Chem. Senses 24: 481–495, 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)

R.G. Vogt, F.E. Callahan[1], M.E. Rogers and J.C. Dickens[2]

Department of Biological Sciences, University of South Carolina, Columbia, SC 29208, [1]United States Department of Agriculture, Agricultural Research Service, Genetics and Precision Agriculture Research Unit, Mississippi State, MS 39762 and [2]United States Department of Agriculture, Agricultural Research Service, Plant Sciences Institute, Vegetable Laboratory, Beltsville, MD 20705, USA

Correspondence to be sent to: Richard G. Vogt, Department of Biological Sciences, University of South Carolina, Columbia, SC 29208, USA. e-mail: vogt@biol.sc.edu

Abstract

Insect odorant binding proteins (OBPs) are thought to deliver odors to olfactory receptors, and thus may be the first biochemical step in odor reception capable of some level of odor discrimination. OBPs have been identified from numerous species of several insect orders, including Lepidoptera, Diptera, Coleoptera and Hymenoptera; all are holometabolous insects belonging to the monophyletic division of insects known as the Endopterygota. Recently, an antennal protein with OBP-like properties was identified from *Lygus lineolaris*, a hemipteran insect representing the Hemipteroid Assemblage, a sister division to the Endopterygota. The full length sequence of *Lygus* antennal protein (LAP) is presented in this report. *In situ* hybridization analysis revealed LAP expression in cell clusters associating with olfactory sensilla; expression was adult-specific, initiating in developing adult tissue during the transitional period that precedes the actual adult molt. Sequence analysis confirmed that LAP is homologous with the OBP-related protein family, and most similar to the OS-E and OS-F proteins of *Drosophila*, the ABPX proteins of Lepidoptera and the OBPRP proteins of the Coleoptera. Assuming that the OBP-related proteins represent one homologous family, the identification of LAP significantly expands the phylogenetic depth of that family and its underlying role in odor detection to encompass all members of the Endopterygota and Hemipteroid Assemblage, which comprise >90% of all insect species.

Introduction

For insects, odorant binding proteins (OBPs) may be the first specific biochemical step in odor reception. OBPs are small, water soluble proteins, expressed in the support cells of olfactory sensilla and secreted into the aqueous fluid surrounding the olfactory neurons at concentrations as high as 10 mM (Vogt and Riddiford, 1981; Vogt, 1987, 1995; Prestwich *et al.*, 1995; Pelosi and Maida, 1995; Pelosi, 1996; Steinbrecht, 1996; Carlson, 1996; Breer, 1997; Kaissling, 1998). OBPs have also been identified in vertebrates (Pelosi *et al.*, 1982; Pevsner *et al.*, 1988; Pelosi, 1996); however, the vertebrate and insect OBPs appear unrelated by sequence. In insects, OBPs are thought to facilitate the transport of lipophilic odorants through the aqueous fluid from the sensilla cuticle to membrane-bound receptors in the olfactory neurons; several variations of this scheme currently under investigation are detailed in Figure 1. Multiple OBPs have been identified in single species and have been shown to associate differentially with functionally distinct classes of olfactory sensilla (Vogt *et al.*, 1991b; Steinbrecht *et al.*, 1992, 1995; Laue *et al.*, 1994; Laue and Steinbrecht, 1997). Several reports have demonstrated selective binding of odorants to different OBPs derived from a given species (Vogt *et al.*, 1989; Du and Prestwich 1995; Prestwich *et al.*, 1995; Feng and Prestwich, 1997). The diversity of OBPs and their differential expression has led to the suggestion that OBPs might act as selective filters, influencing the range of odor molecules which can gain access to the olfactory receptors of a given sensillum (Vogt *et al.*, 1991b; Steinbrecht, 1996). OBPs have also been suggested to have roles in signal termination (Kaissling, 1998).

The OBPs present a narrow gateway through which odors must pass before being recognized; therefore OBP variety (paralogy) and variation (orthology) may reflect the diversity of chemosensory behavior among the insects. But are OBPs common features of insects? More specifically, what is the distribution of OBPs among the insect orders?

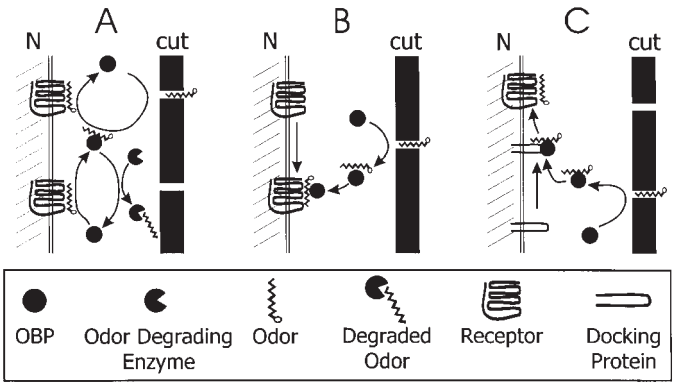


Figure 1 Models of OBP action. The three schemes represent the general features of models currently under investigation (Vogt *et al.*, 1985; Vogt, 1995; Prestwich *et al.*, 1995; Steinbrecht, 1996; Ziegelberger, 1996; Kaissling, 1998). In all three schemes, OBPs bind odor molecules as they enter the sensillum via pores through the outer cuticle (cut); the OBPs facilitate the transport of the lipophilic odor molecules through the aqueous fluid surrounding the neurons (N). In **(A)**, OBPs transport odor molecules both to and from membrane bound odor receptors; the OBP–odor complex is not stable and odorants interact directly with receptors. In **(B)** the OBP–odor complex is stable and interacts with the receptor as a complex. In **(C)**, a variation of (A), OBPs interact with a membrane bound docking protein, possibly SNMP1 (Rogers *et al.*, 1997), enhancing the off-loading of odor molecules near the odor receptors. In all cases, antennal specific odor degrading enzymes participate in the signal termination of the odor molecules—shown in (A) only (Vogt and Riddiford, 1981; Vogt *et al.*, 1985; Rybczynski *et al.*, 1989, 1990; Kaissling, 1998; Rogers *et al.*, 1999).

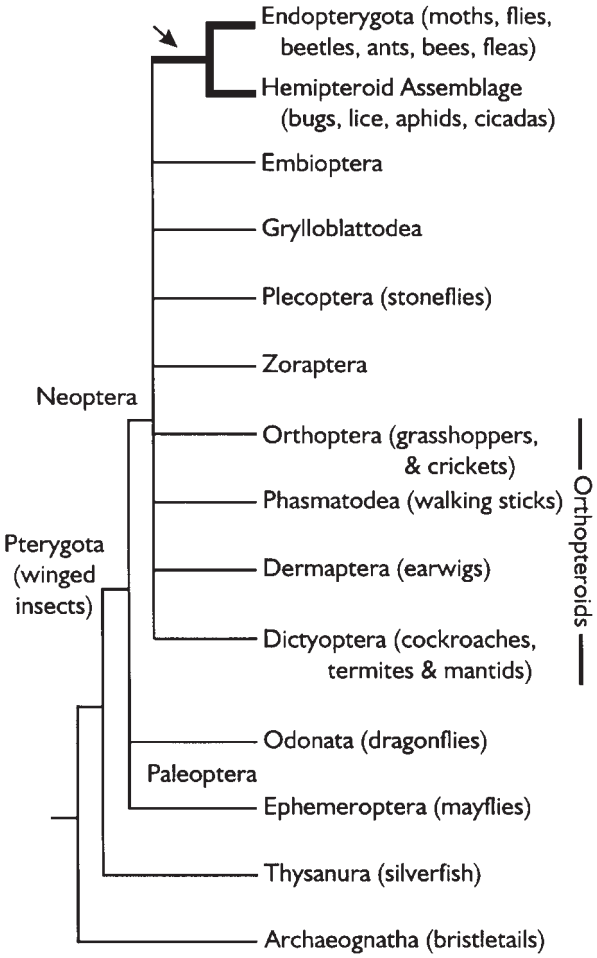


Figure 2 Phylogenetic relationships within the Insecta. The relationships of orders and major divisions are shown; representative familiar insect names are indicated. The tree is the conservative view of Kristensen (Kristensen, 1991), also presented in 'The Tree of Life' (Maddison and Maddison, 1998). Neoptera and Paleoptera form the two major groupings of winged insects (Pterygota). Within the Neoptera, there is strong consensus that the Endopterygota and Hemipteroid Assemblage are sister groups of a single monophyletic division (indicated by the arrow); there is less consensus for the relationships of the other groups shown (Boudreaux, 1979; Hennig, 1981). Previously, OBP-related proteins have been reported only from species belonging to orders of the Endopterygota. LAP was identified from *L. lineolaris*, a hemipteran insect belonging to an order of the Hemipteroid Assemblage. The orthopteroids include the indicated orders, and are often considered to be included within a monophyletic lineage distinct from the Endopterygota and Hemipteroid lineage; this orthopteroid lineage is not so represented because of ongoing debate on the placement of other neopteran orders (Boudreaux, 1979; Hennig, 1981; Kristensen, 1991; Whiting *et al.*, 1997).

Insects comprise the largest number of extant animal species. There are currently ~1.2 million recognized animal species including >800 000 insects, although the projected estimates of the total number of insect species range from 1.5 to 30 million (Erwin, 1982; Freeman and Herron, 1998). Insects are organized into 29 extant orders (Figure 2), with the majority belonging to the division Neoptera with 25 extant orders, ~98% of species (Borror *et al.*, 1989; Kristensen, 1991). Fossil evidence suggests the Neoptera emerged and its extant orders diverged ~300 million years ago—the Carboniferous Era (Kukalová-Peck, 1991). The diversification of insects has been popularly linked to the evolution of angiosperms (flowering plants). However, the majority of extant insect orders were present well before the angiosperms first appeared ~130 000 million years ago (based on fossil evidence); the diversification of insects at the family level was also independent of the angiosperms (Labandeira and Sepkoski, 1993). Insects were apparently preadapted to exploit the ascendancy of the angiosperms, resulting in enormous diversification at the species level within successful, but preexisting, orders (Labandeira and Sepkoski, 1993; Freeman and Herron, 1998).

