Insect Molecular Biology

Insect Molecular Biology (2014) **23**(1), 78–97 doi: 10.1111/imb.12064

Cloning and expression profiling of odorant-binding proteins in the tarnished plant bug, *Lygus lineolaris*

J. J. Hull[1]*, O. P. Perera[1]† and G. L. Snodgrass†

**USDA-ARS Arid Land Agricultural Center, Maricopa, AZ, USA; and †Southern Insect Management Research Unit, USDA-ARS, Stoneville, MS, USA*

Abstract

In insects, the perception and discrimination of odorants requires the involvement of odorant-binding proteins (OBPs). To gain a better molecular understanding of olfaction in the agronomic pest *Lygus lineolaris* (the tarnished plant bug), we used a transcriptomics-based approach to identify potential OBPs. In total, 33 putative OBP transcripts, including the previously reported *Lygus* antennal protein (LAP), were identified based on the characteristic OBP Cys signature and/or sequence similarity with annotated orthologous sequences. The *L. lineolaris* OBP (LylinOBP) repertoire consists of 20 'classic' OBPs, defined by the spacing of six conserved Cys residues, and 12 'Plus-C' OBPs, defined by the spacing of eight conserved Cys and one conserved Pro residue. Alternative splicing of OBP genes appears to contribute significantly to the multiplicity of LylinOBP sequences. Microarray-based analysis of chemosensory tissues (antennae, legs and proboscis) revealed enrichment of 21 LylinOBP transcripts in antennae, 12 in legs, and 15 in proboscis, suggesting potential roles in olfaction and gustation respectively. PCR-based determination of transcript abundance for a subset of the LylinOBP genes across multiple adult tissues yielded results consistent with the hybridization data.

Keywords: *Lygus lineolaris*, odorant-binding proteins, olfaction, real-time PCR, Plus-C OBP, classic OBP, alternative splice variants, next-generation sequencing.

First published online 14 November 2013.

Correspondence: J. Joe Hull, USDA-ARS Arid Land Agricultural Research Center, 21881 N. Cardon Lane, Maricopa, AZ 85138, USA. Tel.: +1 520 316 6334; fax: +1 520 316 6330; e-mail: joe.hull@ars.usda.gov

[1]These authors contributed equally.

Introduction

The tarnished plant bug, *Lygus lineolaris,* is a piercing-sucking polyphagous pest found throughout agricultural regions in the USA and Canada. They have been reported to feed on >300 plant species, including a number of important crops such as soybean, strawberry, alfalfa and cotton (Young, 1986). Management practices have traditionally relied on broad-spectrum insecticides; however, field resistance to many of the commonly used chemistries has been reported (Snodgrass, 1996; Snodgrass & Scott, 2000, 2002; Snodgrass *et al.*, 2009). The decreased efficacy of these traditional chemical approaches, coupled with the success of transgenic crops against lepidopteran pests, has resulted in the elevation of *Lygus* pest status, in particular in the southeastern USA (http://www.entomology.msstate.edu/resources/croplosses/2011loss.asp). Further compounding its pest status, the development of alternative molecular-based control strategies have been hampered by a lack of knowledge concerning the genetics governing *Lygus* biology and physiology, in particular how this pest species interacts with and responds to its external environment.

Like most insects, the foraging and reproductive behaviors of *L. lineolaris* and its sister species (e.g. *Lygus hesperus*, *Lygus elisus*, *Lygus shulli*, and *Lygus rugulipennis*) are strongly influenced by environmentally defined chemical cues (Blackmer *et al.*, 2004; Innocenzi *et al.*, 2005; Frati *et al.*, 2008) that trigger antenna-derived neuronal responses (Chinta *et al.*, 1994; Ho & Millar, 2002; Innocenzi *et al.*, 2004; Frati *et al.*, 2009; Williams *et al.*, 2010). As the primary sensory organs of olfaction, insect antennae typically contain, depending on the species, several hundred to several thousand sensilla hairs and their associated sensory neurons (Zacharuk, 1980). The antennae of *L. lineolaris* adults are four-segmented appendages that contain ~2000 sensilla hairs corresponding to six sensillar types representing sensilla trichodea,

sensilla chaetica, and sensilla basiconica (Chinta *et al*., 1997). Activation of the neurons housed within these sensilla triggers the stereotypical behavioural responses associated with avoidance, host recognition, oviposition and mating (Martin *et al*., 2011). Odorants gain access to the neurons via cuticular pores that line the various sensilla. Odorant-binding proteins (OBPs) facilitate the movement of these predominantly lipophilic compounds through the aqueous sensillar lymph to specific olfactory receptors (ORs) in the neuronal membrane (Leal, 2005; Zhou, 2010). OBPs have also been implicated in indirect OR activation with the odorant-bound OBP conformation serving as the ligand that interacts with the OR-binding pocket (Zhou, 2010; Sachse & Krieger, 2011). Consequently, OBPs are generally thought to comprise the initial signal filtering mechanism in olfaction. Some OBPs have also been implicated in odorant signal sequestration/termination (Steinbrecht, 1998).

Insect OBPs are a family of relatively small (~150 amino acid; 15–20 kDa) water-soluble proteins characterized by six highly conserved Cys residues (C1-C6) in which the number of amino acids between C2-C3 (three residues) and C5-C6 (eight residues) is invariant. OBPs can be further subgrouped based on the presence or absence of Cys residues: 'classic' OBPs are characterized by six Cys residues; 'dimer' OBPs are characterized by two 'classic' Cys signature motifs; 'Plus-C' OBPs have a highly conserved Pro residue and usually two additional Cys residues; 'Minus-C' OBPs lack two of the six highly conserved Cys; and 'atypical OBPs' have 9–10 Cys residues with an extended carboxyl terminus (Zhou, 2010). Aside from the Cys spacing, OBPs are poorly conserved across species and have no homology with vertebrate OBPs (Zhou, 2010). Consequently, the identification and annotation of putative insect OBPs has relied extensively on the characteristic Cys signature (Zhou *et al*., 2004, 2010, 2008; Li *et al*., 2005; Forêt & Maleszka, 2006; Jordan *et al*., 2008; Gong *et al*., 2009; Liu *et al*., 2009; Xu *et al*., 2009, 2010; Gotzek *et al*., 2011; Gu *et al*., 2011a; Mitaka *et al*., 2011). OBPs have been identified from more than 40 insect species representing 10 different orders with many of the transcripts predominantly expressed in olfactory tissues (Zhou, 2010). Surprisingly, non-olfactory OBP transcripts, which presumably have potential roles outside of olfaction, have also been reported (Gong *et al*., 2009; Pelletier & Leal, 2009, 2011; Vogel *et al*., 2010; Del Campo *et al*., 2011; Sun *et al*., 2012). As elaborated by Leal (2005), a major limitation to structural-based OBP gene annotation is that the term OBP has expanded from proteins specifically involved in olfaction (e.g. lepidopteran pheromone-binding proteins) to include proteins that may have diverse functions (Paesen & Happ, 1995; Crampton *et al*., 1998; Paiva-Silva *et al*., 2002; Contreras *et al*., 2013). Consequently, to assess the role of OBPs in olfaction, *in vitro* binding assays have been used to examine the binding capacity of recombinantly expressed OBPs for specific odorant ligands (Zhou, 2010). *In vivo* studies, however, provide more concrete evidence regarding the physiological significance of putative OBPs. Mutational studies in *Drosophila melanogaster* have implicated an OBP termed LUSH in male detection of the sex pheromone, 11-cis-vaccenyl acetate (Xu *et al*., 2005; Laughlin *et al*., 2008), whereas studies using interspecies hybrids of *Drosophila* showed that OBP57d-emediated *Drosophila sechiella* attraction to and *D. melanogaster* avoidance of odours from a *D. sechiella* host plant (Matsuo *et al*., 2007). Systematic RNA interference-mediated suppression of OBP expression has since demonstrated the relevance of OBPs in mediating the detection and behavioural responses of *D. melanogaster* to a host of ecologically relevant odours (Swarup *et al*., 2011). OBPs have also been linked to the detection of olfactory cues and blood-feeding behaviour in mosquitos (Biessmann *et al*., 2010; Pelletier *et al*., 2010; Sim *et al*., 2012). Further evidence for the biological importance of OBPs in olfaction was demonstrated by the significant enhancement in OR sensitivity in various expression systems after the addition of specific OBPs (Syed *et al*., 2006; Forstner *et al.* 2009).

While the molecular basis of insect olfaction has been extensively examined in holometabolous insects (e.g. lepidopterans and dipterans), elucidation of the molecular components and mechanisms that comprise the hemipteran olfactory system has not been as fully developed. An initial effort using a novel computational algorithm with deposited express sequence tag (EST) sequences from multiple hemipteran species identified a relatively small number of putative OBPs (Xu *et al*., 2009). Those data have since been supplemented with recent genome and EST-based OBP projects in aphids (Zhou *et al*., 2010), the lucerne plant bug (*Adelphocoris lineolatus*) (Gu *et al*., 2011a), the green plant bug (*Apolygus lucorum*) (Gu *et al*., 2011a; Hua *et al*., 2012), and the brown planthopper (*Nilaparvata lugens* Stål) (Noda *et al*., 2008; Xu *et al*., 2009; He *et al*., 2011). Our understanding of the molecular components comprising the *Lygus* olfactory system is even more incomplete with sequence and expression data currently limited to *Lygus* antennal protein (LAP), a 'classic' OBP preferentially expressed in adult male *L. lineolaris* and *L. hesperus* antennae (Dickens *et al*., 1995, 1998; Vogt *et al*., 1999), and the recently identified *Lygus* OR coreceptor (Hull *et al*., 2012). To derive a better molecular understanding of olfaction in *L. lineolaris* (and *Lygus* spp. in general) we sought to generate a full transcriptomic profile of OBP expression based on next-generation sequencing of whole body *L. lineolaris* libraries. Using previously established criteria for OBP

classification and bioinformatics analyses, we identified 33 OBP-like transcripts (including the previously identified LAP). To begin to assess the putative biological functionality of these gene products, we used tissue-specific microarrays and quantitative real-time-PCR (qRT-PCR) to examine their expression profiles.

Results

Identification of putative OBP transcripts

To facilitate the identification of OBP-like transcripts in *L. lineolaris,* we used a whole body *L. lineolaris* transcriptome assembled from multiple short sequence reads derived from a non-normalized whole body *L. lineolaris* cDNA library comprising tissues from both sexes across all developmental stages (accession numbers SRX039411 and SRX041504). In total, 33 putative OBP transcripts were identified based on the presence of the characteristic OBP Cys signature and/or sequence similarity with annotated orthologous sequences. To confirm the validity of the assembled transcripts, each putative OBP was cloned from *L. lineolaris* tissues and sequenced. The resulting Sanger sequences have been deposited with GenBank (accession numbers KF240735-KF240770).

Among the 33 *L. lineolaris* OBP (LylinOBP) transcripts annotated was the previously identified LAP (Dickens *et al*., 1995). To facilitate a more uniform nomenclature, we have re-termed this transcript LylinOBP1. BLAST analyses (Table 1) revealed that the LylinOBP transcripts were similar to putative OBP sequences identified from other hemipteran pests. The highest degree of sequence similarity was seen with putative OBPs from two other plant bug species, *A. lucorum* and *A. lineolatus* (Gu *et al*., 2011a), which accounted for most of the orthologous sequences. Sequence similarity was also observed with two *Rhodnius prolixus* salivary gland OBPs (Ribeiro *et al*., 2004). The high degree of sequence conservation between LylinOBPs and those in other plant bugs is consistent with the interspecies sequence similarity observed in aphids (Zhou *et al*., 2010) and was not unexpected given *L. lineolaris*, *A. lucorum* and *A. lineolatus* have broadly overlapping host ranges.

Visual inspection of the predicted LylinOBP sequences indicated that most (32 of 33) contained a significant portion of the characteristic OBP Cys signature. Based on

Table 1. Results from a tBLASTn analysis of putative *Lygus lineolaris* odorant-binding protein transcripts

