ionotropic receptors mediate *Drosophila* oviposition preference through sour gustatory receptor neurons. *Curr. Biol.* 27, 2741–2750.e4. doi: 10.1016/j.cub.2017.08.003

Cheng, D., Lu, Y., Zeng, L., Liang, G., and He, X. (2015). Si-CSP9 regulates the integument and moulting process of larvae in the red imported fire ant, *Solenopsis invicta*. *Sci. Rep.* 5:9245. doi: 10.1038/srep09245

Foret, S., and Maleszka, R. (2006). 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

Galindo, K., and Smith, D. P. (2001). A large family of divergent *Drosophila* odorant-binding proteins expressed in gustatory and olfactory sensilla. *Genetics* 159, 1059–1072.

Guo, M., Krieger, J., Große-Wilde, E., Mißbach, C., Zhang, L., and Breer, H. (2013). Variant ionotropic receptors are expressed in olfactory sensory neurons of coeloconic sensilla on the antenna of the desert locust (*Schistocerca gregaria*). *Int. J. Biol. Sci.* 10, 1–14. doi: 10.7150/ijbs.7624

Guo, W., Wang, X., Ma, Z., Xue, L., Han, J., Yu, D., et al. (2011). CSP and takeout genes modulate the switch between attraction and repulsion during behavioral phase change in the migratory locust. *PLoS Genet.* 7:e1001291. doi: 10.1371/journal.pgen.1001291

Hansson, B. S., and Stensmyr, M. C. (2011). Evolution of insect olfaction. *Neuron* 72, 698–711. doi: 10.1016/j.neuron.2011.11.003

Hassanali, A., Njagi, P. G. N., and Bashir, M. O. (2005). Chemical ecology of locust and related acridids. *Annu. Rev. Entomol.* 50, 223–245. doi: 10.1146/annurev.ento.50.071803.130345

Jeong, Y. T., Shim, J., Oh, S. R., Yoon, H. I., Kim, C. H., Moon, S. J., et al. (2013). An odorant-binding protein required for suppression of sweet taste by bitter chemicals. *Neuron* 79, 725–737. doi: 10.1016/j.neuron.2013.06.025

Jiang, Q. Y., Wang, W. X., Zhang, Z., and Zhang, L. (2009). Binding specificity of locust odorant binding protein and its key binding site for initial recognition of alcohols. *Insect Biochem. Mol. Biol.* 39, 440–447. doi: 10.1016/j.ibmb.2009.04.004

Jiang, X., Krieger, J., Breer, H., and Pregitzer, P. (2017). Distinct subfamilies of odorant binding proteins in locust (Orthoptera, Acrididae): molecular evolution, structural variation, and sensilla-specific expression. *Front. Physiol.* 8:734. doi: 10.3389/fphys.2017.00734

Jiang, X., Pregitzer, P., Grosse-Wilde, E., Breer, H., and Krieger, J. (2016). Identification and characterization of two "sensory neuron membrane proteins" (SNMPs) of the desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae). *J. Insect Sci.* 16:33. doi: 10.1093/jisesa/iew015

Jin, X., Brandazza, A., Navarrini, A., Ban, L., Zhang, S., Steinbrecht, R. A., et al. (2005). Expression and immunolocalisation of odorant-binding and chemosensory proteins in locusts. *Cell. Mol. Life Sci.* 62, 1156–1166. doi: 10.1007/s00018-005-5014-6

Joseph, R. M., and Carlson, J. R. (2015). *Drosophila* chemoreceptors: a molecular interface between the chemical world and the brain. *Trends Genet.* 31, 683–695. doi: 10.1016/j.tig.2015.09.005

Larter, N. K., Sun, J. S., and Carlson, J. R. (2016). Organization and function of *Drosophila* odorant binding proteins. *eLife* 5:e20242. doi: 10.7554/eLife.20242

Leal, W. S. (2013). 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

Maleszka, J., Forêt, S., Saint, R., and Maleszka, R. (2007). 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

Montell, C. (2009). A taste of the *Drosophila* gustatory receptors. *Curr. Opin. Neurobiol.* 19, 345–353. doi: 10.1016/j.conb.2009.07.001

Nomura, A., Kawasaki, K., Kubo, T., and Natori, S. (1992). Purification and localization of p10, a novel protein that increases in nymphal regenerating legs of *Periplaneta americana* (American cockroach). *Int. J. Dev. Biol.* 36, 391–398. doi: 10.1387/IJDB.1445782