Until now, insect OBPs have only been observed in one major division of the Neoptera, the Endopterygota (Figure 2) (Kristensen, 1991; Whiting *et al.*, 1997). The Endopterygota (11 extant orders, ~83% of species) are also referred to as holometabolous insects, because of their unique development; all undergo a complete metamorphosis from the non-reproductive larval stage to the reproductive adult stage. The Endopterygota include the orders Coleoptera (beetles, >300 000 named species), Lepidoptera (moths and butterflies, >150 000 named species), Hymenoptera (bees, wasps and ants, >100 000 named species), Diptera (flies and mosquitoes, >90 000 named species) and Siphonaptera

(fleas, ~2300 named species) (Borror *et al.*, 1989; Kristensen, 1991). All have relatively simple worm-like or grub-like juvenile stages (larvae) in contrast to their more complex adult forms.

The remaining neopteran insects are referred to as hemimetabolous, because the non-reproductive juvenile stages resemble the reproductive adult stage; the final developmental molt from juvenile to adult primarily involves the addition of adult characters to the adult-like juvenile form. One major hemimetabolous group is the orthopteroids, which includes the neopteran orders Orthoptera (grasshoppers and crickets, >12 000 named species), Phasmatodea (walking sticks, >2000 named species), Dermaptera (earwigs, >1100 named species) and the monophyletic Dictyoptera (three orders, cockroaches, termites and mantids, >7500 named species) (Borror *et al.*, 1989; Kristensen, 1991). Many recent phylogenies suggest the orthopteroids are a monophyletic group, sharing a common ancestor and forming a separate lineage from the Endopterygota (Hennig, 1981; Kristensen, 1991; Whiting *et al.*, 1997). This orthopteroid lineage is not so represented in Figure 2 because of ongoing debate on the relationships of some of the additional Neoptera orders (Maddison and Maddison, 1998).

A third major neopteran group is the Hemipteroid Assemblage (~11% of species). These are also hemimetabolous insects; however, both morphological (Hennig, 1981; Kristensen, 1991; Whiting *et al*., 1997) and molecular data (Whiting *et al.*, 1997) suggest they are a sister group to the Endopterygota. These two lineages, the Endopterygota and the Hemipteroid Assemblage, are viewed as forming a single clade sharing a common ancestor distinct from the other Neoptera orders. The Hemipteroids include the orders Hemiptera (true bugs, cicadas and aphids, >25 000 named species), Phthiraptera (lice, >5500 named species) and Thysanoptera (thrips, >4000 named species) (Borror *et al.*, 1989; Kristensen, 1991).

The identification of an OBP homologue within the Hemipteroid Assemblage would suggest that OBP-related genes were present in the species ancestral to both the Endopterygota and Hemipteroids, and that OBPs may be distributed throughout the species of these two groups unless they were secondarily lost. Recently, an antennal protein LAP (*Lygus* antennal protein) was identified from the hemipteran insect *Lygus lineolaris* (Dickens *et al.*, 1995, 1998; Dickens and Callahan, 1996). *Lygus lineolaris* is a polyphagous insect broadly distributed throughout North America and an important pest to many crops including tomatoes, soybeans and cotton (Snodgrass *et al.*, 1984; Young 1986). LAP was suggested to have an OBP-related function based on its antennal specific expression, small size (~15–17 kDa), and extracellular location within olfactory sensilla in the aqueous fluid surrounding the olfactory neurons (Dickens and Callahan, 1996; Dickens *et al.*, 1998). In the study reported here, LAP was cloned, fully sequenced and confirmed to be homologous with several OBP-related proteins. The diversity of the OBP-related proteins and their distribution among the Neoptera insects was characterized. This study strongly suggests a widespread distribution of OBP-related genes throughout the Endopterygota and Hemipteroid species, a distribution which may parallel both the extraordinary success and the elaboration of olfactory based behaviors of these groups.

Materials and methods

Animals and tissue

Lygus lineolaris (Palisot de Beauvois) were obtained as fifth instar nymphs from a laboratory colony annually infused with feral insects and maintained at the USDA–ARS Southern Insect Management Laboratory, Stoneville MS, USA (Snodgrass and McWilliams, 1992). Animals were raised to adults on a diet of fresh green beans at 25°C and a photoperiod of 14 h:10 h L:D (Dickens *et al.*, 1998). When necessary, developmental stages were determined based on established morphological criteria (Schwartz and Foottit, 1992). For RNA isolation, antennae were removed from adult males upon emergence and immediately frozen on dry CO_2 and stored at –70°C until use. For histology, animals were decapitated with antennae attached, the majority of head material was trimmed away and the terminal antennal segment cut; the resulting material was fixed overnight at 4°C in 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS; 150 mM NaCl, 10 mM Na-phosphate, pH 7.0) containing 0.1% Tween-20 (PBST). Tissue was subsequently washed several times in PBST, dehydrated to 70% methanol in PBST and stored at –20°C until use.

Cloning and Sequencing LAP

Total RNA was extracted from 600 adult male antennae; frozen tissue was homogenized in guanidinium thiocyanate (600 µl) under liquid nitrogen and processed through an acid–phenol extraction and isopropanol precipitation (Chomczynski and Sacchi, 1987). Complementary DNA was synthesized from 30% of the extraction product (20 µl reaction using Superscript II reverse transcriptase, GIBCO-BRL, following recommended protocols). Aliquots of the crude reaction product were used undiluted for subsequent polymerase chain reactions (PCR).

An initial round of PCR was performed on antennal cDNA using oligo(dT) as an antisense primer and two degenerate sense primers encoding the same region of the previously sequenced N-terminus (Dickens *et al.*, 1998) (LAP-S1a, GARYTNCCNGARGAAATG; LAP-S1b, GARYTNCCNGARGAGATG; Figure 3). This PCR was performed under the following final conditions: 1× buffer (50 mM KCl, 10 mM Tris–HCl, pH 9.0, 1% Triton X-100), 1.5 mM $MgCl_2$, 0.2 mM dNTP, 1.5 µM LAP sense primer, 0.5 µM oligo(dT) antisense primer, 0.03 U/µl Taq DNA polymerase (Promega) in 100 µl reaction against 20% of the

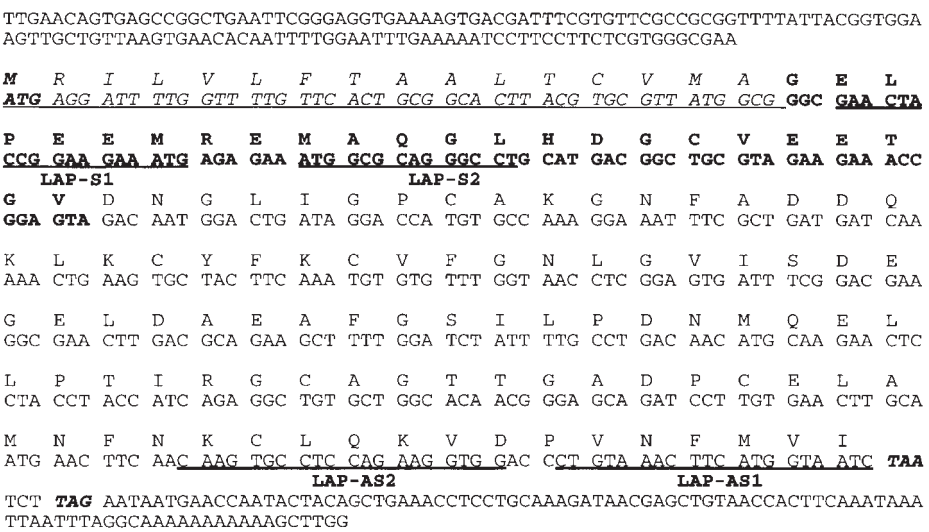


Figure 3 LAP sequence. The cDNA and deduced amino acid sequence of LAP is shown. The previously determined N-terminal sequence is in bold type; the suggested leader sequence is in italics. The suggested start ATG and stop TAA and TAG codons are in bold italics. The positions of sense (LAP-S1 and LAP-S2) and antisense (LAP-AS1 and LAP-AS-2) oligonucleotides used for PCR are indicated.

cDNA reaction product. Reactions were performed on a Cetus PCR1000 thermocycler: 94°C (2.5 min); five cycles of 94°C (30 s), 37°C (2 min), 2 min ramp, 74°C (3 min); 35 cycles of 94°C (30 s), 47°C (2 min), 2 min ramp, 74°C (3 min); 74°C (15 min).

To further select cDNA for cloning, an aliquot of the initial reaction was re-amplified using three different anchored oligo(dTs) ($HT_{11}G$, $HT_{11}A$ and $HT_{11}C$) as antisense primers and two degenerate sense primers encoding a single region of the N-terminus, downstream of the initial LAP-S1 primers (LAP-S2a, AGTGCNCARGGNCT; LAP-S2b AGTGCNCARGGNTT; Figure 3). Aliquots of all PCR reactions (primary and secondary) were analyzed on a 1.5% agarose gel.

For cloning, the product generated by first round amplification with LAPS1A and second round amplification by LAPS2A vs $HT_{11}G$ was re-amplified using the LAPS2A and $HT_{11}G$ primers under the following conditions: 1× buffer (50 mM Tris–HCl, pH 8.3, 2.5 mg/ml BSA, 1% Ficoll, 1 mM tartrazine); 2 mM $MgCl_2$, 0.2 mM dNTP; 2 µM sense primer; 0.5 µM antisense primer; 0.03 U/µl Taq DNA polymerase (Promega) in 50 µl reaction against a 0.01 aliquot of the primary band elution. Reactions were performed in sealed glass capillaries on an Idaho Technology Thermocycler: 94°C (60 s); 40 cycles of 94°C (15 s), 45°C (15 s), 74°C (30 s). The products of four such reactions were pooled (200 µl), and purified by extraction in phenol–chloroform followed by chloroform. DNA was precipitated by addition of 1/10 volume of 10× STE buffer (1M NaCl, 200 mM Tris–HCl, pH 7.5, 100 mM EDTA), 1 volume of 4 M ammonium acetate and 2.5 volumes of EtOH (room temperature) and immediate centrifugation at room temperature for 1 h (12 000 *g*). The resulting pellet was washed in 70% ethanol, dried, and dissolved in 10 µl TE (10 mM Tris–HCl, 1 mM EDTA, pH 7.5). DNA ends were polished with Pfu polymerase (Stratagene) and ligated into pCR-Script (Stratagene), following recommended protocols.