Gene	Accession	E Value	Description	% Identity
LylinOBP1	KF240735	1.01E-89	*Lygus lineolaris* antennal protein LAP mRNA (AF091118)	100.0%
LylinOBP2	KF240736	1.24E-95	*Apolygus lucorum odorant binding protein 8 mRNA (JQ675725)*	93.3%
LylinOBP3	KF240737	3.25E-46	*Adelphocoris lineolatus odorant-binding protein 2 mRNA (GQ477023*)	58.6%
LylinOBP4	KF240738	9.88E-80	*Adelphocoris lineolatus odorant-binding protein 2 mRNA (GQ477023*)	90.6%
LylinOBP5	KF240739	1.15E-91	*Adelphocoris lineolatus odorant binding protein 12 mRNA (GQ477033*)	83.2%
LylinOBP6	KF240740	5.71E-45	*Adelphocoris lineolatus odorant-binding protein 1 mRNA (GQ477022*)	57.8%
LylinOBP7	KF240741	2.76E-26	*Adelphocoris lineolatus odorant-binding protein 1 mRNA (GQ477022*)	45.0%
LylinOBP8	KF240742	3.11E-75	*Apolygus lucorum odorant binding protein 7 mRNA (JQ675724)*	82.8%
LylinOBP9	KF240743	2.70E-66	*Adelphocoris lineolatus odorant-binding protein 1 mRNA (GQ477022*)	84.8%
LylinOBP10	KF240744	2.28E-16	*Rhodnius prolixus* odorant-binding protein precursor mRNA (AY340268)	34.3%
LylinOBP11	KF240745	2.29E-49	*Rhodnius prolixus* odorant-binding protein precursor mRNA (AY340268)	55.2%
LylinOBP12	KF240746	2.24E-61	*Adelphocoris lineolatus odorant binding protein 13 mRNA (GQ477034*)	66.0%
LylinOBP13	KF240747	1.47E-10	*Adelphocoris lineolatus odorant-binding protein 3 mRNA (GQ477024*)	27.7%
LylinOBP14	KF240748	4.49E-07	*Adelphocoris lineolatus odorant-binding protein 3 mRNA (GQ477024*)	34.7%
LylinOBP15	KF240749	1.96E-07	*Adelphocoris lineolatus odorant binding protein 11 mRNA (GQ477032*)	28.1%
LylinOBP16	KF240750	4.85E-18	*Apolygus lucorum odorant binding protein 9 mRNA (JQ675726)*	33.8%
LylinOBP17	KF240752	8.51E-04	*Adelphocoris lineolatus odorant-binding protein 3 mRNA (GQ477024*)	25.6%
LylinOBP18	KF240753	8.87E-105	*Apolygus lucorum odorant binding protein 2 (OBP2) mRNA (HQ631398*)	93.1%
LylinOBP19	KF240756	6.31E-67	*Apolygus lucorum odorant binding protein 9 mRNA (JQ675726)*	64.7%
LylinOBP20	KF240757	2.69E-29	*Apolygus lucorum odorant binding protein 6 (OBP6) mRNA (HQ631402*)	46.0%
LylinOBP21	KF240758	1.22E-57	*Apolygus lucorum odorant binding protein 3 (OBP3) mRNA (HQ631399*)	56.6%
LylinOBP22	KF240759	4.24E-18	*Apolygus lucorum odorant binding protein 3 (OBP3) mRNA (HQ631399*)	31.3%
LylinOBP23	KF240760	3.49E-33	*Apolygus lucorum odorant binding protein 3 (OBP3) mRNA (HQ631399*)	32.9%
LylinOBP24	KF240762	1.54E-66	*Apolygus lucorum odorant binding protein 6 (OBP6) mRNA (HQ631402*)	84.8%
LylinOBP25	KF240763	2.27E-72	*Apolygus lucorum odorant binding protein 3 (OBP3) mRNA (HQ631399*)	61.5%
LylinOBP26	KF240764	8.42E-63	*Adelphocoris lineolatus odorant binding protein 13 mRNA (GQ477034*)	66.7%
LylinOBP27	KF240765	6.14E-104	*Apolygus lucorum odorant binding protein 4 (OBP4) mRNA (HQ631400*)	94.3%
LylinOBP28	KF240766	1.95E-104	*Apolygus lucorum odorant binding protein 4 (OBP4) mRNA (HQ631400*)	93.2%
LylinOBP29	KF240767	6.58E-114	*Adelphocoris lineolatus odorant binding protein 7 mRNA (GQ477028*)	91.8%
LylinOBP30	KF240768	7.26E-83	*Apolygus lucorum odorant binding protein 5 (OBP5) mRNA (HQ631401*)	79.5%
LylinOBP31	KF240769	1.31E-15	*Apolygus lucorum odorant binding protein 1 (OBP1) mRNA (HQ631397*)	32.7%
LylinOBP32	KF240770	6.94E-58	*Apolygus lucorum odorant binding protein 3 (OBP3) mRNA (HQ631399*)	56.9%
LylinOBP33	KF240761	3.31E-24	*Apolygus lucorum odorant binding protein 4 (OBP4) mRNA (HQ631400*)	87.9%

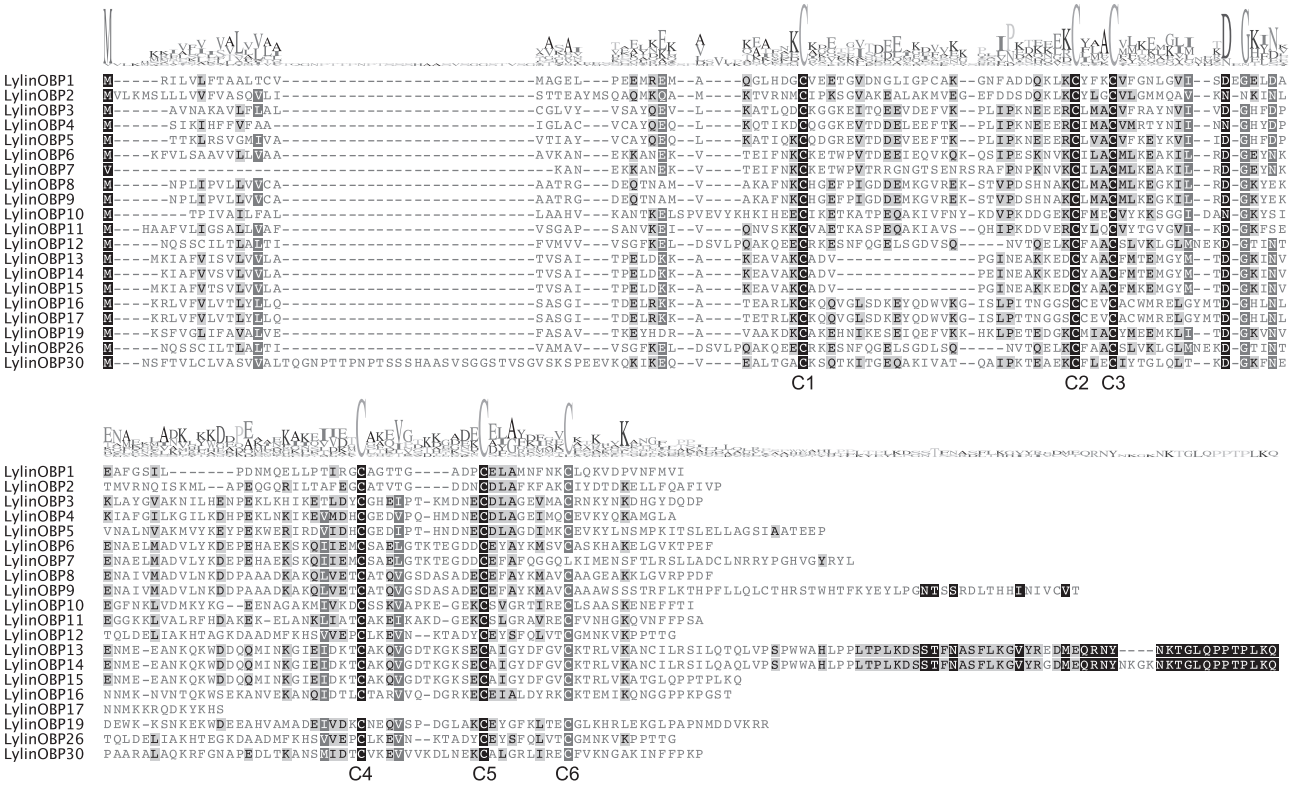


Figure 1. Amino acid alignment of *Lygus lineolaris* 'classic' odorant-binding proteins (OBPs). Sequences were aligned using the L-INS-I strategy in MAFFT (Katoh *et al*., 2005) and rendered in GENEIOUS 6.0.4. The conserved Cys residues (C1-C6) in the 'classic' OBP motif are indicated. Shading represents conservation of sequence identity. In the sequence logo stacks (Crooks *et al*., 2004), the height of each stack corresponds to the degree of sequence conservation at that position.

the hemipteran 'classic' OBP Cys motif (C1-$X_{22–32}$-C2-X_3-C3-$X_{36–46}$-C4-$X_{8–14}$-C5-X_8-C6) (Xu *et al*., 2009), we classified 20 LylinOBP sequences as 'classic' OBPs (Fig. 1). This is somewhat greater diversity than was reported for *A. lineolatus* (14 genes) (Gu *et al*., 2011a), *A. lucorum* (six genes) (Gu *et al*., 2011a), or the pea aphid, *Acyrthosiphon pisum* (13 genes) (Zhou *et al*., 2010). Dipteran species (i.e. drosophilids and mosquitos), in contrast, appear to have 30–40 'classic' OBP genes (Li *et al*., 2005; Vieira *et al*., 2007; Zhou *et al*., 2008; Manoharan *et al*., 2013). Notable exceptions to the 'classic' OBP spacing pattern were observed in LylinOBP7 and LylinOBP17 (Table 2). LylinOBP7 has a leucine at the traditional C6 position; a sixth Cys residue is present, albeit 25 residues downstream of C5. In contrast, the LylinOBP17 transcript appears to be a truncated sequence assembly as it lacks the C4-C6 Cys residues. LylinOBP33 is likewise an incomplete assembly sequence of 339 nucleotides (nt) that lacks both the amino and carboxyl terminal ends. This sequence, however, exhibited significant sequence similarity ($3.31e^{-24}$) with an OBP in *A. lucorum* (AplucOBP4; AEA07662) identified as a 'classic' OBP. We thus grouped LylinOBP33 with the 'classic' OBPs. In addition, we classified 12 of the remaining LylinOBPs as 'Plus-C' (Fig. 2), which are characterized by a Cys spacing pattern that consists of C1-$X_{8–41}$-C2-X_3-C3-$X_{39–47}$-C4-$X_{17–29}$-C4a-X_9-C5-X_8-C6-P-$X_{9–11}$-C6a (Zhou *et al*., 2008). Intriguingly, nine of the sequences had fewer residues (an average of 12 residues) separating C4 and C4a (Table 3) than the reported 17–29 residues. Aphid 'Plus-C' OBPs have 16 residues separating C4 and C4a while 'Plus-C' OBPs from the plant bugs *A. lucorum* and *A. lineolatus* average 13 residues. The 'Plus-C' Cys spacing pattern was derived exclusively from dipteran (*D. melanogaster*, *Drosophila pseudoobscura*, and *Anopheles gambiae*) sequences (Zhou *et al*., 2008), suggesting that the Cys spacing pattern may be order specific. Further exceptions to the 'Plus-C' spacing pattern were also found in LylinOBP20, LylinOBP22, LylinOBP27 and LylinOBP28 (Table 3). The LylinOBP20 sequence appears to be derived from a 5′ truncated transcript that is missing the C1 site. LylinOBP22 lacks both the C6 and C6a sites, but maintains the 'Plus-C' spacing pattern of the other Cys sites. LylinOBP27 and LylinOBP28 deviate more significantly from the typical 'Plus-C' pattern with 45 residues between C1-C2, eight residues between C4a-C5 and no conserved Pro residue after C6, and they are missing C6a (Table 3). The lack of the conserved Pro is not without precedence

Table 2. Cys spacing in *Lygus lineolaris* odorant-binding protein transcripts

	C1–C2	C2–C3	C3–C4	C4–C5	C5–C6
'classic' OBP*	**22–32**	**3**	**36–46**	**8–14**	**8**
LylinOBP1	26	3	36	8	8
LylinOBP2	27	3	40	8	8
LylinOBP3	26	3	42	11	8
LylinOBP4	26	3	42	11	8
LylinOBP5	26	3	42	11	8
LylinOBP6	27	3	42	12	8
LylinOBP7	28	3	42	12	24
LylinOBP8	27	3	42	12	8
LylinOBP9	27	3	42	12	8
LylinOBP10	27	3	41	11	8
LylinOBP11	26	3	41	11	8
LylinOBP12	23	3	44	10	8
LylinOBP13	13	3	41	12	8
LylinOBP14	13	3	41	12	8
LylinOBP15	13	3	41	12	8
LylinOBP16	27	3	43	11	8
LylinOBP17	27	3	*na*†	*na*	*na*
LylinOBP19	27	3	41	11	8
LylinOBP26	23	3	44	10	8
LylinOBP30	26	3	42	12	8

	C1–C2	C2–C3	C3–C4	C4–C4a	C4a–C5	C5–C6	C6–C6a
'Plus-C' OBP‡	**8–41**	**3**	**39–47**	**17–29**	**9**	**8**	**10–12**
LylinOBP18	11	3	42	17	9	8	10
LylinOBP20	*na*	3	41	12	9	8	10
LylinOBP21	32	3	41	13	9	8	10
LylinOBP22	43	3	41	13	9	*na*	*na*
LylinOBP23	43	3	41	13	9	8	10
LylinOBP24	19	3	41	10	9	8	10
LylinOBP25	32	2	41	13	9	8	10
LylinOBP27	45	3	42	21	8	8	*na*
LylinOBP28	45	3	42	21	8	8	*na*
LylinOBP29	23	3	42	14	9	8	10
LylinOBP31	32	3	41	13	9	8	10
LylinOBP32	16	3	41	8	9	8	10

*Xu *et al.*, 2009.
†Not available.
‡Zhou *et al.*, 2008.

as two *Culex quinquefasciatus* 'Plus-C' OBPs (CquiOBP + C5 and CquiOBP + C13) (Pelletier & Leal, 2011) and OBP5 from the aphids *A. pisum* (CAR85632), *Metopolophium dirhodum* (CAR85641), *Sitobion avenae* (CAX63250), *Megoura viciae* (CAR85652), *Nasonovia ribis-nigri* (CAX63258) and *Rhopalosiphum padi* (CAX63254) likewise lack the last two Cys residues. As with the classic OBPs, the diversity of the 'Plus-C' LylinOBPs was significantly larger than that reported for *A. lineolatus* (two genes) (Gu *et al.*, 2011a), *A. lucorum* (four genes)(Gu *et al.*, 2011a), or the pea aphid, *A. pisum* (two genes) (Zhou *et al.*, 2010), but was similar to that reported for *D. melanogaster* (Hekmat-Scafe *et al.*, 2002), *An. gambiae* (Xu *et al.*, 2003; Zhou *et al.*, 2004; Manoharan *et al.*, 2013), *Aedes aegypti* (Zhou *et al.*, 2008; Manoharan *et al.*, 2013) and *C. quinquefasciatus* (Pelletier & Leal, 2011; Manoharan *et al.*, 2013).

In addition to the Cys signature, a distinguishing feature of OBPs is the presence of a signal peptide (Zhou, 2010). Analysis of the predicted LylinOBP amino acid sequences with signal peptide prediction algorithms (Bendtsen *et al.*, 2004; Petersen *et al.*, 2011) suggested that eight of the putative LylinOBP sequences lacked a definable signal peptide sequence (Table 3). This deficiency could be the result of amino terminally truncated transcripts, low sequence coverage or assembly errors that prevented accurate assembly of the complete sequences.