Ochieng, S. A., Hallberg, E., and Hansson, B. S. (1998). Fine structure and distribution of antennal sensilla of the desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae). *Cell Tissue Res.* 291, 525–536. doi: 10.1007/s004410051022

Ochieng, S. A., and Hansson, B. S. (1999). Responses of olfactory receptor neurones to behaviourally important odours in gregarious and solitarious desert locust, *Schistocerca gregaria*. *Physiol. Entomol.* 24, 28–36. doi: 10.1046/j.1365-3032.1999.00107.x

Pelosi, P., Iovinella, I., Felicioli, A., and Dani, F. R. (2014). Soluble proteins of chemical communication: an overview across arthropods. *Front. Physiol.* 5:320. doi: 10.3389/fphys.2014.00320

Pelosi, P., Iovinella, I., Zhu, J., Wang, G., and Dani, F. R. (2017). Beyond chemoreception: diverse tasks of soluble olfactory proteins in insects. *Biol. Rev.* 93, 184–200. doi: 10.1111/brv.12339

Pelosi, P., Zhou, J. J., Ban, L. P., and Calvello, M. (2006). Soluble proteins in insect chemical communication. *Cell. Mol. Life Sci.* 63, 1658–1676. doi: 10.1007/s00018-005-5607-0

Pener, M. P., and Yerushalmi, Y. (1998). The physiology of locust phase polymorphism: an update. *J. Insect Physiol.* 44, 365–377. doi: 10.1016/S0022-1910(97)00169-8

Pregitzer, P., Jiang, X., Grosse-Wilde, E., Breer, H., Krieger, J., and Fleischer, J. (2017). In search for pheromone receptors: certain members of the odorant receptor family in the desert locust *Schistocerca gregaria* (Orthoptera: Acrididae) are co-expressed with SNMP1. *Int. J. Biol. Sci.* 13, 911–922. doi: 10.7150/ijbs.18402

Qiao, H., He, X., Schymura, D., Ban, L., Field, L., Dani, F. R., et al. (2011). Cooperative interactions between odorant-binding proteins of *Anopheles gambiae*. *Cell. Mol. Life Sci.* 68, 1799–1813. doi: 10.1007/s00018-010-0539-8

Rimal, S., and Lee, Y. (2018). The multidimensional ionotropic receptors of *Drosophila melanogaster*. *Insect Mol. Biol.* 27, 1–7. doi: 10.1111/imb.12347

Rytz, R., Croset, V., and Benton, R. (2013). Ionotropic receptors (IRs): chemosensory ionotropic glutamate receptors in *Drosophila* and beyond. *Insect Biochem. Mol. Biol.* 43, 888–897. doi: 10.1016/j.ibmb.2013.02.007

Saini, R. K., Rai, M. M., Hassanali, A., Wawiye, J., and Odongo, H. (1995). Semiochemicals from froth of egg pods attract ovipositing female *Schistocerca gregaria*. *J. Insect Physiol.* 41, 711–716. doi: 10.1016/0022-1910(95)00016-N

Sandler, B. H., Nikonova, L., Leal, W. S., and Clardy, J. (2000). Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. *Chem. Biol.* 7, 143–151. doi: 10.1016/S1074-5521(00)00078-8

Schultze, A., Pregitzer, P., Walter, M. F., Woods, D. F., Marinotti, O., Breer, H., et al. (2013). The co-expression pattern of odorant binding proteins and olfactory receptors identify distinct trichoid sensilla on the antenna of the malaria mosquito *Anopheles gambiae*. *PLoS One* 8:e69412. doi: 10.1371/journal.pone.0069412

Scott, K. (2018). Gustatory processing in *Drosophila melanogaster*. *Annu. Rev. Entomol.* 63, 15–30. doi: 10.1146/annurev-ento-020117-043331

Shanbhag, S. R., Hekmat-Scafe, D., Kim, M. S., Park, S. K., Carlson, J. R., Pikielny, C., et al. (2001). Expression mosaic of odorant-binding proteins in *Drosophila* olfactory organs. *Microsc. Res. Tech.* 55, 297–306. doi: 10.1002/jemt.1179

Steinbrecht, R. A. (1996). Structure and function of insect olfactory sensilla. *Ciba Found. Symp.* 200, 158–174. doi: 10.1002/9780470514948.ch13