To obtain the *N*-terminal and 5′ cDNA sequence, 5′ RACE was performed using the 5′/3′ RACE Kit (Boehringer Mannheim) following recommended protocols. The remainder of the cDNA that yielded the initial LAP clone (above) was desalted using an Ultrafree-MC 30K spin filter (Millipore), and an aliquot was A-tailed at its 5′ end, and amplified by PCR using anchored oligo(dT) sense primer (5′/3′ RACE kit) vs a LAP specific antisense primer (LAP-AS1; GATTACCATGAAGTTTACAG; Figure 3). A 10 µl PCR reaction was performed using the Expand Long Template PCR System (Boehringer Mannheim) and an Idaho Technology Thermocycler using recommended primer concentrations, Expand Long Template enzyme, supplied buffer No. 2, and the following reaction conditions: 94°C (1 min); 10 cycles of 94°C (1 s), 55°C (10 s), 68°C (1 min); 10 cycles of 94°C (1 s), 55°C (10 s), 68°C (2 min); 10 cycles of 94°C (1 s), 55°C (10 s), 68°C (3 min). The resulting product was analyzed on a 1.5% agarose–TAE gel and appeared as a smear extending above 600 bp; a gel fragment corresponding to the upper 25% of this range was isolated, and DNA was eluted by freezing followed by centrifugation. Eluted DNA was re-amplified in 50 µl reactions using a specific sense primer (supplied) vs an internal LAP specific antisense primer (LAP-AS2; CCACCTTCTGGAGG-CACTTG; Figure 3), and conditions similar to those used for the primary cloning: Idaho Technologies Thermocycler; Taq DNA polymerase; reaction conditions: 94°C (2 min); 35 cycles of 94°C (15 s), 55°C (15 s), 74°C (40 s); 74°C (2 min). The resulting PCR products were ligated into pCRScript

(Stratagene) as described above. Cloned inserts were analyzed by PCR using vector specific primers; several clones containing the largest inserts were sequenced.

Plasmid DNAs were purified (Qiagen) and sequenced by ABI Prism Dye Terminator cycle sequencing (Applied Biosystems; Florida DNA Sequencing Core Laboratory, Gainesville FL, USA). All sequences were confirmed in both directions. Sequences were initially analyzed for possible homologues the NCBI (National Center for Biotechnology Information) BLAST network server (Altschul *et al.*, 1997).

In Situ Hybridization

Digoxigenin labeled RNA probes were used for *in situ* hybridization studies following protocols modified from Byrd *et al.* (Byrd *et al.*, 1996) and Rogers *et al.* (Rogers *et al.*, 1997). In brief, the initial LAP clone was linearized and RNAs (antisense and sense) were synthesized using T7 or T3 RNA polymerase (Stratagene) following recommended protocols (Genius System, Boehringer Mannheim) and in the presence of 40 units RNasin (Promega). For *in situ* hybridization studies, RNA was alkaline degraded to ~150 base length (Byrd *et al.*, 1996). Probe quality was confirmed under denaturing conditions by formaldehyde agarose gel electrophoresis (Maniatis *et al.*, 1982); probes were stored at –70°C until use.

For sectioning, fixed tissue (stored in 70% methanol at –20°C, see above) was transferred to 70% ethanol, dehydrated though a graded series of ethanol and toluene, and incubated in melted paraffin (Periplast +) for 2–4 h before being embedded in plastic molds. Antennal tissue was oriented using the attached heads. Longitudinal and cross-sections (10 μm) of the antennae were transferred to water drops on electrostatically charged microscope slides (SuperFrost II, Fisher); slides were dewaxed in xylene. *In situ* hybridization steps were as described in Rogers *et al.*, (Rogers *et al.*, 1997). Antisense or sense LAP probes were applied at 100 ng/ml in hybridization solution at 45°C, following prehybridization. Post-hybridization washes and staining were as described elsewhere (Rogers *et al.*, 1997). Tissue was photographed to color transparencies which were digitized and processed using Adobe Photoshop.

Phylogenetic analysis

OBP-related sequences were initially identified using the NCBI BLAST network server, and retrieved using NCBI Entrez from GenBank, Swiss-Prot or EMBL data bases, or directly from publications for unsubmitted sequences. The *Manduca sexta* sequences credited to Robertson and co-workers (Robertson *et al.*, 1998) were obtained from GenBank as nucleic acid sequences and translated for this analysis. Sequences were aligned in Clustal X (Thompson *et al.*, 1994). Select sequences identified as most similar by BLAST analysis were prealigned using the multiple alignment function in Clustal X; these groups were then aligned to each other using the profile alignment function. The prealigned groups were (A) PBP, (B) GOBP1, (C) GOBP2, (D) ABPX, (E) B1-Tmol, B2-Tmol, Lipocalin-Gmel, and (F) all the remaining sequences shown. These groups were then profile aligned in the order: B + C, BC + A, D + F, DF + E, BCA + EDF.

An unrooted neighbor joining tree (Saitou and Nei, 1987) was constructed using PAUP (Version 4.0b1 for Macintosh), based on mean character difference (distance). The data matrix was trimmed to exclude leader sequences which would be cleaved during protein secretion; several additional bases were removed to even the 3′ end. All other characters were included; the program calculated pairwise differences, ignoring missing characters resulting from alignment gapping. Bootstrap support values were determined based on 1000 neighbor joining replicates, again using the PAUP program. The tree presented only includes nodes with 50% or higher bootstrap support; branch lengths are proportional and indicate mean distance (percentage difference) between the sequences.

Nomenclature

In general, the names of proteins discussed in this manuscript are those used in the original publications. The abbreviation translations and their relevant publications are as follows: PBP, pheromone binding protein (Vogt and Riddiford, 1981); GOBP, general odorant binding protein (Vogt *et al.*, 1991b); OS-E, OS-F, olfactory specific-X (McKenna *et al.*, 1994); PBPRP, pheromone binding protein related protein (Pikielny *et al.*, 1994); LAP, *Lygus* antennal protein (Dickens *et al.*, 1995); LUSH, a gene in *Drosophila* which, when mutated, results in increased behavioral affinity to alcohol (Kim *et al.*, 1998); CSRBP, chemosensory related binding protein (Ozaki *et al.*, 1995); ASP2, antennal specific protein (Danty *et al.*, 1997).

The first member identified of this family was called PBP because of its demonstrated interaction with sex-pheromone (Vogt and Riddiford, 1981). With the identification of the GOBP homologues (Vogt *et al.*, 1991b), the more general designation of ‘odorant binding protein’ (OBP) was applied to the entire family with subclasses given consistent but functionally relevant names (Vogt *et al.*, 1991b). Additional proteins were subsequently reported and arbitrarily named in the absence of functional or sequence data. In this report we have taken the liberty of referring to several of these additional proteins as OBPRP, or odorant binding protein related proteins, on the basis of their sequences and tissue distributions. We have also taken the liberty of indicating the species of origin following the designations, so that the functional/structural class of protein can more easily be identified. Finally, for the purpose of clarity, we have assigned numbers to the PBPs recently identified from *M. sexta* by Robertson and co-workers ((Robertson *et al.*, 1998), assigning PBP1 to the protein initially cloned by

Györgyi *et al.* (Györgyi *et al.*, 1988), which is also the most abundant of the PBPs in this species.

Results

LAP sequence analysis

The full length sequence of the LAP protein is shown in Figure 3. The N-terminal sequence of the mature protein was previously determined by direct amino acid sequencing (Dickens *et al.*, 1995). The 5′-Race results independently identified the cDNA sequence encoding this N-terminus, verifying the cDNA identity as that of LAP. The location of the start methionine is suggested by an in-frame ATG 16 codons upstream of the mature N-terminus; this length is consistent with leader sequences of secreted proteins including those of other insect OBPs (Vogt *et al.*, 1991a). The location of the 3′ terminus is suggested by the presence of two in-frame stop codons immediately downstream from this site. These results predict a mature secreted protein of 116 amino acids and a mass of 12.5 kDa, somewhat smaller than the 15–17 kDa previously estimated by SDS–PAGE (Dickens and Callahan, 1996); the overall length and size of the predicted protein are consistent with other similar OBP-related proteins.

A data base search using the NCBI BLAST network server indicated that LAP shared significant sequence similarity with several OBP-related proteins, including: the OS-E and OS-F proteins previously cloned from *Drosophila melanogaster* (McKenna *et al.*, 1994; Pikielny *et al.*, 1994); several antennal proteins designated as ABPX from Lepidoptera, for example ABPX-Hvir (Krieger *et al.*, 1996); and two antennal proteins from Coleoptera, for example OBPRP-Pjap (Wojtasek *et al.*, 1998). Probability values derived from this search ranged toward 5×10^{-19}, where values <0.05 are considered statistically significant (Karlin and Altschul, 1990) (Table 1). An alignment of LAP with several of these proteins is shown in Figure 4; percentage identities between LAP and these proteins ranged from 31 to 37% (Table 1). Of particular note is the presence of six conserved cysteines; both the presence and spatial distribution of these cysteines is a consistent hallmark of all the proteins so far identified as OBP-related insect proteins (Breer *et al.*, 1990; Vogt *et al.*, 1991a; McKenna *et al.*, 1994).

Phylogenetic analysis of LAP with other OBP-related sequences

All currently available OBP-related sequences were identified and collected using NCBI Entrez and BLAST network servers, and aligned in Clustal X (Table 2, Figure 5). Figure 6 shows an unrooted neighbor joining tree derived from this character matrix, representing the relationships of all these OBP-related proteins. This analysis suggests that the OBP-related proteins can be organized into several different classes, some of which appear monophyletic with respect to orders or species of origin based on currently available data, while others are clearly polyphyletic. For example, the OS-E/OS-F group is so far a species-specific class, unique to *Drosophila*; however, a survey of additional dipteran species would likely broaden this representation. The lepidopteran PBP and GOBP classes remain order-specific classes identified from multiple moth species; no closely related sequences have been identified outside the Lepidoptera.

Table 1 Comparisons of OBP-related proteins

Taxon	%[a]	Prob.[b]	Taxon	%	Prob.[b]
OBPRP-Pjap	37.2	5e-19	GOBP2-Bmor	16.1	>1
ABPX-Hvir	37.5	5e-14	OBPRP-Msex	15.8	>1
OS-F-Dmel	32.2	3e-11	PBPRP1-Dmel	15.4	>1
B1-Tmol	30.9	3e-6	GOBP1-Msex	14.1	>1
CSRBP-Preg	23.2	0.003	ASP2-Amel	13.7	>1
PBPRP5-Dmel	22.8	0.071			
LUSH-Dmel	23.5	0.47			
Sericotropin-Gmel	19.3	0.14	OSF: PBPLdis1	14.5	0.003
PBPRP2-Dmel	16.7	>1	OSF: PBPLdis2	17.9	0.003
PBP-Apol	16.7	>1	OSF: PBPAper1	17.2	0.004

[a]Percentage of amino acid identities with LAP, except for OSF–PBP comparisons.