Comparative analysis of the 33 LylinOBPs revealed that sequence conservation across the transcripts ranged from 19.1 to 97.1% (Fig. S1) with a number of the sequences sharing significant stretches of coding sequence, suggesting that they may be splice variants. LylinOBP27 and LylinOBP28 appear to differ from an alternative 5′ splice site as the first 69 nt of the coding sequences

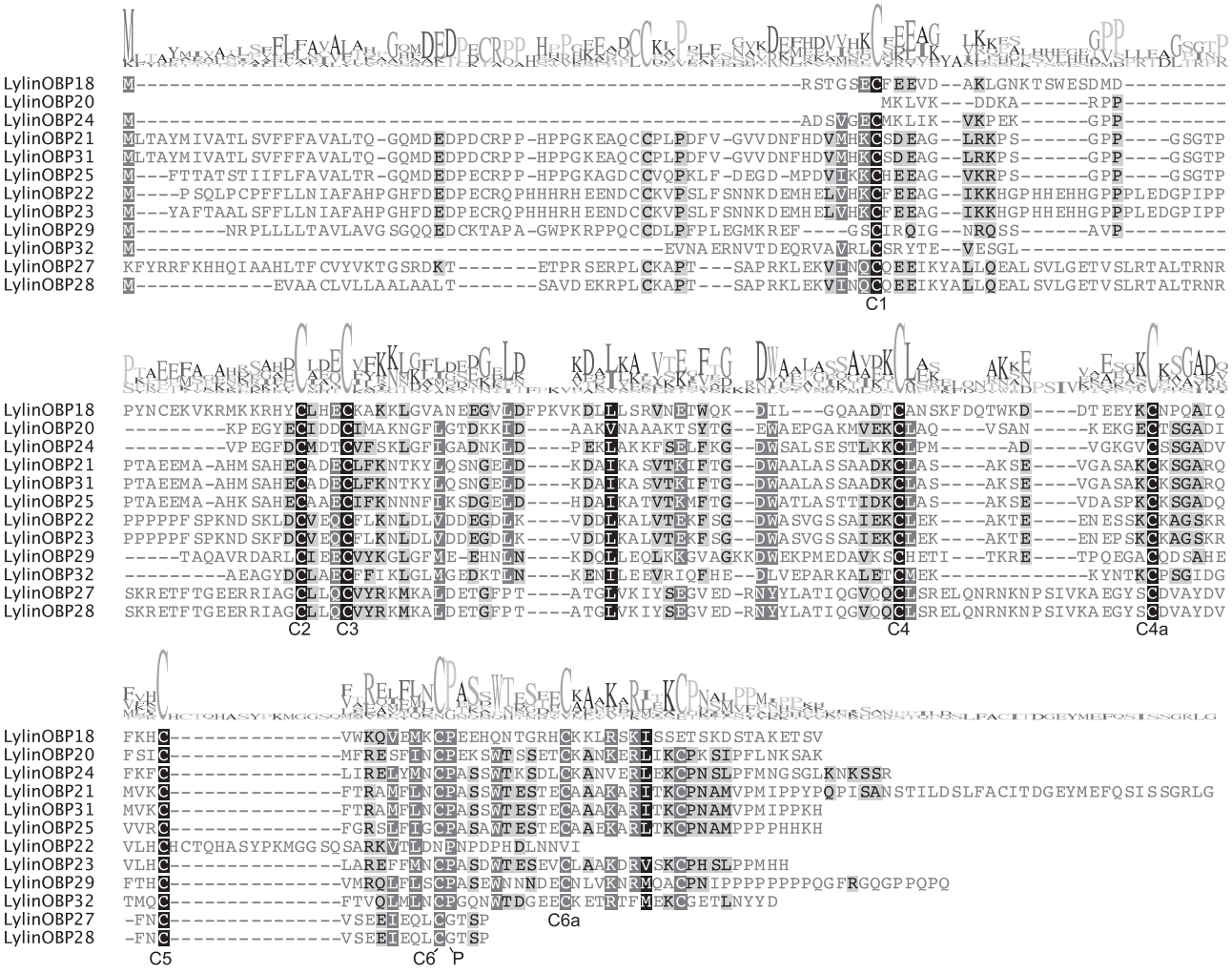


Figure 2. Amino acid alignment of *Lygus lineolaris* 'Plus-C' odorant-bindng proteins (OBPs). Sequences were aligned using the L-INS-I strategy in MAFFT (Katoh *et al*., 2005) and rendered in GENEIOUS 6.0.4. The conserved residues in the 'Plus-C' OBP motif are indicated. Shading represents conservation of sequence identity. In the sequence logo stacks (Crooks *et al*., 2004), the height of each stack corresponds to the degree of sequence conservation at that position.

exhibit a high degree of variability (29% identity), while the terminal 477 nt are essentially identical (Fig. 3). Sequence variations observed in LylinOBP6/7, 8/9, 16/17, 22/23, 12/26 and 21/31 are consistent with differential use of a 3′ splice site (Fig. 3). Variations seen in LylinOBPs13-15 appear to be the result of exon skipping as both LylinOBP13 and LylinOBP15 are deletion variants of LylinOBP14. LylinOPB15 is differentiated from LylinOBP14 by a 171 nt (57 amino acids) deletion between nt 373–545 of LylinOBP14, while LylinOBP13 has a 12 nt (four amino acids) deletion between 532–543 of LylinOBP14. Despite these deletions, both LylinOBP13 and LylinOBP15 retain the same carboxyl coding sequence as LylinOBP14 (Fig. 3). These sequence variations are not the result of inaccuracies in assembly of the short-read sequences as Sanger sequencing of multiple clones confirmed the deletions and retention of the reading frame. While the coding regions of LylinOBP12 and LylinOBP26 are 97.7% identical with four nonsynonymous mutations, the respective 5′ untranslated regions are different, suggesting that these two OBP genes may have arisen from a gene duplication event.

Phylogenetic analysis of LylinOBP sequences

To more accurately assess the relationships amongst the *L. lineolaris* OBP sequences, we performed a neighbour-joining phylogenetic analysis (note: the incomplete LylinOBP33 sequence was not included). The resulting tree (Fig. 4) further highlighted the diversity of the *L. lineolaris* OBP family with the sequences clustering in six central clades, four corresponding to 'classic' OBPs

Table 3. Prediction of signal peptide sequences in *Lygus lineolaris* odorant-binding protein transcripts

Gene	Signal peptide prediction*
LylinOBP1	Cleavage site between pos. 16 and 17: VMA-GE D† = 0.833
LylinOBP2	Cleavage site between pos. 24 and 25: TEA-YM D = 0.746
LylinOBP3	Cleavage site between pos. 21 and 22: VSA-YQ D = 0.869
LylinOBP4	Cleavage site between pos. 20 and 21: VCA-YQ D = 0.780
LylinOBP5	Cleavage site between pos. 21 and 22: VCA-YQ D = 0.636
LylinOBP6	Cleavage site between pos. 19 and 20: VKA-NE D = 0.813
LylinOBP7	no cleavage site
LylinOBP8	Cleavage site between pos. 18 and 19: TRG-DE D = 0.596
LylinOBP9	Cleavage site between pos. 18 and 19: TRG-DE D = 0.596
LylinOBP10	Cleavage site between pos. 18 and 19: VKA-NT D = 0.800
LylinOBP11	Cleavage site between pos. 19 and 20: VSG-AP D = 0.891
LylinOBP12	Cleavage site between pos. 22 and 23: VSG-FK D = 0.837
LylinOBP13	Cleavage site between pos. 19 and 20: VSA-IT D = 0.865
LylinOBP14	Cleavage site between pos. 18 and 19: VSA-IT D = 0.848
LylinOBP15	Cleavage site between pos. 19 and 20: VSA-IT D = 0.825
LylinOBP16	Cleavage site between pos. 19 and 20: ASG-IT D = 0.872
LylinOBP17	Cleavage site between pos. 19 and 20: ASG-IT D = 0.874
LylinOBP18	no cleavage site
LylinOBP19	Cleavage site between pos. 19 and 20: ASA-VT D = 0.805
LylinOBP20	no cleavage site
LylinOBP21	Cleavage site between pos. 23 and 24: TQG-QM D = 0.844
LylinOBP22	no cleavage site
LylinOBP23	Cleavage site between pos. 18 and 19: AFA-HP D = 0.811
LylinOBP24	no cleavage site
LylinOBP25	Cleavage site between pos. 20 and 21: TRG-QM D = 0.767
LylinOBP26	Cleavage site between pos. 22 and 23: VSG-FK D = 0.814
LylinOBP27	no cleavage site
LylinOBP28	Cleavage site between pos. 19 and 20: TSA-VD D = 0.870
LylinOBP29	Cleavage site between pos. 17 and 18: GSG-QQ D = 0.666
LylinOBP30	Cleavage site between pos. 19 and 20: TQG-NP D = 0.812
LylinOBP31	no cleavage site
LylinOBP32	Cleavage site between pos. 23 and 24: TQG-QM D = 0.844
LylinOBP33	no cleavage site

*SignalP 4.0 (http://www.cbs.dtu.dk/services/SignalP/).
†Discrimination score – a weighted average score that discriminates signal peptides from nonsignal peptides.

and two to 'Plus-C' OBPs. As expected, the greatest degree of relatedness was seen with the putative OBP splice variants. LylinOBP1 and LylinOBP2, which clustered together with significant bootstrap support, surprisingly segregated away from the other classic OBPs. This suggests that they underwent a separate gene duplication event and diverged from the ancestral gene much earlier than the others. High bootstrap values supporting the sequence relationship of nonvariant LylinOBP genes (i.e. those presumably not generated through alternative splicing mechanisms) were also seen with the clustering of LylinOBPs 3, 4, and 5, as well as LylinOBP11 with LylinOPB30 and LylinOBP20 with LylinOBP24. Expansion of the neighbour-joining phylogenetic analysis to include 92 additional OBP sequences from multiple aphid species, as well as *N. lugens*, *R. prolixus*, *Euschistus heros (*neotropical brown stinkbug*)* and the two plant bugs, *A. lineolatus* and *A. lucorum*, generated a tree with 'classic' OBP and 'Plus-C' OBP sequences segregating into unique clades (Fig. 5). The consensus tree suggests five central clusters in the 'classic' OBP family and two clusters in the 'Plus-C' OBP family. The 'classic' OBP gene family generated a much more expansive tree with multiple bootstrap supported (1000 iterations) subdivisions suggesting that classic hemipteran OBP genes have undergone extensive gene duplication and divergence. The diversity of the OBP gene family is further demonstrated by the segregated clustering of aphid OBP sequences from the other hemipteran OBP genes, and is consistent with previous reports regarding the lack of overall OBP sequence conservation (Vogt *et al*., 1999; Hekmat-Scafe *et al*., 2002; Xu *et al*., 2003; Zhou *et al*., 2004, 2010; Pelosi *et al*., 2006; Gu *et al*., 2011a). Potential LylinOBP orthologous sequences were largely limited to OBPs derived from other plant bugs. In addition to the previously reported clustering of LylinOBP1 (i.e. LAP) with AdlinOBP4 (ACZ58030) (Gu *et al*., 2011a), related plant bug genes were identified for five 'classic' LylinOBP sequences: LylinOBP2 with AdlinOBP5 (ACZ58031), LylinOBP4 with AdlinOBP2 (ACZ58028), LylinOBP5 with AdlinOBP12 (ACZ58083), LylinOBP19 with AdlinOBP11 (ACZ58082), and LylinOBP30 with AplucOBP5 (AEA07663). Five potentially orthologous 'Plus-C' OBP sequences were identified: LylinOBP18 with AplucOBP2 (AEA07706), LylinOBP24 with AplucOBP6 (AEA07664), LylinOBP28 with AplucOBP4 (AEA07662), LylinOBP31 with Apluc OBP1 (AEA07705), and LylinOBP29 with AdlinOBP7 (ACZ58085). Intriguingly, LylinOBP27 and LylinOBP28, both of which exhibited an atypical 'Plus-C' motif, clustered with a group of aphid OBP genes that also lack features of this motif. LylinOBPs 6, 7, 10, 11, and 13–17 appear to be unique to *Lygus* as they did not segregate with any of the

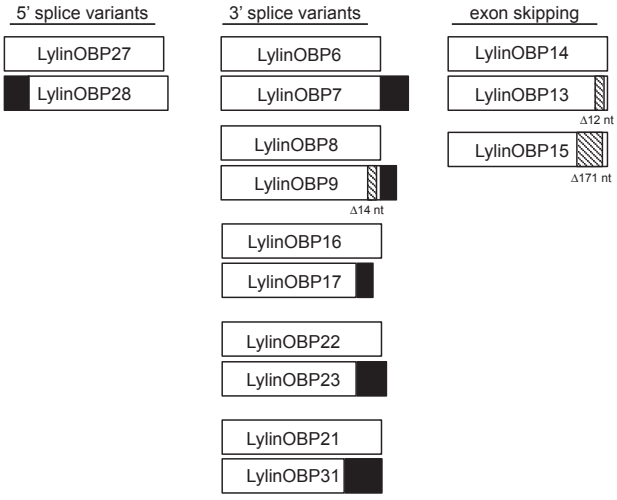


Figure 3. Diagram of LylinOBP splice variants. Figures represent the predicted coding sequence of the transcripts with the putative variants clustered together. Portions of the respective sequences that are identical are depicted by white boxes, dark shading indicates regions of sequence divergence, and hatch-marking indicates deletions. Transcripts are not shown to scale.