Suh, E., Bohbot, J. D., and Zwiebel, L. J. (2014). Peripheral olfactory signaling in insects. *Curr. Opin. Insect Sci.* 6, 86–92. doi: 10.1016/j.cois.2014.10.006

Tegoni, M., Campanacci, V., and Cambillau, C. (2004). Structural aspects of sexual attraction and chemical communication in insects. *Trends Biochem. Sci.* 29, 257–264. doi: 10.1016/j.tibs.2004.03.003

Vieira, F. G., and Rozas, J. (2011). Comparative genomics of the odorant-binding and chemosensory protein gene families across the Arthropoda: origin and evolutionary history of the chemosensory system. *Genome Biol. Evol.* 3, 476–490. doi: 10.1093/gbe/evr033

Vogt, R. G., Callahan, F. E., Rogers, M. E., and Dickens, J. C. (1999). Odorant binding protein diversity and distribution among the insect orders, as indicated by LAP, an OBP-related protein of the true bug *Lygus lineolaris* (Hemiptera, heteroptera). *Chem. Senses* 24, 481–495. doi: 10.1093/chemse/24.5.481

Vogt, R. G., Große-Wilde, E., and Zhou, J. J. (2015). 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

Vogt, R. G., and Riddiford, L. M. (1981). Pheromone binding and inactivation by moth antennae. *Nature* 293, 161–163. doi: 10.1038/293161a0

Wang, X., and Kang, L. (2014). Molecular mechanisms of phase change in locusts. *Annu. Rev. Entomol.* 59, 225–244. doi: 10.1146/annurev-ento-011613-162019

Wang, Z., Yang, P., Chen, D., Jiang, F., Li, Y., Wang, X., et al. (2015). Identification and functional analysis of olfactory receptor family reveal unusual characteristics of the olfactory system in the migratory locust. *Cell. Mol. Life Sci.* 72, 4429–4443. doi: 10.1007/s00018-015-2009-9

Wheeler, W. C., Whiting, M., Wheeler, Q. D., and Carpenter, J. M. (2001). The phylogeny of the extant Hexapod orders. *Cladistics* 17, 113–169. doi: 10.1111/j.1096-0031.2001.tb00115.x

Xu, P. X., Zwiebel, L. J., and Smith, D. P. (2003). Identification of a distinct family of genes encoding atypical odorant-binding proteins in the malaria vector

mosquito, *Anopheles gambiae*. *Insect Mol. Biol.* 12, 549–560. doi: 10.1046/j.1365-2583.2003.00440.x

Xu, Y. L., He, P., Zhang, L., Fang, S. Q., Dong, S. L., Zhang, Y. J., et al. (2009). Large-scale identification of odorant-binding proteins and chemosensory proteins from expressed sequence tags in insects. *BMC Genomics* 10:632. doi: 10.1186/1471-2164-10-632

Yang, Y., Krieger, J., Zhang, L., and Breer, H. (2012). The olfactory co-receptor Orco from the migratory locust (*Locusta migratoria*) and the desert locust (*Schistocerca gregaria*): identification and expression pattern. *Int. J. Biol. Sci.* 8, 159–170. doi: 10.7150/ijbs.8.159

Yu, F., Zhang, S., Zhang, L., and Pelosi, P. (2009). Intriguing similarities between two novel odorant-binding proteins of locusts. *Biochem. Biophys. Res. Commun.* 385, 369–374. doi: 10.1016/j.bbrc.2009.05.074

Zhang, L., Li, H., and Zhang, L. (2017). Two olfactory pathways to detect aldehydes on locust mouthpart. *Int. J. Biol. Sci.* 13, 759–771. doi: 10.7150/ijbs.19820

Zhou, J. J., Huang, W., Zhang, G. A., Pickett, J. A., and Field, L. M. (2004). 'Plus-C' odorant-binding protein genes in two *Drosophila* species and the malaria mosquito *Anopheles gambiae*. *Gene* 327, 117–129. doi: 10.1016/j.gene.2003.11.007

Zhou, S. H., Zhang, S. G., and Zhang, L. (2009). The chemosensilla on tarsi of *locusta migratoria* (Orthoptera: Acrididae): distribution, ultrastructure, expression of chemosensory proteins. *J. Morphol.* 270, 1356–1363. doi: 10.1002/jmor.10763

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.