[b]Probability values from LAP homology search using NCBI BLAST network server; values <0.05 are considered statistically significant (Karlin and Altschul, 1990).

LAP appears to belong to a strongly polyphyletic class of proteins which includes the lepidopteran ABPX proteins (Krieger *et al.*, 1996) and the coleopteran OBPRP proteins (Wojtasek *et al.*, 1998). The percentage identities between LAP and representative sequences from this analysis are shown in Table 1. The only proteins which show significant similarities with LAP are the lepidopteran ABPX proteins, the coleopteran proteins, and the dipteran proteins OS-E, OS-F and CSRBP. There was no significant direct support for a relationship between LAP and the lepidopteran PBP/GOBP proteins based on sequence alone. However, this relationship is weakly supported through the dipteran OS-F protein; NCBI BLAST comparison of OS-F and certain lepidopteran PBPs yield probability values of 0.003–0.004 which are considered significant (Karlin and Altschul, 1990), although the percentage identity values between OS-F and these PBPs are very low (<20%, Table 1).

In situ hybridization studies of LAP expression

In situ hybridization analysis was performed using a digoxigenin incorporated LAP antisense-RNA probe against tissue sections of adult male and female *L. lineolaris* antennae, as well as final instar nymph antennae. Localized hybridization was observed within the olfactory epithelia of both adult male and female antennae (Figure 7A, B), in a pattern similar to that previously observed using antiserum prepared against synthetically produced LAP N-terminus—see Figure 7D (Dickens and Callahan, 1996; Dickens *et al.*,

```
                                1                                 2   3
LAP-Lygus     GELP-EEMREMAQGLHDGCVEETGVDNGLIGPCAKGN-FADDQKLKCYFKCVFGNLGVIS
ABPX-Hvir     AVAMDEDMAELARMVRENCAAETGADVALVERVNAGADLMPDDKLKCYIKCTMETAGMMA
OBPRP-Pjap    ---MSEEMEELAKQLHDDCVSQTGVDEAHITTVKDQKGFPDDEKFKCYLKCLMTEMAIVG
OSF-Dmel      ENYPPPGILKMAKPFHDACVEKTGVTEAAIKEFSDGE-IHEDEKLKCYMNCFFHEIEVVD
OSE-Dmel      GEWPPPAILKLGKHFHDICAPKTGVTDEAIKEFSDGQ-IHEDEALKCYMNCLFHEFEVVD
                 * * * * *   ** &*  &&**   *     *     & **&&& *&      *
                                          4       5       6
LAP-Llin      DEGELDAEAFGSILP--DNMQELLPTIRGCAGTTGADPCELAMNFNKCLQKVDPVNFMVI
ABPX-Hvir     D-GEVDIEAVLALLP-PELAEHNAPSLRACGTVRGADHCDTAFRTQQCWQNANKADYFLI
OBPRP-Pjap    DDGVVDVEAAVGVLP-DELKAKAEPIMRKCGFKPGANPCDNVYQTHKCYYETDAQSYMIV
OSF-Dmel      DNGDVHLEKLFATVP-LSMRDKLMEMSKGCVHPEGDTLCHKAWWFHQCWKKADPKHYFLP
OSE-Dmel      DNGDVHMEKVLNAIPGEKLRNIMMEASKGCIHPEGDTLCHKAWWFHQCWKKADPVHYFLV
              & *  * &*    *&         *  **&    &*  &  *     &  * **
```

Figure 4 LAP alignment with OBP-related proteins. LAP is shown aligned with sequences identified as significantly similar using the NCBI BLAST network server. The positions of six conserved cysteines are indicated by numbers. LAP amino acids which are conserved throughout these sequences are indicated by an ampersand (&), and those which are represented in at least two additional sequences are indicated by an asterisk (*). The alignment was done in Clustal X; sequence names are identified in Table 1.

1998). Close examination revealed LAP hybridization in clusters of cells which associated with sensory hairs or sensory hair-related structures (Figure 8A–C). The cell clusters appeared to include three or four cells; each cell is discernible as a ring of stained cytoplasm surrounding a negatively staining nucleus. These cellular distributions were spatially consistent with the distribution of LAP protein visualized using the LAP-antiserum (Figure 8D). The LAP expressing cells are presumably the sensilla support cells, by analogy to the OBP expressing cells observed in Lepidoptera and Diptera (Steinbrecht *et al.*, 1992, 1995; Hekmat-Schafe *et al.*, 1997). However, there appear to be more cells within each LAP cluster than is typically observed for OBP expression in Lepidoptera or Diptera.

Lygus lineolaris is a hemimetabolous insect with five juvenile or nymph stages all of which resemble the adult stage in general appearance. The olfactory antennae become larger with each stage with an accompanying increase in the number of sensilla; an adult antenna has ~300% more sensilla than that of a fifth (last)-stage nymph (Chinta *et al.*, 1997). *In situ* hybridization studies were performed on fifth-stage nymphs collected during the active feeding period when olfactory activity would be high. No LAP expression was observed in fifth-stage nymphal antennae (data not shown), supporting a previous report based on Western blot analysis that LAP expression is adult specific (Dickens and Callahan, 1996).

In preparation for the molt from the fifth-stage nymph to adult, the adult antennae must pre-form within the nymph antenna; at the molt from nymph to adult, the outer cuticle of the nymph antenna is shed, revealing the adult antenna. *In situ* hybridization studies of such transitional antennae revealed LAP expression in localized patterns which resembled the adult pattern (Figure 8E). Between the epithelium expressing LAP and the outer pigmented cuticle, an unpigmented cuticle complete with sensilla was clearly observed; this unpigmented cuticle was the pre-forming adult cuticle. These results indicate LAP expression initiates in adult tissue but prior to the nymph-adult molt, suggesting that *L. lineolaris* may emerge as adult animals with a fully functional olfactory system. Pre-molt expression of OBPs has also been observed during adult development of several lepidopteran species (Vogt *et al.*, 1989, 1993), although adult antennal development in these holometabolous insects is markedly different than that in the hemimetabolous *L. lineolaris*.

Discussion

LAP and the OBP-related proteins

LAP was previously purified from the adult antennae of the hemipteran insect *L. lineolaris* (Dickens *et al.*, 1995). A partial amino acid sequence was obtained, the N-terminus was synthesized and polyclonal antiserum was generated against the synthetic peptide. Western blot analysis indicated that LAP was uniquely expressed in adult antennae (Dickens and Callahan, 1996). Immunocytochemistry at the electron-microscopic level showed that LAP was localized within the olfactory sensilla, in the fluid surrounding the olfactory neurons (Dickens *et al.*, 1998). LAP was consequently proposed to have an OBP-like function based on its size, relative abundance and localized expression. This view is supported by the current study, where the complete LAP sequence has been deduced and shown to be significantly similar to other OBP-related proteins, including the *Drosophila* OS-E and OS-F proteins (McKenna *et al.*, 1994; Pikielny *et al.*, 1994), the lepidopteran ABPX proteins (Krieger *et al.*, 1996, 1997; Robertson *et al.*, 1998), and the coleopteran proteins OBPRP-Pjap and OBPRP-Aosa (Wojtasek, *et al.*, 1998). The *in situ* hybridization results confirmed the adult specific expression of LAP and strongly suggest that the function of LAP associates with adult specific behavior, possibly reproduction.

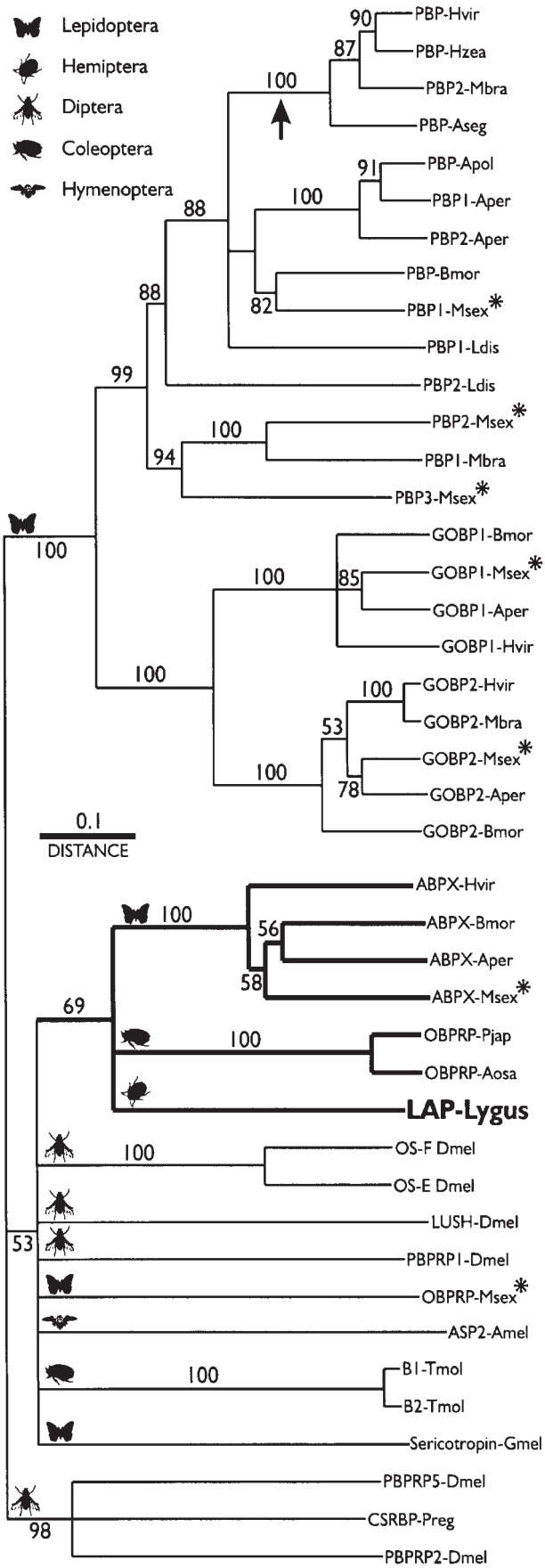
The application of the term OBP to these proteins derives from the initial demonstration that sex-pheromone bound to a male antenna-specific protein that was localized within the pheromone-specific sensilla of the silk moth *Antheraea polyphemus*; this protein was consequently called phero-