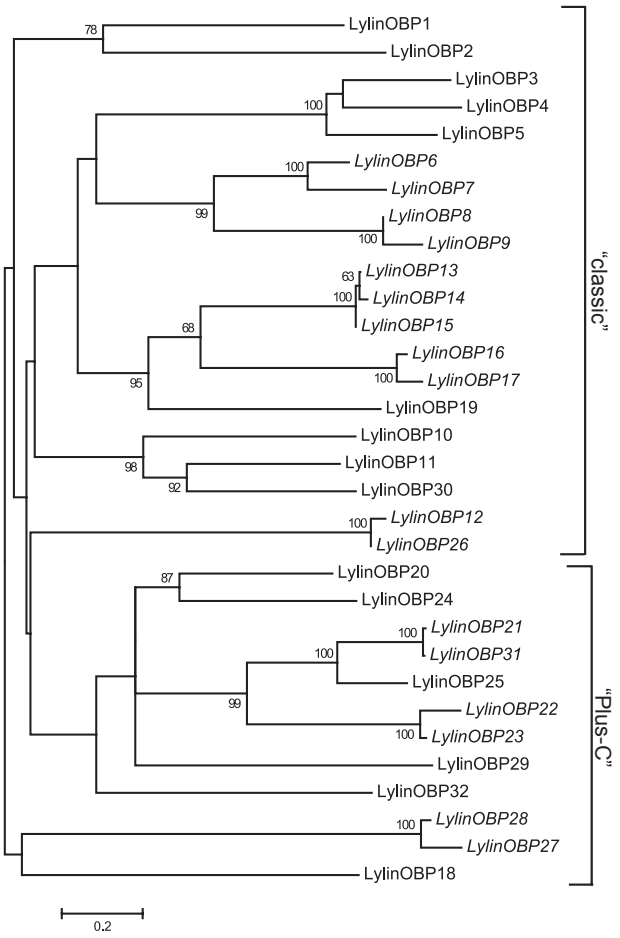


Figure 4. Phylogenetic relationship of LylinOBPs. The unrooted neighbour-joining tree was constructed following alignment with MAFFT and the phylogenetic tree constructed using MEGA5 (Tamura *et al.*, 2011) with bootstrap support based on 1000 iterations (only bootstrap values >60% are shown). Putative splice variants are indicated in italics.

other potentially orthologous hemipteran sequences currently available in the databases.

Microarray-based detection of LylinOBP expression

To further characterize the LylinOBPs, we generated two custom microarrays to examine the relative expression level of the transcripts in tissues frequently associated with chemosensory perception: antennae, legs and proboscis. The arrays were generated using sequence information derived from the assembled short sequence reads with multiple sequences printed for the putative transcripts such that most of the OBPs were represented multiple times on each slide. Notable exceptions were the LylinOBP31 and LylinOBP32 sequences, which were assembled independently of the other Lylin sequences after the arrays had been prepared. To provide insight into potential sex-biased expression, hybridizations were made using antisense RNA (aRNA) prepared from male and female tissues. To assess potential dye affects, equal amounts of Alexa Fluor-555 and Alexa Fluor-647-labelled aRNA were hybridized.

Signals for most of the LylinOBP transcripts were detected in the chemosensory appendages with transcript abundance generally highest in antennae (Table 4). A significant amount of heterogeneity in transcript distribution, however, was observed amongst the three tissues, probably reflecting the diverse behaviours associated with those tissues. The expression of 21 LylinOBP transcripts was >5-fold higher in antennae relative to the control tissue (female body). Similarly elevated LylinOBP transcript expression was also seen with 12 genes in legs and 15 genes in proboscis. The most abundant transcripts in antennae (defined as >100-fold higher expression level relative to control) were LylinOBP1, LylinOBP2 (male), LylinOBP4, LylinOBP5, LylinOBP8, LylinOBP9, LylinOBP12 and LylinOBP26 (male). The most abundant transcripts in legs (defined as >25-fold higher expression level relative to control) were LylinOBP8 (female), LylinOBP9 (female), LylinOBP10, LylinOBP11 and LylinOBP19. In proboscis they were LylinOBP8 (male), LylinOBP9 (male), LylinOBP15, LylinOBP18, LylinOBP19, LylinOBP22, LylinOBP23 (female), LylinOBP25, and LylinOBP26. Surprisingly, the degree of overlap in the expression profile of the high abundance transcripts was limited with LylinOBP8 and LylinOBP9 highly expressed in all three tissues of one or both insect genders. The overlap of the high abundance transcripts between legs and proboscis, which are more often associated with contact/gustatory perception, included LylinOBP8, LylinOBP9 and LylinOBP19, which was the most abundant transcript in both tissues.

Statistically significant sex-biased differences, which we have defined as a >2-fold difference in expression with 95% confidence, were observed with 12 of the LylinOBP transcripts expressed at higher levels in male antennae (Table 4). Sex-biased transcripts were likewise detected in the proboscis and legs; however, the number of genes affected was limited: LylinOBP10 in legs and LylinOBP4 and LylinOBP6 in the proboscis and antenna. In antennae, the bias was towards male dominant expression, whereas in legs and the proboscis higher expression was observed in females (Table 4).

Quantitative real-time PCR-based determination of LylinOBP expression

To validate the microarray data, we used qRT-PCR to assess the transcript abundance of nine LylinOBP genes (LylinOBP1, LylinOBP2, LylinOBP5, LylinOBP8, LylinOBP15, LylinOBP19, LylinOBP24, LylinOBP26 and

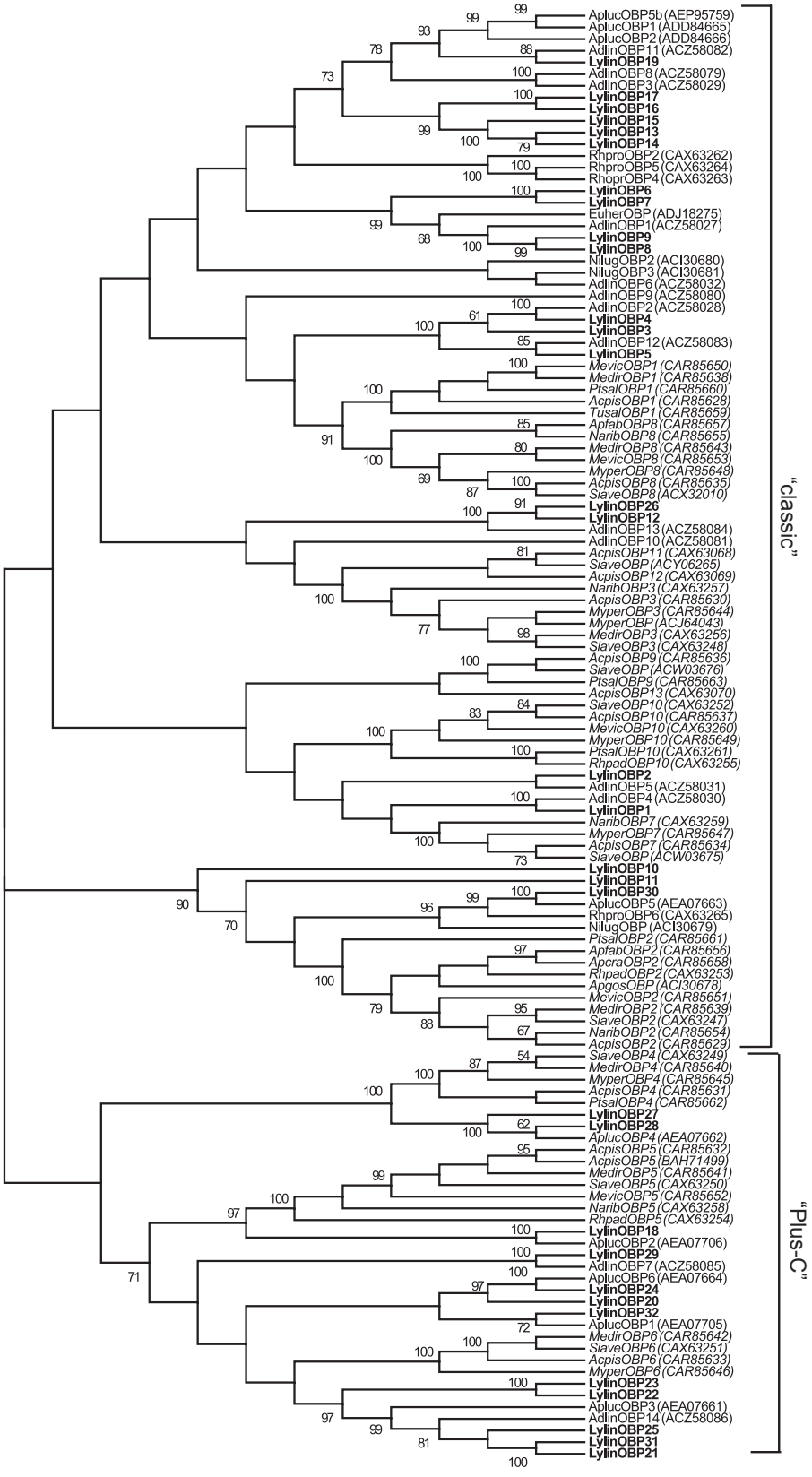

AplucOBP5b (AEP95759)
AplucOBP1 (ADD84665)
AplucOBP2 (ADD84666)
AdlinOBP11 (ACZ58082)
LylinOBP19
AdlinOBP8 (ACZ58079)
AdlinOBP3 (ACZ58029)
LylinOBP17
LylinOBP16
LylinOBP15
LylinOBP13
LylinOBP14
RhproOBP2 (CAX63262)
RhproOBP5 (CAX63264)
RhoprOBP4 (CAX63263)
LylinOBP6
LylinOBP7
EuherOBP (ADJ18275)
AdlinOBP1 (ACZ58027)
LylinOBP9
LylinOBP8
NilugOBP2 (ACI30680)
NilugOBP3 (ACI30681)
AdlinOBP6 (ACZ58032)
AdlinOBP9 (ACZ58080)
AdlinOBP2 (ACZ58028)
LylinOBP4
LylinOBP3
AdlinOBP12 (ACZ58083)
LylinOBP5
MevicOBP1 (CAR85650)
MedirOBP1 (CAR85638)
PtsalOBP1 (CAR85660)
AcpisOBP1 (CAR85628)
TusalOBP1 (CAR85659)
ApfabOBP8 (CAR85657)
NaribOBP8 (CAR85655)
MedirOBP8 (CAR85643)
MevicOBP8 (CAR85653)
MyperOBP8 (CAR85648)
AcpisOBP8 (CAR85635)
SiaveOBP8 (ACX32010)
LylinOBP26
LylinOBP12
AdlinOBP13 (ACZ58084)
AdlinOBP10 (ACZ58081)
AcpisOBP11 (CAX63068)
SiaveOBP (ACY06265)
AcpisOBP12 (CAX63069)
NaribOBP3 (CAX63257)
AcpisOBP3 (CAR85630)
MyperOBP3 (CAR85644)
MyperOBP (ACJ64043)
MedirOBP3 (CAX63256)
SiaveOBP3 (CAX63248)
AcpisOBP9 (CAR85636)
SiaveOBP (ACW03676)
PtsalOBP9 (CAR85663)
AcpisOBP13 (CAX63070)
SiaveOBP10 (CAX63252)
AcpisOBP10 (CAR85637)
MevicOBP10 (CAX63260)
MyperOBP10 (CAR85649)
PtsalOBP10 (CAX63261)
RhpadOBP10 (CAX63255)
LylinOBP2
AdlinOBP5 (ACZ58031)
AdlinOBP4 (ACZ58030)
LylinOBP1
NaribOBP7 (CAX63259)
MyperOBP7 (CAR85647)
AcpisOBP7 (CAR85634)
SiaveOBP (ACW03675)
LylinOBP10
LylinOBP11
LylinOBP30
AplucOBP5 (AEA07663)
RhproOBP6 (CAX63265)
NilugOBP (ACI30679)
PtsalOBP2 (CAR85661)
ApfabOBP2 (CAR85656)
ApcraOBP2 (CAR85658)
RhpadOBP2 (CAX63253)
ApgosOBP (ACI30678)
MevicOBP2 (CAR85651)
MedirOBP2 (CAR85639)
SiaveOBP2 (CAX63247)
NaribOBP2 (CAR85654)
AcpisOBP2 (CAR85629)
SiaveOBP4 (CAX63249)
MedirOBP4 (CAR85640)
MyperOBP4 (CAR85645)
AcpisOBP4 (CAR85631)
PtsalOBP4 (CAR85662)
LylinOBP27
LylinOBP28
AplucOBP4 (AEA07662)
AcpisOBP5 (CAR85632)
AcpisOBP5 (BAH71499)
MedirOBP5 (CAR85641)
SiaveOBP5 (CAX63250)
MevicOBP5 (CAR85652)
NaribOBP5 (CAX63258)
RhpadOBP5 (CAX63254)
LylinOBP18
AplucOBP2 (AEA07706)
LylinOBP29
AdlinOBP7 (ACZ58085)
AplucOBP6 (AEA07664)
LylinOBP24
LylinOBP20
LylinOBP32
AplucOBP1 (AEA07705)
MedirOBP6 (CAR85642)
SiaveOBP6 (CAX63251)
AcpisOBP6 (CAR85633)
MyperOBP6 (CAR85646)
LylinOBP23
LylinOBP22
AplucOBP3 (AEA07661)
AdlinOBP14 (ACZ58086)
LylinOBP25
LylinOBP31
LylinOBP21
"classic"
"Plus-C"

Table 4. Fold increase in LylinOBP transcript levels in various chemosensory tissues compared with control tissue*