Table 2 OBP-related proteins used in phylogenetic analysis

Order	Species	Name	Accession no.	Reference
Lepidoptera (moths)	*Antheraea polyphemus*	PBP-Apol	EMBL-X17559	(Raming *et al.*, 1989)
	Antheraea pernyi	PBP1-Aper	EMBL-X96773	(Raming *et al*., 1990)
		PBP2-Aper	EMBL-X96860	(Krieger *et al.*, 1991)
		GOBP1-Aper	EMBL-Y10970	(Mameli *et al.*, 1997)
		GOBP2-Aper	EMBL-X96772	(Breer *et al.*, 1990)
		ABPX-Aper	EMBL-AJ002519	(Krieger *et al.*, 1997)
	Bombyx mori	PBP-Bmor	SP-P34174	(Krieger *et al.*, 1996)
		GOBP1-Bmor	SP-P34170	(Krieger *et al.*, 1996)
		GOBP2-Bmor	SP-P34171	(Krieger *et al.*, 1996)
		ABPX-Bmor	EMBL-X94990	(Krieger *et al.*, 1996)
	Manduca sexta	PBP1-Msex	GB-M21798	(Györgyi *et al.*, 1988)
		PBP2-Msex	GB-AI172627	(Robertson *et al.* , 998)
		PBP3-Msex	GB-AI172639	(Robertson *et al.* , 998)
		GOBP1-Msex	GB-M73797	(Vogt *et al.*, 1991a)
		GOBP2-Msex	GB-M73798	(Vogt *et al.*, 1991a)
		ABPX-Msex	GB-AI187633	(Robertson *et al.* , 998)
		OBPRP-Msex	GB-AI172726	(Robertson *et al.* , 998)
	Heliothis virescens	PBP-Hvir	SP-Q27388	(Krieger *et al.*, 1993)
		GOBP1-Hvir	SP-Q27226	(Krieger *et al.*, 1993)
		GOBP2-Hvir	SP-Q27288	(Krieger *et al.*, 1993)
		ABPX-Hvir	EMBL-AJ002518	(Krieger *et al.*, 1997)
	Helicoverpa zea	PBP-Hzea	GB-AF090191	(Callahan *et al.*, 1998)
	Mamestra brassicae	PBP1-Mbra	GB-AF051143	(Maibeche-Coisne *et al.*, 1998a)
		PBP1-Mbra	GB-AF051142	(Maibeche-Coisne *et al.*, 1998a)
		GOBP2-Mbra	GB-AF051144	(Maibeche-Coisne *et al.*, 1998b)
	Lymantria dispar	PBP1-Ldis	GB-AF007867	
		PBP2-Ldis	GB-AF007868	(Merritt *et al.*, 1998)
	Agrotis segetum	PBP-Aseg	GB-1345183	(Prestwich *et al.*, 1995)
	Galleria mellonella	Sericotropin (non-olfactory)	GB-L41640	(Filippov *et al.*, 1995)
Coleoptera (beetles)	*Popillia japonica*	OBPRP-Pjap	GB-AF031491	(Wojtasek *et al.*, 1998)
	Anomala osakana	OBPRP-Aosa	GB-AF031492	(Wojtasek *et al.*, 1998)
	Tenebrio molitor	B1-Tmol	GB-M97916	
		B2-Tmol (non-olfactory)	GB-M97917	(Paesen and Happ, 1995)
Diptera (flies)	*Drosophila melanogaster*	OS-E Dmel	GB-U02543	(McKenna *et al.*, 1994)
		OS-F Dmel	GB-U02542	(McKenna *et al.*, 1994)
		PBPRP1-Dmel	GB-U05980	(Pikielny *et al.*, 1994)
		PBPRP2-Dmel	GB-U05981	(Pikielny *et al.*, 1994)
		PBPRP3-Dmel	GB-U05982 (OS-F)	(Pikielny *et al.*, 1994)
		PBPRP5-Dmel	GB-U05985	(Pikielny *et al.*, 1994)
		LUSH-Dmel	GB-AF001621	(Kim *et al.*, 1998)
	Phormia regina	CSRBP-Preg (taste)	PIR-S65458	(Ozaki *et al.*, 1995)
Hymenoptera (honeybee)	*Apis mellifera*	ASP2		(Danty *et al.*, 1997)
Hemiptera (True Bugs)	*Lygus lineolaris*	LAP-Lygus	GB-AF091118	this publication

mone binding protein or PBP (Vogt and Riddiford, 1981). The ability of these proteins to discriminate odors was subsequently demonstrated for the PBPs of the gypsy moth *Lymantria dispar* (Vogt *et al.*, 1989; Prestwich, 1991) and the silk moths *Bombyx mori* (Krieger *et al.*, 1992) and A*ntheraea pernyi* (Du and Prestwich, 1995). Two additional classes of Lepidoptera OBP were identified as the general odorant binding proteins GOBP1 and GOBP2 (Vogt and Lerner, 1989; Breer *et al.*, 1990; *Vogt et al.*, 1991b); these associate with plant-volatile-sensitive olfactory sensilla (Vogt *et al.*, 1991a, 1991b; Laue and Steinbrecht, 1997). A protein identified as GOBP2 in *A. polyphemus* (Vogt *et al.*, 1991b) was previously observed to bind a sex-pheromone odorant (Vogt and Riddiford, 1981), and interactions between a variety of odorants and the *M. sexta* GOBP2 were recently characterized in detail (Feng and Prestwich, 1997).

Figure 5 Character matrix used to construct neighbor joining tree. Alignments were done in Clustal X, as described in Materials and methods. Sequence names, accession numbers and sources are identified in Table 1.



The first non-Lepidoptera OBP-related proteins identified were the *Drosophila* antennal proteins OS-E and OS-F (McKenna *et al.*, 1994), and PBPRP 1, 2, 3 and 5 (Pikielny *et al.*, 1994). NCBI BLAST analysis indicates that OS-F has a weakly significant probability of sequence relationship with three lepidopteran PBPs (0.003–0.004, Table 1), but otherwise these *Drosophila* proteins share little sequence identity with the lepidopteran PBP and GOBP proteins and are consistently smaller (see Figure 5). However, the *Drosophila* proteins do contain six cysteines which are distributed in a similar spatial pattern to the lepidopteran proteins, and appear to be a hallmark feature of the entire group of OBP-related proteins (see Figures 4 and 5). The *Drosophila* proteins were suggested to be related to the lepidopteran OBPs on the basis of size, abundance, antennal specific expression, and presence of the six hallmark cysteines (McKenna *et al.*, 1994; Pikielny *et al.*, 1994). Recent histological analysis has shown that OS-E and OS-F are localized within the sensilla fluid surrounding the olfactory neurons (Hekmat-Schafe *et al.*, 1997), in a manner consistent with that of the lepidopteran OBPs (Vogt and Riddiford, 1981; Steinbrecht *et al.*, 1992, 1995; Laue *et al.*, 1994; Laue and Steinbrecht, 1997). This OS-E/OS-F related group is now known to be broadly distributed, with homologues identified in Lepidoptera (e.g. ABPX), Coleoptera (for example, OBPRP-Pjap) and now Hemiptera (LAP).

Recently, the olfactory role of the insect OBPs was directly affirmed using the LUSH protein in *Drosophila* (Kim *et al.*, 1998). An enhancer trap induced mutation was identified that caused an increased behavioral attraction to alcohols such as ethanol and propanol; wild type flies normally avoid these alcohols. LUSH uniquely expresses in the olfactory tissue of both larvae and adults, associating with olfactory sensilla. NCBI BLAST analysis indicates LUSH shares significant sequence similarity with the coleopteran OBPRP and lepidopteran ABPX proteins (probability values = 0.0025–0.0003). Because LUSH appears to be an OBP-related protein uniquely expressed in olfactory sensilla, the alcohol avoidance defect was interpreted to be a failure to effectively detect the alcohols, which might be consistent with the proposed odor transport function of the OBPs. In another dipteran system, antibody

Figure 6 Neighbor joining tree of OBP-related sequences. Branch lengths are proportional and the scale of distance is indicated. Bootstrap support values (%) based on 1000 replicates are indicated. Seven *M. sexta* sequences are indicated by an asterisk (*). In an earlier publication, the *L. dispar* PBP sequences were shown as a monophyletic group within the PBP group marked by the arrow (Aseg, Hvir, Hzea and Mbra) (Merritt *et al.*, 1998). The more basal position of the *L. dispar* PBPs in this tree is most likely due to their considerable sequence difference from the other PBPs in this group. The earlier interpretation is more parsimonious and consistent with the species phylogeny and may suggest *L. dispar* is undergoing a more rapid evolution than the other species (Merritt *et al.*, 1998), a view which is currently under investigation.

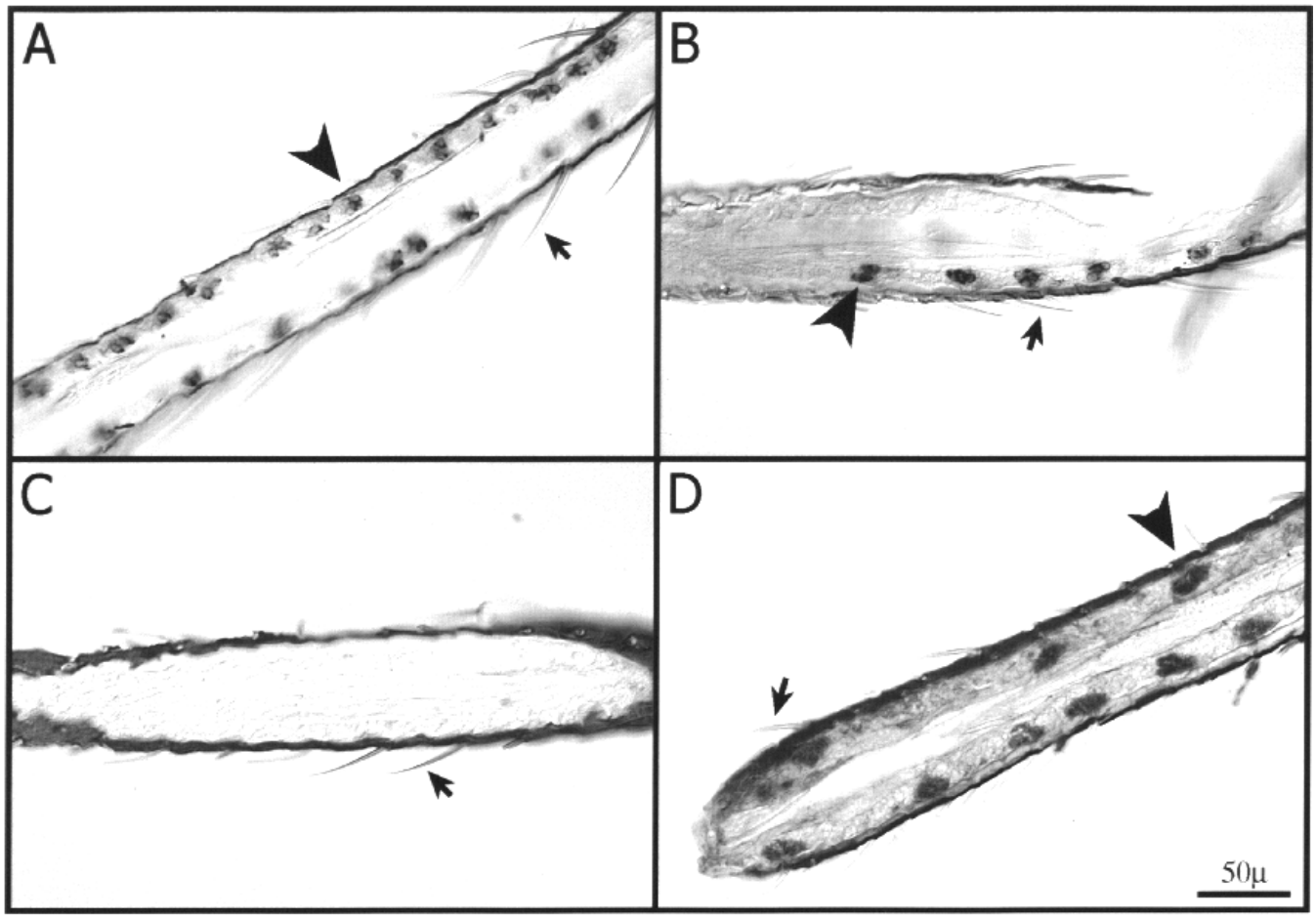


Figure 7 *In situ* hybridization and immunohistological analysis of LAP expression in adult tissue. **(A)** Male antenna; **(B)** female antenna hybridized with antisense LAP probe. **(C)** Negative control, hybridized with a probe encoding the zebrafish odor receptor 1A (Byrd *et al.*, 1996). **(D)** Immunohistological localization of LAP protein in a female antenna, after Dickens *et al.* (Dickens *et al.*, 1998). Large arrows point to representative staining, while small arrows point to the outer cuticular portions of the representative sensilla. All images are at the same magnification.

blocking experiments to the OBP-related CSRBP, which associates with taste sensilla in the fly *Phormia regina*, resulted in decreased electrophysiological response to odorant-like stimulants, offering further direct support for a functional role in stimulant detection for the OBP-related proteins (Ozaki, *et al.*, 1995).

Multiple classes of insect OBPs

The tree analysis presented here suggests a strongly supported separation of insect OBP-related proteins into multiple classes. There is strong support (100% bootstrap) for a distinct Lepidoptera PBP/GOBP class and respective subclasses. There is also strong support for the CSRBP/PBPRP2/PBPRP5 class (98% bootstrap), but only weak support that this should be considered distinct from the other non-PBP/GOBP OBP-like sequences (53% bootstrap). While many of the major branches define species or order specific lineages (OS-E and OS-F, B1 and B2, OBPRP-Pjap and OBPRP-Aosa, PBPs/GOBPs, ABPX), these divisions may also indicate functionally distinct classes of OBPs. For example, in *Drosophila*, there appear to be strong separations between respective branches containing OS-F/OS-E, LUSH, PBPRP1, and PBPRP2/PBPRP5. Some of these divisions may collapse as sequences from other dipteran species become available, but they suggest different functional roles within *Drosophila* for the respective groups. Among the Lepidoptera, there is strong support for multiple classes consisting of PBP, GOBP1, GOBP2 and ABPX (100% bootstrap). To date, *M. sexta* has yielded the largest number of OBP sequences, including three PBPS, GOBP1, GOBP2, ABPX and the OBPRP-Msex sequence (see Table 2). Among the PBPs, the strongly supported branch which includes PBP2-Msex, PBP3-Msex and PBP1-Mbra may define a distinct class of PBPs, separate from those previously identified. Ongoing histological analysis is indicating that many of these proteins are differentially expressed in

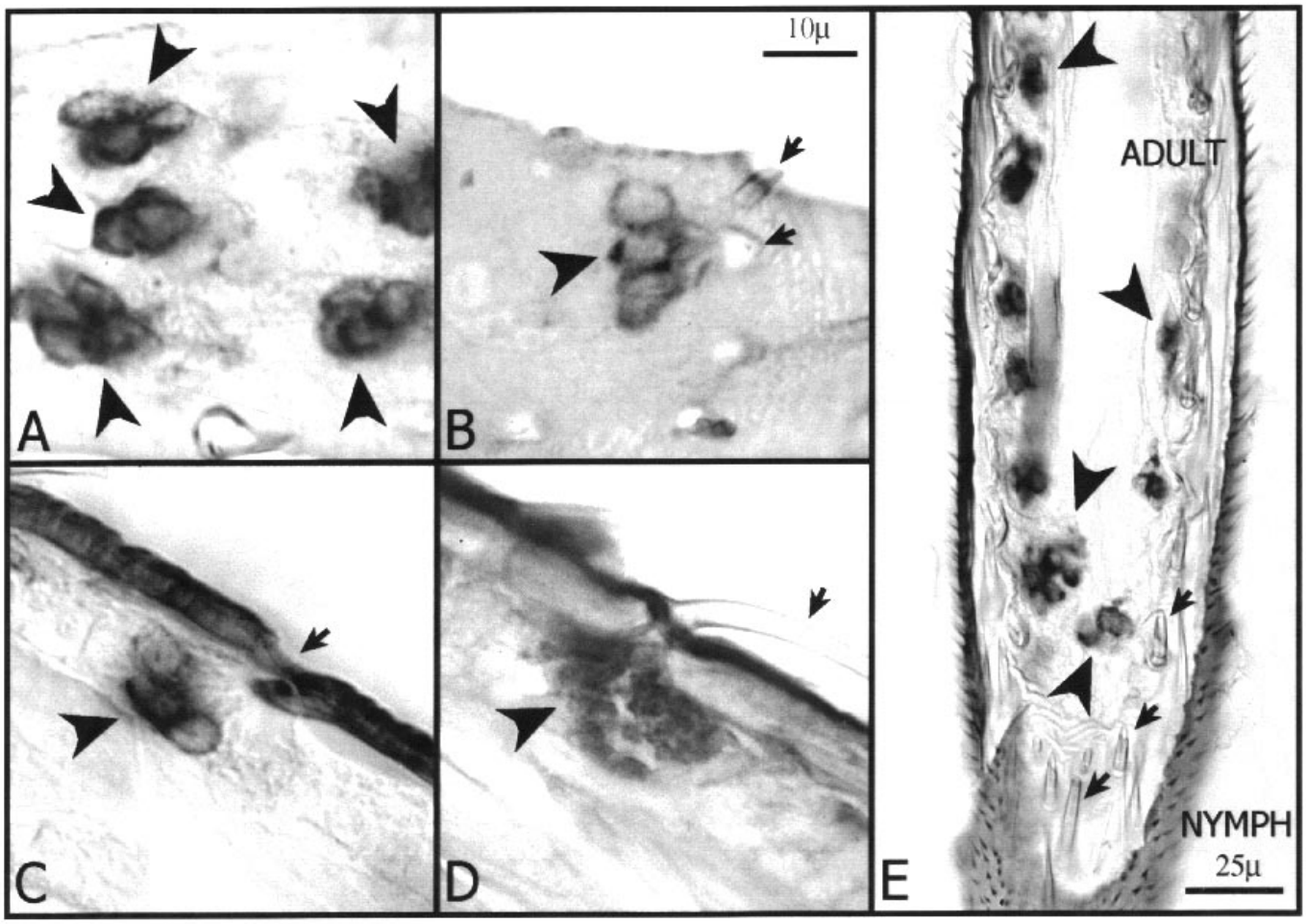


Figure 8 *In situ* hybridization and immunohistological analysis of LAP expression in adult and molting tissue. **(A)** Adult male antenna; **(B)** and **(C)** adult female antennae hybridized with antisense LAP probe. **(D)** Immunohistological localization of LAP protein in an adult female antenna. **(E)** Transitional female antenna, in the process of preparing to molt, hybridized with antisense LAP probe. In **(E)**, NYMPH identifies the outer and darker cuticle of the 5th stage nymph antenna, and ADULT identifies the inner and unpigmented cuticle of the developing adult antenna. Large arrows point to representative staining, while small arrows point to cuticular components of the sensilla. A–D are at the same magnification; scales are indicated.

association with functionally distinct types of olfactory sensilla, supporting the view that different OBPs associate with sensilla mediating different olfactory based behaviors (Vogt *et al.*, 1991b; Steinbrecht, 1996; Laue and Steinbrecht, 1997).

If the collection of sequences gathered here do belong to a single homologous group, derived from a common ancestral gene, then LAP indicates the phylogenetic depth of this gene family. Figure 2 shows a conservative view of the cladistic relationship of the insect orders, and illustrates the monophyletic relationship of the sister groups Endopteryota and Hemipteroid Assemblage. All of the OBP-related sequences presented here, except for LAP, derive from species (moths, flies, beetles and bees) belonging to the Endopterygota. The identification of LAP as an OBP-related protein from the Hemipteroid Assemblage argues that the insect OBP family was represented in the ancestral form that was the common origin of these two divisions. This further argues that OBP-related genes should be represented throughout the species belonging to the Endopterygota and Hemipteroid Assemblage, unless they were secondarily lost.

It is curious that the PBPs and GOBPs are so strongly supported as a lepidopteran specific group (Figure 6). It may be that close homologues have simply not yet been identified in non-lepidopteran orders. However, the apparent diversification of PBPs and GOBPs within the Lepidoptera may reflect the great extent to which these animals rely on highly specific olfactory cues in coordinating their behavior. This would suggest that the lepidopteran OBPs should make a very useful model system for investigating the diversification and evolution of chemosensory based behavior.

OBP-like proteins outside the Endopteryogota and Hemipteroid Assemblage

To date, no OBP-related sequences have been identified outside the Endopterygota and Hemipteroid groups. A

thorough survey of antennal proteins from walking-sticks (Phasmatodea) identified several abundant proteins from chemosensory organs, but these proteins have no supportable sequence relationship to the OBP-related proteins discussed here (Mameli *et al.*, 1996; Pelosi, 1996). However, some of the walking-stick antennal proteins do share significant sequence similarity to a *Drosophila* antennal protein referred to as OS-D (McKenna *et al.*, 1994, SP-Q27377). OS-D related proteins have also been identified in Lepidoptera, for example: *M. sexta* (Robertson *et al.*, 1998, GB-AI172733); *M. brassicae* (Bohbot *et al.*, 1998); Orthoptera—locust (Angeli *et al.*, 1998, GB-AF070961); and Dictyoptera—cockroach (Nomura *et al.*, 1992, GB-AF030340). Recently, the lepidopteran OS-D related protein from *M. brassicae* was shown to bind several odorants, suggesting that these proteins may have an OBP-like function (Bohbot *et al.*, 1998).