	Antennae				Legs				Proboscis			
Gene	Female	SD (+/–)†	Male	SD (+/–)	Female	SD (+/–)	Male	SD (+/–)	Female	SD (+/–)	Male	SD (+/–)
LylinOBP1	105.72	17.06	**198.76**§	14.87	1.68	9.24	1.27	2.14	19.65	5.88	20.19	6.36
LylinOBP2	82.58	14.21	110.42	23.73	12.08	4.63	2.47	39.38	15.64	4.92	16.51	6.54
LylinOBP3	22.45	7.78	**52.94**§	3.75	3.22	3.99	3.22	9.21	4.72	4.10	5.62	9.23
LylinOBP4	226.07	15.94	224.22	14.60	0.00	0.00	0.00	0.00	**19.93****	3.72	1.28	0.00
LylinOBP5	248.98	19.30	**488.93**§	46.50	2.62	9.71	0.41	0.00	7.31	27.58	3.25	1.46
LylinOBP6	42.95	4.29	66.30	13.04	2.40	3.64	1.03	2.23	**3.09****	0.40	4.53	0.00
LylinOBP7	1.21	2.58	4.25	2.94	1.36	7.51	1.02	2.12	1.32	1.84	1.45	1.23
LylinOBP8	480.42	61.93	**775.02**§	54.21	26.95	5.89	20.11	6.18	22.09	7.43	30.27	6.88
LylinOBP9	179.38	17.80	**407.16**§	22.57	28.28	7.98	20.22	4.06	22.21	7.17	28.19	9.13
LylinOBP10	70.22	6.16	57.88	5.75	**60.70**¶	8.63	27.17	4.57	10.56	4.57	12.95	2.47
LylinOBP11	12.10	1.57	14.19	0.46	40.57	8.40	33.44	7.43	8.59	5.24	5.86	5.71
LylinOBP12	107.43	14.86	**280.33**§	15.82	1.72	5.68	1.30	10.00	3.48	2.53	2.94	3.51
LylinOBP13	2.88	3.36	7.76	91.77	1.63	5.96	1.24	11.81	21.42	4.87	12.54	3.60
LylinOBP14	4.26	5.98	8.02	8.96	2.79	8.92	2.82	1.09	24.83	4.17	23.99	4.24
LylinOBP15	6.78	10.77	12.11	565.53	4.48	1670.00	5.69	5.44	49.74	57.44	79.82	81.98
LylinOBP16	0.18	0.22	0.36	0.82	0.11	10.26	0.07	0.01	0.80	0.50	0.45	0.48
LylinOBP17	0.52	1.36	0.60	0.76	0.39	2.57	0.60	0.04	1.08	0.36	0.61	0.83
LylinOBP18	49.01	6.81	42.07	4.13	11.87	2.41	4.82	2.04	58.74	15.87	51.63	13.86
LylinOBP19	60.42	6.23	87.40	8.20	82.84	17.69	46.12	16.94	258.55	21.44	291.05	26.55
LylinOBP20	5.91	6.24	**71.82**§	8.43	2.66	7.63	1.40	2.02	14.13	12.32	3.53	5.57
LylinOBP21	1.77	2.66	**30.04**§	6.06	0.00	0.00	0.00	0.00	6.99	21.35	0.00	0.00
LylinOBP22	20.52	5.13	27.30	1.70	19.71	8.86	11.84	5.85	32.02	6.24	30.08	5.96
LylinOBP23	19.70	2.18	25.33	2.04	14.98	3.06	10.26	3.14	25.78	5.63	24.49	2.70
LylinOBP24	10.19	3.31	**30.37**§	5.23	18.19	4.38	15.06	9.59	12.29	6.06	5.61	0.00
LylinOBP25	17.55	2.07	**41.69**§	5.87	10.55	7.00	8.43	3.95	55.60	6.61	55.00	8.08
LylinOBP26	76.90	6.98	**140.51**§	11.25	11.73	3.48	4.73	11.83	32.13	4.12	25.29	0.00
LylinOBP27	4.23	4.49	6.55	3.83	1.80	9.30	1.00	8.40	3.30	2.25	2.06	3.96
LylinOBP28	4.25	4.81	6.49	3.69	1.67	8.85	1.02	9.31	3.21	2.50	2.10	4.28
LylinOBP29	20.56	1.83	**41.95**§	4.44	2.78	66.73	2.67	2.14	6.80	1.45	5.89	1.01
LylinOBP30	1.10	0.93	0.47	1.06	0.03	0.06	0.01	0.01	0.16	0.13	0.02	0.02
LylinOBP31	*nd*‡	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*
LylinOBP32	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*
LylinOBP33	2.35	3.43	6.71	3.09	1.68	9.24	1.27	2.14	2.62	2.03	2.79	1.68

*Expression relative to control tissue (female gut) based on microarray hybridization experiments (see Materials and Methods).
†Standard deviation.
‡Not determined.
§Statistically significant difference in expression compared with female antennae using an unpaired t-test for each odorant-binding protein with significance defined as $P < 0.05$.
¶Statistically significant difference in expression compared with male leg; significance determined as previously.
**Statistically significant difference in expression compared with male proboscis significance determined as previously.

LylinOBP29) that showed varying degrees of expression in antennae, proboscis and legs. Transcript abundance for each OBP was determined for multiple tissues (antenna, proboscis, leg, head, body, gut and fat body) relative to standard curves prepared from serially diluted plasmid DNAs containing a portion of the LylinOBP sequence of interest. Consistent with the microarray data, we found that transcript abundance was heterogeneous amongst the tissues we examined (Fig. 6); however, there was clear preferential expression ($P < 0.001$) of LylinOBP1, LylinOBP2, LylinOBP5, LylinOBP8 and LylinOBP26 in antennae, suggesting a potential role in the detection and discrimination of odorants. Furthermore, our findings that LylinOBP1 transcripts were highly expressed in antennae with a male-dominant expression profile ($P < 0.001$) were consistent with previous studies (Dickens *et al.*, 1995,

Figure 5. Phylogenetic relationship of 124 odorant-bindng proteins (OBPs) from various hemipteran species. Sequences were aligned using MAFFT and the phylogenetic tree constructed using MEGA5 (Tamura *et al.*, 2011) with bootstrap support based on 1000 iterations (only bootstrap values >60% are shown). The GenBank accession numbers are shown in parentheses. LylinOBP sequences are shown in bold. Aphid sequences are shown in italics. Species abbreviations: Acpis – *Acyrthosiphon pisum* (pea aphid); Adlin – *Adelphocoris lineolatus* (alfalfa plant bug); Apcra – *Aphis craccivora* (cowpea aphid); Apfab – *Aphis fabae* (black bean aphid); Apgos – *Aphis gossypii* (cotton aphid); Apluc – *Apolygus lucorum* (green plant bug); Euher – *Euschistus heros* (neotropical brown stinkbug); Lylin – *Lygus lineolaris* (tarnished plant bug); Medir – *Metopolophium dirhodum* (rose-grain aphid); Mevic – *Megoura viciae* (vetch aphid); Myper – *Myzus persicae* (green peach aphid); Narib – *Nasonovia ribis-nigri* (lettuce aphid); Nilug – *Nilaparvata lugens* (brown planthopper); Ptsal – *Pterocomma salicis* (black willow aphid); Rhpad – *Rhopalosiphum padi* (bird cherry-oat aphid); Rhpro- *Rhodnius prolixus* (Triatomid bug); Siave – *Sitobion avenae* (English grain aphid); Tusal – *Tuberolachnus salignus* (giant willow aphid).

Figure 6. Quantitative real-time PCR-based analysis of select LylinOBP transcript levels across various adult tissues. Transcript levels were determined relative to standard curves generated using serially diluted plasmid DNAs containing a portion of the gene of interest. Tissues examined: antenna (A), proboscis (P), leg (L), head (H), body (B), gut (G) and fat body (FB). Expression of the control gene, *Lygus lineolaris* β-tubulin (DQ471301), was included as an indicator of cDNA quality. A no template control (NTC) was included as a negative control. Female-derived tissues are shown in light grey bars, male-derived tissues are shown in dark grey bars. Data represent mean ± SEM of technical triplicates from two biological replicates. Statistically significant differences in transcript abundance were determined using ANOVA with Tukey's multiple comparison test. Grey asterisks indicate differences between female tissues ($P < 0.001$ and $P < 0.05$), black asterisks indicate differences between male tissues ($P < 0.001$), and the underlined asterisk indicates differences between males and females ($P < 0.001$).

1998; Vogt *et al*., 1999). Similar to other reports regarding the ubiquity of 'Plus-C' OBP expression (Zhou *et al*., 2004; Biessmann *et al*., 2005; Li *et al*., 2005; Pelletier & Leal, 2011), we found that the 'Plus-C' OBPs examined (i.e. LylinOBP24 and LylinOBP29) were expressed in most tissues. The distribution of LyinOBP8 (antenna > leg > proboscis) and LylinOBP19 (proboscis > leg > antenna) amongst the chemosensory tissues was consistent with that observed in the microarray experiments. While not as pronounced as that seen in the hybridization experiments, our qRT-PCR data also demonstrated a certain degree of sex-biased OBP expression in the chemosensory tissues (Fig. 6), in particular in male antennae for LylinOBP1 and LylinOBP2 ($P < 0.001$). Surprisingly, greater sex bias was seen for some of the transcripts in non-olfactory tissues such as LylinOBP2 in male bodies and fat body and LylinOBP24 in male bodies and female heads ($P < 0.001$).

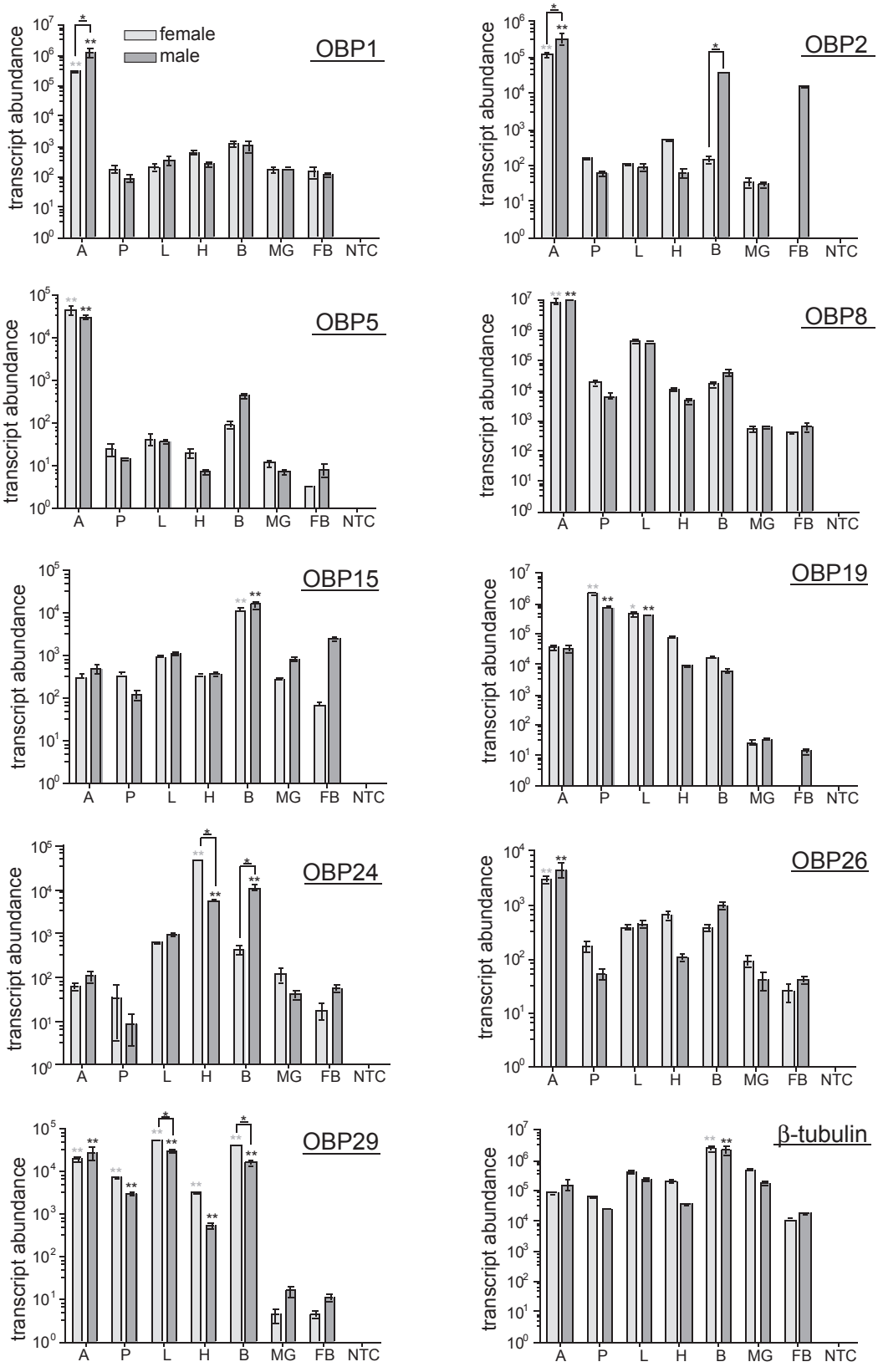
The most significant deviation from the microarray data was the expression profile of LylinOBP15. In the array hybridization experiments, LylinOBP15 appeared to be up-regulated in the proboscis relative to other tissues (Table 4), but in the qRT-PCR experiments LylinOBP15 expression was fairly uniform amongst the chemosensory tissues (Fig. 6). This discrepancy is probably indicative of the position of the nt sequences used in the hybridization experiments. As mentioned previously, LylinOBP13-15 appear to be splice variants of the same gene with LylinOBP15 differentiated from LylinOBP13 and LylinOBP14 by deletions in the coding sequence (Fig. 3). Consequently, hybridization may have involved portions of all three genes, an effect that may account for the high standard errors seen in the microarray data. The qRT-PCR primers were designed to amplify across the putative splice site in LylinOBP15 such that only that variant was amplified. As a result, we saw much greater uniformity of transcript abundance amongst the chemosensory tissues with highest expression in the body (Fig. 6).

As a measure of cDNA template integrity, we also examined the expression profile of β-tubulin. While variation was seen in the tissues examined with highest expression in body ($P < 0.001$), the overall profile differed from that seen with the OBP transcripts. This suggested that that the observed LylinOBP transcript expression profiles were related to the specific enrichment of those transcripts in various tissues and were not PCR artifacts generated from unequal template amounts.