In the context of chemoreception, there are two puzzling features of the OS-D related proteins. First, they are often not antennal specific in their expression; a homologue from the cockroach *Periplaneta americana* expresses in regenerating leg tissue—leg regeneration protein p10, GB-AF030340 (Nomura *et al.*, 1992) and another from *Drosophila* expresses in the ejaculatory bulb (Dyanov, H.M., direct submission, GB-U08281). Second, the OS-D related proteins are highly conserved across distantly related insect orders. For example, pairwise identity comparisons between OS-D homologues of different orders range from 43 to 59% (Diptera—*Drosophila*, OS-D, SP-Q27377; Orthoptera—*Schistocerca gregaria*, CSP-sg1, GB-AF070961; Dictyoptera—*P. americana*, GB-AF030340; Lepidoptera—*Cactoblastis cactorum*, GB-U95046). These values seem high if the role of the proteins is related to chemosensory function. Chemosensory function is often adaptive with life history, and the life histories of the various insect groups are quite different. Greater sequence divergence might be expected to match the life-history differences, as is observed for the OBP-related proteins; pairwise identity values range from 10 to 38% between PBP1-Msex, ABPX-Msex, OBPRP-Pjap, OS-E and LAP. Certainly, additional functional and expression analyses are necessary to clarify the role of the OS-D related proteins. If the OS-D related proteins prove to have an OBP-like function, it will be useful to distinguish them from the proteins discussed here, perhaps as OBP-Type 1 (PBPs, GOBPs, OS-E, etc.) and OBP-Type 2 (OS-D, etc.).

Many questions remain regarding the function and activity of the OBP-related proteins in olfactory processing, especially concerning the nature of the OBP–odorant interaction with respect to on–off rates, the specificity of OBP–odorant interactions, the spatial and developmental patterns of expression of multiple OBPs, and the regulation of OBP expression in developmentally and phenotypically appropriate contexts. Also, in light of the recent identification of presumptive odor receptors from *Drosophila* (Clyne *et al.*, 1999; Vosshall *et al.*, 1999), characterizing the interactions between OBPs and such receptors may prove important.

Acknowledgements

The authors thank Dr Gordon Snodgrass, USDA–ARS, Southern Insect Management Laboratory, Stoneville, MS, USA for graciously supplying *Lygus lineolaris*, Dr Joseph M. Quattro and Mr Thomas J.S. Merritt for discussions and advice regarding phylogenetic analyses, M. Sun for histological assistance and the following agencies for their support: National Institutes of Health (R.G.V., NICDC DC-00588); National Science Foundation (R.G.V., IBN9731005), United States Department of Agriculture (R.G.V., CGRP 94-37302-0615; F.E.C., CRIS 6406-22000-015-00D) and Cotton Incorporated (J.C.D., Coop. Agreement No. 96-388). Mention of a trademark, product, or vendor does not constitute a guarantee or warranty of the product by USDA, and does not imply its approval to the exclusion of other products or vendors that may be suitable.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Shang, Z., Miller, W. and **Lipman, D.J.** (1997) *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res., 25, 3389–3402.

Angeli, S., Ceron, F., Monteforti, G., Scaloni, A., Minnocci, A. and **Pelosi, P.** (1998) Unpublished sequence, GenBank accession no. AF070964.

Bohbot, J., Sobrio, F., Lucas, P. and **Nagnan-Le Meillour, P.** (1998) *Functional characterization of a new class of odorant-binding proteins in the moth* Mamestra brassicae. Biochem. Biophys. Res. Commun., 253, 489–494.

Borror, D.J., Triplehorn, C.A. and **Johnson, N.F.** (1989) An Introduction to the Study of Insects, 6th edn. Harcourt Brace, Orlando, FL.

Boudreaux, H.B. (1979) Arthropod Phylogeny with Special Reference to Insects. Wiley, New York.

Breer, H. (1997) *Molecular mechanisms of pheromone reception in insect antennae*. In Cardé, R.T. and Minks, A.K. (eds), Insect Pheromone Research: New Directions. Chapman & Hall, New York, pp. 115–130.

Breer, H., Krieger J. and **Raming, K.** (1990*) A novel class of binding proteins in the antennae of the silkmoth* Antheraea pernyi. Insect Biochem., 20, 735–740.

Byrd, C.A., Jones, J.T., Quattro, J.M., Rogers, M.E., Brunjes, P.C. and **Vogt, R.G.** (1996) *Ontogeny of odorant receptor gene expression in zebrafish,* Danio rerio. J. Neurobiol., 29, 445–458.

Callahan, F.E., Dickens, J.C., Tucker, M.L., Vogt, R.G. and **Mattoo, A.K.** (1998) *Pheromone binding proteins in noctuid moths*. Unpublished sequence, GenBank accession no. AF090191.

Carlson, J.R. (1996) *Olfaction in* Drosophila*: from odor to behavior*. Trends Genet., 12, 175–180.

Chinta, S., Dickens, J.C. and **Baker, G.T.** (1997*) Morphology and distribution of antennal sensilla of the tarnished plant bug,* Lygus lineolaris *(Palisot de Beauvois) (Hemiptera: Miridae).* Int. J. Insect Morphol. Embryol., 26, 21–26.

Chomczynski, P. and **Sacchi, N.** (1987) *Single-step method of RNA isolation by acid guanidium thiocyanate–phenol–chloroform extraction*. Anal. Biochem., 162, 156–159.

Clyne, P., Warr, C., Freeman, M., Lessing, D., Kim, J. and **Carlson, J.**

(1999) *A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in* Drosophila. Neuron, 22, 327–338.

Danty, E., Michard-Vanhee, C., Huet, J.C., Genecque, E., Pernollet, J.C. and **Masson, C.** (1997) *Biochemical characterization, molecular cloning and localization of a putative odorant-binding protein in the honey bee* Apis mellifera *L. (Hymenoptera: Apidea).* FEBS Lett., 414, 595–598.

Dickens, J.C. and **Callahan, F.E.** (1996) *Antennal-specific protein in tarnished plant bug,* Lygus lineolaris*: production and reactivity of antisera*. Ent. Exp. Appl., 80, 19–22.

Dickens, J.C., Callahan, F.E., Wergin, W.P. and **Erbe, E.F.** (1995) *Olfaction in a hemimetabolous insect: antennal-specific protein in adult* Lygus lineolaris *(Heteroptera: Miridae).* J. Insect Physiol., 41, 857–867.

Dickens, J.C., Callahan, F.E., Wergin, W.P., Murphy, C.A. and **Vogt, R.G.** (1998) *Intergeneric distribution immunolocalization of a putative odorant-binding protein in true bugs (Hemiptera, Heteroptera).* J. Exp. Biol., 201, 33–41.

Du, G. and **Prestwich, G.D.** (1995) *Protein structure encodes the ligand binding specificity in pheromone binding proteins.* Biochemistry, 34, 8726–8732.

Erwin, T.L. (1982) *Tropical forests: their richness in Coleoptera and other arthropod species*. Coleop. Bull., 36, 74–75.

Feng, L. and **Prestwich, G.D.** (1997) *Expression and characterization of a lepidopteran general odorant binding protein*. Insect Biochem. Mol. Biol. 27, 405–412.

Filippov, V.A., Filippova, M.A. and **Sehnal, F.** (1995) *Lipocalin-like brain-specific protein*. Unpublished sequence, GenBank accession no. L41640.

Freeman, S. and **Herron, J.C.** (1998) Evolutionary Analysis. Prentice Hall, Upper Saddle River, NJ.

Györgyi, T.K., Roby-Shemkovitz, A.J. and **Lerner, M.R.** (1988) *Characterization and cDNA cloning of the pheromone-binding protein from the tobacco hornworm,* Manduca sexta*: a tissue-specific developmentally regulated protein.* Proc. Natl Acad. Sci. USA, 85, 9851–9855.

Hekmat-Schafe, D.S., Steinbrecht, R.A. and **Carlson, J.R.** (1997) *Coexpression of two odorant-binding protein homologs in* Drosophila*: implications for olfactory coding*. J. Neurosci., 17, 1616–1624.

Hennig, W. (1981) Insect Phylogeny. Wiley, New York.

Kaissling, K.-E. (1998*) Pheromone deactivation catalyzed by receptor molecules: a quantitative kinetic model*. Chem. Senses, 23, 385–395.

Karlin, S. and **Altschul, S.F.** (1990) *Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes*. Proc. Natl Acad. Sci. USA, 87, 2264–2268.

Kim, M.S., Repp, A. and **Smith, D.P.** (1998) *LUSH odorant-binding protein mediates chemosensory responses to alcohols in* Drosophila melanogaster. Genetics, 150, 711–721.

Krieger, J., Raming, K. and **Breer, H.** (1991) *Cloning of genomic and complementary DNA encoding insect pheromone binding proteins: evidence for microdiversity*. Biochim. Biophys. Acta, 1088, 277–284.

Krieger, J., Raming, K., Prestwich, G.D., Frith, D., Stabel, S. and **Breer, H.** (1992) *Expression of a pheromone-binding protein in insect cells using a baculovirus vector*. Eur. J. Biochem., 203, 161–166.

Krieger, J., Gänßle, H., Raming, K. and **Breer, H.** (1993) *Odorant binding proteins of* Heliothis virescens. Insect Biochem. Mol. Biol., 23, 449–456.

Krieger, J., von Nickisch-Rosenegk, E., Mameli, M., Pelosi, P. and **Breer, H.** (1996) *Binding proteins from the antennae of* Bombyx mori. Insect Biochem. Mol. Biol., 26, 297–307.

Krieger, J., Mameli, M. and **Breer, H.** (1997) Unpublished sequences, EMBL accession nos AJ002518 and AJ002519.

Kristensen, N.P. (1991) *Phylogeny of extant hexapods*. In CSIRO (eds), The Insects of Australia: A Textbook for Students and Research Workers. Melbourne University Press, Melbourne, pp. 125–140.

Kukalová-Peck, J. (1991) *Fossil history and the evolution of hexapod structures*. In CSIRO (eds), The Insects of Australia: A Textbook for Students and Research Workers. Melbourne University Press, Melbourne, pp. 141–179.