Discussion

As the first step in elucidating the molecular basis of olfaction in *L. lineolaris,* we used a transcriptomic approach to identify genes encoding proteins exhibiting OBP-like features. Based on the characteristic OBP Cys signature and sequence similarities with other OBPs, we identified 33 OBP-like transcripts, 32 of which represent new *L. lineolaris* sequences. While this degree of OBP heterogeneity is less than the 50–60 genes seen in many dipteran species (Hekmat-Scafe *et al*., 2002; Xu *et al*., 2003; Zhou *et al*., 2008; Pelletier & Leal, 2009), it is significantly higher than that previously reported for any hemipteran species: 15 genes in the pea aphid (Zhou *et al*., 2010), five genes in the human body louse (Kirkness *et al*., 2010), 14 genes in the lucerne plant bug (Gu *et al*., 2011a) and six genes in the green plant bug (Gu *et al*., 2011a). Zhou *et al.* (2010) suggested that the relatively small number of OBPs in the pea aphid could potentially be attributable to its parasitic lifestyle and specialized ecology. This, however, is probably not the case for *A. lucorum* and *A. lineolatus*, both of which have broad host ranges that overlap with that of *L. lineolaris*. We speculate that our use of a high throughput sequencing approach using whole insects across multiple life stages accounts for the discrepancy in the number of identified plant bug OBP genes. The AdlinOBP sequences, in contrast, were identified from ~2900 ESTs derived from an antennal cDNA library (Gu *et al*., 2011a). While that approach has the advantage of identifying transcripts specifically expressed in the primary olfactory tissue, the number of ESTs that can be screened for OBP characteristics is orders of magnitude less than that which can be achieved using next-generation sequencing methods. As a consequence, the probability of identifying rare or low abundance transcripts with the traditional EST method is low.

One of the most critical mechanisms underlying diversity at the protein level is alternative splicing, a molecular process that makes use of different exon splice sites in a gene to generate multiple transcripts encoding varying protein isoforms (Licatalosi & Darnell, 2010; Nilsen


female
male
OBP1
OBP2
OBP5
OBP8
OBP15
OBP19
OBP24
OBP26
OBP29
β-tubulin
transcript abundance
A P L H B MG FB NTC

& Graveley, 2010; Sánchez-Pla *et al*., 2012). It has been estimated that nearly one-third of all genes in *D. melanogaster* undergo some form of alternative splicing (Daines *et al*., 2011). Because many of these splice variants are frequently low abundance transcripts they have been poorly represented in traditional EST libraries. The advent of high throughput transcriptomic sequencing has overcome this deficiency and it has led to the identification of previously unidentified splice variants (Graveley *et al*., 2011; Sze *et al*., 2012; Venables *et al*. 2012). In further support of this, a number of the OBPs identified in our transcriptomic-based screen appear to be splice variants (Fig. 3). OBP splice variants have previously been described (Lacazette *et al*., 2000; Zhou *et al*., 2004). In *An. gambiae*, AgamOBPjj5A and AgamOBP5479 have the same signal peptide and 5′ coding sequence through residue 135 at which point the sequences diverge, suggestive of differential 3′ splicing (Zhou *et al*., 2004). We observed similar sequence divergence with LylinOBP6/7, LylinOBP8/9, LylinOBP13/15, LylinOBP16/17, LyliOBP22/23, as well as LylinOBP21 and LylinOBP31. While the specific biochemical roles of these variants remain to be determined, accumulating evidence has shown that alternative splicing profoundly affects biological complexity in terms of tissue- and stage-dependent expression and functionality (Hartmann & Valcárcel, 2009; Licatalosi & Darnell, 2010; Nilsen & Graveley, 2010). It is clear from structural studies that relatively small variations in the lengths of 'classic' OBP carboxyl terminal segments can significantly impact the conformational character of the ligand-binding pocket (Tegoni *et al*., 2004). Indeed, AgamOBP47, a 'Plus-C' OBP in *An. gambiae* that is distinguished from its ancestral 'classic' OBP precursor by disulphide-stabilized amino and carboxyl terminal extensions, lacks the buried internal cavity intrinsic to all known OBP structures (Lagarde *et al*., 2011). This difference effectively alters the physiochemical nature of the substrates bound by the 'Plus-C' OBP compared with the ancestral gene.

While the degree of sequence conservation amongst OBPs can vary widely even within species, a certain degree of sequence segregation is apparent in species within the same order and family (Pelosi *et al*., 2006). Consistent with this, we found that a number of LylinOBP sequences segregated phylogenetically with known *A. lucorum* and *A. lineolatus* OBPs, both of which belong to the same family (Miridae) as *L. lineolaris*. Although different methods were used [microarray hybridization and absolute qRT-PCR in the present study compared with relative qRT-PCR by Gu *et al*. (2011a)], potentially orthologous transcripts exhibited similar tissue expression profiles, suggesting that these OBPs may mediate similar functions in the respective species. LylinOBP1/AdlinOBP4, LylinOBP2/AdlinOBP5, LylinOBP4/AdlinOBP2, LylinOBP5/AdlinOBP12, LylinOBP8-9/AdlinOBP1, and LylinOBP12/AdlinOBP13 are predominantly expressed in antennae, LylinOBP19/AdlinOPB11 are highly expressed in legs, and LylinOBP29/AdlinOBP7 are relatively ubiquitous. The specific head expression of AdlinOBP14 is consistent with our finding that LylinOBP25 is highly expressed in proboscis. Intriguingly, AdlinOBP13, which clusters phylogenetically with the splice variants LylinOBP12 and LylinOBP26, was female-dominant while the *L. lineolaris* sequences were male dominant (Table 4).

In our expression profiling (Table 4), we found that transcripts for a number of the putative 'Plus-C' OBPs were enriched in antennae. While 'Plus-C' OBPs generally do not predominate in olfactory tissues (Zhou *et al*., 2004; Biessmann *et al*., 2005; Li *et al*., 2005; Pelletier & Leal, 2011), antennal enrichment is not without precedent. In *C. quinquefasciatus*, Culex OBP + C1 is the fourth most abundant transcript in antennae (Pelletier & Leal, 2011) and its putative orthologue in *An. gambiae*, AgamOBP48, is the second most abundant transcript in female antennae (Biessmann *et al*., 2005). These high levels of expression suggest a potential biological function in olfaction. Andronopoulou *et al*. (2006) demonstrated that AgamOBP48 can heterodimerize with OBPs that are congruently expressed in the antennae. This expands the number of potential substrates recognizable by the suite of antennal OBPs. Evidence supporting the combinatorial nature of olfaction has recently been reported in *D. melanogaster* (Swarup *et al*., 2011). We speculate that LylinOBPs may fulfill a similar function; however, this remains to be experimentally determined.

Basic clues regarding the potential biological significance of OBPs can be derived from their transcript expression profile. It is clear that high antennal expression is correlated with a role in olfaction. Enrichment in the proboscis (i.e. LylinOBPs 8, 9, 15, 19, 22, 23, 25 and 26), an appendage frequently associated with taste perception, could be indicative of a role in gustation, as OBP expression in proboscis/maxillary palp sensilla has been associated with gustatory responses (Shanbhag *et al*., 2001; Del Campo *et al*., 2011). Proboscis expression, however, could also be associated with olfaction as multiple ORs and their associated neurons have been identified in the mosquito proboscis (Kwon *et al*., 2006). Similar to the gustatory aspects of the proboscis, OBP transcript expression in legs may be indicative of gustatory function. Females frequently use taste sensilla distributed on legs to determine the suitability of host plants prior to egg laying. Consequently, elevated OBP expression in legs could be indicative of a role in oviposition, as was shown for the drosophilid OBPs Obp57d and Obp57e (Matsuo *et al*., 2007). Unequivocal determination of the physiological functions mediated by LylinOBPs will require

a clear demonstration of ligand binding within a biological context. Binding specificity studies have been performed using two recombinant AdlinOBPs, AdlinOBP1 and AdlinOBP3 (Gu *et al*., 2010, 2011b). AdlinOBP1 tightly bound the putative attractants ethyl butyrate and trans-2-hexenyl butyrate as well as the cotton-derived volatiles myrcene, β-ionone, and β-caryophyllene (Gu *et al*., 2011b). In our phylogenetic analysis, AdlinOBP1 clustered with LylinOBP8 and LylinOBP9. Given the sequence similarities it is possible that the LylinOBPs may exhibit similar binding specificities. Furthermore, trans-2-hexenyl butyrate has been shown to elicit a strong antennal electrophysiological response in male *L. lineolaris* (Chinta *et al*., 1994). AdlinOBP3, in contrast, was found to only weakly bind α-phellandrene, a cotton volatile (Gu *et al*., 2010). That study was constrained by a limited number of compounds assayed, which probably contributed to the poor determination of a *bona fide* AdlinOBP3 substrate. Similar studies have been conducted using recombinant AplucOBPs (Hua *et al*., 2012). Intriguingly, the OBPs assayed appeared to exhibit higher affinities for a number of secondary cotton metabolites than for plant volatiles, suggesting a potential role in gustation.

Given the critical role that olfaction has in governing a number of essential behaviours, it is clear that targeted disruption of key components of the insect olfactory system has enormous potential for the development of novel insect strategies. In order to fully develop these new approaches, however, it is critical to understand the molecular basis underlying odorant perception and discrimination. Using a transcriptomics-based approach, we have expanded our understanding of the *Lygus* olfactory system by identifying 32 new OBPs and have highlighted the potential role alternative splicing may play in generating OBP sequence diversity (a mechanism that potentially provides the means of fine-tuning the odorant signal detection system). Based on the results of the transcript expression profile, we will focus future studies on characterizing the role that preferentially expressed OBP genes have in antennal and proboscis-based olfaction as well as what role OBPs may have in gustatory-driven behaviours.

Experimental procedures

Insect rearing

Lygus lineolaris were obtained from a laboratory colony (USDA-ARS Southern Insect Management Research Unit, Stoneville, MS, USA) maintained at 25 °C under 20% humidity and a 14 h light: 10 h dark photoperiod. Insects were reared on green beans and an artificial diet mix in disposable packs as described by Debolt (1982) and Patana (1982). At varying points, the colony has been periodically outbred with locally caught conspecifics.

Microarray analyses

A *L. lineolaris* transcriptome was assembled using 606 381 Roche 454 sequence reads and >100 million 36–100 nt reads generated using the Illumina GA and HiSeq 2000 platforms. Sequencing libraries were constructed using mRNA purified from a pool of total RNA extracted from all life stages from eggs to adult. All total RNA extractions were carried out using TriZol reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions. All sequence reads were submitted to the National Center for Biotechnology Information SRA (short read archive) database SRA029359 under accession numbers SRX039411 and SRX041504. Transcriptome sequence assembly was performed using SeqMan NGEN4.0 (DNAStar, Inc., Madison, WI, USA) software and automated annotation was performed using BLAST2GO software (Conesa *et al*., 2005; Götz *et al*., 2008). A microarray containing 62976 probes (60 nt in length) was designed using Agilent eArray 4.0 platform (Agilent Technologies, Santa Clara, CA, USA). Each microarray contained a duplicate set of 20 860 probes designed from all annotated ESTs in the *L. lineolaris* transcriptome, 20 replicates of 186 probes designed to selected OBP transcripts, five replicates of 15 antisense probes designed from randomly selected transcripts, and 1319 Agilent quality control probes. Unassigned array features were filled using the probes designed from the ESTs and all probes were distributed randomly across the microarray.

Total RNA was extracted using TriZol reagent (Invitrogen) from legs, antennae and proboscis of 8-day old males and females (experimental) and the head, thorax and abdomen of females without any appendages (control). Two replicates were generated for each tissue set with tissues from 10 insects (i.e. one replicate) pooled in a single extraction. Messenger RNA was then purified from the total RNA samples using a Poly A(+)-tract mRNA isolation system (Promega, Madison, WI, USA). A 40-fold dilution of Spike-In RNA mixtures A and B (Agilent) containing various proportions of control RNA was mixed with 100 ng of mRNA from control and experimental samples, respectively. cDNA synthesis and labelled aRNA production was carried out using Agilent's aRNA production kit. Alexa Fluor-555 and Alexa Fluor-647 fluorescent dyes were used for labelling control and experimental aRNA, respectively. The efficiency of labelling was verified using a NanoDrop instrument to obtain absorbance values for nucleic acids as well as Alexa Fluor-555 and Alexa Fluor-647 dyes. Equal amounts of aRNA (~900 ng) from female body (control) were mixed with labelled aRNA from antennae, proboscis or legs from males or females. The control and experimental aRNA mixtures were fragmented at 65 °C using Agilent's RNA fragmentation reagent. At the end of the 30-min incubation, the RNA samples were mixed with an equal volume of 2X hybridization buffer and the arrays were hybridized at 65 °C for 18 h in an array hybridization incubator (Agilent). Duplicate hybridizations were performed for male and female antennae, proboscis and legs. Arrays were disassembled and washed using wash buffers 1 and 2 (Agilent) as recommended. Slides were scanned immediately using an array scanner (Agilent) and fluorescence data were extracted using the Feature Extraction Software V 9.3.4 (Agilent) with the standard quality control and normalization settings.