Labandeira, C.C. and **Sepkoski, J.J., Jr** (1993) *Insect diversity in the fossil record.* Science, 261, 310–315.

Laue, M. and **Steinbrecht, R.A.** (1997) *Topochemistry of moth olfactory sensilla.* Int. J. Insect Morphol. Embryol., 26, 217–228.

Levine, R.B., Morton, D.B. and **Restifo, L.L.** (1995*) Remodeling of the insect nervous system.* Curr. Opin. Neurobiol., 5, 28–35.

McKenna, M.P., Hekmat-Schafe, D.S., Gaines, P. and **Carlson, J.R.** (1994) *Putative* Drosophila *pheromone-binding proteins expressed in a subregion of the olfactory system*. J. Biol. Chem., 269, 16340–16347.

Maddison, D.R. and **Maddison, W.P.** (1998) The Tree of Life. A multi-authored, distributed Internet project containing information about phylogeny and biodiversity. Internet address: http://phylogeny.arizona.edu/tree/phylogeny.html

Maibeche-Coisne, M., Jacquin-Joly, E., Francois, M.C. and **Nagnan-Le Meillour, P.** (1998a) *Molecular cloning of two pheromone binding proteins in the cabbage armyworm* Mamestra brassicae. Insect Biochem. Mol. Biol., 28, 815–818.

Maibeche-Coisne, M., Jacquin-Joly, E., Francois, M.C. and **Nagnan-Le Meillour, P.** (1998b) Unpublished sequence, GenBank accession no. AF051144.

Mameli, M., Tuccini, A., Mazza, M., Petacchi, R. and **Pelosi, P.** (1996) *Soluble proteins in chemosensory organs of phasmids*. Insect Biochem. Mol. Biol., 26, 875–882.

Mameli, M., Kreiger, J. and **Breer, H.** (1997) Unpublished sequences, EMBL accession no. Y10970.

Maniatis, T., Fritsch, E.F. and **Sambrook, J.** (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Merritt, T.J., LaForest, S., Prestwich, G.D., Quattro, J.M. and **Vogt, R.G.** (1998) *Patterns of gene duplication in lepidopteran pheromone binding proteins*. J. Mol. Evol., 46, 272–276.

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.

Ozaki, M., Morisaki, K., Idei, W., Ozaki, K. and **Tokunaga, F.** (1995) *A putative lipophilic stimulant carrier protein commonly found in the taste and olfactory systems. A unique member of the pheromone-binding protein superfamily*. Eur. J. Biochem., 230, 298–308.

Paesen, G.C. and **Happ, G.M.** (1995) *The B proteins secreted by the tubular accessory sex glands of the male mealworm beetle,* Tenebrio molitor*, have sequence similarity to moth pheromone-binding proteins.* Insect Biochem. Mol. Biol., 25, 401–408.

Pelosi, P. (1996) *Perireceptor events in olfaction.* J. Neurobiol., 30, 3–19.

Pelosi, P. and **Maida, R.** (1995) *Odorant-binding proteins in insects.* Comp. Biochem. Physiol., 111B, 503–514.

Pelosi, P., Baldaccini, N.E. and **Pisanelli, A.M.** (1982) *Identification of a specific olfactory receptor for 2-isobutyl-3-methoxypyrazine*. Biochem. J., 201, 245–248.

Pevsner, J., Hwang, P.M., Sklar, P.B., Venable, J.C. and **Snyder, S.H.** (1988) *Odorant-binding protein and its mRNA are localized to lateral nasal gland implying a carrier function*. Proc. Natl Acad. Sci. USA, 85, 2383–2387.

Pikielny, C.W., Hasan, G., Rouyer, F. and **Rosbash, M.** (1994) *Members of a family of* Drosophila *putative odorant-binding proteins are expressed in different subsets of olfactory hairs*. Neuron, 12, 35–49.

Prestwich, G.D. (1991) *Photoaffinity-labeling and biochemical-characterization of binding proteins for pheromones, juvenile hormones, and peptides*. Insect Biochem., 21, 27.

Prestwich, G.D., Du, G. and **LaForest, S.** (1995) *How is pheromone specificity encoded in proteins?* Chem. Senses, 20, 461–469.

Raming, K., Krieger, J. and **Breer, H.** (1989) *Molecular cloning of an insect pheromone-binding protein.* FEBS Lett., 256, 215–218.

Raming, K., Krieger, J. and **Breer, H.** (1990) *Primary structure of a pheromone-binding protein from* Antheraea pernyi*: homologies with other ligand-carrying proteins*. J. Comp. Physiol. B, 160, 503–509.

Robertson, H.M., Martos, R., Sears, C. and **Nardi, J.B.** (1998) *Expressed sequence tags from the antennae of male* Manduca sexta *reveal diversity of odorant binding proteins*. Unpublished sequences, Genbank accession nos AI172627 (PBP2), AI172639 (PBP3), AI187633 (ABPX), AI172726 (OBPRP), AI172668 (OS-D homologue), AI172734 (OS-D homologue), AI187476 (OS-D homologue).

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.

Rogers, M.E., Jani, M.K. and **Vogt, R.G.** (1999) *An olfactory specific glutathione-*S*-transferase in the sphinx moth* Manduca sexta. J. Exp. Biol., 202, 1625–1637.

Rybczynski, R., Reagan, J. and **Lerner, M.R.** (1989) *A pheromone-degrading aldehyde oxidase in the antennae of the moth* Manduca sexta. J. Neurosci., 9, 1341–1353.

Rybczynski, R., Vogt, R.G. and **Lerner, M.R.** (1990) *Antennal-specific pheromone-degrading aldehyde oxidases from the moths* Antheraea polyphemus *and* Bombyx mori. J. Biol. Chem., 32, 19712–19715.

Saitou, N. and **Nei, M.** (1987) *The neighbor joining method: a new method for reconstructing phylogenetic trees*. Mol. Biol. Evol., 4, 406–425.

Schwartz, M.D. and **Foottit, R.G.** (1992) Lygus bugs on the Prairies: Biology, Systematics and Distribution. Agric. Can. Tech. Bull. 1992-4. Internet address: http://res.agr.ca/brd/lygus/title.html

Snodgrass, G.L. and **McWilliams, J.M.** (1992) *Rearing of the tarnished plant bug (Heteroptera: Miridae) using a tissue paper oviposition site*. J. Econ. Ent., 85, 1162–1166.

Snodgrass, G.L., Scott, W.P. and **Smith, J.W.** (1984) *Host plants and seasonal distribution of the tarnished plant bug (Hemiptera: Miridae) in the delta of Arkansas, Lousiana and Mississippi.* Env. Ent., 13, 110–116.

Steinbrecht, R.A. (1996) *Are odorant-binding proteins involved in odorant discrimination?* Chem. Senses, 21, 718–725.

Steinbrecht, R.A., Ozaki, M. and **Ziegelberger, G.** (1992) *Immuno-cytochemical localization of pheromone-binding protein in moth antennae*. Cell Tissue Res., 270, 287–302.

Steinbrecht, R.A., Laue, M. and **Ziegelberger, G.** (1995) *Immuno-localization of pheromone-binding protein and general odorant-binding protein in olfactory sensilla of the silk moths* Antheraea *and* Bombyx. Cell Tissue Res., 282, 203–217.

Thompson, J.D., Higgins, D.G. and **Gibson, T.J.** (1994) *Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice.* Nucleic Acids Res., 22, 4673–4680.

Vogt, R G. (1987) *The molecular basis of pheromone reception: its influence on behavior.* In Prestwich, G.D. and Blomquist, G.J. (eds), Pheromone Biochemistry. Academic Press, New York, pp. 385–431.

Vogt, R.G. (1995) *Molecular genetics of moth olfaction: a model for cellular identity and temporal assembly of the nervous system*. In Goldsmith, M.R. and Wilkins, A.S. (eds), Molecular Model Systems in the Lepidoptera.. Cambridge University Press, Cambridge, pp. 341–367.

Vogt, R.G. and **Lerner, M.R.** (1989) *Two groups of odorant binding proteins in insects suggest specific and general olfactory pathways*. Neurosci. Abstr., 15, 1290.

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

Vogt, R.G., Riddiford, L.M. and **Prestwich, G.D.** (1985) *Kinetic properties of a pheromone degrading enzyme: the sensillar esterase of* Antheraea polyphemus. Proc. Natl Acad. Sci. USA, 82, 8827–8831.

Vogt, R.G., Köhne, A.C., Dubnau, J.T. and **Prestwich, G.D.** (1989) *Expression of pheromone binding proteins during antennal development in the gypsy moth* Lymantria dispar. J. Neurosci., 9, 3332–3346.

Vogt, R.G., Prestwich, G.D. and **Lerner, M.R.** (1991a) *Molecular cloning and sequencing of general-odorant binding proteins GOBP1 and GOBP2 from tobacco hawk moth* Manduca sexta*: comparisons with other insect OBPs and their signal peptides*. J. Neurosci., 11, 2972–2984.

Vogt, R.G., Prestwich, G.D. and **Lerner, M.R.** (1991b) *Odorant-binding-protein subfamilies associate with distinct classes of olfactory receptor neurons in insects.* J. Neurobiol., 22, 74–84.

Vogt, R.G., Rybczynski, R., Cruz, M. and **Lerner, M.R.** (1993) *Ecdysteroid regulation of olfactory protein expression in the developing antenna of the tobacco hawk moth,* Manduca sexta. J. Neurobiol., 22, 581–597.

Vosshall, L., Amrein, H., Morozov, P., Rzhetsky, A. and **Axel, R.** (1999*) A spatial map of olfactory receptor expression in the* Drosophila *antenna.* Cell, 96, 725–736.

Whiting, M.F., Carpenter, J.C., Wheeler, Q.D. and **Wheeler, W.C.** (1997) *The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology*. Systematics Bull., 46, 1–68.

Wojtasek, H., Hansson, B.S. and **Leal, W.S.** (1998) *Attracted or repelled?—A matter of two neurons, one pheromone binding protein, and a chiral center*. Biochem. Biophys. Res. Commun., 250, 217–222.

Young, O.P. (1986) *Host plants of the tarnished plant bug,* Lygus lineolaris *(Heteroptera: Miridae).* Ann. Ent. Soc. Am., 79, 747–762.

Ziegelberger, G. (1996) *The multiple role of the pheromone-binding protein in olfactory transduction. olfaction in mosquito–host Interactions*. Chiba F. Symp., 200, 267–280.

Accepted May 21, 1999