GENESPRING v 9.3 software (Agilent) was used to analyse the data. Replicate data were averaged. Intensity dependent (Lowess) normalization was used for each spot per chip, followed by per chip normalization to the 50th percentile. Features with raw

Table 5. Oligonucleotide primers used in LylinOBP cloning and sequencing

Primer	Sequence (5′-3′)
LylinOBP1-F1	ATGAGGATTTTGGTTTTG
LylinOBP2-F1	ATGGTGCTGAAAATGAGC
LylinOBP3-F1	ATGGCCGTCAACGCGAAAG
LylinOBP4-F1	ATGTCGATCAAAATCCATTT
LylinOBP5-F1	ATGACGACTAAACTCCGT
LylinOBP6-F1	ATGAAGTTCGTACTTTCAG
LylinOBP7-F1	AAGGCCAACGAGAAGAAAG
LylinOBP8-F1	ATGAATCCACTCATTCCCG
LylinOBP9-F1	ATGAATCCACTCATTCCC
LylinOBP10-F1	ATGACCCCTATTGTCGCTA
LylinOBP11-F1	ATGCACGCCGCATTCGTTC
LylinOBP12-F1	ATGGTTGCGGAATGCCCC
LylinOBP13-F1	ATGATGAAAATCGCATTCG
LylinOBP14-F1	ATGAAGATCGCATTCGTAG
LylinOBP15-F1	ATGATGAAAATCGCATTCG
LylinOBP16-F1	ATGAAGAGATTAGTGTTTG
LylinOBP17-F1	ATGAAGAGATTAGTGCTTG
LylinOBP18-F1	GTGAAGAGAATGAAGAAAAGAC
LylinOBP19-F1	ATGAAGTCCTTTGTAGGT
LylinOBP20-F1	ATGAAACTGGTGAAAGATG
LylinOBP21-F1	ATGGTTGGGAGAAAGTTGAC
LylinOBP22-F1	ATGCCTTCACAGCTGCCTTG
LylinOBP23-F1	ATGTATGCCTTCACAGCTG
LylinOBP24-F1	ATGGCTGATTCTGTGGGAG
LylinOBP25-F1	ATGTTCACCACCGCCACTTC
LylinOBP26-F1	ATGACTAAACGAAACATA
LylinOBP27-F1	GTGAAAACGGGAAGTCGTG
LylinOBP28-F1	ATGGAGGTTGCAGCTTGCC
LylinOBP29-F1	ATGAACCGTCCTCTTCTTT
LylinOBP30-F1	ATGAACAGCTTCACCGTC
LylinOBP31-F1	ATGAAGTTCGTACTTTCAGCC
LylinOBP32-F1	AGTGTGGCATTCCTCGGCAGTT
LylinOBP1-F2	CGGTGGAAGTTGCTGTTAAGTG
LylinOBP2-F2	CGTTCAGTCAAGACTTGGTTTCAC
LylinOBP3-F2	ATCAGTTCCATACAGAATCTCTCG
LylinOBP4-F2	CGGAAACCAGTGATACCAACGA
LylinOBP5-F2	CCGTTCGTTAGAGTCCATTTCAGAG
LylinOBP6-F2	GGACAAGTGTTAGCCCAGGA
LylinOBP6-F3	ACGGTATCAACGCAAGAGTACGC
LylinOBP7-F2	CAGAAGCAGAGCATTCCCG
LylinOBP7-F3	AATGTCAAGTGCATCCTCGC
LylinOBP8-F2	GCCAGTATCAACGCAGAGTACG
LylinOBP9-F2	TGTCCATGGTAGATCACTGGTGC
LylinOBP10-F2	GCTCAGTTCCCATCTCTACTGC
LylinOBP11-F2	TCGGAACACCAGTCGTAGGCTC
LylinOBP12-F2	GTACTCTTGGGTGTGTCAGTCAAC
LylinOBP13-F2	GGTAGTTTGTCCTCAGAACTTCG
LylinOBP14-F2	CTCAATAGGTAGTTTGWCCTCACAA
LylinOBP16-F2	CGGTYGTCGCYAAACAACTCAT
LylinOBP19-F2	GTCGTGTGAAGTGTACGATTATCAAC
LylinOBP20-F2	CATCGAAAGAAGTATCAGCTTTCCAG
LylinOBP21-F2	CTTTGCTGTCCGCCCTAACT
LylinOBP21-F3	TTATGGATGCTCACTGCTTACATG
LylinOBP21-F4	CATTTTTATGGATGCTCACTGCTTAC
LylinOBP23-F2	CAGTGTGTTGTTTAGTGCTCATTCA
LylinOBP24-F2	ATTGATGCAGTCTTCGACGAG
LylinOBP25-F2	GGGAACTAGCCGTAATCATTTGC
LylinOBP26-F2	CGGCAAATCACTCTCATCGTAG
LylinOBP27-F2	TGGCAGTCTCAGACGAGCGG
LylinOBP28-F2	GTTACGAAAGGCCGTCAGAAGG
LylinOBP29-F2	AATTCCCTGAGGTCAGAGTACGC
	GGTAATACGACTCACTATAGGGAGAAGAGGCGAGCACAGAA
T7-OligodT	TTAATACGACTTTTTTTTTTTTTTTTTTV
RACE-3′-Out	GCGAGCACAGAATTAATACGACT

expression values >50 were filtered and subjected to a one-way ANOVA test. Features with statistically significant differences were filtered using a parametric test, variances not assumed equal (Welch ANOVA) with a Benjamini and Hochberg false discovery rate *P* value cutoff set to 0.05 for multiple testing correction. Under the selection criteria, about 5.0% of the identified genes would be expected to pass the restriction by chance.

Bioinformatics analyses

The BLASTX similarity search algorithm (http://blast.ncbi.nlm.nih.gov/) was used to identify contig sequences with sequence relatedness to other OBP sequences. In general, *e* values <1.0 e^{-5} were considered to be significant. For comparative purposes, sequences were aligned using the L-INS-I strategy in MAFFT (Katoh *et al.*, 2005) in GENEIOUS 6.0.4. Phylogenetic analyses of the putative *L. lineolaris* OBP sequences were performed alone or in conjunction with other hemipteran OBP sequences identified from BLAST analyses. The phylogenetic evolutionary history was inferred using the neighbour-joining method implemented in MEGA 5 (Tamura *et al.*, 2011) with default settings and bootstrap support based on 1000 iterations. Identification of predicted signal peptide sequences was performed using the SIGNALP 4.0 server (http://www.cbs.dtu.dk/services/SignalP/) (Petersen *et al.*, 2011).

Cloning and re-sequencing of OBPs

OBP transcripts identified in high-throughput sequence assembly were used as reference sequences in template-driven assemblies to generate full-length transcripts. Roche-454 sequence reads (85% match) and Illumina short reads (93% match) were used in separate assemblies using the 'templated assembly' option of SEQMAN NGEN v4.0 software. This process extends the 5′- and 3′-ends of the reference sequences using the high-throughput sequence reads. Open reading frames and untranslated sequences at 5′- and 3′-ends were identified using Vector NTI advance v11.5 software. Oligonucleotide primers were designed to the 5′-ends of each OBP sequence identified in the high-throughput sequence assemblies. Equal amounts of total RNA extracted from 2- to 8-day old adult males and females were pooled and 5 μg of this RNA pool was used in the first-strand cDNA synthesis using Invitrogen first-strand cDNA synthesis kit (Invitrogen). The oligo d(T) primer provided in the kit was replaced with a custom designed anchored oligo d(T) primer containing an anchor sequence complementary to the universal primer used for 3′-end amplifications. All primer sequences are given in Table 5.

Quantitative real-time PCR analysis

Total RNA was isolated using TriZol (Invitrogen) from various tissues (antenna, proboscis, leg, head devoid of chemosensory appendages, body devoid of head and legs, gut, and fat body) harvested from 2-day-old adult male and female *L. lineolaris.* Total RNA was also isolated from two biological replicates. First-strand cDNA synthesis was performed using the first-strand cDNA synthesis reagent kit (Invitrogen). PCR primers were designed to the coding sequence of LylinOBP1, LylinOBP2, LylinOBP5, LylinOBP8, LylinOBP15, LylinOBP19, LylinOBP24, LylinOBP26, LylinOBP29, and *L. lineolaris* β-tubulin using the Primer3 (Rozen & Skaletsky, 2000) module implemented in GENEIOUS v5.6. Primers (Table 6) were designed to amplify a 100–150-bp fragment of each gene. Amplification of single discrete products from the cDNAs described above was confirmed by end-point PCR with Sapphire Amp Fast PCR Master Mix (Clonetech Laboratories Inc., Mountain View, CA, USA). PCR was performed using 0.4 μL template in a 15 μL reaction with thermocycler conditions consisting of: 95 °C for 2 min followed by 35 cycles at 94 °C for 20 s, 61 °C for 15 s, and 72 °C for 15 s. The transcript abundance of the respective LylinOBP transcripts was assessed quantitatively by qRT-PCR using standard curves prepared from serially diluted plasmid DNAs harbouring a fragment of the gene of interest. Plasmid DNAs were sequence verified and linearized with KpnI (New England Biolabs, Ipswich, MA, USA). The linearized DNA was gel purified and DNA concentration determined in duplicate on a Synergy H4 (BioTek Instruments Inc., Winooski, VT, USA). qRT-PCR was performed on a Bio-Rad CFX96 real-time cycler (Bio-Rad Laboratories Inc, Hercules, CA, USA) in a 10-μL volume using a Kapa SYBR Fast qPCR kit (Kapa Biosystems, Woburn, MA, USA) with a 1:10 dilution of cDNA templates. PCR conditions were optimized for each gene set. Transcript abundance was determined from standard curves using the Bio-Rad CFX manager software and an online calculator for determining template copy number (http://cels.uri.edu/gsc/cndna.html), which uses the equation $(6.022 \times 10^{23} \times \text{DNA amount ng})/(\text{plasmid size} \times 10^{9} \times \text{average nt weight} \times 2)$ (Whelan *et al.*, 2003). Transcript abundance ± SEM was determined based on triplicate values from two biological replicates with statistical analysis (ANOVA with Tukey's multiple comparison test, $P < 0.05$) performed using GraphPad Prism v6 (GraphPad Software, La Jolla, CA, USA).

Table 6. Oligonucleotide primers used in quantitative real-time PCR

Primer	Sequence (5′-3′)
LylinOBP1rt F	TGGTTTTGTTCACTGCGGCACTT
LylinOBP1rt R	CCGGTTTCTTCTACGCAGCCGT
LylinOBP2rt F	TGGCATGATGCAAGCTGTGAAAAAC
LylinOBP2rt R	TCTCCAGTGACGGTGGCGCA
LylinOBP5rt F	ACCGACGATGAGGTGGAGGAGT
LylinOBP5rt R	ACAGGGTCGAAGTGGCCGTCG
LylinOBP8rt F	CGCTGCTGCGACCAGGGGAG
LylinOBP8rt R	TGGCGTTGTGGGAGTCTGGA
LylinOBP15rt F	TGCGCAAAGCAAGTCGGGGA
LylinOBP15rt R	GGGTGGCTGGAGTCCCGTCG
LylinOBP19rt F	TCGCCGTGGCTCTCGTTGAGT
LylinOBP19rt R	TGCTCTTTGGCGCACTTGTCCT
LylinOBP24rt F	ACCCATGGCCGATGTGGGCA
LylinOBP24rt R	TGTCCATGACGATGCAGGGCAGT
LylinOBP26rt F	AGGCCTAAGGAGGAGCAAGAAGC
LylinOBP26rt R	TGTCCTTCCGATGCAGGGGTTCG
LylinOBP29rt F	AACTGCCGTTCTGGCCGTGGG
LylinOBP29rt R	TCGCAACATTGGGGCGGCCT
Lylin βtubulin rt F	CGGCATCCTCGTCGGCCGTG
Lylinβtubulin rt R	TGCATTGGTACACTGGCGAGGGC

Acknowledgements

We thank Charles Lanford (USDA-ARS SIMRU) for assistance with insect rearing and Calvin A Pierce III (USDA-ARS SIMRU) for RNA extractions and cDNA synthesis.

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. USDA is an equal opportunity provider and employer.

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.

Bendtsen, J.D., Nielsen, H., von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**: 783–795.

Biessmann, H., Nguyen, Q.K., Le, D. and Walter, M.F. (2005) Microarray-based survey of a subset of putative olfactory genes in the mosquito *Anopheles gambiae*. *Insect Mol Biol* **14**: 575–589.

Biessmann, H., Andronopoulou, E., Biessmann, M.R., Douris, V., Dimitratos, S.D., Eliopoulos, E. *et al*. (2010) The *Anopheles gambiae* odorant binding protein 1 (AgamOBP1) mediates indole recognition in the antennae of female mosquitoes. *PLoS ONE* **5**: e9471.

Blackmer, J., Rodriguez-Saona, C., Byers, J., Shope, K. and Smith, J. (2004) Behavioral response of *Lygus hesperus* to conspecifics and headspace volatiles of alfalfa in a Y-tube olfactometer. *J Chem Ecol* **30**: 1547–1564.

Chinta, S., Dickens, J. and Aldrich, J. (1994) Olfactory reception of potential pheromones and plant odors by tarnished plant bug, *Lygus lineolaris* (Hemiptera: Miridae). *J Chem Ecol* **20**: 3251–3267.

Chinta, S., Dickens, J. and Baker, G. (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.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M. and Robles, M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674–3676.

Contreras, E., Rausell, C. and Real, M.D. (2013) Proteome response of *Tribolium castaneum* larvae to *Bacillus thuringiensis* toxin producing strains. *PLoS ONE* **8**: e55330.

Crampton, J.M., James, A.A., Thymlanou, S., Mavroidis, M., Kokolakis, G., Komitopoulou, K. *et al*. (1998) Cloning and characterization of a cDNA encoding a male-specific serum protein of the Mediterranean fruit fly, *Ceratitis capitata*, with sequence similarity to odourant-binding proteins. *Insect Mol Biol* **7**: 345–353.

Crooks, G.E., Hon, G., Chandonia, J.-M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res* **14**: 1188–1190.

Daines, B., Wang, H., Wang, L., Li, Y., Han, Y., Emmert, D. *et al*. (2011) The *Drosophila melanogaster* transcriptome by paired-end RNA sequencing. *Genome Res* **21**: 315–324.

Debolt, J.W. (1982) Meridic diet for rearing successive generations of *Lygus hesperus*. *Ann Entomol Soc Am* **75**: 119–122.

Del Campo, M., Palmer, S. and Caillaud, M. (2011) Characterization of a new odorant binding protein gene in gustatory organs of *Manduca sexta* larvae (Lepidoptera: Sphingidae). *Ann Entomol Soc Am* **104**: 319–325.

Dickens, J., Callahan, F., Wergin, W. and Erbe, E. (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 and immunolocalization of a putative odorant-binding protein in true bugs (Hemiptera, Heteroptera). *J Exp Biol* **201**: 33–41.

Forêt, 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., Breer, H. and Krieger, J. (2009) A receptor and binding protein interplay in the detection of a distinct pheromone component in the silkmoth *Antheraea polyphemus*. *Int J Biol Sci* **7**: 745–757.

Frati, F., Salerno, G., Conti, E. and Bin, F. (2008) Role of the plant–conspecific complex in host location and intra-specific communication of *Lygus rugulipennis*. *Physiol Entomol* **33**: 129–137.

Frati, F., Chamberlain, K., Birkett, M., Dufour, S., Mayon, P., Woodcock, C. *et al*. (2009) *Vicia faba-Lygus rugulipennis* interactions: induced plant volatiles and sex pheromone enhancement. *J Chem Ecol* **35**: 201–208.

Gong, D.-P., Zhang, H.-J., Zhao, P., Xia, Q.-Y. and Xiang, Z.-H. (2009) The odorant binding protein gene family from the genome of silkworm, *Bombyx mori*. *BMC Genomics* **10**: 332.

Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J. *et al*. (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* **36**: 3420–3435.

Gotzek, D., Robertson, H.M., Wurm, Y. and Shoemaker, D. (2011) Odorant binding proteins of the red imported fire ant, *Solenopsis invicta*: an example of the problems facing the analysis of widely divergent proteins. *PLoS ONE* **6**: e16289.

Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L. *et al*. (2011) The developmental transcriptome of *Drosophila melanogaster*. *Nature* **471**: 473–479.

Gu, S., Sun, Y., Ren, L., Zhang, X., Zhang, Y., Wu, K. *et al*. (2010) Cloning, expression and binding specificity analysis of odorant binding protein 3 of the lucerne plant bug, *Adelphocoris lineolatus* (Goeze). *Chin Sci Bull* **55**: 3911–3921.

Gu, S.-H., Wang, S.-P., Zhang, X.-Y., Wu, K.-M., Guo, Y.-Y., Zhou, J.-J. *et al*. (2011a) Identification and tissue distribution of odorant binding protein genes in the lucerne plant bug *Adelphocoris lineolatus* (Goeze). *Insect Biochem Mol Biol* **41**: 254–263.

Gu, S.-H., Wang, W.-X., Wang, G.-R., Zhang, X.-Y., Guo, Y.-Y., Zhang, Z. *et al*. (2011b) Functional characterization and immunolocalization of odorant binding protein 1 in the lucerne plant bug, *Adelphocoris lineolatus* (Goeze). *Arch Insect Biochem Physiol* **77**: 81–99.

Hartmann, B. and Valcárcel, J. (2009) Decrypting the genome's alternative messages. *Curr Opin Cell Biol* **21**: 377–386.

He, P., Zhang, J., Liu, N., Zhang, Y. and Yang, K. (2011) Distinct expression profiles and different functions of odorant binding proteins in *Nilaparvata lugens* Stål. *PLoS ONE* **6**: e28291.

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.

Ho, H. and Millar, J. (2002) Identification, electroantennogram screening, and field bioassays of volatile chemicals from *Lygus hesperus* Knight (Heteroptera: Miridae). *Zool Stud* **41**: 311–320.

Hua, J.-F., Zhang, S., Cui, J.-J., Wang, D.-J., Wang, C.-Y., Luo, J.-Y. *et al*. (2012) Identification and binding characterization of three odorant binding proteins and one chemosensory protein from *Apolygus lucorum* (Meyer-Dur). *J Chem Ecol* **38**: 1163–1170.

Hull, J.J., Hoffmann, E.J., Perera, O.P. and Snodgrass, G.L. (2012) Identification of the western tarnished plant bug (*Lygus hesperus*) olfactory co-receptor Orco: expression profile and confirmation of atypical membrane topology. *Arch Insect Biochem Physiol* **81**: 179–198.

Innocenzi, P.J., Hall, D.R., Cross, J.V., Masuh, H., Phythian, S.J., Chittamaru, S. *et al*. (2004) Investigation of long-range female sex pheromone of the European tarnished plant bug, *Lygus rugulipennis*: chemical, electrophysiological, and field studies. *J Chem Ecol* **30**: 1509–1529.

Innocenzi, P.J., Hall, D., Cross, J.V. and Hesketh, H. (2005) Attraction of male European tarnished plant bug, *Lygus rugulipennis* to components of the female sex pheromone in the field. *J Chem Ecol* **31**: 1401–1413.

Jordan, M.D., Stanley, D., Marshall, S.D.G., De Silva, D., Crowhurst, R.N., Gleave, A.P. *et al*. (2008) Expressed sequence tags and proteomics of antennae from the tortricid moth, *Epiphyas postvittana*. *Insect Mol Biol* **17**: 361–373.

Katoh, K., Kuma, K.-I., Toh, H. and Miyata, T. (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* **33**: 511–518.

Kirkness, E.F., Haas, B.J., Sun, W., Braig, H.R., Perotti, M.A., Clark, J.M. *et al*. (2010) Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proc Natl Acad Sci U S A* **107**: 12168–12173.

Kwon, H.-W., Lu, T., Rützler, M. and Zwiebel, L.J. (2006) Olfactory responses in a gustatory organ of the malaria vector mosquito *Anopheles gambiae*. *Proc Natl Acad Sci U S A* **103**: 13526–13531.

Lacazette, E., Gachon, A.M. and Pitiot, G. (2000) A novel human odorant-binding protein gene family resulting from genomic duplicons at 9q34: differential expression in the oral and genital spheres. *Hum Mol Genet* **9**: 289–301.

Lagarde, A., Spinelli, S., Qiao, H., Tegoni, M., Pelosi, P. and Cambillau, C. (2011) Crystal structure of a novel type of odorant binding protein from *Anopheles gambiae*, belonging to the C-plus class. *Biochem J* **437**: 423–430.

Laughlin, J.D., Ha, T.S., Jones, D.N.M. and Smith, D.P. (2008) Activation of pheromone-sensitive neurons is mediated by conformational activation of pheromone-binding protein. *Cell* **133**: 1255–1265.

Leal, W. (2005) Pheromone reception. In *The Chemistry of Pheromones and Other Semiochemicals II* (Schulz, S., ed.), pp. 341–360. Springer, Berlin.

Li, Z.-X., Pickett, J.A., Field, L.M. and Zhou, J.-J. (2005) Identification and expression of odorant-binding proteins of the malaria-carrying mosquitoes *Anopheles gambiae* and *Anopheles arabiensis*. *Arch Insect Biochem Physiol* **58**: 175–189.

Licatalosi, D.D. and Darnell, R.B. (2010) RNA processing and its regulation: global insights into biological networks. *Nat Rev Genet* **11**: 75–87.

Liu, R., Lehane, S., He, X., Lehane, M., Hertz-Fowler, C., Berriman, M. *et al*. (2009) Characterisations of odorant-binding proteins in the tsetse fly *Glossina morsitans morsitans*. *Cell Mol Life Sci* **67**: 919–929.

Manoharan, M., Chong, M.N.F., Vaïtinadapoulé, A., Frumence, E., Sowdhamini, R. and Offmann, B. (2013) Comparative genomics of odorant binding proteins in *Anopheles gambiae*, *Aedes aegypti*, and *Culex quinquefasciatus*. *Genome Biol Evol* **5**: 163–180.

Martin, J.P., Beyerlein, A., Dacks, A.M., Reisenman, C.E., Riffell, J.A., Lei, H. *et al*. (2011) The neurobiology of insect olfaction: sensory processing in a comparative context. *Prog Neurobiol* **95**: 427–447.

Matsuo, T., Sugaya, S., Yasukawa, J., Aigaki, T. and Fuyama, Y. (2007) Odorant-binding proteins OBP57d and OBP57e affect taste perception and host-plant preference in *Drosophila sechellia*. *PLoS Biol* **5**: e119.

Mitaka, H., Matsuo, T., Miura, N. and Ishikawa, Y. (2011) Identification of odorant-binding protein genes from antennal expressed sequence tags of the onion fly, *Delia antiqua*. *Mol Biol Rep* **38**: 1787–1792.

Nilsen, T.W. and Graveley, B.R. (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**: 457–463.

Noda, H., Kawai, S., Koizumi, Y., Matsui, K., Zhang, Q., Furukawa, S. *et al*. (2008) Annotated ESTs from various tissues of the brown planthopper *Nilaparvata lugens*: a genomic resource for studying agricultural pests. *BMC Genomics* **9**: 117.

Paesen, G. and Happ, G. (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.

Paiva-Silva, G., Sorgine, M., Benedetti, C., Meneghini, R., Almedia, I., Machado, E. *et al*. (2002) On the biosynthesis of *Rhodnius prolixus* heme-binding protein. *Insect Biochem Mol Biol* **32**: 1533–1541.

Patana, R. (1982) Disposable diet packet for feeding and oviposition of *Lygus hesperus* (Hemiptera: Miridae). *J Econ Entomol* **75**: 668–669.

Pelletier, J. and Leal, W.S. (2009) Genome analysis and expression patterns of odorant-binding proteins from the southern house mosquito *Culex pipiens quinquefasciatus*. *PLoS ONE* **4**: e6237.

Pelletier, J. and Leal, W.S. (2011) Characterization of olfactory genes in the antennae of the southern house mosquito, *Culex quinquefasciatus*. *J Insect Physiol* **57**: 915–929.

Pelletier, J., Hughes, D.T., Luetje, C.W. and Leal, W.S. (2010) An odorant receptor from the southern house mosquito *Culex pipiens quinquefasciatus* sensitive to oviposition attractants. *PLoS ONE* **5**: e10090.

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.

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

Ribeiro, J.M.C., Andersen, J., Silva-Neto, M.A.C., Pham, V.M., Garfield, M.K. and Valenzuela, J.G. (2004) Exploring the sialome of the blood-sucking bug *Rhodnius prolixus*. *Insect Biochem Mol Biol* **34**: 61–74.

Rozen, S. and Skaletsky, H. (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**: 365–386.

Sachse, S. and Krieger, J. (2011) Olfaction in insects. *E-Neuroforum* **2**: 49–60.

Sánchez-Pla, A., Reverter, F., Ruíz De Villa, M.C. and Comabella, M. (2012) Transcriptomics: mRNA and alternative splicing. *J Neuroimmunol* **248**: 23–31.

Shanbhag, S.R., Park, S.-K., Pikielny, C.W. and Steinbrecht, R.A. (2001) Gustatory organs of *Drosophila melanogaster*: fine structure and expression of the putative odorant-binding protein PBPRP2. *Cell Tissue Res* **304**: 423–437.

Sim, S., Ramirez, J.L. and Dimopoulos, G. (2012) Dengue virus infection of the *Aedes aegypti* salivary gland and chemosensory apparatus induces genes that modulate infection and blood-feeding behavior. *PLoS Pathog* **8**: e1002631.

Snodgrass, G.L. (1996) Insecticide resistance in field populations of the tarnished plant bug (Heteroptera: Miridae) in cotton in the Mississippi Delta. *J Econ Entomol* **89**: 783–790.

Snodgrass, G.L. and Scott, W.P. (2000) Seasonal changes in pyrethroid resistance in tarnished plant bug (Heteroptera: Miridae) populations during a three-year period in the delta area of Arkansas, Louisana, and Mississippi. *J Econ Entomol* **93**: 441–446.

Snodgrass, G.L. and Scott, W.P. (2002) Tolerance to acephate in tarnished plant bug (Heteroptera: Miridae) populations in the Mississippi river delta. *Southwest Entomol* **27**: 191–199.

Snodgrass, G.L., Gore, J., Abel, C. and Jackson, R. (2009) Acephate resistance in populations of the tarnished plant bug (Heteroptera: Miridae) from the Mississippi River Delta. *J Econ Entomol* **102**: 699–707.

Steinbrecht, R. (1998) Odorant-binding proteins: expression and function. *Ann N Y Acad Sci* **855**: 323–332.

Sun, Y.-L., Huang, L.-Q., Pelosi, P. and Wang, C.-Z. (2012) Expression in antennae and reproductive organs suggests a dual role of an odorant-binding protein in two sibling *Helicoverpa* species. *PLoS ONE* **7**: e30040.

Swarup, S., Williams, T.I. and Anholt, R.R.H. (2011) Functional dissection of odorant binding protein genes in *Drosophila melanogaster*. *Genes Brain Behav* **10**: 648–657.

Syed, Z., Ishida, Y., Taylor, K., Kimbrell, D.A. and Leal, W.S. (2006) Pheromone reception in fruit flies expressing a moth's odorant receptor. *Proc Natl Acad Sci U S A* **103**: 16538–16543.

Sze, S.-H., Dunham, J.P., Carey, B., Chang, P.L., Li, F., Edman, R.M. *et al*. (2012) A de novo transcriptome assembly of *Lucilia sericata* (Diptera: Calliphoridae) with predicted alternative splices, single nucleotide polymorphisms and transcript expression estimates. *Insect Mol Biol* **21**: 205–221.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.

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

Venables, J. P., Tazi, J. and Juge, F. (2012) Regulated functional alternative splicing in *Drosophila*. *Nucleic Acids Res* **40**: 1–10.

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.

Vogel, H., Heidel, A.J., Heckel, D.G. and Groot, A.T. (2010) Transcriptome analysis of the sex pheromone gland of the noctuid moth *Heliothis virescens*. *BMC Genomics* **11**: 29.

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.

Whelan, J.A., Russel, N.B. and Whelan, M.A. (2003) A method for the absolute quantification of cDNA using real-time PCR. *J Immunol Methods* **278**: 261–269.

Williams, L., Blackmer, J.L., Rodriguez-Saona, C. and Zhu, S. (2010) Plant volatiles influence electrophysiological and behavioral responses of *Lygus hesperus*. *J Chem Ecol* **36**: 467–478.

Xu, P., Atkinson, R., Jones, D. and Smith, D. (2005) Drosophila OBP LUSH is required for activity of pheromone-sensitive neurons. *Neuron* **45**: 193–200.

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.

Xu, W., Cornel, A.J. and Leal, W.S. (2010) Odorant-binding proteins of the malaria mosquito *Anopheles funestus sensu stricto*. *PLoS ONE* **5**: e15403.

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.

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

Zacharuk, R.Y. (1980) Ultrastructure and function of insect chemosensilla. *Annu Rev Entomol* **25**: 27–47.

Zhou, J.-J. (2010) Odorant-binding proteins in insects. *Vitam Horm* **83**: 241–272.

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., He, X.-L., Pickett, J.A. and Field, L.M. (2008) Identification of odorant-binding proteins of the yellow fever mosquito *Aedes aegypti*: genome annotation and comparative analyses. *Insect Mol Biol* **17**: 147–163.

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.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Comparison of predicted *Lygus lineolaris* OBP sequences based on a percent identity matrix. The matrix, which includes partial sequences, was generated from a MAAFT alignment and indicates the percent identity across the LylinOBP protein sequences. Cell shading denotes the degree of identity with black indicating the highest percent identity and white the lowest.